Contemporary Mathematics
https://ojs.wiserpub.com/index.php/CM/

Research Article

Advanced Radar Signal Processing Using Adaptive Threshold Convolution Neural Network

Bozhi Qiu[1], Sheng Li[1*], Ying Li[2]

[1] School of Electronic Information, Xijing University, Xi'an, China
[2] Shaanxi Huanghe Group Co., Ltd., Xi'an, China
E-mail: sheng@mail.xjtu.edu.cn

Received: 12 May 2025; **Revised:** 18 June 2025; **Accepted:** 26 June 2025

Abstract: Radar Automatic Target Recognition (RATR) is critical for surveillance in complex electromagnetic environments. Traditional methods struggle with interference suppression, while existing deep learning approaches lack adaptability to dynamic Signal-to-Interference Ratios (SIR). This paper proposes an Adaptive Threshold Convolutional Neural Network (ATCNN) featuring three innovations: a convolution unit that dynamically adjusts activation thresholds using real-time SIR to suppress noise and enhance feature extraction; a multi-scale framework with varied kernel sizes to capture global and local patterns in High-Resolution Range Profiles; and a channel attention mechanism fused with residual connections to prioritize salient features while preserving data integrity. Evaluations on measured High-Resolution Range Profile (HRRP) datasets show ATCNN's superior accuracy over existing methods. It maintains stable performance under diverse interference conditions across SIR levels, outperforming baseline models. Ablation studies confirm each module's necessity, with performance dropping significantly when core components are removed. The framework's environmental adaptability and balanced feature extraction advance target recognition in real-world scenarios, offering scalable solutions for autonomous systems application.

***Keywords*:** radar automatic target recognition, adaptive threshold convolution, multi-scale feature fusion, channel attention, anti-interference robustness

MSC: 68T07, 94A12, 68T10

1. Introduction

Radar is a form of early warning and detection that uses electromagnetic waves to measure features such as range and azimuth of targets. Radar detects, identifies, and tracks targets, and is an indispensable tool for various sensing applications, air traffic control, and identification of targets [1]. In recent years, radar technology has advanced on this basis due to the rapid development of microelectronics, digital circuits, and other fields. The application of radar in modern information systems is no longer limited to simple detection and ranging. Instead, the methodology aims to derive discriminative target-specific attributes from electromagnetic backscattering signatures captured by the receiving aperture, thereby facilitating robust feature extraction from complex radar returns, so as to detect the target and determine the target type and other information [2].

DOI: https://doi.org/10.37256/cm.6520257217

High-Resolution Range Profile (HRRP) represents the coherent superposition of electromagnetic backscattering contributions from discrete target elements aligned along the radar line-of-sight vector, encoding substantial structural and geometrical target information. HRRP signatures offer significant operational advantages over two-dimensional radar imaging modalities in terms of acquisition efficiency, computational overhead, and data storage requirements [3]. The proliferation of advanced machine learning paradigms has catalyzed significant research activity in Radar Automatic Target Recognition (RATR) methodologies, yielding promising empirical outcomes. Practical implementation scenarios in civilian systems pose significant technical challenges due to complex electromagnetic environments, dynamic target behaviors, and multiple interference sources, collectively constraining radar automatic target recognition reliability [4].

Recent advances in radar target classification demonstrate a clear trend toward integrated network architectures, where diverse neural frameworks collaborate to enhance both feature extraction and classification performance. Study [5] developed a hybrid approach utilizing autoencoders for dimensionality reduction and feature learning, combined with Convolutional Neural Networks (CNNs) for classification tasks, leveraging the deep learning structure's capability to model complex patterns. Zhang et al. [6] introduced an innovative end-to-end system incorporating multiple inputs into convolutional gated recurrent units, designed to capture both temporal and spectral information from radar signals, effectively modeling time-dependent patterns in HRRP sequences. Pan et al. [7] proposed a sophisticated CNN-BIRNN system enhanced with attention mechanisms, integrating adaptive layers, convolutional networks, and feature enhancement modules within a bi-directional recurrent neural framework to process radar echo data more effectively. A contrasting methodology by He et al. [8]. implemented pyramid-structured one-dimensional convolutional architectures to extract multi-scale features, while employing depth-separable convolutions to minimize computational demands through more efficient parameter utilization. More recently, Wu et al. [9] developed a CNN-BiLSTM-SE model specifically for ship HRRP target recognition that demonstrates robust performance against decoy jamming techniques.

Previous neural network frameworks largely function under the unrealistic premise of clean radar echo data-a scenario seldom found in field deployments where radar signals frequently encounter various forms of electromagnetic interference. To address this practical challenge, Su et al. [10] developed a robust recognition system utilizing dynamic noise variance estimation for model adaptation, coupled with temporal correlation analysis to strengthen performance stability in noisy environments. In parallel, Liu et al. [11] formulated a combined approach integrating signal cleaning and pattern recognition functions within a single trainable architecture, achieving significant performance improvements in low signal-to-noise environments. It bears noting that these research efforts typically simulate interference using elementary noise models-primarily Gaussian or Rayleigh distributions-applied directly to original HRRP data. Such simplified approaches inadequately represent the actual complexity of real-world electromagnetic environments where various interference sources can impact signal quality.

To solve the above problems, this paper also proposes an adaptive threshold convolution neural network. The core module of the ATCNN network is the Adaptive Threshold Convolution Unit (ATCU), which is used to achieve accurate and efficient feature extraction. In ATCU, an adaptive threshold function is used as the activation function to automatically adjust the threshold to face the data with a different signal-to-interference ratio. Multiple convolution kernels of different scales were used to capture the regional differences in HRRP data. The channel attention mechanism is introduced to optimize the network structure. After testing by numerical experiments, the proposed ATCNN shows very good recognition accuracy and anti-interference robustness.

In summary, the contributions of this paper are as follows:

(1) This paper presents the adaptive threshold convolution neural network, featuring the adaptive threshold convolution unit. The ATCU employs an adaptive threshold function as the activation mechanism, adjusting thresholds for data with different signal-to-interference ratios. Utilizing multiple convolution kernels of various scales, it captures regional differences in HRRP data, enhancing feature extraction accuracy and robustness against interference.

(2) This paper introduces a channel attention mechanism to optimize the ATCNN network structure. By enhancing the extraction of key features, this mechanism improves the processing of radar echo signals, resulting in superior recognition accuracy and anti-interference capability. The integration allows ATCNN to selectively focus on important information while suppressing less relevant data, leading to more precise and reliable target identification.

(3) This paper incorporates residual connections within the ATCNN architecture, ensuring the integrity of the original features throughout the network layers. By directly adding the original features to the processed ones, the network facilitates more effective information transfer and alleviates gradient vanishing or overfitting issues. This design choice improves training stability and overall robustness in radar signal recognition under varying conditions.

2. Related work

2.1 *Introduction of radar signal model*

High-Resolution Range Profile represents the magnitude of coherently combined complex returns from various target scattering points within each range bin. This one-dimensional signature effectively captures the distribution of radar echo intensity across the target's spatial extent, as illustrated in Figure 1.

