# A Hierarchical Attention Model for Local and Global Feature Integration in RCS Classification

Yida Wu*, Caiyun Wang*‡, Jianing Wang†, Xiaofei Li† and Ying Nan*
* College of Astronautics, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China.
† Beijing Institute of Electronic System Engineering, Beijing 100854, China.
‡ Corresponding author: wangcaiyun@nuaa.edu.cn

***Abstract*—Radar Cross Section (RCS) target recognition plays a vital role in radar automatic target recognition (RATR) systems. As a critical electromagnetic signature, RCS encodes unique target characteristics include geometric profiles, structural features, and material properties. However, the complex non-linear relationship between RCS patterns and target attributes introduces significant challenges for accurate target identification. To address this, this paper proposes a hierarchical attention model for local and global feature integration, named Local-Global Efficient Channel Attention (LGECA), which enhances RCS feature representation through modeling both local-specific interactions and global contextual dependencies. Moreover, an improved multi-scale convolutional neural network is proposed as the backbone architecture, leveraging the LGECA module to establish new state-of-the-art performance in RCS classification. The comprehensive experimental results validate the effectiveness and robustness of the proposed module against conventional approaches.**

## I. Introduction

Radar is a commonly used sensor characterized by weather-proof capability, long-range detection, and the ability to detail the specifics of the target. It is widely applied in target detection, tracking, and recognition. Among the various data for Radar Automatic Target Recognition (RATR), Radar Cross Section (RCS) plays a crucial role due to its inherent ability to depict target properties such as geometry, structural, and material through electromagnetic scattering phenomena [1], [2]. With its superior discriminative capability, RCS has been widely adopted in multiple classification tasks [3], [4].

As Convolutional Neural Networks (CNNs) have yielded remarkable success across diverse domains such as image classification, target localization, signal processing, and semantic segmentation [5]–[7], their demonstrated abilities of feature extraction and pattern recognition are particularly suited for RCS classification tasks [8]–[11]. To further enhance the capability of CNNs, attention mechanisms have become an essential component, significantly improving model performance. By enabling the network to learn what and where to focus, attention effectively strengthens feature representation [12].

Depending on the application domain, attention mechanism can be divided into channel attention mechanism, spatial attention mechanism, and hybrid attention mechanism [13]. The channel attention mechanism serves as a fundamental component for feature refinement, whereas spatial attention mechanism operates in a complementary capacity to boost the performance. First introduced in [14], the Squeeze-and-Excitation (SE) mechanism has become a cornerstone of channel attention design, inspiring numerous variants while remaining widely implemented in its original form. It captures the channel information after global average pooling (GAP) through two fully connected layers (FCs), thereby generating the channel attention weights. ECA [15] improves SE by replacing FCs with 1D convolution, showing that avoiding dimensionality reduction is vital for effective attention learning. Meanwhile, it also demonstrates that moderate cross-channel interaction can greatly reduce computational complexity while preserving performance. However, the limited channel interaction captured by 1D convolution inevitably leads to the omission of some information in feature learning. Due to its sliding mechanism, the 1D convolution maintains global processing consistency while unavoidably overlooking certain distinctive local channel interactions. Through group-wise channel processing, this issue can be avoided to some extent. SGE [16] divides channels into multiple groups, treating each group as a representative of a semantic feature and calculating its attention. Similarly, the SA [17] module also employs grouping but further divides each channel group to separately calculate both spatial and channel attention. In addition, a shuffle unit is added at the end to shuffle the channels based on their groups. In fact, the result calculated by SGE and SA is the attention of each group, as they treat each group as a single semantic feature and process it as a whole. Their performance improvement is largely due to the higher local feature granularity brought about by channel grouping. In contrast, SGCA [18] applies convolution operations to each channel group to facilitate local interaction, but the dependencies between group-wise channels are overlooked.

