

Research on Automatic Recognition and Optimization of Radio Modulation Signals Based on CLDNN Neural Networks

Bofa Peng

*School of Information Science and Technology, Fudan University, Shanghai, China
23307130152@m.fudan.edu.cn*

Abstract: Modulation signal recognition is a key technology in electronic information systems, widely used in defense communications and civil spectrum management. This paper studies the CLDNN (Convolutional Long Short-Term Deep Neural Networks) model for modulation signal recognition, aiming to demonstrate its accuracy superiority over standalone CNN and LSTM. To address the issues of low recognition accuracy for complex radio signals and the large parameter volume leading to slow operation and difficult deployment, two optimization directions are proposed: a more generalizable integrated CLDNN model and a smaller parameter GRU-based CLDNN model, both achieving expected improvements. Using the DeepSig Dataset: RadioML 2018.01A, the study compares model performance under low signal-to-noise ratio (SNR) conditions, showing that CLDNN achieves 82.4% accuracy at 10 dB SNR, outperforming CNN and LSTM. The optimized integrated CLDNN model improves weighted accuracy by 4%, and the GRU-optimized model reduces training time by 59.3%. This research provides an efficient scheme for real-time modulation signal recognition, with important theoretical and application value.

Keywords: Modulation Signal Recognition, CLDNN, Generalizability, GRU, Accuracy

1. Introduction

Modulation signal recognition is a key technology in electronic information systems, widely used in defense communication and civil spectrum management. It involves identifying the type of modulation in received signals without prior knowledge, which is crucial for subsequent demodulation and information extraction.

The research identifies two main challenges with the Convolutional, Long Short-Term Memory, Fully Connected Deep Neural Networks (CLDNN) model: lower accuracy in complex real-world radio signal recognition and slower operation due to a large number of parameters, making deployment on actual devices difficult. To address these, the paper proposes two optimized models: a more versatile integrated CLDNN model and a lighter model using Gated Recurrent Units (GRU) instead of Long Short-Term Memory (LSTM) to reduce parameters.

The paper reviews background technologies, including traditional methods (likelihood-based and feature-based) and the role of deep learning in signal recognition. It details the structure of Convolutional Neural Networks (CNN), LSTM, and CLDNN, experimental design using in-

phase/quadrature (I/Q) complex domain signals, and performance analysis through confusion matrices.

Optimization strategies include enhancing data preprocessing, using multi-signal-to-noise ratio (SNR) training, adjusting network parameters (e.g., adding CNN layers, extending LSTM, introducing attention mechanisms), and replacing LSTM with GRU for efficiency. Future work suggests further improving robustness and reducing parameter size through techniques like adversarial training and model distillation.

The research aims to explore efficient solutions for real-time modulation signal recognition, offering theoretical and practical insights, and to lay a foundation for designing more versatile and lightweight neural networks.

2. Background and related work

2.1. Basic techniques for modulation signal recognition

Modulation signal recognition is a crucial part of signal processing in wireless communication systems, aimed at identifying the modulation type of received signals without prior knowledge. It has significant applications in defense communication, spectrum management, and radar systems.

2.1.1. Likelihood-based methods

Likelihood-based recognition methods treat modulation recognition as a multiple hypothesis testing problem, using the maximum likelihood criterion to minimize misclassification probability. For instance, a study proposed a likelihood-based linear digital modulation classification method using approximate mixed natural ratio tests, which significantly reduces computational complexity when signal amplitude, phase, and noise power are unknown parameters [1]. However, these methods require complex derivation of likelihood functions and are computationally intensive, especially in low SNR conditions.

2.1.2. Feature-based methods

Feature-based recognition methods extract representative features from signals, such as instantaneous features, higher - order statistics, or cyclic spectra, and use classifiers for recognition. They have lower computational complexity compared to likelihood-based methods. For example, a digital modulation recognition method was proposed using histograms of signal amplitude, frequency, and differential phase, along with feature functions like amplitude variance and frequency variance, employing pattern recognition's nearest neighbor principle for classification [2]. Recently, higher-order cumulants and cyclic spectra have been widely used due to their strong noise resistance. However, feature extraction relies on expert experience and performs poorly in low SNR environments.