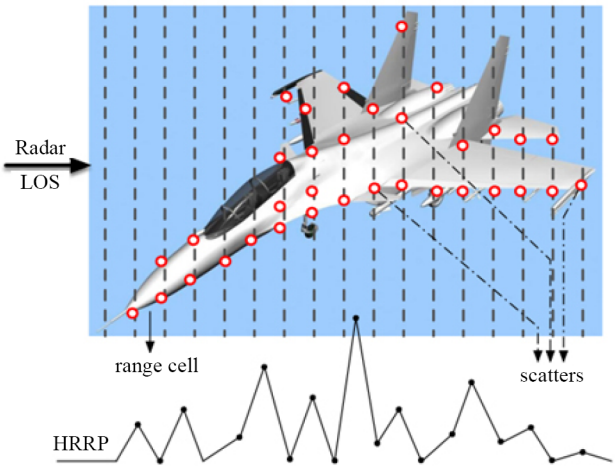

Figure 1. Illustration of an HRRP sample from an aircraft target

High-Resolution Range (HRR) radar systems typically operate in microwave frequency bands, where target dimensions substantially exceed the radar wavelength. When imaging complex objects like aircraft, these systems effectively partition the target into multiple range "cells" along the radar line-of-sight. Electromagnetic returns from all scattering centers residing within an individual range cell undergo coherent summation, producing a consolidated signature for that specific cell. As documented in previous research literature, this approach enables detailed structural characterization of complex targets [12, 13]. Given $s(\cdot)$ represents the complex signal envelope and f_c denotes the carrier angular frequency, the transmitted radar waveform can be expressed by $s(t) * e^{j*2\pi*f_c*t}$. Subsequently, the complex baseband echo return corresponding to the len th range resolution cell (where $len = 1, \cdots LEN$) can be mathematically approximated through the following expression:

$$x_{len}(t,\ LEN) \approx e^{j*\theta(n)} \sum_{i=1}^{V_{len}} {\sigma_{len}}^{i} * e^{j*{\varphi_{len}}^{i}(n)} \tag{1}$$

where $\theta(n) = -\frac{4*\pi}{\alpha} * R(n)$ signifies the beginning phase element of the nth echo return, directly correlating to both target range and electromagnetic wavelength, while V_{len} characterizes the total count of scattering centers contained within the len distance cell. The parameter ${\sigma_{len}}^{i}$ expresses the reflective strength coefficient of the ith scattering element positioned

in the len range bin, and ${\varphi_{len}}^{i}(n)$ denotes the remaining phase component associated with the ith reflection point situated in the len resolution cell of the nth sequential radar return. Then, the n complex HRRP can be represented as follows:

$$X(n) = e^{j*\theta(n)} * [\widetilde{x}_1(n),\ \widetilde{x}_2(n),\ \cdots,\ \widetilde{x}_{len}(n),\ \cdots,\ \widetilde{x}_{LEN}(n)]^T \tag{2}$$

$$\widetilde{x}_{len}(n) = \sum_{i=1}^{V_{len}} {\sigma_{len}}^{i} * e^{j*{\varphi_{len}}^{i}(n)} \tag{3}$$

where $\widetilde{x}_{len}(n)$ represents the expanded form of the len signal and $X(n)$ is the n complex HRRP.

The n time domain HRRP named $X_{\text{time}}(n)$, which is obtained by taking absolute value of, which can be represented as:

$$X_{\text{time}}(n) = [|\widetilde{x}_1(n)|,\ \cdots,\ |\widetilde{x}_{len}(n)|,\ \cdots,\ |\widetilde{x}_{LEN}(n)|]^T \tag{4}$$

where $|\cdot|$ denotes the amplitude extraction operation. This profile $X_{\text{time}}(n)$ characterizes the distribution of radar echo intensity across range cells along the radar line-of-sight direction. The energy contained within the nth range cell component $\widetilde{x}_{len}(n)$ of profile $X_{\text{time}}(n)$ can be quantified through the following mathematical formulation:

$${\tilde{x}_{len}}^{2}(n) = \sum_{i=1}^{V_{len}} \left({\sigma_{len}}^{i}\right)^2 e^{j*{\varphi_{len}}^{i}(n)} + 2\sum_{i=1}^{V_{len}}\sum_{k=1}^{i-1} {\sigma_{len}}^{i}{\sigma_{len}}^{k}\left[{\varphi_{len}}^{i}(n) - {\varphi_{len}}^{k}(n)\right] \tag{5}$$

As evidenced in Equation (3), the "speckle effect" emerges from the complex interplay between amplitude variations, phase fluctuations, and coherent interference among discrete scattering centers. Infinitesimal modifications in target aspect angle induce significant phase perturbations across individual scattering elements, resulting in substantial amplitude modulations at intensity maxima (characterized as amplitude instability phenomena). Paradoxically, despite these pronounced intensity fluctuations, the spatial coordinates of these maxima exhibit remarkable invariance (demonstrating structural persistence properties). This characteristic dual behavior of HRRP signatures-amplitude variability coupled with positional stability.

This study deliberately focuses on the amplitude characteristics of time-domain HRRP signatures and their associated spectral representations. This approach is particularly advantageous because these signal properties remain stable despite variations in initial phase components. By emphasizing amplitude-based features rather than phase-dependent characteristics, the extraction process maintains reliability even when confronted with the phase variability inherently present in raw radar returns. This strategic limitation in analysis scope significantly enhances the robustness of the proposed recognition methodology.

2.2 *Radar recognition*

In recent years, deep learning network models for target recognition have increasingly adopted composite structures, combining multiple network configurations to enhance performance. For instance, Kong et al. [3] integrates an autoencoder with a CNN for feature extraction and classification, leveraging the deep network's complex function representation. An end-to-end multi-input convolution gated recurrent unit neural network was proposed in [4], utilizing gated units to extract time and frequency domain features from the HRRP sequence. Pan et al. [7] introduced a stacked CNN-BIRNN model with an attention mechanism, incorporating a dynamic adjustment layer, CNN, and squeeze excitation block into a bidirectional RNN to improve radar echo signal feature extraction. He et al. [8] used CNN and

bidirectional long short-term memory networks to extract envelope and time features, further processed by a squeeze excitation block to enhance key information. Additionally, Su et al. [10] proposed a one-dimensional pyramid convolution neural network that employs pyramid convolution for local feature extraction and depth separable convolution to reduce parameters. With the advancement of electronic countermeasure technology, radar signals have evolved from simple to complex formats.

Researchers are focusing on using neural networks to extract reliable intra-pulse modulation features from signals [14]. Yan et al. [15] employ a stacked denoising autoencoder to learn the robust representation of the original HRRP. Feng et al. [16] employ the average profile as the correction terms and stack a series of corrective autoencoders to extract features from HRRP. However, these two models are based on fully connected nets and may not capture the structural information among the range cells of HRRP layer by layer, since HRRP reflects the distribution of scatterers in the target along the range dimension. Differently, CNN explicitly exploit structural locality in the feature space and can extract more descriptive features that exist in signals [17].

3. Recognition method based on ATCNN

3.1 *Preprocessing and average profile*

The employment of unprocessed HRRP signatures as direct neural network inputs is predicated on their comprehensive preservation of structural information integrity. Several fundamental challenges warrant consideration when implementing HRRP-based radar target classification systems. Foremost among these is the temporal displacement sensitivity inherent to HRRP. To optimize computational efficiency, HRRP extraction typically involves a selective range window that encompasses the target signature within the broader radar return. Consequently, the positional registration of target features within the HRRP frame exhibits measurement-dependent variability. Given that effective feature learning necessitates parameter uniformity across the training corpus, absolute alignment methodologies present advantageous solutions through their capacity for independent processing of individual HRRP instances. Comparative evaluation of alternative alignment techniques [18] revealed negligible performance differentials between methodologies. Accordingly, this investigation adopts centroid-based alignment for temporal displacement compensation due to its implementation simplicity and operational efficiency.

The amplitude scaling variability characteristic of HRRP derives from multiple factors influencing return intensity, including transmission power, target-radar separation, antenna gain parameters, receiver sensitivity, and system-specific losses. Consequently, HRRP measurements obtained across different radar platforms or operational conditions manifest inconsistent amplitude scaling. To mitigate this variability, each HRRP undergoes normalization through L2-norm division. The resultant processed signature can be mathematically expressed as:

$$h_n = \frac{1}{LEN} \sum_{len=1}^{LEN} |h_{len}(n)|^T \tag{6}$$

$$H_{\text{total}} = \{h_n\}_{n=1}^{N},\ h_n = [|h_1(n)|, \cdots, |h_{len}(n)|, \cdots, |h_{LEN}(n)|]^T \tag{7}$$

where H_{total} are HRRP data after L2-norm, with n HRRP sample. N represents the dimension of the HRRP sequence and h_n represent the amplitude of N range cells of the HRRP. The illustration demonstrates that individual HRRP instances exhibit substantial variability across significant target-aspect angle differences, whereas the composite average profile manifests a more continuous and refined signal morphology, thereby providing enhanced representation of the target's electromagnetic scattering characteristics within the specified aspect-angle sector.