To learn more accurate channel attention, we proposed a Local Global Efficient Channel Attention (LGECA) module, through which finer local channel interactions and global channel interactions can be captured in a more efficient way. As illustrated in Fig. 1, after channel-wise global average pooling (GAP), our LGECA first captures local intra-group interactions interaction using group convolution. After local enhancement, the features are shuffled according to the grouping order [17], and global interactions are explored based on this.

The contributions of this paper are summarized below. (1) We dissected existing modules and empirically demonstrated that special local interactions of channels are crucial to learning effective channel attention. (2) Based on this, we propose a

Local-Global Efficient Channel Attention (LGECA) module, which excavates special local and global interactions to make attention learning more accurate. (3) We also proposed a multi-scale attention convolutional neural network (MSACNN) with LGECA for the classification of RCS sequences. (4) The experimental results demonstrate that our method achieves highly competitive performance compared to state-of-the-art methods.

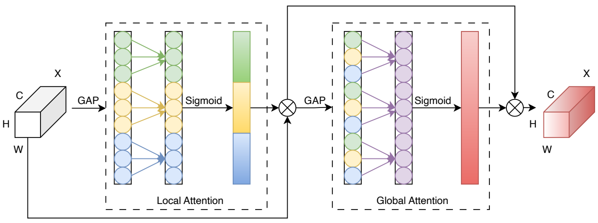



Fig. 1. Diagram of local-global efficient channel attention (LGECA) Module.

## II. Related works

### A. RCS classification

RCS time series is the most common data that radars can obtain and is one of the key data types for ballistic target classification. Traditional RCS classification methods mainly involve extracting various features, such as statistical features [19] and transformation features [20]–[23], and then combining them with classifiers such as KNN, SVM and others. Recently, image encoding methods have also been applied to RCS classification, such as Gramian angular field (GAF), Markov transition field (MTF), and recurrence plot (RP) [24].

With the rapid development of deep learning, end-to-end models have been employed for RCS classification. Chen [8] developed RCSnet with CNN to classify space target. Wang [25] used CNN for the classification of rotating targets. As a time series, the temporal features of RCS are crucial. RNNs, including LSTM and GRU, have also been utilized for this purpose [26]–[28].

### B. Attention mechanism

Recently, the attention mechanism has been shown to significantly enhance the performance of deep convolutional networks by guiding the network to focus on more important features, thus improving its feature representation. SE-Net [14] first introduced a method to learn the attention of the channel, significantly increasing the representational capacity of CNNs. This approach to attention learning has since been adopted and continuously refined, such as A2-Nets [29], ECA [15], CBAM [30], GCNet [31], and ECSA [32].

In contrast to the methods mentioned above, some approaches take into account the local heterogeneity of channels. The SGE module [16] groups the input feature maps and treats each group as a representative of a semantic feature, allowing each individual group to independently enhance its learned representation. Similarly to SGE, the SA module [17] treats each channel group as a sub-feature and independently enhances the features in both the channel and spatial dimensions for each group. Then, a Shuffle Unit is applied to shuffle the channel groups, facilitating subsequent interactions between different groups. However, SGCA [18] does not treat channel groups as sub-features but as sub-domains of the channels. It uses 1D-CNN to extract features within each group, mining deeper features. The grouping operation allows it to preserve the distinct local characteristics between channels.

## III. Proposed method

In this section, we first revisit the typical channel attention module in SE and ECA to analyze the fundamental theory behind channel attention computation. Then we discuss the improvements to channel attention computation which motivate us to propose our LGECA method. In addition, we develop the MSCNN model with our LGECA module for the classification of RCS series.

