Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Modulation Recognition Method with Dual-Branch Feature Fusion Based on Lightweight CNN and Dual-Attention Network

Xiuping Song

410441527@qq.com

Chongqing Polytechnic University of Electronic Technology

Junxia Wang

Chongqing Polytechnic University of Electronic Technology

Youlong Gong

Chongqing Open University

Xianyou Yao

Chongqing Polytechnic University of Electronic Technology

Research Article

Keywords: AMR, Dual-branch Network, Dual-attention Mechanism, Feature Fusion, Low-SNR

Posted Date: October 19th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10668892/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Modulation Recognition Method with Dual-Branch Feature Fusion Based on Lightweight CNN and Dual-Attention Network

Xiuping Song[1*], Junxia Wang[1†], Youlong Gong[2†] and Xianyou Yao[1†]
[1]Communication Engineering Department, Chongqing Polytechnic University of Electronic Technology, No.76, Daxuecheng East Road, Shapingba District, Chongqing, China.
[2]Intelligent Manufacturing Department, Chongqing Open University, No.1 Hualong Avenue, Science and Technology Park, Jiulongpo District, Chongqing, China.
[*]Corresponding author.E-mail: 410441527@qq.com
[†]These authors contributed equally to this work.

Abstract

Aiming at the problems of high-order MPSK/MQAM signals being easily confused, poor robustness under low signal-to-noise ratio, and incomplete coverage of modulation types in automatic modulation recognition under complex electromagnetic environments, a dual-branch feature fusion modulation recognition method based on lightweight CNN and IQ-2AttNet is proposed. First, a unified mathematical model of MPSK, MQAM, and signals without IQ components is constructed to analyze the differences in IQ characteristics of different modulation signals. Then, a dual-branch parallel architecture of "wide-coverage preliminary screening + high-precision fine recognition" is designed, where the lightweight CNN performs fast preliminary screening for 12 signal types, and IQ-2AttNet focuses on fine recognition of 6 types of core orthogonal modulation signals. Finally, confidence-weighted decision fusion is adopted to output the optimal results. Ablation experiments show that the dual-branch structure and decision fusion strategy can effectively improve recognition accuracy and generalization ability. Comparative experiments demonstrate that the proposed method outperforms existing typical modulation recognition models under the same conditions. Experimental results indicate that the overall recognition accuracy of the proposed method reaches 86.1% over the SNR range from -10 dB to 25 dB. Especially over the low-SNR range from -8 dB to 0 dB, the average recognition rate is 71.8%, which is 1.5% to 17.5% higher than the comparison models, showing strong noise robustness.

Keywords: AMR, Dual-branch Network, Dual-attention Mechanism, Feature Fusion, Low-SNR

1. INTRODUCTION

Automatic Modulation Recognition is a critical technology in communication reconnaissance, electronic countermeasures, and spectrum monitoring, playing an important role in both military and civilian fields[1]. With the development of 5G/6G technologies, high-order modulation schemes such as MPSK and MQAM have been widely used. However, in complex electromagnetic environments with low-SNR, signal features are easily submerged by noise, leading to a significant decrease in recognition accuracy. Furthermore, in practical complex communication environments, in addition to digital modulated signals with IQ components such as MPSK/MQAM, there also exist amplitude-modulated signals such as PAM4 and SSB-AM, as well as frequency-modulated signals such as B-FM and GFSK. The diverse signal system poses higher requirements on the generalization ability of AMR technologies.

Traditional modulation recognition methods are divided into likelihood-based estimation and

manual feature-based categories. The former relies on prior information and is difficult to implement in practice, while the latter requires manual feature design with poor anti-noise and anti-channel distortion capabilities. In recent years, deep learning has been widely used due to its advantage of automatic feature extraction. However, most existing models adopt a single-channel structure, which is difficult to balance both coverage and recognition accuracy. Especially, the intra-class and inter-class recognition effects for high-order QAM/PSK signals are unsatisfactory, with obvious limitations.

For MQAM signal recognition, existing methods have room for improvement: Reference [3] systematically summarizes the mainstream models, datasets, and key issues of deep learning-driven AMR, providing an important reference for subsequent research; Reference [4] realizes MQAM recognition through fourth-order cumulants and instantaneous amplitude compactness without complex iteration, but the clustering radius relies on prior information, limiting engineering practicability; Reference [2] achieves blind MQAM recognition based on template matching with optimized constellation distribution entropy, but it highly depends on the pre-processing accuracy of carrier frequency and symbol rate with poor fault tolerance; Reference [5] recognizes MQAM signals combined with joint distribution adaptation and BP neural network, with an average accuracy of only 73.25% without considering the influence of SNR, leading to insufficient practicability; Reference [6] designs hybrid high-order moment features combined with DNN to resist Rician fading, although improving recognition performance at low-SNR, the feature calculation complexity is high, which is not conducive to real-time deployment. For joint recognition of MPSK and MQAM, relevant schemes still need improvement: Reference [7] realizes joint recognition based on constellation phase statistics and clustering, with recognition rate exceeding 94% at low-SNR, but only adapts to the specific TETRA system with insufficient generalization ability; Reference [8] applies AlexNet to unsynchronized MPSK/MQAM signal recognition, although effectively adapting to asynchronous scenarios, only 4 modulation signals are verified, and the model requires large-scale training samples and high inference latency, which cannot meet real-time requirements; Reference [9] combines high-order cumulants with improved constellation clustering, with recognition rate exceeding 90% at 10 dB SNR, but relies on accurate feature extraction and clustering parameter optimization, and the robustness needs to be improved.

Therefore, this paper proposes a dual-branch feature fusion modulation recognition method based on lightweight CNN and IQ-2AttNet, focusing on improving the recognition robustness in low-SNR environments. The main contributions are as follows:

(1) A dual-branch parallel architecture of “wide-coverage preliminary screening + high-precision fine recognition” is designed, which divides 12 modulation signals into 6 core orthogonal modulation signals with IQ components and 6 amplitude or frequency modulation signals without IQ components,

effectively solving the problem of easy confusion of high-order MPSK/MQAM signals.

(2) A lightweight CNN is constructed to realize rapid preliminary screening of all types of signals, with an overall recognition rate of 74.8% over the wide SNR range from -10 dB to 25 dB. After removing easily confused core signals, the recognition rate of other signals is about 92%, both the coverage and recognition robustness are better than those of existing references [7] and [9].

(3) An improved IQ-2AttNet dual attention network is proposed to conduct fine feature learning for 6 core easily confused MPSK/MQAM signals, with an average recognition rate of 73.6% (excluding BPSK), which significantly alleviates the bottleneck of poor generalization and low recognition accuracy of high-order signals in traditional schemes.