3.2 *Adaptive threshold convolution neural network*

The feature extraction unit designed in this paper, namely the ATCU, as shown in Figure 2 serves as the core of the ATCNN, and its structure is shown in Figure 3. ATCU first uses three convolution kernels of different sizes (1, 3, and 5) for multi-scale feature extraction. The features are normalized by the batch normalization layer to stabilize the gradient propagation and alleviate the overfitting problem. The adaptive threshold function is used to realize anti-interference feature extraction and increase the nonlinear ability of data. On this basis, the channel attention mechanism layer is introduced to weigh the features, and the feature importance between channels is adaptively learned. ATCU also employs residual connections to protect the integrity of the original features and enhance the network performance. By adding the original features directly to the processed features, additional paths can be provided, allowing more adequate transfer of information. In addition, ATCU can be set to not perform data downsampling, to avoid the problem that the feature map size is too small due to the deep network, which limits the feature extraction.

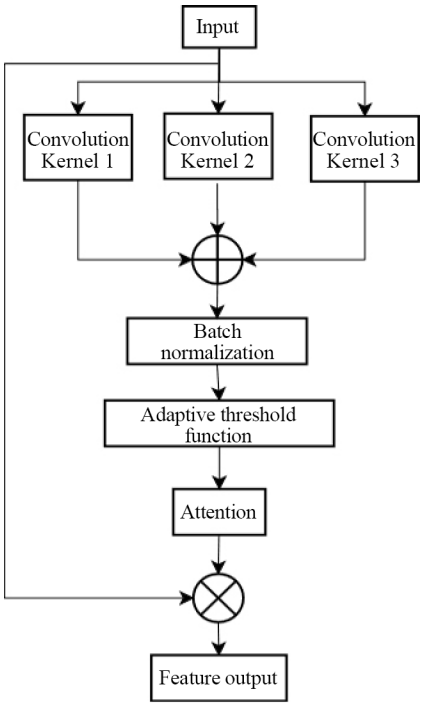

Figure 2. Adaptive threshold convolution unit

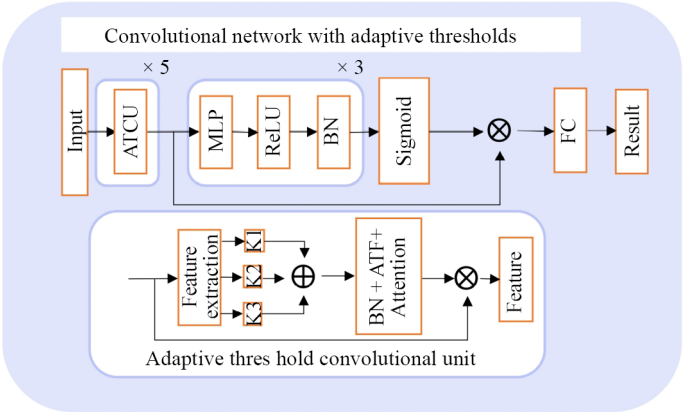

Figure 3. Adaptive threshold convolution neural network

3.2.1 *Adaptive threshold function*

In the radar signal recognition task, the soft threshold function is usually embedded into the deep learning network model as a nonlinear activation function. However, using the soft threshold function as the activation function has some problems, that is, its nonlinear characteristics are too weak.

The expression of the soft threshold function is given by:

$$f(x) = sgn(x) * \max\left[|x| - \text{threshold},\ 0\right] \tag{8}$$

where sgn$(\cdot)$ is a sign function, *threshold* is the threshold value greater than 0. When the threshold value threshold is close to 0, the soft threshold function degenerates into a linear function, which is contrary to the original purpose of the activation function, which is to introduce nonlinear properties. To enhance the nonlinear ability of the model, the following adaptive threshold activation function is designed by referring to Leaky-Rectified Linear Unit (ReLU) activation function:

$$g(x) = \begin{cases} x - \text{threshold}, & \text{threshold} \le x \\ 0, & -\text{threshold} \le x \le \text{threshold} \\ \alpha * (x + \text{threshold}), & x \le \text{threshold} \end{cases} \tag{9}$$

where threshold is an adaptive threshold, which can be learned through the network. α is a constant and the derivative function of the adaptive threshold function is 1 or 0 or α, which can effectively prevent the gradient disappearance and gradient explosion problems. As shown in Figure 4, a comparison between the Rectified Linear Unit (ReLU) function and Adaptive Threshold Functions (ATF) with different parameter values. The X-axis represents the input values ranging from -10 to 10, while the Y-axis represents the function values within the same range. The ReLU function, shown by a solid blue line, stays at 0 for negative inputs and increases linearly for positive inputs. In contrast, the ATF functions, represented by dashed red, dash-dotted green, and dotted orange lines for values of 1, 0.5, and 0.1 respectively, exhibit varying degrees of slope and smooth transitions around the origin, indicating different levels of sensitivity and threshold adaptation compared to the ReLU function. In this paper, α is set to 0.5, and threshold is set to 2.5.

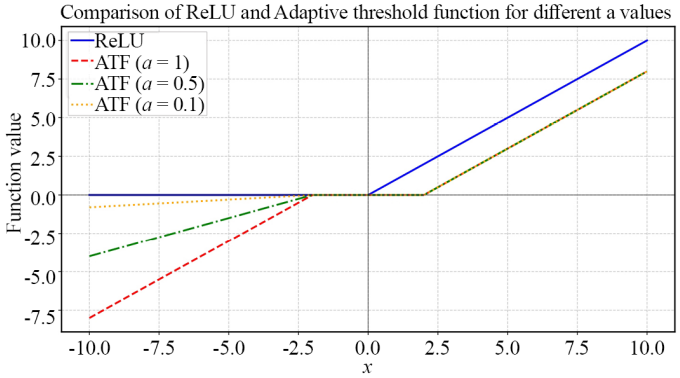

Figure 4. Adaptive threshold function

The adaptive threshold function presents several key innovations that significantly improve the performance of the ATCNN. First, its dynamic threshold adapts based on the target Signal-to-Interference Ratio, allowing for effective filtering and denoising of radar signals in real-time. This dynamic adjustment helps reduce the impact of interference signals, enhancing the model's ability to recognize targets even in challenging environments with varying levels of noise. Second, compared to traditional activation functions like ReLU, which can lead to "dead neurons" when output values are set to zero, the adaptive threshold function avoids this issue by maintaining the flexibility for neurons to be reactivated. This ensures the network's robustness and prevents the model from losing its capacity to learn from data, which is particularly important in complex radar signal environments where features may change over time. Lastly, when compared to the soft threshold activation function, the adaptive threshold function exhibits more pronounced nonlinear characteristics. These enhanced nonlinearities enable the model to better capture complex relationships in radar data, making the adaptive threshold function more suitable for use in signal processing tasks, where intricate patterns and variations in the data need to be accounted for. By incorporating these dynamic and nonlinear elements, the proposed activation function provides a more powerful mechanism for feature extraction and classification, particularly in the presence of interference.

3.2.2 *Multi-scale convolution layers*

Different sizes of convolution kernels will extract different feature information for the same input image. Generally, large-size convolution kernels can pay more attention to the global information of the image, which can capture the saliency and large-scale context information of the target, and also pay attention to the connection between adjacent data. On the contrary, small-size convolution kernels are more suitable for highlighted details and can locate small-scale features of object boundaries more accurately. Therefore, the combination of convolution kernels of different sizes can extract richer features of the data.