### A. Channel Attention Mechanism

Let the input of the attention module be $\chi \in \mathbb{R}^{W,H,C}$, where $W$, $H$ and $C$ are width, height and channel dimension. In the SE module, the weights of the channels can be computed as

$$\omega_{SE} = \sigma(\mathbf{W_2} ReLU(\mathbf{W_1 y})) \tag{1}$$

where $\mathbf{y} = g(\chi) = \frac{1}{WH}\sum_{i=1,j=1}^{W,H} \chi_{ij}$ is channel-wise global average pooling (GAP) and $\sigma$ is a Sigmoid function. For RCS time series, the processing differs only in the GAP section, that is, $\mathbf{y} = g(\chi) = \frac{1}{L}\sum_{i=1}^{L} \chi_i$, where $L$ is the length of feature extracted from RCS series. ReLU indicates the Rectified Linear Unit. To avoid disrupting the direct relationships and weights between channels, the ECA module replaces the weight computation process as

$$\omega_{ECA} = \sigma(\mathbf{W_k y}) \tag{2}$$

where $\mathbf{W_k}$ is the weight matrix as follows

$$\mathbf{W_k} = \begin{bmatrix} w_1 & \cdots & w_k & 0 & 0 & \cdots & \cdots & 0 \\ 0 & w_1 & \cdots & w_k & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 & w_1 & \cdots & w_k \end{bmatrix} \tag{3}$$

where the weight of $y_i, i = 1, 2, \cdots, k$ is calculated by only considering interaction between $y_i$ and its $k$ neighbors, i.e.,

$$\omega_i = \sigma\left(\sum_{j=1}^{k} w_j y_i^j\right), y_i^j \in \Omega_i^k \tag{4}$$

where $\Omega_i^k$ indicates the set of $k$ adjacent channels to $y_i$. It can be seen that the ECA module calculates attention based only on the interaction between neighboring $k$ channels to reduce computational complexity, while interactions between distant channels are not considered. Moreover, since only a single filter is used, the captured local interactions tend to favor global consistency, thereby suppressing potentially important local-specific interactions, which are also crucial for attention learning.

We believe that local-specific interactions (local channel attention) should be captured and preserved, while interactions between distant channels (global channel attention) are also essential. Therefore, we develop the LGECA to learn more comprehensive attention from both aspects mentioned above.

### *B. Local Channel Attention*

In LGECA, we divide the channels into multiple groups using a grouping method, where each group preserves its own unique internal interactions. Each channel group is processed with an independent convolutional filter. The weights of local channel attention can be computed as

$$\omega_{local} = \sigma\left(\mathbf{W}_{local}\mathbf{y}\right) \tag{5}$$

where $\omega_{\mathbf{local}}$ is local channel attention, $\mathbf{W}_{local}$ is the weights matrix which can be represented as

$$\mathbf{W}_{local} = \begin{bmatrix} \mathbf{W}_1 & \cdots & 0 \\ \vdots & \mathbf{W}_g & \vdots \\ 0 & \cdots & \mathbf{W}_G \end{bmatrix} \tag{6}$$

where $G$ is the number of channel groups, and the weights matrix $\mathbf{W}_g$ of each group is consistent with Equ. 3.

This processing strategy enables the independent modeling of intra-group interactions while systematically preserving local cross-channel correlations, thereby maintaining comprehensive feature relationships.

### *C. Global Channel Attention*

Although channel grouping preserves the specificity of local interactions, interactions between different groups are overlooked. Therefore, we further capture inter-group information through interactions between channel groups, while also enabling long-range channel interactions i.e., global channel attention. The weights of global channel attention can be computed as

$$\omega_{global} = \sigma\left(\mathbf{W}_{global}\mathbf{z}\right) \tag{7}$$

where $\mathbf{z} = g\left(\omega_{local}\chi\right)$, the weight of $z_i, i = 1, 2, \cdots, k$ is calculated as

$$\omega_i = \sigma\left(\sum_{j=1}^{G} w_j z_i^j\right), z_i^j \in \Omega_i^G \tag{8}$$

where $\Omega_i^G$ indicates the set of channels at each corresponding position in the channel groups. Note that global channel attention should be performed after local attention enhancement for the following reasons: 1)

1) Global attention is based on local attention, and different local attention results in variations in global attention.
2) To avoid disrupting direct relationships and weights between distant channels, global attention should process the features after local enhancement rather than directly handling local channel interactions.

### *D. MSACNN*

To better incorporate attention modules into RCS classification, we propose a multi-scale attention convolutional neural network (MSACNN), which consists of multiple multi-scale blocks (MS Block). The structure of the proposed model is illustrated below. The symbol C refers to the output channels of each multi-scale block.