(4) The probability-weighted decision fusion strategy is adopted to integrate the dual-branch output, realizing the global optimal recognition of 12 signals. The overall recognition accuracy reaches 86.1% over the wide SNR range from -10 dB to 25 dB, and the performance remains stable in low-SNR environments, which greatly improves the adaptability and engineering practicability of the algorithm in complex communication scenarios.

2. Signal Modeling and IQ Feature Analysis

2.1. MPSK Signal Model

Multi-ary Phase Shift Keying (MPSK) carries digital information through multiple discrete phase states of the carrier, and its time-domain expression is shown in Formula 1:

$$S_{\mathrm{MPSK}}(t)=\sum_{n} g(t-nT_b)\cos(\omega_c t+\varphi_n) \tag{1}$$

Expanding via trigonometric identities yields Formula 2.

$$S_{\mathrm{MPSK}}(t)=\cos\omega_c t\sum_{n}\cos\varphi_n g(t-nT_b)-\sin\omega_c t\sum_{n}\sin\varphi_n g(t-nT_b) \tag{2}$$

Let $a_n=\cos\varphi_n$, $b_n=\sin\varphi_n$

The in-phase component (I-channel signal) and quadrature component (Q-channel signal) are defined as Formula 3 and Formula 4, respectively.

$$I(t)=\sum a_n g(t-nT_b)=\sum\cos\varphi_n g(t-nT_b) \tag{3}$$

$$Q(t)=\sum_{n} b_n g(t-nT_b)=\sum_{n}\sin\varphi_n g(t-nT_b) \tag{4}$$

Substituting Formula 3 and Formula 4 into Formula 2 yields the equivalent quadrature representation of the MPSK signal, i.e., the time-domain expression of the combined IQ signal, as shown in Formula 5:

$$S_{\mathrm{MPSK}}(t)=I(t)\cos\omega_c t-Q(t)\sin\omega_c t \tag{5}$$

where $g(t)$ is a rectangular pulse function with a height of 1 and a width of T_b; T_b is the duration of the M-ary symbol; ω_c is the carrier angular frequency; φ_n is the phase corresponding to the n-th symbol, and its value is given by Formula 6:

$$\varphi_n=\frac{2\pi(i-1)}{M},\quad i=1,2,\ldots,M \tag{6}$$

where M is the modulation order, which determines the total number of phase states.

2.2. MQAM Signal Model

Multi-ary Quadrature Amplitude Modulation (MQAM) is a high-order modulation scheme based on combined amplitude and phase keying, which can achieve high spectral efficiency through multi-level quadrature amplitude keying. Its specific implementation is as follows: two orthogonal carrier signals are amplitude-modulated in a certain manner, where the amplitude variation on each carrier carries part of the transmitted data information. By combining these components, the data transmission rate can be significantly improved without increasing the bandwidth. Formula 7 is the time-domain expression of MQAM:

$$S_{\mathrm{MQAM}}(t)=\sum_{n}A_{n}g(t-nT_{s})\cos(\omega_{c}t+\phi_{n}) \tag{7}$$

where A_n is the amplitude keying component and ϕ_n is the phase keying component. Expanding Formula 7 yields Formula 8, which is the quadrature representation.

$$S_{\mathrm{MQAM}}(t)=\left[\sum_{n}A_{n}g(t-nT_{s})\cos\phi_{n}\right]\cos\omega_{c}t-\left[\sum_{n}A_{n}g(t-nT_{s})\sin\phi_{n}\right]\sin\omega_{c}t \tag{8}$$

Let $X_n=A_n\cos\phi_n=c_nA$, $Y_n=A_n\sin\phi_n=d_nA$

The in-phase component (I-channel signal) and quadrature component (Q-channel signal) are defined as Formula 9 and Formula 10, respectively.

$$I(t)=X(t)=\sum_{n}X_{n}g(t-nT_{s}) \tag{9}$$

$$Q(t)=Y(t)=\sum_{n}Y_{n}g(t-nT_{s}) \tag{10}$$

Substituting Formula 9 and Formula 10 into Formula 8 yields the equivalent quadrature representation of the MPSK signal, i.e., the time-domain expression of the combined IQ signal, as shown in Formula 11:

$$S_{\mathrm{MQAM}}(t)=I(t)\cos\omega_{c}t-Q(t)\sin\omega_{c}t \tag{11}$$

In the above formulas, A is the fixed amplitude, c_n and d_n are determined by the input data, and determine the coordinates of the modulated QAM signal in the signal space. It can be seen from the above formula that $S_{MQAM}(t)$ is the sum of two orthogonal amplitude-shift keying signals, i.e., the combined IQ signal formed by superimposing the I-channel signal and Q-channel signal after carrier modulation. For MQAM, the modulation order M determines the total number of constellation points, that is, the number of distinguishable amplitude-phase combinations in the signal space, making it the most difficult modulation type to recognize at low-SNR.

2.3. Description of IQ Components for Other Modulated Signals

Specifically, the six modulated signals, namely PAM4, SSB-AM, DSB-AM, B-FM, GFSK, and CPFSK, have no independent IQ components and exhibit high feature discrimination. However, they can all be converted into

complex forms to facilitate processing by the IQ-2AttNet model, and finally output a unified IQ combined signal with a shape of (2, 128), which is suitable for rapid recognition using a lightweight CNN. The characteristics of the 12 modulation signals are shown in Table 1.

Table 1 Features of Modulated Signals

Modulation Type	In-Phase component	Quadrature component	Features
BPSK	±1	0	Quadrature Modulation,Q=0
QPSK	±1	±1	Quadrature Modulation,I/Q Independent
8PSK	$\cos(\phi_i)$	$\sin(\phi_i)$	Quadrature Modulation,8 Phase Points
MQAM	Normalized Constellation Points	Normalized Constellation Points	Quadrature Modulation, Amplitude and Phase
PAM4	Normalized Amplitude	0	Baseband,Q=0
SSB-AM/DSB-AM	$m_i \cos(2\pi f_c t_i)$	$m_i \sin(2\pi f_c t_i)$	Single carrier,complex representation
B-FM/GFSK/CPFSK	$\cos(\phi_i)$	$\sin(\phi_i)$	Single carrier,phase accumulation

3. Dual-Branch Feature Fusion Recognition Model

The proposed dual-branch feature fusion model in this paper mainly consists of three parts: a lightweight CNN for preliminary screening branch, an IQ-2AttNet for fine recognition branch, and a decision fusion module. The overall structure is shown in Figure 1. The received signal is fed into the two recognition branches in parallel. The first branch uses a lightweight CNN to perform fast coarse classification on all 12 modulation signals, efficiently distinguishing easily recognizable signals and reducing subsequent computational load.The second branch locks onto six types of highly confused high-order PSK and QAM core modulation signals through an IQ signal extraction module, and then IQ-2AttNet conducts fine-grained feature learning and accurate recognition in this high-confusion region. Finally, the output results of the two branches are jointly sent to the decision fusion module, where an adaptive decision is made via a probability weighting strategy to output the optimal recognition result.

