Lightweight Deep Learning Model for SAR Image Recognition: Survey

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: SAR imagery can perform all-weather imaging in the defence and environmental monitoring arena. Although deep learning is a turning point in image recognition, its application in SAR ATR is presently thwarted by major technical bottlenecks. The state-of-the-art models presented in these papers are mostly based on heavy Convolutional Neural Networks (CNNs) that were originally designed for optical imagery. The high parameter counts and memory usage of these architectures make them unsuited for resource-constrained edge devices such as UAVs. Training complex models on scarce labelled SAR datasets also tends to result in overfitting. To deal with these challenges, there is an urgent need for both efficiency and scalability to meet future demand. We introduce the AutoSAR framework, an approach that unifies lightweight architectures (e.g., MobileNetV3), Neural Architecture Search (NAS), and modern model compression techniques such as pruning and quantization. This work synthesizes the current state of the art in the field and provides a starting point for model development, with the aim of maximizing size reduction while maintaining a high level of recognition accuracy for real-time processing of SAR data on operational platforms.

Keywords: Synthetic Aperture Radar (SAR), Automatic Target Recognition (ATR), Lightweight Deep Learning Models, Neural Architecture Search (NAS), Model Compression Techniques

Introduction

SAR (Synthetic Aperture Radar) imaging is one of the most significant fields of remote sensing, as it allows images to be acquired both day and night. SAR is therefore used in various applications such as environmental mapping, disaster monitoring, and national security. Recent developments in deep learning have enabled significant progress in image recognition, but application to SAR images still involves technical hurdles that differ from those of optical images. This survey, however, acknowledges some caveats, such as the requirement for specific SAR measurements for prediction, limited generalizability to other sensors, and the balance between high compression and high recognition rate.

SAR Literature Review

The traditional target recognition system is limited by noise and radar properties, and relies on hand-designed feature extraction methods, such as the feature matching method based on SIFT features [2] and the object detection method based on HOG [3]. Before deep learning became prevalent, a popular method was Support Vector Machine (SVM) analysis, which has been widely applied to handcrafted features [15] but is susceptible to feature quality and does not scale well with increasing amounts of data. In addition, statistical models such as the conditionally Gaussian model [16] serve as useful analytic tools but cannot be practically applied in the case of complex targets and ambient parameters. Despite the success of deep learning, some researchers have still advocated for feature-based approaches, such as salient key point descriptors with sparse representation (also referred to as "sparse methods") [25], which are well suited for small datasets; however, they do not scale as well as deep learning architectures.

Hand-crafted features have inherent limitations for target detection in Synthetic Aperture Radar (SAR) systems, such as sensitivity to noise and poor generalization across varied targets, which deep learning techniques help address. Early results [1] indicated that CNNs outperformed traditional recognition approaches on the MSTAR dataset. The development of image techniques such as AlexNet [6] spurred further research into deep architectures. The VGG [7] and Inception [8] models, however, were limited by having too many parameters and being too compute-intensive, making them unsuitable for real-time applications. Following this, more recent techniques like Residual Networks (ResNet) [9] and DenseNet [10] improved accuracy but still consumed substantial resources in real-time and edge SAR applications. Squeeze-and-Excitation Networks (SE-Net) [11] were introduced to enhance feature representation with only a slight increase in computational cost, though their usefulness for SAR target detection has yet to be evaluated.

In one study [4], the authors showed that data augmentation can be applied to CNNs for SAR target recognition to mitigate the problem of limited training data, reducing overfitting and achieving high classification accuracy, but at a high computational cost. Related work [5] highlighted the importance of lightweight and efficient SAR models, and of the MSTAR dataset for SAR data analysis and ATR experimentation, as implemented in the MSTAR toolkit [12]. Other studies [13], [14] show that deep learning models can effectively learn SAR target attributes, but face difficulties in computational complexity, generalization, and model optimization. Furthermore, an adaptive boosting method [17] was proposed to enhance classification accuracy, but it depends on a suitable feature selection method and does not scale well to larger datasets. A more complex but robust decision fusion method [18] was also proposed to achieve better recognition performance in challenging environments.

Many studies have examined deep learning for the classification of Synthetic Aperture Radar (SAR) images, particularly using CNNs. Despite the success of transfer learning methods [19] with small amounts of labelled data, models pre-trained on optical imagery have not been optimally adapted to the particular characteristics of SAR. CNNs have also proven promising in maritime applications, such as ship classification from TerraSAR-X data [20], although performance is still influenced by sea state and weather conditions, and in oceanographic target detection [21], which remains constrained by dataset availability and environmental variability. Separately, CNN-based classification on the MSTAR dataset [22] achieved higher accuracy than traditional feature-based methods, but remained computationally expensive and sensitive to training data quality. A survey focused specifically on SAR ATR [23] confirmed the need for more efficient, less computationally demanding models, while a broader remote sensing review [24] highlighted the impact of deep learning on image analysis but noted that most of this work remains conceptual rather than experimentally validated.