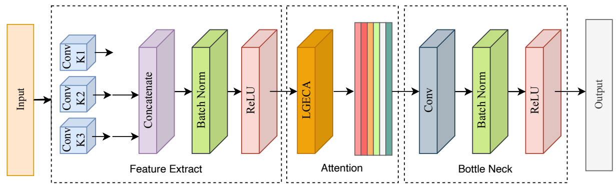



Fig. 2. Structure of multi-scale block (MS Block).

The MSACNN model is an improved version based on MACNN [33], where the attention layer used to improve features in all channels on multiple scales is replaced with LGECA, as shown in Fig. 2. Moreover, we added a bottleneck layer after each attention layer consisting of a convolution layer reducing the dimensionality.

## IV. Experiments

In this section, we evaluate the effectiveness of the proposed LGECA and MSACNN through comparative experiments with other methods. All experiments are conducted on a simulated RCS dataset, which contains RCS series of multiple different targets under various motion parameters.

### *A. Implementation Details*

We adopted the same method as in [8] to simulate the RCS time series of targets. Nine target trajectories were simulated, each with 7 targets of different shapes and micro-motion parameters. The RCS time series were computed at the frequency of 10 GHz with a sampling rate of 100 Hz and a signal signal-to-noise ratio (SNR) of 10 dB. We used a sliding-window approach for sampling, with a window length of 2 seconds and a stride of 1 second. 6 trajectories of all trajectories were used as a training set, and the remaining three were used as a test set.

The parameters of the model are optimized by stochastic gradient descent (SGD) with a weight decay of 0.0001, the learning rate is 0.001 and the mini-batch size of 64. The models are trained within 50 epochs. In MSACNN, the kernel sizes are $[3, 6, 12]$ and the filters of each MS Block are 64, 128, 256, respectively.

### *B. Comparison of different attention mechanisms*

In this part, we compare our LGECA with other state-of-the-art attention modules, including SE [14], ECA [15], SGE [16], CBAM [30], etc. We have modified them to make them suitable for sequence-based feature maps. The results are as follows. Except for the attention module, all other settings remain the same.

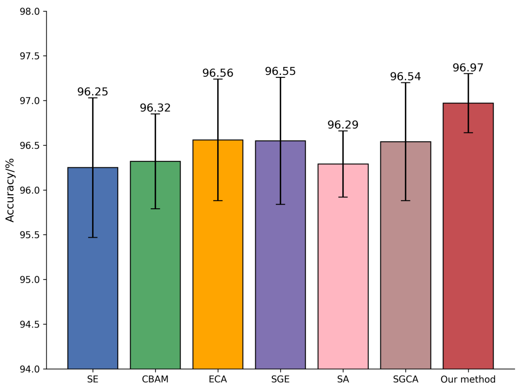
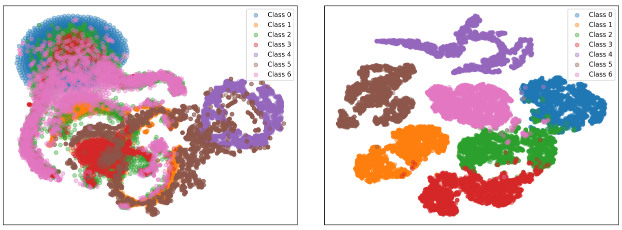



Fig. 3. Comparison of classification accuracy and standard deviation across different attention modules.

As shown in Fig. 3, our LGECA outperforms other attention modules by 0.41% and 0.72% in classification accuracy, demonstrating its effectiveness in learning channel attention. Furthermore, the proposed method exhibits the smallest standard deviation, further validating its superior stability and robustness.



(a) Original Data (b) LGECA

Fig. 4. t-SNE visualization of high-dimensional features.

To further validate the effectiveness of the proposed method, we employ t-SNE to visualize the high-dimensional representations of the original input and our metho. The results show that features learned by our approach exhibit more compact intra-class clustering and clearer inter-class separation, indicating enhanced discriminative capability, as shown in Fig. 4.

