Lightweight Deep Learning Model for SAR Image Recognition

S.Nagarajan[1], Dr.S.K.Manju Bargavi[2]
[1]Lincoln University College, 47301, Petaling Jaya, Selangor Darul Ehsan, Malaysia.
[2]Professor, School of Computer Science and IT, Jain Deemed to be University, Bangalore.
pdf.Nagarajanselvam@lincoln.edu.my, b.manju@jainuniversity.ac.in

Abstract: Synthetic Aperture Radar (SAR) image classification is very important in modern remote sensing applications, such as military surveillance, disaster management, and autonomous vehicle guidance. While traditional deep learning models are accurate, they are generally computationally expensive and have many parameters, making them less applicable in resource-constrained environments. In this paper, we propose a new light-weight architecture based on capsule optimization for SAR target recognition on the MSTAR benchmark dataset, named CPLO-LWRCapsNet. The proposed pipeline involves guided filtering for speckle noise suppression, Discrete Hartley Transform (DHT) for feature enhancement in the frequency domain, KAN-Driven Multiscale Synergy Network (KDMSNet) for accurate target segmentation, and Relative Directional Edge Binary Pattern (RDEBP) descriptors for robust texture feature extraction. The architecture hyperparameters are tuned by the Capsule-based Pareto-Level Optimization (CPLO) algorithm. The experiment results show that the test accuracy is 99.72% with only 210,357 trainable parameters and 172.6 million MACs. This exemplifies the capacity of the model to discover a suitable trade-off between high classification performance and computational efficiency.

Keywords: Synthetic Aperture Radar (SAR); Lightweight Capsule Network; Capsule-based Pareto-Level Optimization (CPLO); KAN-Driven Multiscale Synergy Network (KDMSNet); Relative Directional Edge Binary Pattern (RDEBP).

1. Introduction

Synthetic Aperture Radar (SAR) is a powerful active microwave remote sensing system that can obtain high-resolution images under all weather conditions and any time of the day or night. Unlike optical sensors, SAR emits microwave pulses and measures the amplitude and phase of the backscattered energy to generate two-dimensional images of the scene. Such properties make SAR indispensable for many applications such as terrain mapping, ocean surveillance, urban planning, mineral exploration, disaster assessment and Automatic Target Recognition (ATR) [7].

However, ATR from SAR imagery is still a challenging task due to several image-domain factors such as: the speckle noise inherent in coherent imaging, large target aspect-angle sensitivity, changing depression-angle effects, and the complex scattering mechanisms that can give rise to non-intuitive image signatures [2]. Classical feature-extraction methods relying on hand-crafted descriptors are often incapable of generalizing over these variations, hence motivating the use of data-driven deep learning techniques [10].

Recently, convolutional neural networks (CNNs), attention mechanisms, graph neural networks and capsule networks have been rapidly incorporated into SAR-ATR pipelines. Despite their impressive accuracy gains, a common shortcoming of these methods is their reliance on a large parameter set and a high floating-point operation count, making deployment on embedded or edge platforms impractical. Lightweight design is still an open research problem.

This paper proposes a holistic SAR-ATR framework-CPLO-LWRCapsNet to address the dual goals of accuracy and efficiency. The main contributions are: (i) A Guided Filter preprocessing step, which quantifiably improves Signal-to-Noise Ratio (SNR) and Equivalent Number of Looks (ENL) before feature extraction; (ii) DHT-based frequency-domain transformation that improves information entropy by 143.3%; (iii) KDMSNet segmentation to obtain stable and high-retention foreground isolation; (iv) RDEBP descriptors to capture directional texture and edge patterns; and (v) A CPLO algorithm to efficiently search the architecture search space to generate a minimal and high-performing capsule network[6].

2. Related Work and Background

During the past 20 years, the classification of SAR images has been studied extensively[8]. Early methods relied on statistical models of clutter and physics-based scattering analysis. Machine learning has its origins in support vector machines and sparse representation methods. These methods had better generalization than hand-crafted features but needed careful feature engineering.