Recent advances in deep learning have produced attention mechanisms that enable neural networks to concentrate on relevant information. This approach was first introduced for sequence-to-sequence neural machine translation [26] to improve translation quality by enabling models to focus on relevant details. Attention mechanisms have also been used for image captioning [27]. This technique holds promise for SAR target recognition by highlighting salient radar scattering points, but at the cost of additional processing and with difficulties in adapting to SAR due to fundamental differences in imaging semantics. Temporal modelling methods [28] have contributed to better video description but are limited in their direct applicability to SAR. Trainable visual attention mechanisms [29] are useful for SAR target recognition by focusing on salient parts of images but complicate the training process. Gated recurrent feature pyramids [30] were designed for improved object detection and have shown positive results, but still need to be adapted for SAR-specific use.

Deep learning architectures continue to evolve as researchers develop designs that improve performance while reducing computational requirements. Prominent examples include ResNet [31], which introduced residual learning but faced difficulties with deeper variants in resource-constrained environments. Inception-ResNet architectures [32] combined Inception modules with residual connections and achieved better accuracy, but their complexity makes them unsuitable for low-power applications. Xception [33] achieved Inception-level accuracy with lower complexity by using depthwise separable convolutions, showing promise as a basis for adapting such techniques to SAR systems. A similar depthwise separable approach was shown to generalize beyond vision to sequence modelling tasks such as neural machine translation [34]. MobileNet [35] was proposed as a lightweight architecture for embedded systems, making it suitable for real-time SAR recognition. ShuffleNet [36] and SqueezeNet [37] also proposed lightweight model designs; both achieved satisfactory accuracy with fewer parameters but may require adaptation to be best suited for radar applications.

Class imbalance and reliance on hand-crafted features remain persistent challenges in radar and other imaging applications, even as machine learning methods have advanced. An early study used hand-crafted features and classical machine learning techniques for SAR oil spill detection [38] and showed promising results. Extensive research has addressed the class imbalance problem, including sampling and algorithmic approaches [39], reviews of cost-sampling approaches highlighting the need for domain-specific solutions such as in SAR [41], cost-sensitive loss functions during training [44], earlier cost-sensitive decision tree models predating deep learning [45], and empirical comparisons of sampling versus cost-sensitive methods showing no universally best approach [46]. Further validation on SAR-specific datasets is needed, as a detailed study of class imbalance in CNNs [47] reports significant impacts on performance. Deep learning and machine learning methods have also been successfully applied to other image classification tasks, such as digital pathology imaging with CNNs [43], age and gender classification of facial images with CNNs [42], and breast cancer diagnosis using SVMs trained on ultrasound data [40]. These applications offer transferable approaches for SAR. The OpenSARShip dataset [48] consists solely of ship targets and is used for SAR-specific training and testing. Techniques such as the ReLU activation function [49] have made models easier to train, though a shortage of training data remains a challenge.

Through a gradual process, conventional feature-based methods have evolved into more complex deep learning architectures for SAR target detection. CNN-based, attention-based, and lightweight deep learning models have brought major advances in recognition performance. Computational complexity, model size, dataset size, and class imbalance, however, remain open and unresolved challenges affecting overall performance. Although lightweight models such as MobileNet, ShuffleNet, and SqueezeNet exist, they are not yet widely used in SAR imaging. A SAR target recognition system that achieves strong classification performance while requiring reduced processing resources for real-world deployment is therefore needed.

Research Gap

Although deep learning has improved SAR ATR, existing models are based on complex CNNs designed for optical images, which require large numbers of parameters and substantial memory, making them unsuitable for edge devices such as UAVs and microsatellites with limited resources. These models are also difficult to train given the relatively small amount of SAR data available, and standard feature-extraction methods remain vulnerable to high noise interference. In response to these challenges, this paper introduces AutoSAR, a unified framework that integrates lightweight architectures (MobileNetV3, EfficientNet-lite, etc.), Neural Architecture Search (NAS) for optimization, and model compression techniques (pruning, quantization, knowledge distillation). The aim is to reduce model size by up to 10x while preserving recognition accuracy, as a first step toward scalable, efficient, real-time SAR processing for military and environmental monitoring purposes.

References

[1] S. Chen, H. Wang, F. Xu, and Y.-Q. Jin, "Target Classification Using the Deep Convolutional Networks for SAR Images," IEEE Transactions on Geoscience and Remote Sensing, vol. 54, no. 8, pp. 4806–4817, Aug. 2016.