Therefore, this paper uses the modified multi-scale convolution shown in Figure 5, k represents the size of different convolution filters. Since the input HRRP data is in sequence format, the convolution kernel used is in the form of 1D. Convolution kernels of different sizes are used to construct the convolution layer, from up to down, the kernel size gradually increases and the depth gradually decreases. For large kernels, they pay more attention to the overall characteristics of the data. For each convolution kernel, all input channel data are simultaneously applied to obtain the feature map of the output channel. The feature maps obtained by each convolution kernel are linearly added to obtain the final feature data.

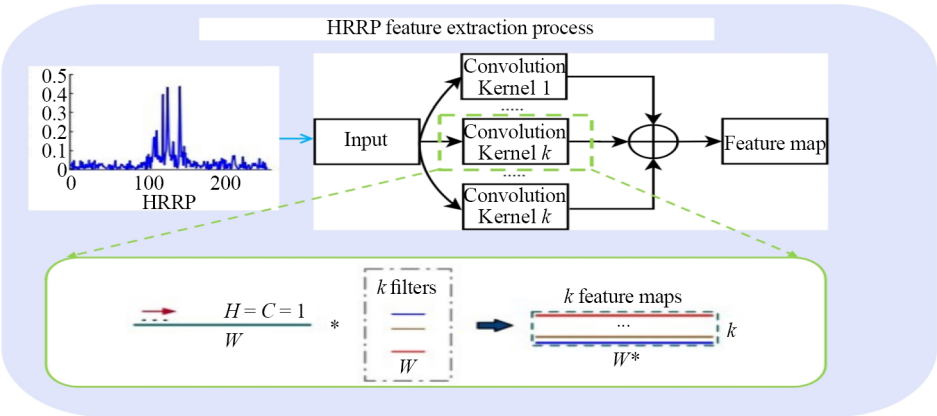

Figure 5. Illustration of HRRP feature extraction

To be specific, assuming that for every single channel of input H_{total} that has the size of $1 \times W_{\text{in}}$, the output channel size of different kernels is set to be W_{out}:

$$F_H = Sum\left(\sum_{k=1}^{K} Conv_{1\times k}\left(H_{\text{total}}\right)\right) \tag{10}$$

where F_H represents the feature map extracted by multiple convolution kernels and $Conv_{1\times k}$ denotes operation of the k filters. To make the feature maps of the output channels can be linearly superimposed on the channel dimension, it is necessary to ensure that the output length W_{out} of different convolution sizes is consistent. So the convolution output length W_{out} is known to be calculated as follows:

$$W_{\text{out}} = \frac{W_{\text{in}} + P - F + 1}{S} \tag{11}$$

where P is the number of zero complements, F is the convolution kernel size, and S is the step size. In this paper, S is set to 1 and P is set to 2.

By using convolution kernels of different sizes, the network can pay attention to the context information of the data and the detailed characteristics of the data at the same time. By fusing feature information of different scales, the network can obtain better feature representation.

3.2.3 *Attention module*

With the wide and successful application of attention mechanisms in the field of deep learning, more researchers have proposed new attention mechanisms, including neighborhood attention [19], fast attention mechanism [20], etc. In this paper, we refer to the channel attention mechanism block proposed by Jie et al. [21] and modify it to be suitable for the one-dimensional convolution network of this paper, as shown in Figure 6.

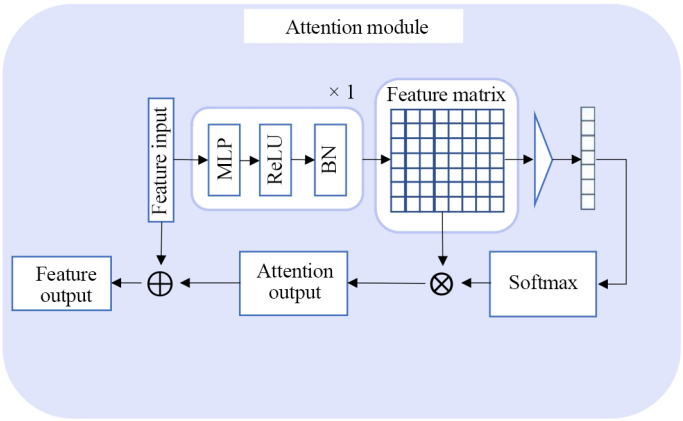

Figure 6. Attention module

Channel attention improves the perception of features by adjusting the weight of each channel so that the model can adaptively learn the importance of each channel, reduce the interference of irrelevant information, and improve the generalization ability and robustness of the model.

Supposed that the feature input F_{in} that has the size of $N \times W_{\text{in}}$, Multi-Layer Perceptron (MLP), ReLU and Batch Normalization (BN) is first applied to get F_m, which the size is $N \times M \times W_{\text{in}}$. Max pooling is then used to extract attention coefficients $A_M = \{a_1, \cdots, a_M\}$, which indicates the importance of sequence. For each different channel,

softmax operation is applied to its attention coefficient and combined with the feature matrix to obtain its weighted feature map.

$$AO = F_m * \text{Softmax}(A_M) \tag{12}$$

$$\text{Softmax}(A_M) = \frac{A_M}{\sum A_M} \tag{13}$$

With the increase of network depth, the data size gradually decreases, the data information is missing, and the features that the network can learn become less, which makes the network more difficult to train and causes problems such as overfitting, gradient disappearance, and gradient explosion. To solve this problem, the convolutional network in this paper uses residual connections [22]. The residual connection can be simply expressed by the formula:

$$F_{\text{out}} = F_{\text{in}} + \text{AttentionModule}(F_{\text{in}}) = F_{\text{in}} + AO \tag{14}$$

3.3 *Metric*

When the final network output is obtained, the results are classified by using FC layers, the whole process can be modeled as:

$$C_{\text{out}} = FC(Net_\theta(H_{\text{total}})) \tag{15}$$

where Net_θ is the Net parametrized by θ. The loss function $Loss$ of the model is cross entropy loss, which can be expressed as:

$$Loss = -\sum_{i=1}^{C} P_i^{GT} * \log(P_i^{pre}) \tag{16}$$

where P_i^{GT} stands for the probability distribution of the Ground Truth (GT) of i^{th} class and P_i^{pre} represents the probability distribution the model generated for the i^{th} class.

In the process of neural network model training, the training effect usually needs to be evaluated by calculating the F1-score, Precision, and recall.

TP means the instance is positive and predicted to be positive, FP means the instance is negative and predicted to be positive, FN means the instance is positive and predicted to be negative, and TN means the instance is negative and predicted to be negative. In the binary classification problem, TP and TN both represent the situation that the prediction result is accurate, so *recall* and *Precision* can be calculated as follows:

$$recall = \frac{TP}{TP + FN} \tag{17}$$

$$Precision = \frac{TP}{TP + FP} \tag{18}$$

and $Accuracy(ACC)$, F_1-score can be calculated by:

$$ACC = \frac{TP + TN}{TP + FP + TN + FN} \tag{19}$$

$$F_1 = 2 * \frac{Precision * recall}{Precision + recall} \tag{20}$$

4. Experimental results

4.1 *Datasets*

Our research includes computational modeling using four polarization modes (HH, HV, VH, VV) and advanced 3D rendering software for electromagnetic signature analysis of three different aircraft (F15, F18, and Indigenous Defense Fighter (IDF)). As shown in Table 1 these aircraft have a significant structural diversity: the F15 platform has a superior size (vertical section: 5.65 m, longitudinal extension: 19.45 m, lateral span: 13.05 m), while the F18 has a medium proportion (vertical section: 4.66 m, longitudinal extension: 17.07 m, lateral span: 11.43 m), the area of the IDF is significantly reduced (vertical section: 4.70 m, longitudinal extension: 14.48 m, transverse span: 8.53 m). The experimental setup uses 9.5-10.5 GHz electromagnetic radiation (equivalent to a wavelength of approximately 3.2-2.9 cm) to step in a sequence of 10 MHz, capturing elevation angles of 75° to 105° at 3° intervals and horizontal angles of 0° to 60° at a gradient of 0.05°. These structural differences manifest as unique electromagnetic reflection patterns under X-band illumination, producing feature-detailed range-resolution features. The experimental framework generates paired data sets for each aerial platform, structured as $2 \times 900 \times 501$ and $2 \times 300 \times 501$ matrices, where the initial value represents the number of elevation views and the middle value (900/300) represents the horizontal views of the strategy sampling (deliberately limited to a combined 1,201 horizontal views to check the limited training data conditions). The terminal value (501) represents the spatial resolution element in each feature profile. The development dataset combines the horizon-horizontal polarization information of the initial elevation view of all horizontal viewpoints, while the evaluation dataset utilizes the horizon-horizontal polarization information of the fifth elevation view of the corresponding horizontal viewpoint, both of which maintain the same dimensional structure. Substantial morphological differences between these aerial platforms provide a solid basis for a comparative analysis approach, with physical properties directly modulating the electromagnetic reflection properties of the entire angular spectrum, ultimately defining the features captured in each resolution profile. Our dataset organization approach leverages these structural differences to build meaningful learning and validation scenarios for an electromagnetic-based platform identification technique, which shows the feature profiles of all aircraft at different elevation angles. The AdamW optimizer is used for backpropagation with an initial learning rate of 0.001, and weight decay, with Cosine Decay. The GPU is RTX A5000 (24 GB).

Table 1. Dataset parameters

Aircraft type	Height (m)	Length (m)	Width (m)
F15	5.65	19.45	13.05
F18	4.66	17.07	11.43
IDF	4.70	14.48	8.53

To evaluate the algorithm's robustness under challenging conditions, we constructed a specialized test dataset by applying systematic signal degradation to the original test samples, as shown in Figure 7. This evaluation framework incorporates two distinct and realistic interference mechanisms: Noise Amplitude Modulation (NAM) and Noise Phase Modulation (NPM), which are designed to more accurately reflect the complexities of electromagnetic interference in

practical radar applications. Unlike conventional models that rely on simplified noise distributions, such as Gaussian or Rayleigh noise, our approach aims to simulate the multifaceted nature of real-world interference, including jamming, clutter, and environmental factors. The experimental design involved selecting 1,400 samples from the original test corpus. Each sample was subjected to both interference types, resulting in a comprehensive evaluation set of 2,800 noise-contaminated radar signatures. This method ensures that the evaluation process better captures the diverse and dynamic nature of interference found in operational settings. The mathematical formulations governing these interference mechanisms are defined as follows:

$$\begin{cases} Noise_a = [U_o + U_n(t)] \cdot \cos(w \cdot t + \varphi) \\ \\ Noise_p = U_p \cdot \cos\left(w \cdot t + K_{PM} u(t) + \varphi\right) \end{cases} \tag{21}$$

where w is the center frequency of the signal, and φ is the uniform distribution between $[0,\ 2\pi]$. As shown in Figure 7, for NAM $Noise_a$, $U_n(t)$ is a generalized stationary random process with zero mean in $[-U_o,\ \text{inf}]$. U_o is a constant and set to be one. For NPM $Noise_a$, U_p is the amplitude constant and K_{PM} is the frequency modulation slope. $u(t)$ is the modulated noise, which is a zero mean generalized stationary random process.

These noise models play a key role in accurately simulating the types of interference encountered in radar signal processing. In contrast to traditional Gaussian or Rayleigh noise models, which simplify the interference to only basic additive noise, the expressions presented here incorporate both amplitude and phase variations. These models more effectively capture the complexity of real-world electromagnetic interference, where radar signals are subject to diverse and dynamic disturbances, such as jamming, clutter, and multipath effects. This makes the models significantly more representative of the actual conditions that modern radar systems face, which are far more complex than the assumptions of conventional noise models.

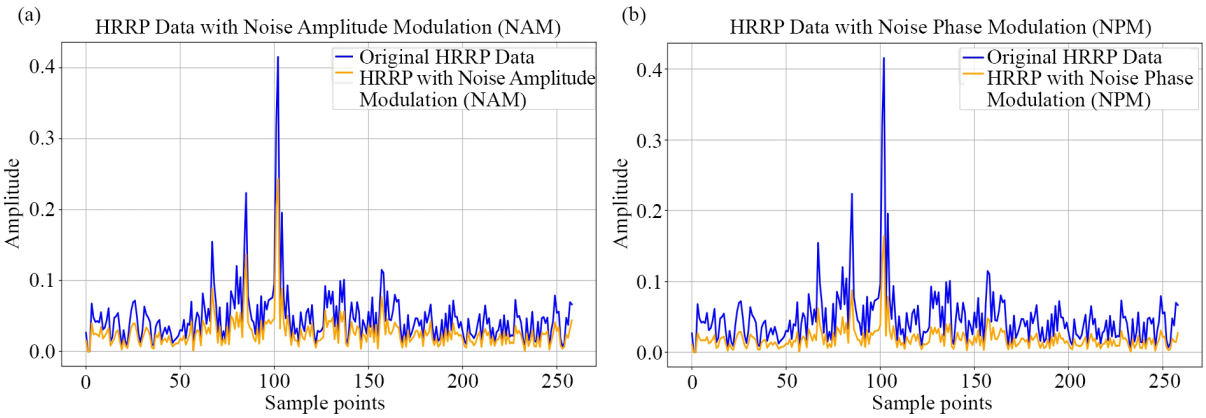

Figure 7. HRRP with noise (a) NAM interference (b) NPM interference

4.2 *Comparative experiments with other methods*

In this section, comparative experiments can be conducted between our proposed ATCNN and Auto-Encoder (AE), CNN-BIRNN, PyDSC and Linear Discriminant Analysis (LDA) based on the above-mentioned dataset.

Feature extraction capabilities were assessed through dimensionality reduction visualization techniques applied to high-dimensional feature representations extracted from penultimate network layers, as illustrated in Figure 8. The visualization reveals that ATCNN-derived features exhibit enhanced spatial compactness coupled with superior inter-class

separability, indicating more effective discriminative feature extraction capabilities. This characteristic translates directly into enhanced classification performance in subsequent decision processes.