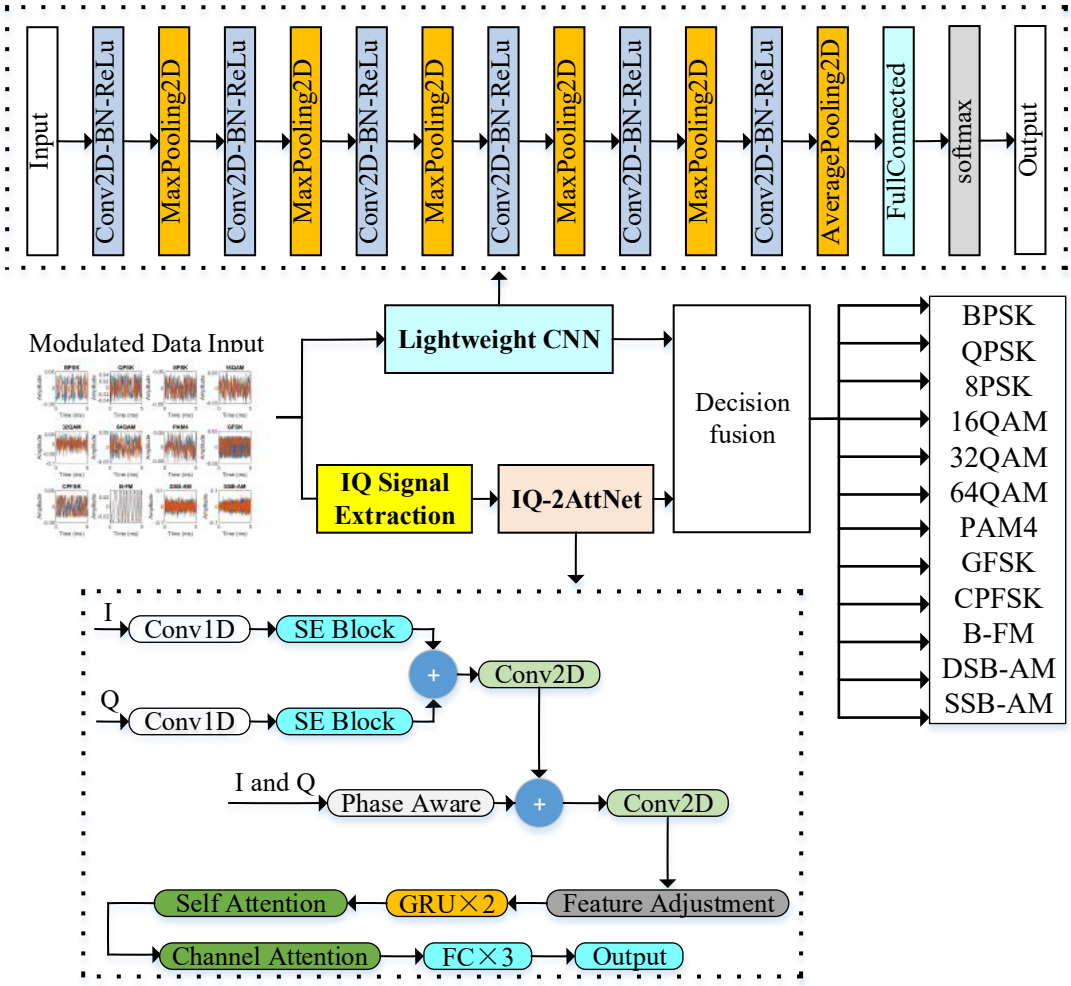


Fig. 1. Recognition Scheme Model Based on Lightweight CNN-IQ-2AttNet Decision Fusion

3.1. Lightweight CNN Network Design

It adopts the structure of "Convolution + BN + ReLU + Average Pooling", and replaces fully connected layers with pooling layers to significantly reduce the number of parameters, meeting the requirements of lightweight design and real-time performance, as shown in Figure 1[10]. Targeted improvements are made around the two core lightweight design principles: parameter reduction and efficiency improvement. First, a batch normalization layer is embedded after each convolutional layer. By standardizing the data distribution, it accelerates model training convergence and alleviates the vanishing gradient problem, improving training stability and feature learning efficiency without introducing extra computational overhead[12]. Second, the traditional multi fully-connected layer feature mapping scheme in CNN is abandoned. Parameter-free pooling operations are used instead of fully connected layers for feature dimension reduction, with only one fully connected layer retained for classification output. This reduces massive redundant parameters caused by fully connected layers and strengthens the lightweight property of the network[13].

3.2. IQ Signal Extraction Module

The IQ signal extraction module takes the in-phase component and quadrature component as input features, and realizes differentiated representation of different modulation types through a multi-branch feature extraction mechanism. For modulated signals with IQ components, the specific implementation process of this module is as follows:

(1)Complex Signal Construction

All modulation signals to be identified are first converted to complex domain representation, whose mathematical form is shown in Formula 12.

$$s_{\text{complex}}(n) = \text{Re}\{s(n)\} + j.\text{Im}\{s(n)\} \tag{12}$$

Where $n = 1, 2, \ldots, N$ is the index of signal sampling points, $N = 128$ is the sampling length of a single signal sample; $\text{Re}\{s(n)\}$ denotes the real part of the signal, $\text{Im}\{s(n)\}$ denotes the imaginary part of the signal, j is the imaginary unit。

(2)I/Q Component Separation

Based on the above complex signal form, the I-channel and Q-channel signals are extracted through real-imaginary decomposition, whose mathematical definition is shown in Formula 13.

$$\begin{cases} s_{\text{I}}(n) = \text{Re}\{s_{\text{complex}}(n)\} \\ s_{\text{Q}}(n) = \text{Im}\{s_{\text{complex}}(n)\} \end{cases} \tag{13}$$

Where, $s_{\text{I}}(n)$ denotes the I-channel (in-phase) signal and $s_{\text{Q}}(n)$ denotes the Q-channel (quadrature) signal at the n-th sample. The separated I/Q components are concatenated along the branch dimension to form the combined IQ signal, with its matrix expression as Formula 14.

$$\mathbf{S}_{\text{IQ}} = \begin{bmatrix} s_{\text{I}}(1) & s_{\text{I}}(2) & \ldots & s_{\text{I}}(128) \\ s_{\text{Q}}(1) & s_{\text{Q}}(2) & \ldots & s_{\text{Q}}(128) \end{bmatrix} \tag{14}$$

$\mathbf{S}_{\text{IQ}}$ is a two-dimensional matrix with dimensions $\mathbf{R}^{2\times128}$, where the first row corresponds to the 128 sampling points of the I-channel signal, and the second row corresponds to the 128 sampling points of the Q-channel signal.