In the deep learning era, CNNs have played an important role in SAR-ATR. Based on optical image recognition networks like VGG, ResNet, etc., these networks have achieved significant improvements in accuracy on MSTAR, but with millions of parameters. Examples are lightweight architectures like MobileNet and ShuffleNet[9], but domain-specific issues like speckle and aspect-angle change, related to SAR images, needed domain-specific modifications.

Hybrid approaches based on the combination of convolutional and attention modules were proposed to focus the feature extraction on the discriminative regions of the target [1]. Graph neural networks and position-aware GNNs show promise in the low-data regime [3], and few-shot SAR recognition has also been studied[11].

To address the drawbacks of CNNs, Sabour et al. [12] proposed another model, capsule networks, which encode the spatial relations among the features via dynamic routing. The capacity of these models to conserve part-whole relations renders them naturally fit for target detection tasks where viewpoint-invariance is a major requirement. Standard capsule networks are computationally expensive. This indicates the necessity of lightweight design of this work.

Different filtering methods have been discussed in the pre-processing section for speckle noise reduction. Guided filtering is particularly attractive due to its ability to suppress noise and preserve structural edges, and it is a suitable pre-processing step for deep feature extraction in SAR imagery [5]. The DHT feature transformation provides a complementary way to enhance the discriminability via entropy amplification in frequency domain.

3. Dataset

All experiments in this paper are conducted on the Moving and Stationary Target Acquisition and Recognition (MSTAR) dataset, a standard benchmark for SAR-ATR research collected using an X-band SAR sensor, consisting of grayscale images with a pixel resolution of 128 × 128; eight target classes are considered, including the 2S1 self-propelled artillery vehicle, D7 bulldozer, T62 main battle tank, ZIL131 military transport truck, BRDM_2 armored reconnaissance vehicle, BTR_60 armored personnel carrier, SLICY calibration target, and ZSU_23_4 self-propelled anti-aircraft gun. The dataset is partitioned into training, validation, and test subsets following an approximately 70/15/15 split, comprising 6,626 training samples, 1,420 validation samples, and 1,420 test samples. All images are standard MSTAR complex-valued SAR chips, converted to magnitude images prior to processing, and no external data augmentation is used; instead, the robustness of the model stems from the pre-processing and architectural choices described in Section 4.

4. Proposed Methodology

The CPLO-LWRCapsNet framework is a five-stage sequential process that includes (1) speckle reduction by Guided Filtering, (2) frequency domain transform with DHT, (3) target segmentation with KDMSNet, (4) descriptor generation with RDEBP, and (5) classification with CPLO-optimized Lightweight Capsule Network.

4.1 Speckle Reduction Using the Guided Filter

Several widely used measures were used to quantify the effectiveness of the Guided Filter. The Equivalent Number of Looks (ENL) indicating smoothness of the image was significantly enhanced (from 2.68 to 4.27, 59.2% increase) in the filtered image as compared to the original image, indicating good speckle suppression. Likewise a Signal to Noise Ratio (SNR) increase from 4.28 dB to 6.30 dB, an increase of 47.2% was achieved. The Structural Similarity Index (SSIM) and the Edge Preservation Index (EPI) of the filtered image were also calculated and found to have values of 0.812 and 0.333 respectively, which indicate good structural preservation and moderate edge preservation.These results demonstrate that the Guided Filter effectively suppresses speckle noise while preserving image structure, as reflected in the high SSIM value (0.812) and the 59.2% improvement in ENL.