[2] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," International Journal of Computer Vision, vol. 60, no. 2, pp. 91–110, 2004.

[3] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), vol. 1, pp. 886–893, 2005.

[4] J. Ding, B. Chen, H. Liu, and M. Huang, "Convolutional Neural Network with Data Augmentation for SAR Target Recognition," IEEE Geoscience and Remote Sensing Letters, vol. 13, no. 3, pp. 364–368, March 2016.

[5] K. El-Darymli, E. W. Gill, P. McGuire, D. Power, and C. Moloney, "Automatic Target Recognition in Synthetic Aperture Radar Imagery: A State-of-the-Art Review," IEEE Access, vol. 4, pp. 6014–6058, 2016.

[6] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in Advances in Neural Information Processing Systems (NeurIPS), vol. 25, pp. 1097–1105, 2012.

[7] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," arXiv preprint arXiv:1409.1556, 2014.

[8] C. Szegedy et al., "Going Deeper with Convolutions," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1–9, 2015.

[9] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778, 2016.

[10] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4700–4708, 2017.

[11] J. Hu, L. Shen, and G. Sun, "Squeeze-and-Excitation Networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 7132–7141, 2018.

[12] J. Wissinger, R. Ristroph, J. R. Diemunsch, W. E. Severson, and E. Freudenthal, "MSTAR's Extensible Search Engine and Model-Based Inferencing Toolkit," in Algorithms for Synthetic Aperture Radar Imagery VI, vol. 3721, pp. 554–570, SPIE, 1999.

[13] H. Wang, S. Chen, F. Xu, and Y.-Q. Jin, "Application of Deep Learning Algorithms to MSTAR Data," in Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 3743–3745, 2015.

[14] A. Profeta, A. Rodriguez, and H. S. Clouse, "Convolutional Neural Networks for SAR Classification," in Proceedings of the IEEE National Aerospace and Electronics Conference (NAECON), pp. 287–291, 2016.

[15] Q. Zhao and J. C. Principe, "Support Vector Machines for SAR Automatic Target Recognition," IEEE Transactions on Aerospace and Electronic Systems, vol. 37, no. 2, pp. 643–654, April 2001.

[16] J. A. O'Sullivan, M. D. DeVore, V. Kedia, and M. I. Miller, "SAR ATR Performance Using a Conditionally Gaussian Model," IEEE Transactions on Aerospace and Electronic Systems, vol. 37, no. 1, pp. 91–108, Jan. 2001.
[17] Y. Sun, Z. Liu, S. Todorovic, and J. Li, "Adaptive Boosting for SAR Automatic Target Recognition," IEEE Transactions on Aerospace and Electronic Systems, vol. 43, no. 1, pp. 112–125, Jan. 2007.
[18] L. Huang and X. Qiu, "Robust Decision Fusion Strategy for SAR Target Recognition," IEEE Geoscience and Remote Sensing Letters, vol. 15, no. 12, pp. 1887–1891, Dec. 2018.
[19] Z. Huang, Z. Pan, and B. Lei, "Transfer Learning with Deep Convolutional Neural Network for SAR Target Classification with Limited Labeled Data," Remote Sensing, vol. 9, no. 9, p. 907, 2017.
[20] C. Bentes, D. Velotto, and B. Tings, "Ship Classification in TerraSAR-X Images with Convolutional Neural Networks," IEEE Journal of Oceanic Engineering, vol. 43, no. 1, pp. 258–266, Jan. 2018.
[21] C. Bentes, D. Velotto, and S. Lehner, "Target Classification in Oceanographic SAR Images with Deep Neural Networks: Architecture and Initial Results," in Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 3703–3706, 2015.
[22] S. Chen and H. Wang, "SAR Target Recognition Based on Deep Learning," in Proceedings of the 2014 International Conference on Data Science and Advanced Analytics (DSAA), Shanghai, China, pp. 541–547, Oct. 2014.
[23] M. Wilmanski, C. Kreucher, and J. Lanskold, "Modern Approaches in Deep Learning for SAR ATR," in Algorithms for Synthetic Aperture Radar Imagery XXIII, vol. 9843, p. 98430N, SPIE, 2016.
[24] J. E. Ball, D. T. Anderson, and C. S. Chan, "Comprehensive Survey of Deep Learning in Remote Sensing: Theories, Tools, and Challenges," Journal of Applied Remote Sensing, vol. 11, no. 4, p. 042609, 2017.
[25] A. Karine, A. Toumi, A. Khenchaf, and M. E. Hassouni, "Radar Target Recognition Using Salient Keypoint Descriptors and Multitask Sparse Representation," Remote Sensing, vol. 10, no. 5, p. 791, 2018.
[26] D. Bahdanau, K. Cho, and Y. Bengio, "Neural Machine Translation by Jointly Learning to Align and Translate," in International Conference on Learning Representations (ICLR), 2015.
[27] K. Xu et al., "Show, Attend and Tell: Neural Image Caption Generation with Visual Attention," in International Conference on Machine Learning (ICML), pp. 2048–2057, 2015.
[28] L. Yao et al., "Describing Videos by Exploiting Temporal Structure," in Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp. 4507–4515, 2015.
[29] V. Mnih, N. Heess, A. Graves, and K. Kavukcuoglu, "Recurrent Models of Visual Attention," in Advances in Neural Information Processing Systems (NeurIPS), vol. 27, pp. 2204–2212, 2014.
[30] H. Shen, S. Lin, C. Yuan, and Z. Zhang, "Gated Recurrent Feature Pyramids for Object Detection," in Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017.
[31] K. He, X. Zhang, S. Ren, and J. Sun, "Identity Mappings in Deep Residual Networks," in Proceedings of the European Conference on Computer Vision (ECCV), pp. 630–645, 2016.
[32] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, "Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning," in Proceedings of the AAAI Conference on Artificial Intelligence, vol. 31, no. 1, 2017.
[33] F. Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1251–1258, 2017.
[34] L. Kaiser, A. N. Gomez, and F. Chollet, "Depthwise Separable Convolutions for Neural Machine Translation," in International Conference on Learning Representations (ICLR), 2018.
[35] A. G. Howard et al., "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," arXiv preprint arXiv:1704.04861, 2017.