(3)Data Format Specification

After processing through the IQ signal extraction module, the data format of a single sample is unified into a 2D tensor of shape $\mathbf{R}^{2\times128}$.

- Dimension 0 corresponds to the in-phase (I-channel) signal sequence, which consists of 128 consecutive sampling points;
- Dimension 1 corresponds to the quadrature (Q-channel) signal sequence, which consists of 128 consecutive sampling points.

(4)Description of IQ Component Characteristics for Different Modulation Signals

From the perspective of modulation principle, the existence of I/Q components depends on whether the signal adopts quadrature carrier modulation: MPSK and MQAM carry I/Q components via two orthogonal carriers respectively, and the combined IQ signal can be constructed through Equations (3)-(5); in contrast, PAM4 is a baseband signal, SSB-AM/DSB-AM are single-carrier amplitude modulation signals, and B-FM/GFSK/CPFSK are single-carrier frequency modulation signals (with CPFSK additionally maintaining phase continuity). Such signals carry information only through discrete changes in amplitude or frequency, without the design of orthogonal components, thus cannot be decomposed into I/Q components, and have no combined IQ signal. This characteristic provides the core theoretical basis for the proposed dual-branch hierarchical processing strategy in this paper: the lightweight CNN branch can efficiently process easily distinguishable signals without I/Q components, while the improved IQ-2AttNet branch is specifically designed to tackle the easily confused

MPSK/MQAM signals that require I/Q features.

3.3. IQ-2AttNet Network Design

The IQ-2AttNet proposed in this paper adopts a dual-branch parallel "structure + temporal attention + channel attention" to enhance the extraction of high-order subtle signal features and improve robustness under low-SNR. Its core design idea is to fully utilize the characteristics of I/Q components of modulated signals, and achieve accurate recognition of high-order modulated signals with I/Q components through the dual-branch parallel structure and dual attention mechanism, whose network structure is shown in Figure 1.

(1)Dual-Branch Parallel Structure

IQ-2AttNet adopts a dual-branch parallel architecture, consisting of an I/Q component feature extraction branch and an IQ combined signal fine recognition branch, to realize hierarchical and parallel learning of basic features and high-order features. The I/Q component feature extraction branch takes the I-channel and Q-channel signals as separate inputs and extracts basic features through lightweight convolution and channel attention. The I and Q signals are respectively processed by Conv1D layers to capture temporal envelope features and quadrature-domain phase offset features, and then channel-wise attention weighting is performed via the Squeeze-and-Excitation (SE) block to enhance the response to critical feature channels. After the two branches of features are fused by element-wise addition, they are fed into a Conv2D layer for spatial feature integration to form a basic feature representation of the modulated signal. The IQ combined signal fine recognition branch targets six typical confusable signals with IQ components such as MPSK and MQAM, and takes the IQ combined signal as input to achieve deep mining and accurate recognition of phase-sensitive features. It first uses a Phase Aware module to specifically extract phase-sensitive features and capture the phase coding rules of high-order modulated signals. The output features of this branch are then concatenated with the fused features from the I/Q feature extraction branch, and further deep feature extraction is conducted through a Conv2D layer. Finally, a Feature Adjustment module is used to regularize and normalize the feature dimensions and distributions, preparing for subsequent temporal feature modeling.

(2)Dual-Attention Fusion Mechanism

After completing the deep feature extraction, IQ-2AttNet introduces a dual-attention mechanism that adaptively weights features from the temporal and channel dimensions, respectively, to significantly enhance the model's characterization capability for confusable signals. Specifically, the temporal self-attention mechanism, built upon two-layer GRU for modeling sequential dependencies, further captures long-range temporal dependencies to strengthen the focus on critical temporal segments and improve the precision of subtle feature perception for high-order modulated signals. Meanwhile, the channel attention mechanism adaptively evaluates the importance of different feature channels, suppresses redundant responses through adaptive weighting, and highlights channels that are highly relevant to modulation types, thereby continuously boosting the discriminative power of feature representations. The fused features from both temporal and channel attention branches are then fed into a three-layer fully connected (FC×3) classification head to output the classification probabilities for the six core confusable signal types.

3.4. Decision Fusion Module

The core rule of the decision fusion module is confidence-based optimal decision: the recognition result with higher confidence from the two branches is selected as the final output. Let the input modulated signal be $s(t)$.The recognition result and confidence of the first branch are ($\hat{k}_{CNN}$, p_{CNN}), while those of the second branch are ($\hat{k}_{IQ2Att}$, p_{IQ2Att}), where $\hat{k}_{CNN}, \hat{k}_{IQ2Att} \in 1,2,\ldots,12$, and $0 \le p_{CNN}, p_{IQ2Att} \le 1$. The final

recognition result $\hat{k}_{final}$ of the fusion module is defined as formula 15:

$$\hat{k}_{final} = \begin{cases} \hat{k}_{CNN}, p_{CNN} > p_{IQ2Att} \\ \hat{k}_{IQ2Att}, p_{IQ2Att} \geq p_{CNN} \end{cases} \tag{15}$$

When the confidences of the two branches are equal, the result of IQ-2AttNet is selected by default to ensure the recognition accuracy of the core confusable MPSK/MQAM signals. This strategy simplifies the fusion logic, balances recognition efficiency and accuracy, and achieves robust recognition of 12 modulation types over the SNR range from -10 dB to 25 dB[11].

4. Experimental Results and Analysis

4.1. Description of the Experimental Dataset

In this paper, a modulated signal dataset is constructed with SNR=[-10:2:18]dB, where the length of each individual signal sample is uniformly set to 128 points. The channel adopts a Rician multipath fading channel with a Rician factor K=4, a maximum Doppler shift of 4 Hz, and a maximum clock offset of 5 ppm. The system sampling rate is 200 kHz, and the length of a single frame of signal is 1024 sampling points; digital modulated signals are mapped at a rate of 8 sampling points per symbol. The network makes independent decisions on a per-frame basis without relying on consecutive multi-frame information. The center frequency of digital modulated signals is set to 900 MHz, and that of analog modulated signals is set to 100 MHz.

4.2. Model Training

All experiments are conducted on a hardware platform equipped with an Intel Core i9-13900H processor, 40 GB of RAM, and an NVIDIA GeForce RTX 4060 graphics card with 8 GB of dedicated memory. The first branch is implemented on the MATLAB R2022b platform, where the dataset is split into training, validation, and test sets with a ratio of 8:1:1. The second branch is built under the environment of CUDA 11.8 and cuDNN 8.7.0, and the dataset is divided into a training-validation set and a test set with a ratio of 8:2. The hyperparameters for network training are listed in Table 2. An early-stopping strategy is adopted during training, which takes the validation loss as the optimization criterion. The patience value is set to 12, meaning that training will be terminated and the best-performing model will be saved if no performance improvement is observed for 12 consecutive epochs.