4.2 Discrete Hartley Transform (DHT)

Denoised and filtered image is transformed from spatial domain to frequency domain using Discrete Hartley Transform (DHT). The DHT works entirely in real domain, thus it is less computational when compared to Discrete Fourier Transform and still contains the same information about a spectrum. The results demonstrate that the DHT transformation is able to increase the spectral entropy of the transformed coefficients by 143.3% (from 5.7531 to 13.9987), which results in a richer and more information dense feature representation that is helpful for subsequent feature extraction. The transformed feature values are additionally normalized into two parallel ranges, [0, 1] and [0, 255], with the latter being more suitable for deep learning model inputs and helping to stabilize gradient flow during training.

4.3 Segmentation of KDMSNet

The image processed by DHT is then passed on to a KAN-Driven Multiscale Synergy Network (KDMSNet) for target area separation from the background noise using multifaceted layers of Kolmogorov-Arnold Network (KAN) and multi-scale feature fusion. The segmentation results show consistently strong performance across all metrics, with a foreground retention rate of 99.25%, mean mask activation of 0.5033, and a narrow activation range (0.4989–0.5106) indicating minimal background leakage and stable, consistent segmentation performance.

4.4 RDEBP Descriptor Generation

The features of texture and edge are extracted with the Relative Directional Edge Binary Pattern (RDEBP) descriptor, which encodes relative directional difference between neighboring pixel pairs and is sensitive to texture and edge information, and is robust to illumination change, represented in binary codes. The statistical characteristics of the descriptors extracted show that the directional texture patterns are balanced, have mean value of 0.4896 and standard deviation of 0.2481, indicating that they are not too strong or too weak throughout the image and that they contain both weak and strong structure.

4.5 CPLO-Optimized Lightweight Capsule Network (LWRCapsNet)

The core is a Lightweight Residual CapsNet developed by tuning hyperparameters with a multi-objective optimization algorithm, which balances between the model size (in terms of number of parameters), the computational cost (in terms of number of MACs) and the accuracy of the model validation. This search ended after approximately 2.95 hours, resulting in a backbone channel width of 45, primary capsule of dimension 4 and 8 primary capsules, class capsule of dimension 8, with a learning rate of 0.005042, a best fitness value of 0.0077, best validation accuracy of 99.80%, and 210,357 parameters, 172,602,688 MACs, and 5 iterations of routing. With only 210,357 parameters and near-perfect validation accuracy, the optimized model is substantially smaller than standard deep networks, demonstrating the efficiency of the CPLO search strategy.

5. Experimental Results

The CPLO-LWRCapsNet was trained for 60 epochs on the MSTAR train set, with a stable training process in which validation accuracy reached a maximum of 99.93% at epoch 57, and final training and validation losses converging at 0.0003 and 0.0028 respectively, indicating successful learning with little overfitting; the final validation accuracy at the end of training was 99.79%. On the held-out test set of 1,420 samples, the model achieved an overall test accuracy of 99.72%, a macro TPR (recall) of 99.56%, and a macro TNR (specificity) of 99.96%, with the near-perfect specificity indicating an extremely low false positive rate and the high recall indicating that almost all true target instances were correctly detected. The class-wise breakdown reveals consistently strong performance across all eight SAR target categories, with BRDM_2 and SLICY achieving perfect recognition on both metrics, while the most challenging target was the D7 class (a civilian bulldozer with atypical SAR signatures), which nonetheless achieved a still-competitive recall of 97.67%; overall, these results demonstrate strong class separation achieved by the feature extraction pipeline.

The proposed CPLO-LWRCapsNet achieved 172.6 million MACs and a model size of 6.8 MB, with an accuracy of 99.72%, precision of 99.68%, recall of 99.56%, F1-score of 99.64%, and specificity of 99.96%. By comparison, R-LRBPNet [4] achieved an accuracy of 78.20% with 225K parameters and 762.5M MACs. The proposed framework therefore improves accuracy by approximately 21.5 percentage points while using a smaller parameter budget and less than a quarter of the computational cost, demonstrating a highly competitive and deployable SAR-ATR solution built on feature isolation via segmentation, frequency-domain transformation, and CPLO-driven architecture search.