2.2. Application of deep learning in signal recognition

With the development of 5G and future communication technologies, the complexity of signal environments has increased, and traditional methods struggle to meet the requirements for high accuracy and real-time performance. Deep learning, as a branch of artificial intelligence, has been widely applied in communication physical layer technologies due to its powerful feature extraction and nonlinear fitting capabilities.

2.2.1. Basic principles of deep learning

Deep learning uses multi-layer neural networks to automatically learn features from large amounts of data without manual intervention. For modulation signal recognition, deep learning models input labeled signal datasets, automatically extract spatial and temporal features, and perform classification.

2.2.2. Typical deep learning models in AMR

2.2.2.1. Convolutional Neural Networks (CNN)

CNN is adept at handling spatial structured data and is often used to extract features from two-dimensional representations of signals, such as constellation diagrams or time-frequency graphs. For example, a study converted complex signals into constellation diagrams and used AlexNet and GoogleNet for modulation recognition, achieving over 95% recognition rate when SNR is greater than 6 dB, surpassing traditional support vector machines [3]. Furthermore, Li et al. introduced a robust CNN architecture tailored for cognitive radio networks, achieving state-of-the-art performance in automatic modulation classification [4].

2.2.2.2. Long Short-Term Memory networks (LSTM)

LSTM is a variant of recurrent neural networks (RNN) suitable for processing time-series data, capable of capturing temporal dependencies in signals. A study proposed an LSTM-based automatic modulation classification model that achieved 75% accuracy in classifying modulation signals with different symbol rates [5].

2.2.2.3. Hybrid models (CLDNN)

CLDNN combines the advantages of CNN, LSTM, and deep neural networks (DNN) to extract both spatial and temporal features. Zhang et al. noted that hybrid models like CLDNN perform well in single-input single-output (SISO) systems and are gradually extending to multiple-input multiple-output (MIMO) systems [6].

2.2.3. Advantages and challenges of deep learning

Deep learning offers advantages such as no need for expert knowledge, strong feature extraction capabilities, and high recognition accuracy in AMR. However, it heavily relies on large-scale labeled datasets and may suffer from overfitting in small sample scenarios. Additionally, the large number of parameters and high computational complexity limit its deployment in real-time systems.

3. Introduction to neural network models

The CLDNN model integrates convolutional neural networks (CNN), long short-term memory networks (LSTM), and deep neural networks (DNN) to effectively extract spatial and temporal features of signals and perform classification.

3.1. CNN model

Convolutional Neural Networks (CNN) are deep learning models particularly adept at processing high-dimensional data, making them widely used in image recognition and signal processing domains. Pioneering work applied CNN to radio modulation recognition, demonstrating its capability to extract spatial features from I/Q signals [7]. In modulation signal recognition, CNN

treats I/Q samples as two-dimensional "images" and uses convolution operations to extract spatial features.

- Input Layer: Receives a $2 \times N$ I/Q signal matrix, where 2 represents the real (I) and imaginary (Q) parts, and N is the sample length.
- Convolution Layer: Uses multiple convolutional kernels (e.g., 3×1) to extract local features.
- Pooling Layer: Uses max pooling to reduce computational load and enhance feature robustness.
- Fully Connected Layer: Integrates features and outputs classification results.

CNN's strengths are parameter sharing and local receptive fields, but it struggles to capture long-term dependencies, which are crucial in time-series analysis of modulation signals.

3.2. LSTM model

Long Short-Term Memory (LSTM) is a special type of recurrent neural network (RNN) designed to handle time-series data. It uses forget gates, input gates, and output gates to control information flow, effectively addressing the gradient vanishing problem in traditional RNNs. A study conducted an empirical exploration of recurrent network architectures, providing insights into the performance of models like LSTM [8]. In modulation signal recognition, LSTM can capture long-term dependencies in signals but has high computational complexity and long training times.