Table 2 Hyperparameter Settings for Network Training

Network	Parameter Name	Value	Network	Parameter Name	Value
Lightweight CNN	Optimizer	sgdm	IQ-2AttNet	Optimizer	adam
	Learning Rate	0.0002		Learning Rate	0.001
	Learning Rate Scheduling Strategy	piecewise		weight_decay	1e-4
	Dropout Rate	0.5		Loss Function	CrossEntropyLoss
	Batch Size	256		Input Layer Size	2×128
	Convolution Kernel Size	[1,8]		Batch Size	64

4.3. Recognition Performance of the Lightweight CNN

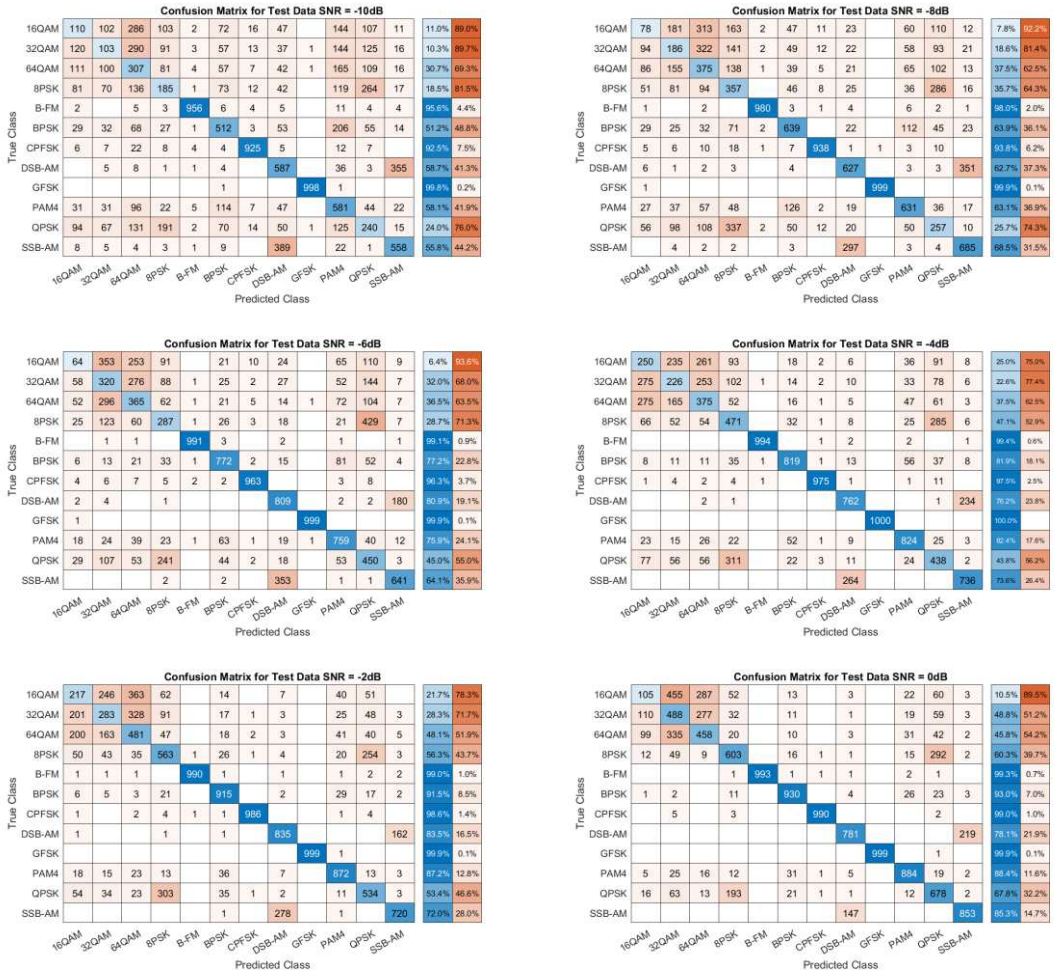
The test set with SNR=[-10:2:18]dB is used to verify the performance of the lightweight CNN, and 25 dB is taken as the high-SNR baseline. The confusion matrices under each SNR are given in Table 3. The overall recognition performance of the lightweight CNN shows a steady upward trend over a wide SNR range, with the

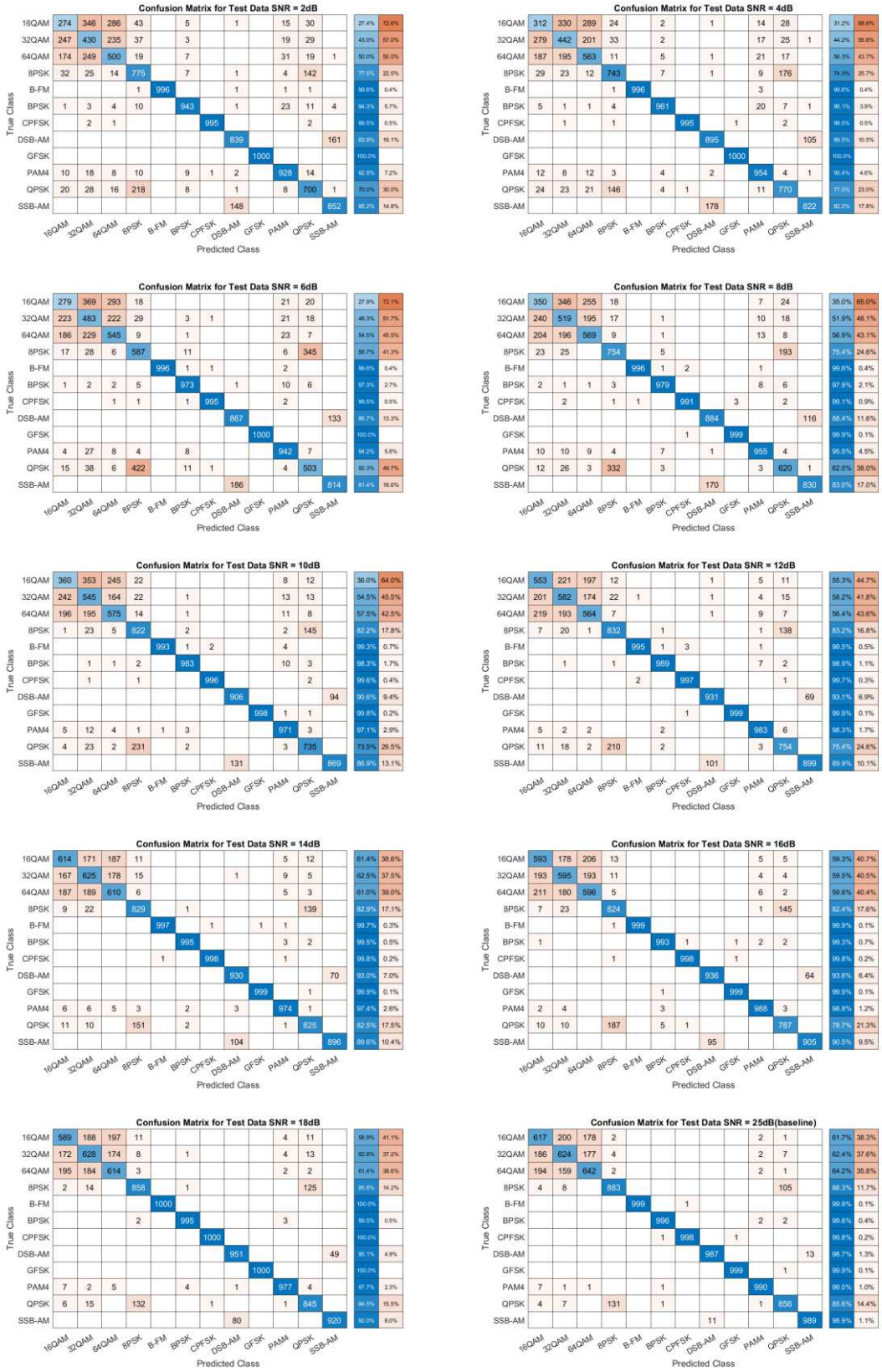
recognition rate improved from 50.5% at -10 dB to 88.2% at 25 dB, and no obvious performance collapse is observed, verifying the rationality and noise robustness of the lightweight structure. Among them, modulated signals without IQ components have high feature discrimination and are less affected by noise. Their recognition rate remains around 90% across the full SNR range and is still above 76% at -10 dB. In particular, the recognition rate of GFSK signals is close to 100%, which is highly consistent with the design goal of the lightweight CNN for rapid preliminary screening.