6. Conclusion

This paper introduces a lightweight yet accurate framework—CPLO-LWRCapsNet—for SAR target recognition, which tackles the fundamental issues in SAR image classification such as speckle noise, limited information, target-background separation, and high computational costs with a series of preprocessing, feature enrichment, segmentation, descriptor extraction, and optimized capsule classification steps. Stable activation of the Target Region preserved by KDMSNet segmentation, increased smoothness (ENL) of 59.2%, and Signal Quality (SNR) of 47.2% were preserved in the full range of values for Balanced Texture information and Edge information. The CPLO-optimized LWRCapsNet achieves a test accuracy of 99.72% with only 210,357 parameters and 172.6 million MACs, which is a good balance between accuracy and efficiency, and future work will include supporting extended operating conditions (differences in depression angles and the configuration of the target), few-shot learning for target classes with limited training data, and compression methods (quantization and pruning for deployment on embedded devices).

References

[1] J. Yoon, J. Song, T. Hussain, S. A. Khowaja, K. Muhammad, and I. H. Lee, "Hybrid Conv-Attention Networks for Synthetic Aperture Radar Imagery-Based Target Recognition," IEEE Access, 2024, doi: 10.1109/access.2024.3387314.

[2] N. Rahman, M. Khan, Md. I. Khan, J. Khan, and Y. Lee, "Robust ensemble classifier for advanced synthetic aperture radar target classification in diverse operational conditions," Scientific Reports, 2025, doi: 10.1038/s41598-025-93536-x.

[3] J. Zheng, M. Li, P. Zhang, Y. Wu, and H. Chen, "Position-Aware Graph Neural Network for Few-Shot SAR Target Classification," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2024, doi: 10.1109/JSTARS.2024.3382158.

[4] G. Gao et al., "R-LRBPNet: A Lightweight SAR Image Oriented Ship Detection and Classification Method," Remote Sensing, 2024, doi: 10.3390/rs16091533.

[5] J. Shi, S. Ji, H. Jin, Y. Zhang, M. Gong, and W. Lin, "Multi-Feature Lightweight DeeplabV3+ Network for Polarimetric SAR Image Classification with Attention Mechanism," Remote Sensing, 2025, doi: 10.3390/rs17081422.

[6] F. Sadeghi, A. Larijani, O. Rostami, D. Martin, and P. Hajirahimi, "A novel multi-objective binary chimp optimization algorithm for optimal feature selection: Application of deep-learning-based approaches for SAR image classification," Sensors, vol. 23, no. 3, pp. 1180, 2023.

[7] S. K. Chaturvedi, "Study of synthetic aperture radar and automatic identification system for ship target detection," Journal of Ocean Engineering and Science, vol. 4, no. 2, pp. 173-182, 2019.

[8] A. Passah, S. N. Sur, B. Paul, and D. Kandar, "SAR image classification: A comprehensive study and analysis," IEEE Access, vol. 10, pp. 20385-20399, 2022.

[9] J. Wang, W. Li, Y. Gao, M. Zhang, R. Tao, and Q. Du, "Hyperspectral and SAR image classification via multiscale interactive fusion network," IEEE Transactions on Neural Networks and Learning Systems, vol. 34, no. 12, pp. 10823-10837, 2022.

[10] Z. Huang, X. Yao, Y. Liu, C. O. Dumitru, M. Datcu, and J. Han, "Physically explainable CNN for SAR image classification," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 190, pp. 25-37, 2022.

[11] C. Wang, H. Gu, and W. Su, "SAR image classification using contrastive learning and pseudo-labels with limited data," IEEE Geoscience and Remote Sensing Letters, vol. 19, pp. 1-5, 2021.

[12] S. Sabour, N. Frosst, and G. E. Hinton, "Dynamic Routing Between Capsules," Advances in Neural Information Processing Systems (NeurIPS), 2017.