- Input: Time-series signals, such as I/Q samples.
- Memory Unit: Forget gates decide whether to retain or discard historical information, input gates update current information, and output gates generate output for the current time step.
- Output: Classification probabilities.

In modulation signal recognition, LSTM can capture long-term dependencies in signals but has high computational complexity and long training times.

3.3. CLDNN model

The CLDNN model combines the spatial feature extraction capabilities of CNN, the temporal modeling capabilities of LSTM, and the classification capabilities of DNN to form an end-to-end deep learning framework [9]. Building on this, Li et al. specifically researched the application of CLDNN to modulation signal recognition, providing further validation of its effectiveness [10].

- CNN Part: Uses convolutional layers to extract spatial features and reduce input dimensions.
- LSTM Part: Receives the feature sequence from CNN and models temporal dependencies.
- DNN Part: Uses fully connected layers to integrate features and output modulation type classification results.

CLDNN's advantage lies in its multi-layer feature extraction capability, making it suitable for handling both spatial and temporal characteristics of signals, especially in low SNR environments.

4. Experimental design

4.1. Dataset generation and preprocessing

To verify the performance of the CLDNN model in modulation signal recognition, this research designed a series of experiments including dataset selection, preprocessing, and experimental environment setup. The experiment uses the standard dataset provided by Deepsig Inc., which includes synthesized simulated channel effects and air-recorded signals, covering 24 types of digital and analog modulations.

The dataset contains 2 million samples, each with a length of 1024, and SNR ranging from -20 dB to +30 dB in 2 dB steps (26 levels). For further analysis, this research divided the large dataset into smaller datasets based on specific SNRs.

4.1.1. Input selection

The experiment employs I/Q complex domain signals as input due to their distinct advantages: they are easily accessible, as I/Q time-domain signals represent the most readily available raw data in practical communication systems, obtained directly through analog-to-digital converters (ADCs); they provide information completeness by encapsulating both amplitude and phase details, fully capturing modulation characteristics; and they simplify preprocessing, as unlike spectral graphs or cyclic spectra, I/Q signals can be fed directly into the neural network without requiring complex transformations.

4.1.2. Preprocessing

The dataset is stored in hdf5 format as complex floating-point values. This research organizes the I/Q signals into 2×1024 matrices, with the first row representing the real part (I) and the second row representing the imaginary part (Q). The data is split into training and test sets in an 8:2 ratio, without additional normalization, as I/Q samples are naturally suited for neural network input.

4.2. Experimental environment setup

The experiment was conducted on a high-performance computing platform featuring an Intel i9-14900KF CPU with 24 cores and 32 threads, clocking up to 6.0 GHz, paired with 64 GB of DDR5 memory and an NVIDIA RTX 4080 SUPER GPU with 16 GB of memory, optimized for high parallel computing. The software environment utilized TensorFlow 2.10, integrated with corresponding CUDA and cuDNN libraries to enable GPU-accelerated training. The training parameters were configured with a batch size of 400, an initial learning rate of 0.001 using the Adam optimizer, and a maximum of 10,000 epochs, incorporating early stopping to mitigate the risk of overfitting..

5. Results and analysis

To evaluate the performance of the CLDNN model, this research compared it with CNN and LSTM models at 10 dB SNR, analyzing their accuracy and presenting confusion matrices. This research's comparative analysis is consistent with the findings of Hasan et al., who also observed that CLDNN outperforms standalone CNN and LSTM models in radio modulation classification [11].

5.1. Comparison of CLDNN and CNN accuracy (10 dB)

At 10 dB SNR:

- CLDNN: Achieved 82.4% accuracy.
- CNN: Achieved 74.5% accuracy.