Among signals with IQ components, BPSK, as the lowest-order MPSK signal, performs prominently. Its recognition rate reaches 51.2% at -10 dB and rises rapidly with the increase of SNR, up to 99.6% at 25 dB, significantly better than other MPSK/MQAM signals. However, the other core MPSK/MQAM signals with IQ components are highly sensitive to noise, with an average recognition rate of only 24.3% at -10 dB, which becomes the main factor restricting the overall performance. In particular, high-order signals such as 16QAM, 32QAM, and 64QAM all have recognition rates lower than 30% at low-SNR, mainly because their joint amplitude-phase coding features are easily corrupted by noise, and the shallow lightweight CNN is difficult to capture subtle constellation differences.

In summary, the lightweight CNN can achieve wide-coverage preliminary screening of 12 modulation signals, with a recognition accuracy of nearly 98% for signals without IQ components. Nevertheless, its recognition ability for high-order and easily confused MPSK/MQAM signals is limited, requiring a dedicated deep network to further refine the recognition performance.

Table 3 Confusion Matrices for Lightweight CNN Recognition





4.4. Recognition Performance of IQ-2AttNet

The IQ-2AttNet is adopted to perform fine-grained recognition on six core types of MPSK/MQAM signals, and verification tests are conducted on the other six signals without IQ components. The recognition accuracy of IQ-2AttNet is shown in Table 4, and the corresponding overall confusion matrix is illustrated in Figure 2.

IQ-2AttNet achieves excellent overall recognition performance for the six core orthogonal modulation signals, with an average recognition rate of 85.7% over the full SNR range. Specifically, the recognition rate of QPSK remains 100% when SNR ≥ 2 dB, and the average recognition rates of 8PSK and 32QAM are 72.6% and 85.7%, respectively, which effectively compensates for the shortcomings of the lightweight CNN in recognizing core easily confused signals. The recognition rate of core signals increases significantly with SNR, most notably in the range from -10 dB to 2 dB, and stabilizes when SNR ≥ 4 dB. When SNR ≥ 10 dB, the recognition rates of high-order signals such as 16QAM, 32QAM, and 64QAM all exceed 95%, indicating that the dual-attention and phase-aware modules can effectively capture subtle features of high-order signals. The recognition performance of MPSK signals is generally better than that of MQAM signals, because MPSK only relies on phase coding with relatively stable features, whereas MQAM adopts joint amplitude-phase coding with denser constellation points, which leads to more severe distortion under noise and greater recognition difficulty.

Notably, BPSK, as the lowest-order MPSK signal, has a recognition rate of nearly 0% in IQ-2AttNet, and most samples are misclassified as SSB-AM. The main reasons are as follows: the constellation points of BPSK are only distributed on the real axis with single phase features, which do not match the high-order MPSK/MQAM phase coding and dense constellation features emphatically learned by the network. Meanwhile, the single-axis IQ component features of BPSK are highly similar to the time-domain features of SSB-AM signals without IQ components, making it impossible for the network to effectively distinguish them and resulting in severe feature confusion. This phenomenon further demonstrates that IQ-2AttNet is a dedicated network customized for high-order easily confused signals and has insufficient adaptability to simple orthogonal modulation signals such as BPSK.

IQ-2AttNet performs poorly in recognizing signals without IQ components. The average recognition rates of CPFSK and DSB-AM are only 2.8% and 2.7%, respectively, because the network is designed specifically for IQ components and constellation features and cannot be effectively adapted to signals without IQ components. This verifies the rationality of dedicated networks for specific types of signals and also provides experimental support for the complementary fusion of the two branches.

Table 4 Recognition Accuracy of IQ-2AttNet Network/%

method / SNR(dB)	BPSK	QPSK	8PSK	16QAM	32QAM	64QAM	B-FM	CPFSK	DSB-AM	GFSK	PAM4	SSB-AM
-10	0	26.5	9	19.5	55.5	11	67	0	10	5.5	33	18
-8	0	28	10.5	28.5	62	15	82.5	0	12.5	12.5	37	31.5
-6	0	37	25.5	19	68.5	25	81.5	0.5	9	18.5	51	47.5
-4	0	44.5	29	26.5	75	20	74	2	9	17	57.5	68
-2	0	71	43	40.5	75.5	28.5	73.5	1.5	2.5	24.5	76	90.5
0	0	93	70	46	83.5	36	66.5	3	1	26.5	91	94.5
2	0	98.5	83	56	86	44	83	5.5	0	24.5	98.5	100
4	0	100	93.5	58	90	66	67	3.5	0	24.5	100	100
6	0	100	98.5	73	93	84	74.5	6.5	0	24	100	100
8	0	100	100	81.5	96	92	61.5	3.5	0	34	100	100
10	0	100	100	98.5	95.5	97	64.5	2.5	0	37.5	100	100
12	0	100	100	100	95.5	99	54.5	2	0	51.5	100	100
14	0	100	100	100	98	99.5	57.5	1.5	0	52	100	100
16	0	100	99.5	100	99	98.5	45.5	3.5	0	52	100	100
18	0	100	100	100	99.5	99.5	36	4.5	0	61	100	100
25	0	100	100	100	99	99.5	41	5.5	0	61.5	100	100