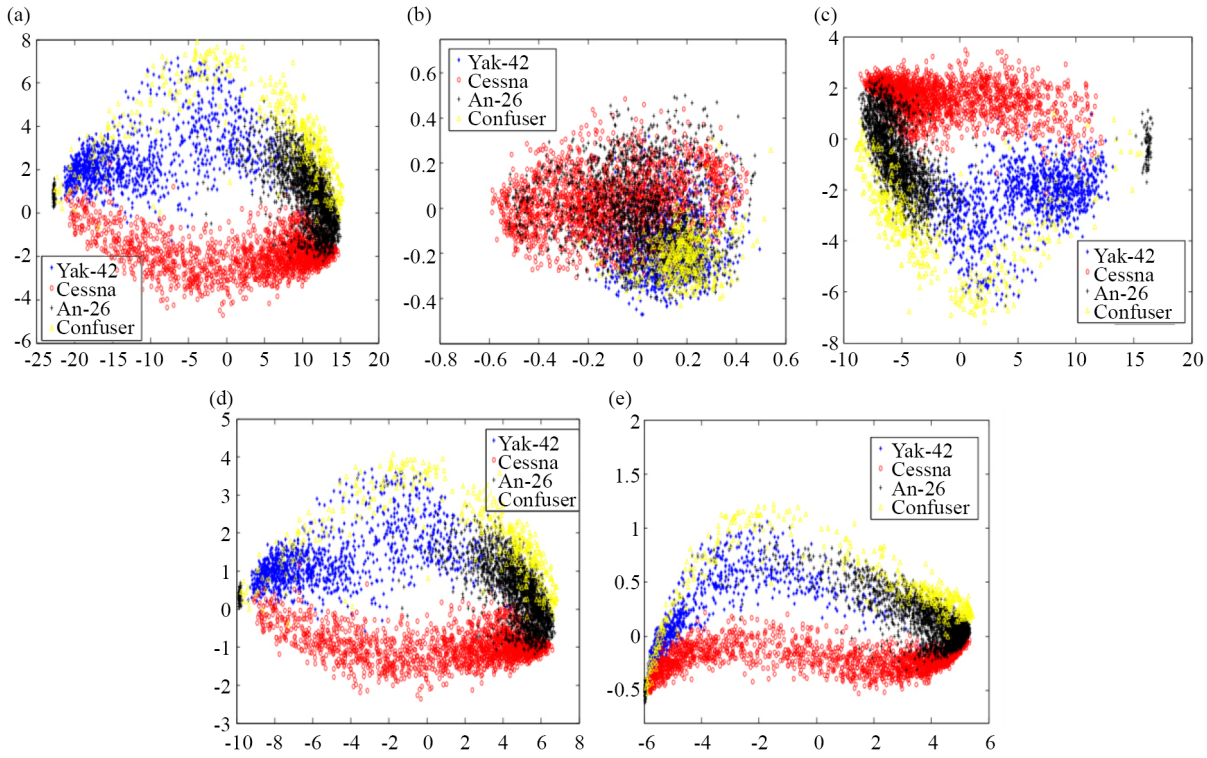

Figure 8. Visualizations of the test HRRP samples and their corresponding features

The quantitative performance metrics presented in Table 2 offer a comprehensive comparative analysis across several key evaluation criteria, including accuracy, F1-score, precision, and recall. The results clearly indicate that the ATCNN consistently outperforms the other methods across all evaluation metrics. Notably, ATCNN achieves the highest classification accuracy of 96.81%, surpassing the other methods by a substantial margin, which highlights the robustness and effectiveness of the proposed model. In terms of recall, ATCNN demonstrates superior performance with 92.84%, indicating its capability to effectively identify true positives in complex radar signature classification tasks. Furthermore, ATCNN achieves the highest precision at 95.54%, reflecting its ability to minimize false positives. These results are complemented by an outstanding F1-score of 0.94, which balances precision and recall, further solidifying the model's reliability. In contrast, the other methods, such as AE, Convolutional Neural Network-Bidirectional Recurrent Neural Network (CNN-BIRNN), Pattern-Discriminative Subspace Clustering (PDSC), and LDA, exhibit lower performance across these metrics. For example, PDSC achieves an accuracy of 93.63%, significantly trailing behind ATCNN. Similarly, other methods show lower precision, recall, and F1-scores, reinforcing the superior stability and performance of ATCNN in handling the challenges of radar signal classification. These findings highlight the methodological advantages of ATCNN in addressing the complexities of radar signature classification and demonstrate its superior efficacy in comparison to existing techniques.

Table 2. Computation time and metric of different methods

Methods	ATCNN	AE	CNN-BIRNN	PyDSC	LDA
Training time (s)	0.09	0.19	0.73	0.80	0.63
Test time (s)	0.08	0.11	0.26	0.51	0.86
Accuracy (%)	96.81	94.30	93.70	93.63	89.05
Recall (%)	92.84	89.71	88.45	85.67	82.58
Precision (%)	95.54	91.14	89.66	87.75	86.96
F1-Score	0.94	0.90	0.885	0.86	0.83

To compare various aspects, we introduce the Receiver Operating Characteristic Curve (ROC) for comparison. ROC is a commonly used tool to evaluate the performance of classification models. It helps to judge the effect of the classification model by showing the relationship between the True Positive Rate (TPR) and False Positive Rate (FPR) under different thresholds. Specifically, the performance of different classification models can be visually compared through the ROC curve. The closer the curve is to the top-left corner, the better the model is at classifying. The specific ROC results are shown in Figure 9, we can notice that the curve of ATCNN is closer to the top-left corner, indicating that ATCNN performs better in classification tasks. A larger Area Under the ROC Curve (AUC) signifies a better classification performance of the model. Therefore, the AUC value of ATCNN may be higher than other comparison methods, further proving the superiority of its classification ability.

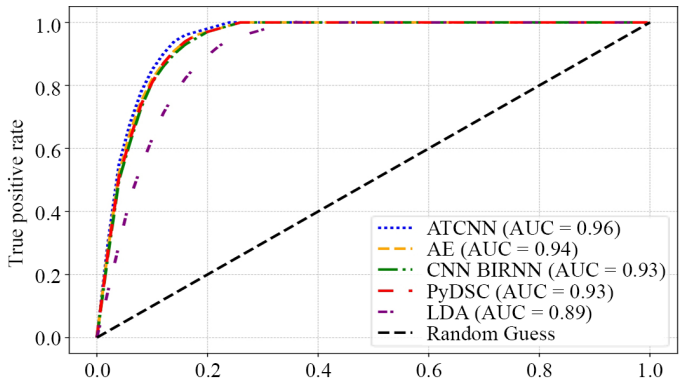

Figure 9. Evolution of ROC for different methods

At the end of this section, this paper tests the ability of different methods to extract data features. Specifically, the classification ability of different algorithms under different sampling rates is analyzed by sampling the data. Usually, the number of azimuth angles will be sampled to form the sampled data set. The purpose of this is to verify the discriminative ability of the model under different azimuth angles. We define the sampling rate as:

$$rate = \log_2(1,024/Num),\ Num = \{16,\ 32,\ 64,\ 128,\ 256,\ 512,\ 1,024\} \tag{22}$$

where Num denotes the number of azimuth angles after sampling. Therefore, we can obtain datasets with different sampling rates. The results of the qualitative analysis are shown in Figure 10. The results of the quantitative analysis are shown in the table below. It can be observed from Figure 10 that the accuracy of various methods decreases as the sampling rate declines. However, ATCNN maintains relatively high accuracy across different sampling rates, with a comparatively smaller decline. This indicates that ATCNN possesses strong capabilities in data feature extraction, even

under conditions of reduced data volume. The results presented in Table 3 further support this conclusion, showing that ATCNN consistently outperforms other comparison methods at varying sampling rates. Especially when the sampling rate is 0 (i.e., using all data), the accuracy of ATCNN reaches 96.75%, significantly higher than the accuracy of other methods. This once again demonstrates the superior performance of ATCNN in radar signal recognition tasks.

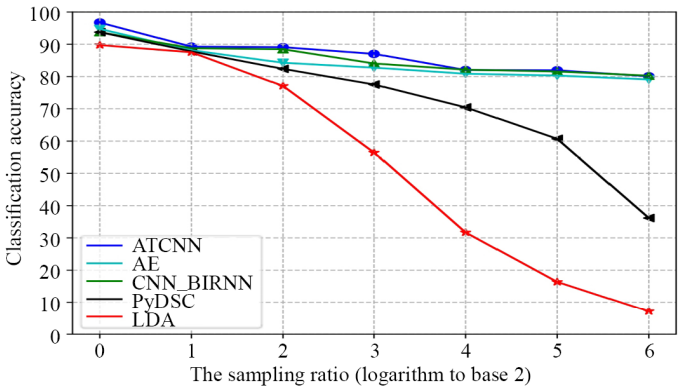

Figure 10. Evolution of classification accuracy for different methods

Table 3. Comparison of accuracy of different methods

rate	ATCNN	AE	CNN-BIRNN	PyDSC	LDA
0	96.75	94.75	93.74	93.75	89.75
1	89.26	88.16	88.78	87.80	87.52
2	89.10	84.27	88.46	82.38	77.09
3	87.02	82.75	84.04	77.53	56.44
4	82.02	80.86	82.11	70.45	31.69
5	81.94	80.31	81.56	60.70	16.35
6	80.07	79.09	80.26	36.16	7.15

4.3 *Based on test in case of noise interference*

In this section, comparative experiments can be conducted between our proposed ATCNN and Auto-Encoder (AE), CNN-BIRNN, PyDSC and Linear Discriminant Analysis (LDA) based on the above-mentioned dataset.