[36] X. Zhang, X. Zhou, M. Lin, and J. Sun, "ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 6848–6856, 2018.
[37] F. N. Iandola et al., "SqueezeNet: AlexNet-level Accuracy with 50x Fewer Parameters and <0.5MB Model Size," arXiv preprint arXiv:1602.07360, 2016.
[38] M. Kubat, R. C. Holte, and S. Matwin, "Machine Learning for the Detection of Oil Spills in Satellite Radar Images," Machine Learning, vol. 30, no. 2, pp. 195–215, 1998.
[39] H. He and E. A. Garcia, "Learning from Imbalanced Data," IEEE Transactions on Knowledge and Data Engineering, vol. 21, no. 9, pp. 1263–1284, Sept. 2009.
[40] R. F. Chang, W. J. Wu, W. K. Moon, and D. R. Chen, "Support Vector Machines for Diagnosis of Breast Tumors on US Images," Academic Radiology, vol. 10, no. 2, pp. 189–197, 2003.
[41] X. Guo, Y. Yin, C. Dong, G. Yang, and G. Zhou, "On the Class Imbalance Problem," in Proceedings of the 4th International Conference on Natural Computation (ICNC), vol. 4, pp. 192–201, 2008.
[42] G. Levi and T. Hassner, "Age and Gender Classification Using Convolutional Neural Networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 34–42, 2015.
[43] A. Janowczyk and A. Madabhushi, "Deep Learning for Digital Pathology Image Analysis: A Comprehensive Tutorial with Selected Use Cases," Journal of Pathology Informatics, vol. 7, p. 29, 2016.
[44] S. H. Khan, M. Hayat, M. Bennamoun, F. A. Sohel, and R. Togneri, "Cost-Sensitive Learning of Deep Feature Representations from Imbalanced Data," IEEE Transactions on Neural Networks and Learning Systems, vol. 29, no. 8, pp. 3573–3587, 2018.
[45] F. Min and W. Zhu, "A Competition Strategy to Cost-Sensitive Decision Trees," in International Conference on Rough Sets and Knowledge Technology, pp. 486–493, Springer, 2012.
[46] G. M. Weiss, K. McCarthy, and B. Zabar, "Cost-Sensitive Learning vs. Sampling: Which is Best for Handling Unbalanced Classes with Unequal Error Costs?" in Proceedings of the International Conference on Data Mining (DMIN), pp. 35–41, 2007.
[47] M. Buda, A. Maki, and M. A. Mazurowski, "A Systematic Study of the Class Imbalance Problem in Convolutional Neural Networks," Neural Networks, vol. 106, pp. 249–259, 2018.
[48] L. Huang, B. Liu, B. Huang, W. Guo, and Z. Zhang, "OpenSARShip: A Dataset Dedicated to Sentinel-1 Ship Interpretation," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 11, no. 1, pp. 195–208, Jan. 2018.
[49] V. Nair and G. E. Hinton, "Rectified Linear Units Improve Restricted Boltzmann Machines," in Proceedings of the 27th International Conference on Machine Learning (ICML), pp. 807–814, 2010.