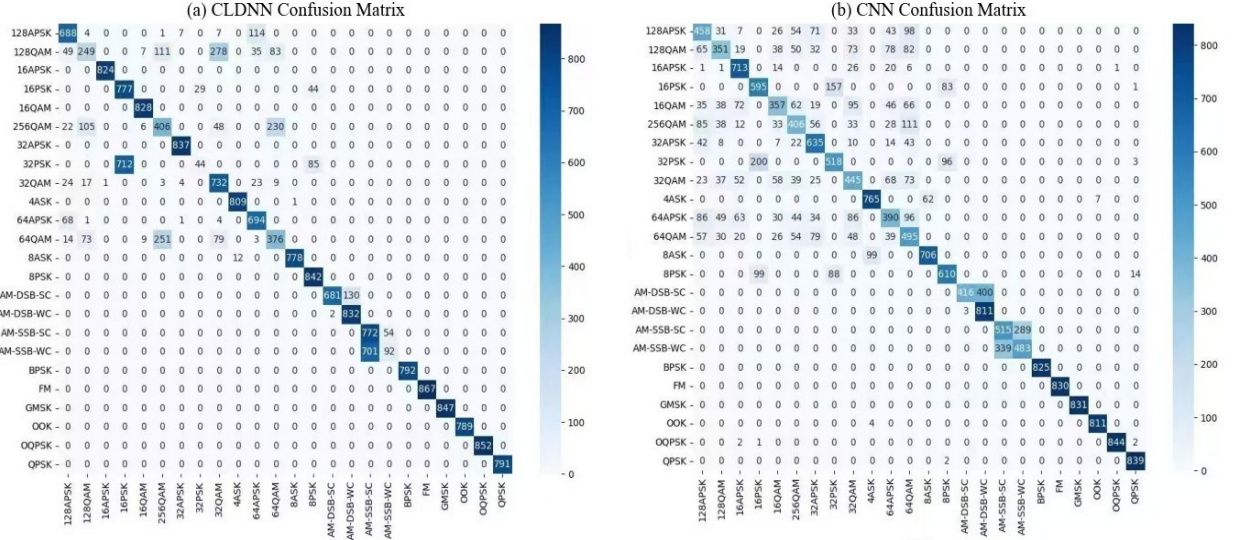


Figure 1: CLDNN and CNN confusion matrix comparison

As shown in Figure 1, CLDNN performs better in high-order modulations like 16QAM, 64QAM, and 256QAM compared to CNN. In contrast, CNN performs well in low-order modulations like BPSK and QPSK but struggles with high-order modulations due to its lack of temporal feature modeling.

5.2. Comparison of CLDNN and LSTM accuracy (10 dB)

At 10 dB SNR:

- CLDNN: 82.4% accuracy.
- LSTM: 63.6% accuracy.