To study the performance of the network model in anti-interference identification, this paper uses MATLAB simulation to obtain three types of suppressed interference signals: noise FM interference, noise AM interference, and noise phase-modulated interference and sets up a total of nine signal-to-density ratio experimental conditions of [0 dB, 5 dB, 10 dB, 15 dB, 20 dB, 25 dB, 30 dB, 35 dB, 40 dB]. Suppressed interference was added to all HRRP raw data. For each HRRP data, the interference is generated independently with some randomness. The accuracy curves of different methods based on NAM are shown in Figure 11a. From the Figure 11a, it can be seen that as the Signal-to-Noise Ratio (SNR) increases, the recognition accuracy of all algorithms gradually rises. However, the recognition accuracy curve of ATCNN consistently lies above those of other algorithms, indicating its higher recognition accuracy and stronger robustness in dealing with NAM interference. Additionally, when the SNR is low (e.g., 0 dB to 10 dB), the gap in recognition accuracy between ATCNN and other algorithms is more significant, further demonstrating the superior performance of ATCNN under low SNR conditions. The accuracy curves of different methods based on Noise Amplitude Perturbation (NAP) are shown in Figure 11b. Similar to Figure 11a, as the SNR increases, the recognition accuracy of all algorithms gradually rises. However, ATCNN maintains its lead in recognition accuracy. This reaffirms the superior

performance of ATCNN in handling NAP interference. Furthermore, from the figure, it can be observed that when the SNR is low, the gap in recognition accuracy between ATCNN and other algorithms is also larger, further validating the robustness and accuracy of ATCNN under low SNR conditions.

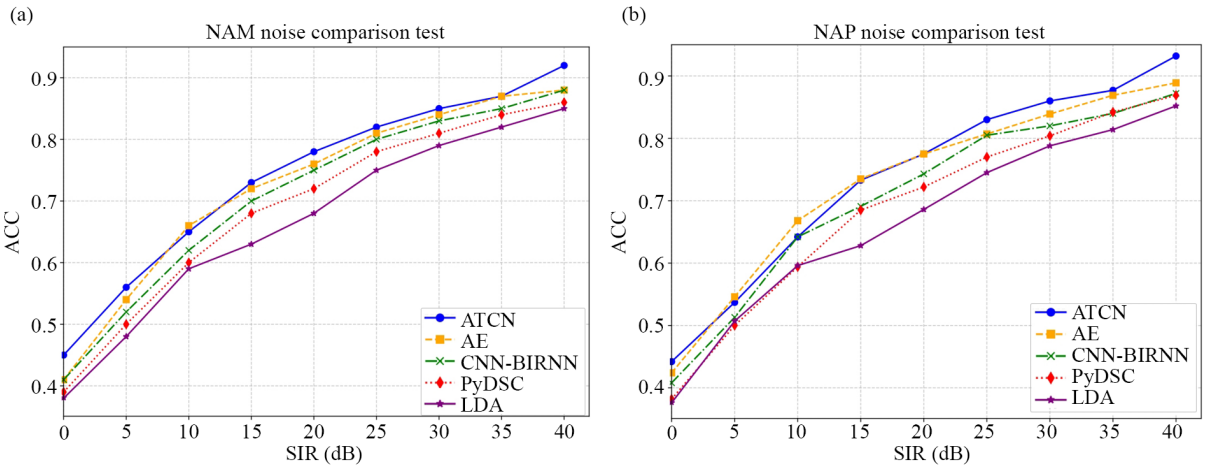

Figure 11. Interference experiments with different signal-to-noise ratios

4.4 *Ablation experiment*

To thoroughly evaluate the contributions of each component in the ATCNN architecture, two comprehensive ablation studies were conducted. These experiments were designed to isolate and assess the impact of individual modules on the overall performance of the model.

The first set of experiments, as presented in Figure 12a, investigates the performance of the ATCNN model with and without its key modules: the adaptive threshold function, attention module, and Multi-scale Convolutional Layers (MCL). The objective of this analysis was to determine the significance of each component in enhancing the model's ability to extract features, maintain robustness, and achieve high classification accuracy. The results demonstrate that the full ATCNN model, which integrates all modules, achieves the highest performance with a recognition accuracy of 0.9675. Notably, the removal of any of these essential modules leads to a significant reduction in accuracy, emphasizing their crucial role in the model's robustness. Specifically, the absence of the ATF impairs the model's ability to effectively manage dynamic signal-to-interference ratios, while the exclusion of the AM hampers the network's capacity to focus on key features, resulting in decreased classification performance. These findings highlight the indispensable role of each component in ensuring accurate and efficient radar signal recognition.

The second set of experiments, shown in Figure 12b, was designed to evaluate the efficacy of the proposed activation function by comparing it to several commonly used activation functions, including Sigmoid, ReLU, and Leaky ReLU. In this experiment, the proposed activation function was substituted with each of these alternative functions, and the effect on model accuracy was assessed. The results clearly indicate that the proposed activation function outperforms the others, achieving the highest accuracy of 0.9675, compared to accuracy values of 0.9511, 0.9482, and 0.9521 for Sigmoid, ReLU, and Leaky ReLU, respectively. This substantiates the superior suitability of the proposed activation function for radar signal recognition tasks, as it effectively mitigates the noise and interference typically encountered in real-world radar data. The superior performance of the proposed activation function can be attributed to its adaptive characteristics, which enable the network to better accommodate varying signal conditions and preserve critical feature information.

In conclusion, these ablation experiments not only demonstrate the critical importance of the individual components within the ATCNN architecture but also validate the superior performance of the proposed activation function in enhancing model accuracy for radar signal recognition.

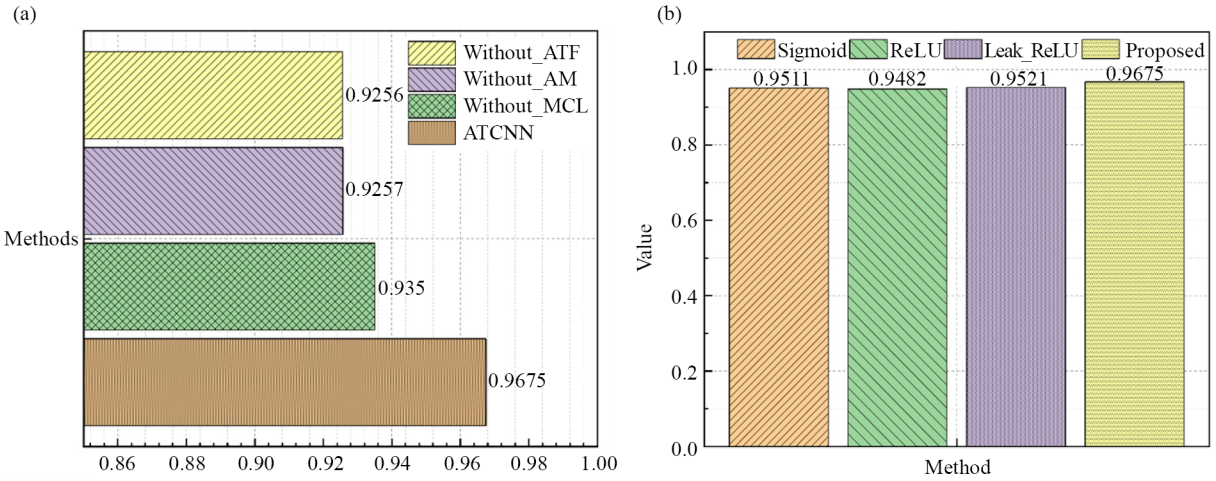

Figure 12. Ablation experiments on ATCNN (a) Module ablation, (b) Activation function comparison

5. Conclusion

In this study, we introduced an adaptive threshold convolutional neural network designed to enhance radar signal recognition performance, particularly in complex electromagnetic environments. The core innovation of the ATCNN is the adaptive threshold convolution unit, which dynamically adjusts thresholds based on varying signal-to-interference ratios. This approach, combined with multi-scale convolution kernels and a channel attention mechanism, significantly improves feature extraction accuracy and robustness against interference.