## C. Ablation study

To further demonstrate the effectiveness of the proposed method, we conducted a series of ablation experiments by selectively enabling or disabling key components of the attention mechanism. We compared the classification accuracy and per-class F1 scores for each attention mechanism to gain deeper insights into both overall classification performance and class-wise discriminability. The results are presented as follows.

As shown in Table. I, the combination of local and global attention achieves better performance than using either local or global attention alone. This demonstrates that local and global attention mechanisms are complementary, and their integration enhances the model's ability to extract both detailed and holistic features. This observation is further supported by the visualization in Fig. 5, where the combined attention mechanism consistently achieves higher scores across most categories, indicating improved class-wise discriminability.

TABLE I
MEAN AND (STANDARD DEVIATION) OF ACCURACY UNDER LOCAL, GLOBAL, AND COMBINED ATTENTION MECHANISM.

| Attention mechanism | Accuracy |
|---|---|
| Local | 96.33 (0.57) |
| Global | 96.28 (0.39) |
| Local & Global | 96.97 (0.33) |

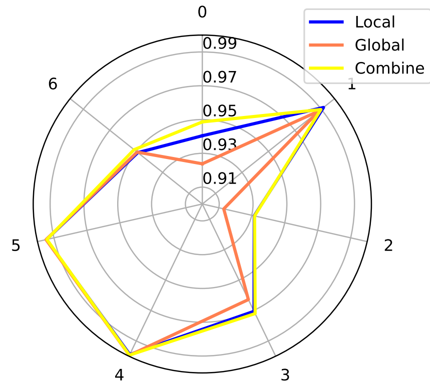



Fig. 5. Per-class F1-scores under different attention mechanisms.

## D. Comparison of different methods

To demonstrate the effectiveness of the proposed MSACNN, we conducted comparative experiments against several existing approaches, including LSTM [28], RCSnet [8], MACNN [33]. Meanwhile, to evaluate the performance of the model under different signal-to-noise ratios (SNR), we add noise to the data with SNR levels ranging from 0 to 20 dB, with a step size of 2. The experimental results of different methods under different SNR are shown below.

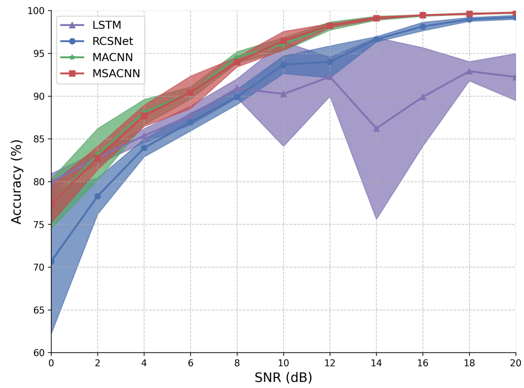



Fig. 6. Accuracy of different methods vs. SNR, the shaded area depicts mean ± the standard deviation over five different random seeds.

As shown in Fig. 6, the performance of all methods improves as the SNR increase; however, the proposed method consistently achieves superior performance throughout various conditions. In addition, the narrower shaded region reflects higher stability and lower variance. Based on this, although MACNN achieves comparable accuracy to the proposed method, the proposed method exhibits superior robustness. These observations validate the effectiveness of the proposed approach in extracting comprehensive features and enhancing feature-level attention.

## V. Conclusion

In this paper, we focus on learning discriminative feature representations from RCS series to advance target recognition performance. To this end, we propose a hierarchical attention model for local and global feature integration called LGECA, which learns better channel attention by capturing both fine-grained local and long-range global channel interactions through group-wise processing. Moreover, we further develop a multi-scale attention convolutional neural network (MSACNN) incorporating LGECA for RCS classification. Experimental results demonstrate that our LGECA significantly enhances feature representation, leading to improved performance across various benchmarks. This validates the effectiveness of our method in capturing both global context and local dependencies through cross-attention mechanisms. In future work, we aim to further expand the applicability of LGECA to broader multi-modal fusion tasks.

