Statistically Principled Deep Learning for SAR Image Segmentation

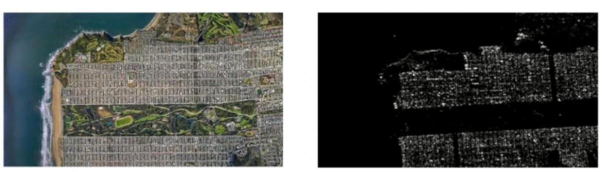
Cassandra Goldberg
Bowdoin College, Brunswick, Maine, USA
cgoldber@bowdoin.edu

Abstract

This paper proposes a novel approach for Synthetic Aperture Radar (SAR) image segmentation by incorporating known statistical properties of SAR into deep learning models. We generate synthetic data using the Generalized Gamma distribution, modify the U-Net architecture to encompass statistical moments, and employ stochastic distance losses for improved segmentation performance. Evaluation against traditional methods will reveal the potential of this approach to advance SAR image analysis, with broader applications in environmental monitoring and general image segmentation tasks.

Background

Remote sensing offers a comprehensive perspective of the Earth's surface from distant vantage points, serving as a valuable tool for environmental monitoring. Synthetic Aperture Radar (SAR), one form of remote sensing, operates by emitting wave signals from an airborne platform and capturing reflections to produce high-resolution images. Since this technique functions regardless of weather conditions and sunlight, it delivers extremely valuable remote sensing insights. An example of SAR data is displayed in Figure 1.



Google Maps Screenshot SAR Image

Figure 1: Google Maps aerial view of San Francisco and SAR data depiction (SAR data from Rodrigues et. al.).

Previous SAR image analysis approaches have typically employed either statistical models or deep learning techniques in isolation. Though deep learning models outperform statistical methods, they present a challenge due to their high demand for extensive training data, which is frequently scarce in SAR applications. [1]

SAR data is often characterized by the Generalized Gamma (G^0) distribution (Frery et. al. 2022). This distribution has gained prominence due to its simplicity in sampling and its ability to generate realistic synthetic SAR imagery (Li, Fan, Farias, Jeova 2023). Furthermore, SAR image data has exhibited a notable capacity for efficient processing through direct application of computational methods to its moments. Prior research indicates that neural network models can effectively exploit these moment-based features, particularly for tasks like parameter estimation (Li, Fan, Farias, Jeova 2023). For segmentation, the most commonly used deep learning architecture is the U-Net (Ronneberger et. al. 2015), which has also been broadly applied to other computer vision tasks on SAR data (Zhu et. al. 2021). In addition, stochastic distances present a robust mathematical framework for comparing samples derived from probability distributions (Nascimento et. al., 2009). Within the SAR context, these distances have found applications in parameter estimation, segmentation, and change detection.

Methodology

This project integrates statistically-sound input data, loss functions, and deep learning model architectures with the collective goal of enhancing SAR image segmentation capabilities. The intersection over union (IOU) metric will be used to evaluate whether the proposed changes outperform typical segmentation techniques.

Input Data

To address the scarcity of labeled SAR images, high-quality synthetic SAR data will be generated. Perlin noise will be used to create ground truth segmentation maps that resemble the terrains depicted in SAR data. Subsequently, pixel intensities will be sampled from the G^0 distribution with distinct parameters to depict different regions, concurrently imparting the speckle noise effect.

Loss Function

For each region in the input SAR images and output segmentation maps, probability distributions of the pixel intensities will be computed. Stochastic distances will then be calculated to assess disparities between class probability distributions. For the supervised approach, the stochastic differences between known classes will be added to the typical mean squared error loss. For the unsupervised approach, it will be assumed that distinct object regions exhibit maximized distribution divergence. Accordingly, the output segmentation can be determined by a network trained to maximize such distances.

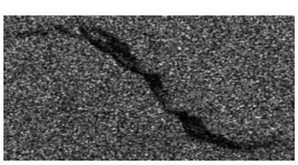
Model Architecture

The typical convolutional layers that compose a U-Net will be replaced with a series of 1x1 convolutional and average pooling layers. The average pooling layer can help enhance local pattern recognition while it computes statistical moments from the data, and the 1x1 convolutions are employed to transform those moments into meaningful segmentations. Further refinements may involve the implementation of max pooling layers with various kernel sizes to improve edge and pattern detection.

Synthetic Data Generation Results

Two hundred segmentation maps were generated using Perlin noise. For each segmentation map, corresponding synthetic data images were then produced by sampling from the G^0 distribution with all permutations of 11 alpha values ranging from -1.5 to -11. The synthetic dataset exhibits a diverse range of visual interpretability, with increased difficulty when alpha values are close together.

A standard U-Net was trained and validated on the synthetic dataset, achieving an average validation IOU of 82.2% and an average inference time of 1.34 ms. This attests to the dataset's efficacy in training a model for segmentation tasks, while also significantly outpacing traditional statistical models. The trained model was then tested on real SAR data, such as the example displayed in Figure 2.



SAR Data

Predicted Segmentation Map

Figure 2: The U-Net trained on synthetic data was tested on a SAR image of an oil spill, producing the corresponding segmentation map (SAR data from Rodrigues et. al.).

The U-Net trained on synthetic data was able to discern the regions of real SAR data in a general sense. However, the granularity of the segmentation map in Figure 2 suggests the need for further improvements, such as incorporating stochastic distance losses to preserve region unity.

Conclusion

This project holds the potential to advance the analysis of SAR images through image segmentation. It introduces deep learning techniques rooted in statistical foundations for neural network training, even in scenarios where abundant data or labeled samples are lacking.

From a broader perspective, this project could significantly impact numerous environmental monitoring endeavors, such as tracking deforestation, oil spills, or carbon-dense regions. Moreover, the exploration of whether decomposition with 1x1 convolutions and average pooling can rival classic convolutional layers has implications for the wider field of AI. The efficiency granted through this architecture, with significantly fewer weights to train and the capacity to run on the GPU, could be applied to address a wide range of computer vision challenges.

Acknowledgements

Thank you to my project advisor, Professor Jeová Farias.

References

Fan, Li, Farias Sales Rocha Neto, Jeova. "Using Neural Networks for Fast SAR Roughness Estimation of High Resolution Images." arXiv preprint arXiv:2309.03351 (2023).

Frery, Alejandro C., Jie Wu, and Luis Gomez. *SAR Image Analysis-A Computational Statistics Approach: With R Code, Data, and Applications*. John Wiley & Sons, 2022.

Nascimento, Abraao DC, Renato J. Cintra, and Alejandro C. Frery. "Hypothesis testing in speckled data with stochastic distances." *IEEE Transactions on geoscience and remote sensing* 48.1 (2009): 373-385.

Rodrigues, FA Ávila, et al. "SAR image segmentation using the roughness information." *IEEE Geoscience and Remote Sensing Letters* 13.2 (2016): 132-136.

Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." *Medical Image Computing and Computer-Assisted Intervention–MICCAI* (2015).

Zhu, Xiao Xiang, et al. "Deep learning meets SAR: Concepts, models, pitfalls, and perspectives." *IEEE Geoscience and Remote Sensing Magazine 9.4* (2021): 143-172.