Our experimental results demonstrated that the ATCNN outperforms traditional methods in both recognition accuracy and anti-interference capability. Specifically, the ATCNN achieved superior performance across various datasets, including those with added noise and interference, validating its robustness in real-world applications. Additionally, the ablation studies confirmed the critical contribution of each component, particularly the ATCU and attention mechanisms, to the overall effectiveness of the network.

In summary, the ATCNN offers a significant advancement in radar signal processing, providing a reliable and efficient solution for radar automatic target recognition in challenging conditions. Future work could explore the integration of Self-Supervised Learning (SSL) [23] techniques into the ATCNN framework. SSL refers to a class of machine learning methods that enable models to learn from unlabelled data by creating supervisory signals from the data itself. Unlike supervised learning, which requires large amounts of labeled data, SSL leverages the inherent structure of the data to generate labels or representations without manual annotation. SSL has gained traction in various fields due to its ability to improve model performance in scenarios where labeled data is scarce or expensive to obtain. For the ATCNN, applying SSL could enhance the model's ability to extract meaningful features from radar signals without relying on extensive labeled datasets. Techniques such as contrastive learning and generative pre-training could allow the ATCNN to learn from vast amounts of unlabelled radar data, improving its adaptability to dynamic environments and enhancing its recognition capabilities.

Acknowledgments

The authors express their thanks to the people assisting this work, and acknowledges the valuable suggestions from the peer reviewers. This research was supported by National Natural Science Foundation of China (No. 11974289): a comparative study of aerodynamic modeling of differences in speech production mechanisms between men and women.

Authors contributions

Methodology, Bozhi Qiu; software, Bozhi Qiu; validation, Bozhi Qiu; formal analysis, Bozhi Qiu; investigation, Bozhi Qiu; resources, Bozhi Qiu; data curation, Bozhi Qiu; writing-original draft preparation, Bozhi Qiu and Ying Li; writing-review and editing, Bozhi Qiu and Sheng Li; visualization, Sheng Li; supervision, Sheng Li; project administration, Ying Li; funding acquisition, Sheng Li. All authors have read and agreed to the published version of the manuscript.

Conflict of interest

The authors declare no competing interests.

References

[1] Jia S, Liao J, Xu M, Li Y, Zhu J, Sun W, et al. 3-D gabor convolutional neural network for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*. 2022; 60: 1-16. Available from: https://doi.org/10.1109/TGRS.2021.3087186.

[2] Rahman MM, Gurbuz SZ, Amin MG. Physics-aware generative adversarial networks for radar-based human activity recognition. *IEEE Transactions on Aerospace and Electronic Systems*. 2023; 59(3): 2994-3008. Available from: https://doi.org/10.1109/TAES.2022.3221023.

[3] Kong Y, Feng D, Zhang J. Radar HRRP target recognition based on composite deep networks. In: *2022 International Applied Computational Electromagnetics Society Symposium (ACES-China)*. Xuzhou, China: IEEE; 2022. p.1-5.

[4] Zeng Z, Sun J, Han Z, Hong W. Radar HRRP target recognition method based on multi-input convolutional gated recurrent unit with cascaded feature fusion. *IEEE Geoscience and Remote Sensing Letters*. 2022; 19: 1-5. Available from: https://doi.org/10.1109/LGRS.2022.3192289.

[5] Wan J, Chen B, Xu B, Liu H, Jin L. Convolutional neural networks for radar HRRP target recognition and rejection. *EURASIP Journal on Advances in Signal Processing*. 2019; 2019: 5. Available from: https://doi.org/10.1186/s13634-019-0603-y.

[6] Zhang Y, Xiao F, Qian F, Li X. VGM-RNN: HRRP sequence extrapolation and recognition based on a novel optimized RNN. *IEEE Access*. 2020; 8: 70071-70081. Available from: https://doi.org/10.1109/ACCESS.2020.2986027.

[7] Pan M, Liu A, Yu Y, Wang P, Li J, Liu Y. Radar HRRP target recognition model based on a stacked CNN-Bi-RNN with attention mechanism. *IEEE Transactions on Geoscience and Remote Sensing*. 2022; 60: 1-14. Available from: https://doi.org/10.1109/TGRS.2021.3055061.

[8] He J, Wang X, Song Y, Xiang Q. A multi-scale radar HRRP target recognition method based on pyramid depthwise separable convolution network. In: *2022 7th International Conference on Image, Vision and Computing (ICIVC)*. Xi'an, China: IEEE; 2022. p.579-585.

[9] Wu L, Hu S, Xu J, Liu Z. Ship HRRP target recognition against decoy jamming based on CNN-BiLSTM-SE model. *IET Radar, Sonar & Navigation*. 2024; 18(2): 361-378. Available from: https://doi.org/10.1049/rsn2.12507.

[10] Su K, Gong L, Wang G, Zhang L. Noise-robust radar HRRP target sequential recognition based on correlative scattering centers. *IEEE Geoscience and Remote Sensing Letters*. 2023; 20: 1-5. Available from: https://doi.org/10.1109/LGRS.2023.3289138.

[11] Liu X, Wang L, Bai X. End-to-end radar HRRP target recognition based on integrated denoising and recognition network. *Remote Sensing*. 2022; 14(20): 5254. Available from: https://doi.org/10.3390/rs14205254.

[12] Du L, Liu H, Bao Z, Zhang J. Radar automatic target recognition using complex high-resolution range profiles. *IET Radar, Sonar & Navigation*. 2007; 1(1): 18-26. Available from: https://doi.org/10.1049/iet-rsn:20050119.

[13] Du L, Liu H, Wang P, Feng B, Pan M, Bao Z. Noise robust radar HRRP target recognition based on multitask factor analysis with small training data size. *IEEE Transactions on Signal Processing*. 2012; 60(7): 3546-3559. Available from: https://doi.org/10.1109/TSP.2012.2191965.

[14] Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks. *Science*. 2006; 313(5786): 504-507. Available from: https://doi.org/10.1126/science.1127647.

[15] Yan H, Zhang Z, Xiong G, Yu W. Radar HRRP recognition based on sparse denoising autoencoder and multi-layer perceptron deep model. In: *2016 Fourth International Conference on Ubiquitous Positioning, Indoor Navigation and Location Based Services (UPINLBS)*. Shanghai, China: IEEE; 2016. p.283-288.

[16] Feng B, Chen B, Liu H. Radar HRRP target recognition with deep networks. *Pattern Recognition*. 2017; 61: 379-393. Available from: https://doi.org/10.1016/j.patcog.2016.08.012.

[17] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Communications of the ACM*. 2017; 60(6): 84-90. Available from: https://doi.org/10.1145/3065386.

[18] Feng B, Du L, Liu H, Li F. Radar HRRP target recognition based on K-SVD algorithm. In: *Proceedings of 2011 IEEE CIE International Conference on Radar*. Chengdu, China: IEEE; 2011. p.642-645.

[19] Hassani A, Walton S, Li J, Li S, Shi H. Neighborhood attention transformer. In: *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Vancouver, Canada: IEEE/CVF; 2023. p.6185-6194.

[20] Wazir S, Fraz MM. HistoSeg: Quick attention with multi-loss function for multi-structure segmentation in digital histology images. In: *2022 12th International Conference on Pattern Recognition Systems (ICPRS)*. Saint-Étienne, France: IEEE; 2022. p.1-7.

[21] Hu J, Shen L, Sun G. Squeeze-and-excitation networks. In: *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. USA: IEEE/CVF; 2018. p.7132-7141.

[22] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. USA: IEEE; 2016. p.770-778.

[23] Liu X, Zhang F, Hou Z, Wang Z, Mian L, Zhang J. Self-supervised learning: Generative or contrastive. *IEEE Transactions on Knowledge and Data Engineering*. 2023; 35(1): 857-876. Available from: https://doi.org/10.1109/TKDE.2021.3090866.