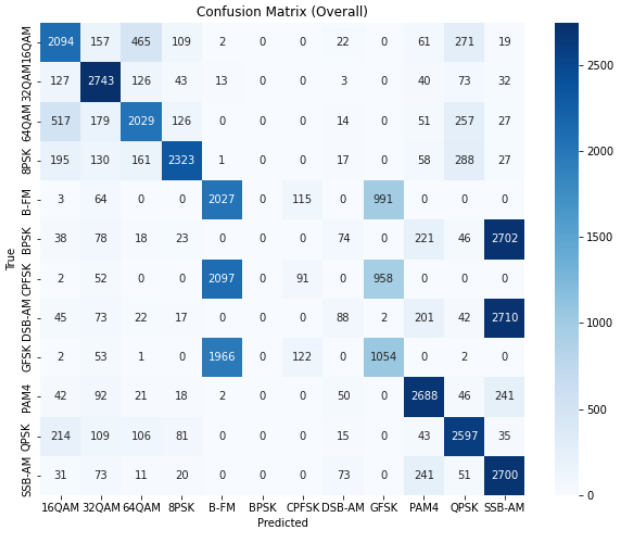


Fig. 2. Overall confusion matrix of IQ-2AttNet recognition

4.5. Recognition Results of Decision Fusion

Based on the "optimal result selection" decision fusion strategy shown in formula 6, this paper comprehensively judges the classes and confidences output by the two branches, and selects the result with higher confidence as the final output. When the confidences are equal, the recognition result of IQ-2AttNet is adopted by default. The fusion recognition results of the 12 modulation signals are listed in Table 5.

(1) Significant improvement in overall recognition performance.The fusion algorithm unifies the advantages of "wide-coverage preliminary screening + high-precision fine recognition". The overall average recognition rate reaches 86.1% over the SNR range from -10 dB to 25 dB. The average recognition rate of the six core MPSK/MQAM signals is increased by 12.3 percentage points compared with the single lightweight CNN, which fully validates the effectiveness of the complementary dual-branch fusion.

(2) Greatly enhanced robustness at low-SNR.The proposed method exhibits significantly improved robustness in low-SNR environments. The overall average recognition rate at −10 dB is 55.2%, and the recognition rate of signals without IQ components remains above 76.8%. When SNR ≥ 8 dB, the recognition rate of all signal types stabilizes above 90%, and the average recognition rate of core signals reaches 98.5%. Besides, the method requires no manual prior parameters or high-precision preprocessing, and its anti-noise capability outperforms traditional algorithms.

(3) Greatly improved recognition of high-order modulation signals.After fusion, the recognition rates of high-order signals such as 16QAM, 32QAM, and 64QAM are greatly improved compared with the single lightweight CNN. When SNR ≥ 12 dB, the recognition rate of 16QAM reaches 100%, and those of 32QAM and 64QAM stabilize above 95%. The generalization capability is significantly better than that of the single lightweight CNN.

(4) Stronger engineering practicability and scenario adaptability.The proposed method covers four mainstream modulation schemes and a wide dynamic SNR range. The decision fusion logic is simple with low computational cost and no complex iterations. It can well adapt to 5G/6G adaptive modulation and coding scenarios, and can be directly deployed in practical engineering applications such as communication reconnaissance and spectrum monitoring.

Table 5 Recognition Accuracy of 12 Modulated Signals/%

method / SNR(dB)	BPSK	QPSK	8PSK	16QAM	32QAM	64QAM	B-FM	CPFSK	DSB-AM	GFSK	PAM4	SSB-AM

-10	51.2	26.5	18.5	19.5	55.5	30.7	95.6	92.5	58.7	99.8	58.1	55.8
-8	63.9	28	35.7	28.5	62	37.5	98	93.8	62.7	99.9	63.1	68.5
-6	77.2	45	28.7	19	68.5	36.5	99.1	96.3	80.9	99.9	75.9	64.1
-4	81.9	44.5	47.1	26.5	75	37.5	99.4	97.5	76.2	100	82.4	73.6
-2	91.5	71	56.3	40.5	75.5	48.1	99	98.6	83.5	99.9	87.2	90.5
0	93.0	93	70	46	83.5	45.8	99.3	99	78.1	99.9	91	94.5
2	94.3	98.5	83	56	86	50	99.6	99.5	83.9	100	98.5	100
4	96.1	100	93.5	58	90	66	99.6	99.5	89.5	100	100	100
6	97.3	100	98.5	73	93	84	99.6	99.5	86.7	100	100	100
8	97.9	100	100	81.5	96	92	99.6	99.1	88.4	99.9	100	100
10	98.3	100	100	98.5	95.5	97	99.3	99.6	90.6	99.8	100	100
12	98.9	100	100	100	95.5	99	99.5	99.7	93.1	99.9	100	100
14	99.5	100	100	100	98	99.5	99.7	99.8	93	99.9	100	100
16	99.3	100	99.5	100	99	98.5	99.9	99.8	93.6	99.9	100	100
18	99.5	100	100	100	99.5	99.5	100	100	95.1	100	100	100
25	99.6	100	100	100	99	99.5	99.9	99.8	98.7	99.9	100	100

4.6. Comparative Experiments with Different Neural Network Models

To verify the effectiveness and superiority of the proposed dual-branch feature fusion method, comparative experiments are conducted with four mainstream modulation recognition models, namely LSTM [12], CGDNet [13], PET-CGDNN [14], and DualSR-Net [15]. The accuracy data of all comparison methods are taken from the published results of the corresponding literature with consistent evaluation metrics and calculation criteria. The average recognition accuracy is calculated over three intervals: −8~0 dB, above 0 dB, and full SNR (SNR = [−8:2:18] dB). The results are listed in Table 6. The proposed method outperforms the comparison models significantly in recognition performance over the low-SNR range, medium–high-SNR range, and full SNR range. In the critical low-SNR range from -8 dB to 0 dB, the average recognition accuracy of the proposed method is 71.8%, which is 1.5% to 17.5% higher than that of the comparison models. In the medium-high-SNR range above 0 dB, the recognition accuracy stabilizes at 94.6%, which is 1.8% to 11% higher than that of the comparison models. The average recognition rate over the full SNR range reaches 87.3%. Experimental results demonstrate that the proposed dual-branch structure and decision fusion strategy can effectively improve the recognition robustness in complex electromagnetic environments, and the overall performance is superior to existing mainstream deep learning models.