## Acknowledgment

This work was supported by the National Natural Science Foundation of China (Grant No. 61301211), the China Scholarship Council (No. 202506830119), and the Postgraduate Education and Teaching Reform Project of Nanjing University of Aeronautics and Astronautics (No. 2025YJXGG-C12).

## References

[1] H. He, L. Wu, and F. Lu, "Radar autonomous recognition technology of sea target," in *2021 IEEE 2nd International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA)*, vol. 2, 2021, pp. 451–456.

[2] M. Zhu, J. Jiang, X. Qiao, and Z. Zhao, "Flight target recognition method based on rcs sequenceed," in *2024 8th International Conference on Electrical, Mechanical and Computer Engineering (ICEMCE)*, 2024, pp. 1873–1876.

[3] A. Rawat, A. Sharma, and A. Awasthi, "Machine learning based non-cooperative target recognition with dynamic rcs data," in *2023 IEEE Wireless Antenna and Microwave Symposium (WAMS)*, 2023, pp. 1–5.

[4] C. Wang, H. Zhang, J. Wang, Y. Wu, and Y. Chang, "Ballistic midcourse target rcs recognition based on dtcwt-vae," *Systems Engineering and Electronics*, vol. 46, no. 7, pp. 2269–2275, 2024.

[5] Y. Xie, W. Sun, M. Ren, S. Chen, Z. Huang, and X. Pan, "Stacking ensemble learning models for daily runoff prediction using 1d and 2d cnns," *Expert Systems with Applications*, vol. 217, pp. 119 469–, 2023.

[6] J. Chen, T. H. Teo, C. L. Kok, and Y. Y. Koh, "A novel single-word speech recognition on embedded systems using a convolution neuron network with improved out-of-distribution detection," *Electronics*, vol. 13, no. 3, p. 16, 2024.

[7] I. Rakhmatulin, M. S. Dao, A. Nassibi, and D. Mandic, "Exploring convolutional neural network architectures for eeg feature extraction," *Sensors*, vol. 24, no. 3, p. 39, 2024.

[8] Chen, Jian, Shiyou, and Zengping, "Convolutional neural network for classifying space target of the same shape by using rcs time series," *IET Radar, Sonar Navigation*, 2018.

[9] E. Wengrowski, M. Purri, K. Dana, and A. Huston, "Deep cnns as a method to classify rotating objects based on monostatic rcs," *Radar, Sonar & Navigation, IET*, vol. 13, no. 7, pp. 1092–1100, 2019.

[10] L. Zhang, B. Xue, Y. Zhou, Y. Wang, and X. Pan, "Intelligent classification of spatial targets based on rcs sequences," in *2023 International Applied Computational Electromagnetics Society Symposium (ACES-China)*, 2023, pp. 1–3.

[11] Q. Lv, H. Fan, Y. Zhao, Y. Quan, and M. Xing, "Agile frequency rcs-based deep fusion network for ship and corner reflector identification," *IEEE Geoscience and Remote Sensing Letters*, vol. 21, pp. 1–5, 2024.

[12] B. Zhao, X. Wu, J. Feng, Q. Peng, and S. Yan, "Diversified visual attention networks for fine-grained object classification," *IEEE Transactions on Multimedia*, vol. 19, no. 6, pp. 1245–1256, 2017.

[13] J. Pan, H. Hu, A. Liu, Q. Zhou, and Q. Guan, "A channel-spatial hybrid attention mechanism using channel weight transfer strategy," in *2022 26th International Conference on Pattern Recognition (ICPR)*, 2022, pp. 2524–2531.

[14] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 7132–7141.

[15] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, "Eca-net: Efficient channel attention for deep convolutional neural networks," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 11 531–11 539.

[16] Y. Li, X. Li, and J. Yang, "Spatial group-wise enhance: Enhancing semantic feature learning in cnn," in *Computer Vision – ACCV 2022*, L. Wang, J. Gall, T.-J. Chin, I. Sato, and R. Chellappa, Eds., Cham: Springer Nature Switzerland, 2023, pp. 316–332.