Table 6 Average Accuracy of Different Models/%

SNR(dB) / Module	−8~0	>0	[-8:2:18]
LSTM	62.5	89.7	60.1
CGDNet	61.1	84.5	57.2
PET-CGDNN	63.3	90.2	60.7
DualSR-Net	70.7	92.9	63.7
Proposed Method	71.8	94.6	87.3

4.7. Comparative Experiments with Different Input Configurations

To further verify the contribution of each module in the dual-branch architecture, comparative experiments with different branch combinations are carried out in this section. The recognition performance of three configurations is tested separately: lightweight CNN only, IQ-2AttNet only, and the proposed dual-branch fusion

method. The calculation methods and evaluation metrics are kept consistent, and the results are shown in Table 7. Using only lightweight CNN or only IQ-2AttNet both present obvious performance bottlenecks: the lightweight CNN has limited recognition ability for high-order modulation signals, while IQ-2AttNet has poor adaptability to signals without IQ components and simple modulation signals.

The proposed dual-branch feature fusion method achieves the highest accuracy in all three evaluation intervals, i.e., −10~0 dB, above 0 dB, and full SNR, through complementary advantages, with significant improvements over single structures. Especially in the low-SNR interval of −10~0 dB, the accuracy after fusion reaches 69.0%, which fully validates the necessity and effectiveness of the dual-branch architecture of "wide-coverage preliminary screening + high-precision fine recognition".

Table 7 Average Accuracy of Different Inputs Structures/%

Module \ SNR(dB)	−10~0	>0	all
Lightweight CNN Only	62.8	81.2	74.8
IQ-2AttNet Only	35.0	63.1	53.3
Dual-Branch Fusion Method	69.0	95.1	86.1

5. CONCLUSION

Aiming at the low recognition accuracy of communication signal modulation schemes under low-SNR environments, this paper proposes a modulation recognition method based on dual-branch feature fusion and an IQ dual-path attention mechanism. This method extracts signal amplitude and phase features separately through two parallel IQ branches, strengthens the weights of critical features with the attention mechanism, and realizes final recognition via decision-level fusion, which effectively improves the robustness under low-SNR.

Experimental results show that the overall recognition accuracy of the proposed method reaches 86.1% over the SNR range from -10 dB to 25 dB, and the accuracy over the low-SNR range from -8 dB to 0 dB is 71.8%, which is significantly improved compared with the single CNN branch and traditional recognition models. Under weak signal conditions, the IQ dual-path attention structure can retain signal detailed features more completely and reduce feature loss caused by noise interference, enabling the model to maintain stable output in low-SNR environments.

Without significantly increasing network complexity, the proposed method achieves an effective balance between feature learning and noise suppression, making it suitable for communication scenarios with complex electromagnetic environments and poor signal quality. The overall scheme features a clear structure and easy deployment, which can provide reliable technical support for modulation recognition in practical engineering applications such as wireless communication, electromagnetic perception, and non-cooperative communication.

ACKNOWLEDGMENTS

Na.

Authors' contributions

Xiuping Song: Conceptualization, investigation, resources, methodology. Junxia Wang: Visualization, writing—original draft and revised draft. Youlong Gong and Xianyou Yao: Supervision, project administration, investigation, visualization, writing—original draft and revised draft. All authors have read and approved the final manuscript.

Funding

Supported by the Science and Technology Research Program of Chongqing Municipal Education Commission(Grant No. KJZD-K202503104).

Data availability

All data used shall be made available upon request.

DECLARATIONS

Competing interests

The authors declare no competing interests.

REFERENCES

[1] J. Yang, H. Liu, X. Y. Bu, et al., Modulation Recognition of Communication Signals: Principles and Algorithms. Beijing: Posts & Telecom Press, 2014.

[2] X. Han, J. Zheng, X. Feng, et al., “A blind MQAM modulation recognition method based on constellation template matching,” Electronic Measurement Technology, vol. 48, no. 11, pp. 123-130, 2025.

[3] F. X. Zhang, C. B. Luo, J. L. Xu, et al., “Deep learning based automatic modulation recognition: models, datasets, and challenges,” Digital Signal Processing, vol. 129, pp. 103-115, 2022.

[4] H. D. Zhang and H. X. Lou, “Automatic recognition method for MQAM signal modulation,” Journal on Communications, vol. 40, no. 8, pp. 200-211, 2019.

[5] C. C. Zhang, X. M. Li, J. L. Li, et al., “MQAM signal modulation recognition based on JDA-BP network,” Experimental Technology and Management, vol. 41, no. 5, pp. 31-37, 2024.

[6] P. C. Hu, W. D. Yang, P. Li, “Automatic modulation recognition of MQAM signals over Rician fading channels,” Journal of Army Engineering University, vol. 1, no. 3, pp. 22-28, 2022.

[7] T. T. Wang and X. F. Gong, “Application of joint recognition algorithm for PSK and QAM signals based on constellation diagram,” Application Research of Computers, vol. 32, no. 7, pp. 2116-2118, 2015.

[8] Y. H. Pei, Y. Qu, J. M. Li, et al., “Modulation recognition of MPSK and MQAM signals based on AlexNet,” Laser Journal, vol. 39, no. 10, pp. 75-78, 2018.

[9] G. Jajoo, Y. Kumar, and S. K. Yadav, “Blind signal PSK/QAM recognition using clustering analysis of constellation signature in flat fading channel,” IEEE Communications Letters, vol. 23, no. 99, pp. 1853-1856, 2019.

[10] L. W. Li, J. Lu, H. Liu, “Modulation recognition method based on complex convolution and neural network,” Computer Engineering and Design, vol. 46, no. 8, pp. 2211-2218, 2025.

[11] H. Yang, et al., “SMTrans: An efficient automatic modulation recognition network based on the scale-aware modulation transformer,” Physical Communication, vol. 74, p. 102966, 2026.

[12] W. Long, et al., “A parallel CNN-LSTM automatic modulation recognition network,” Applied Sciences, vol. 16, no. 1, p. 2076, 2026.

[13] J. N. Njoku, M. E. Morocho-Cayamcela, and W. Lim, “CGDNet: Efficient hybrid deep learning model for robust automatic modulation recognition,” IEEE Networking Letters, vol. 3, no. 2, pp. 47-51, 2021.

[14] F. Zhang, C. Luo, J. Xu, et al., “An efficient deep learning model for automatic modulation recognition based on parameter estimation and transformation,” IEEE Communications Letters, vol. 25, no. 10, pp. 3287-3290, 2021.

[15] P. H. Wang, D. He, M. Y. You, et al., “Modulation recognition method based on stochastic resonance and multi-scale feature fusion,” Journal of System Engineering and Electronics, 2025, pp. 1-9.