[17] Q. Zhang and Y. Yang, "Sa-net: Shuffle attention for deep convolutional neural networks," *arXiv preprint arXiv:2102.00240*, 2021.

[18] Z.-c. Zheng, C. Fan, MiaoWang, C.-q. Wang, X.-l. He, and X.-w. He, "Spatial group and cross-channel attention: Make smaller models more effective, focus on high-level semantic features," in *Advanced Intelligent Computing Technology and Applications*, D.-S. Huang, Y. Pan, and J. Guo, Eds., Singapore: Springer Nature Singapore, 2024, pp. 66–77.

[19] X. Lei, X. Fu, C. Wang, and M. Gao, "Statistical feature selection of narrowband rcs sequence based on greedy algorithm," in *Proceedings of 2011 IEEE CIE International Conference on Radar*, vol. 2, 2011, pp. 1664–1667.

[20] X. Xiang and X. Xu, "Feature extraction for radar target recognition using time sequences of radar cross section measurements," in *2013 6th International Congress on Image and Signal Processing (CISP)*, vol. 03, 2013, pp. 1583–1587.

[21] G. Delisle, Z. Zebbani, C. Charrier, and F. Cote, "A novel approach to complex target recognition using rcs wavelet decomposition," *IEEE Antennas and Propagation Magazine*, vol. 47, no. 1, pp. 35–55, 2005.

[22] T. Wang, W. Bi, Y. Zhao, and W. Xue, "Target recognition algorithm based on rcs observation sequence — set-valued identification method," in *Proceedings of the 33rd Chinese Control Conference*, 2014, pp. 881–886.

[23] W. Wenbo and C. Hongjun, "Narrowband radar ballistic missile target recognition technology based on svm," *Electronic Science and Technology*, 2016.

[24] X. Xu, C. Feng, and L. Han, "Classification of radar targets with micro-motion based on rcs sequences encoding and convolutional neural network," *REMOTE SENSING*, vol. 14, no. 22, Nov. 2022.

[25] Y. Wang, C. Feng, Y. Zhang, and S. He, "Space precession target classification based on radar high-resolution range profiles," *International Journal of Antennas and Propagation*, vol. 2019, pp. 1–9, 2019.

[26] J. Mansukhani, D. Penchalaiah, and A. Bhattacharyya, "Rcs based target classification using deep learning methods," in *2021 2nd International Conference on Range Technology (ICORT)*, 2021, pp. 1–5.

[27] B. Sehgal, H. S. Shekhawat, and S. K. Jana, "Automatic radar target identification using radar cross section fluctuations and recurrent neural networks," in *TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON)*, 2019, pp. 2490–2495.

[28] Y. Wang, B. Long, and F. Wang, "Rcs statistical feature extraction for space target recognition based on bi-lstm," in *IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium*, 2023, pp. 6049–6052.

[29] Y. Chen, Y. Kalantidis, J. Li, S. Yan, and J. Feng, *$A^2$-nets: Double attention networks*, 2018. arXiv: 1810.11579 [cs.CV]. [Online]. Available: https://arxiv.org/abs/1810.11579.

[30] S. Woo, J. Park, J. Y. Lee, and I. S. Kweon, "Cbam: Convolutional block attention module," *Springer, Cham*, 2018.

[31] Y. Cao, J. Xu, S. Lin, F. Wei, and H. Hu, "Gcnet: Non-local networks meet squeeze-excitation networks and beyond," in *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*, 2019, pp. 1971–1980.

[32] H. Liu, Y. Zhang, and Y. Chen, "A symmetric efficient spatial and channel attention (esca) module based on convolutional neural networks," *SYMMETRY-BASEL*, vol. 16, no. 8, Aug. 2024.

[33] W. Chen and K. Shi, "Multi-scale attention convolutional neural network for time series classification," *Neural networks : the official journal of the International Neural Network Society*, vol. 136, pp. 126–140, 2021.