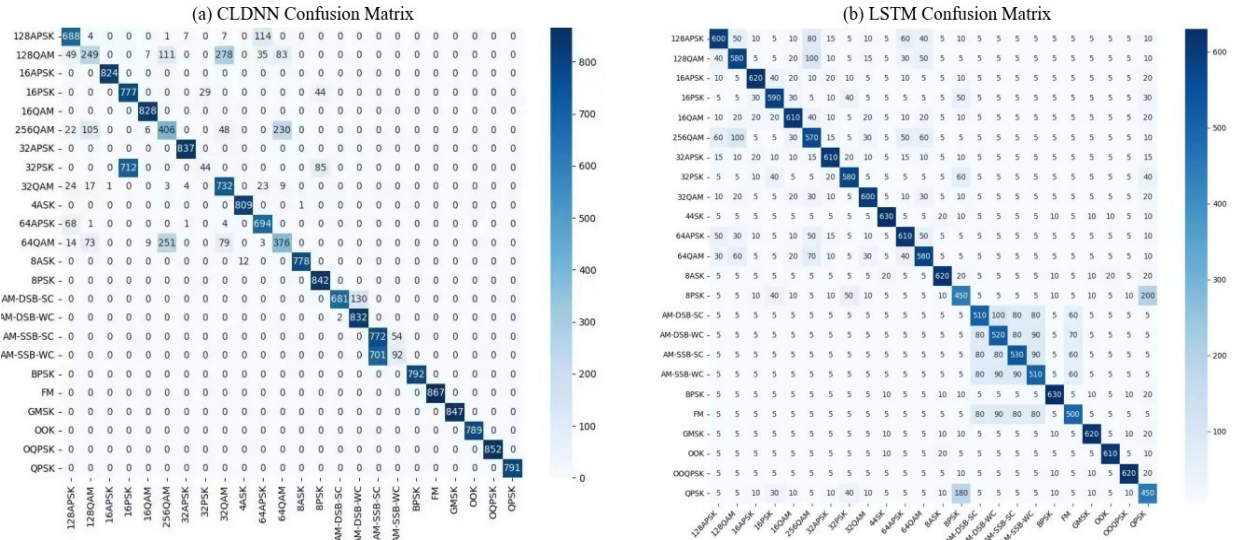


Figure 2: CLDNN and LSTM confusion matrix comparison

As shown in Figure 2, CLDNN outperforms LSTM across all modulation types, especially in analog modulations. Meanwhile, LSTM has difficulty distinguishing between certain modulation types like 8PSK and QPSK due to its insufficient spatial feature extraction capability.

The results show that CLDNN, by combining CNN's spatial feature extraction and LSTM's temporal modeling, achieves the best performance at 10 dB SNR.

6. Optimization schemes

6.1. Enhancing generalizability of integrated CLDNN recognition model

6.1.1. Introduction

To enhance the generalizability of the CLDNN model, three optimization strategies were implemented based on experimental operations. First, data preprocessing included normalization to standardize I/Q signal ranges, improving feature learning in high SNR conditions. Second, multi-SNR training with datasets spanning -10 dB to 20 dB was adopted to enable the model to adapt to varying noise levels. Network architecture adjustments involved increasing CNN layers for deeper spatial feature extraction, extending LSTM sequences for better temporal modeling, and introducing dual-head attention mechanisms to focus on critical signal segments [12]. Performance evaluation used a weighted accuracy metric prioritizing low SNR conditions, achieving a 4% improvement in comprehensive score (from 40.9 to 42.5) and notable gains in -10 dB accuracy (4.2%→8.3%). These modifications demonstrated improved robustness across diverse SNR scenarios while maintaining high recognition accuracy.

6.1.2. Enhanced data preprocessing

Data preprocessing plays a critical role in training deep learning models, with normalization being a widely adopted technique that scales data to a standard range, ensuring equitable contribution of all features to the model's learning process. In high SNR conditions, normalization enhances model performance by standardizing data ranges, thereby facilitating more effective learning of signal features; however, in low SNR conditions, where noise predominates, its impact is limited, as normalization does not mitigate the underlying interference caused by noise.

6.1.3. Multi-SNR data input and network parameter adjustment

To enhance the model's generalizability across different SNR conditions, this research trains it with a dataset covering multiple SNRs (e.g., -10 dB, 0 dB, 10 dB, 20 dB). This allows the model to learn features under various noise levels.

Multi-SNR training offers significant benefits by enabling the model to learn robust features capable of effectively handling signal waveform variations across a range of signal-to-noise ratios (SNRs), thereby enhancing its adaptability and performance in diverse and challenging environments.

To enhance the performance of the network, several parameter adjustments were proposed: increasing the number of CNN layers to extract deeper spatial features, extending the length of the LSTM to better capture temporal dependencies, and introducing dual-head attention mechanisms to selectively focus on different parts of the input sequence, thereby improving the model's ability to process complex data effectively [13].

6.1.4. Comprehensive performance comparison

To assess the model's generalizability, this research uses a weighted average accuracy method, giving higher weights to low SNR conditions.

- Weighted Accuracy Method:

- Weights: -10 dB: 0.4, 0 dB: 0.3, 10 dB: 0.2, 20 dB: 0.1
- Comprehensive score = (accuracy at -10 dB * 0.4 + accuracy at 0 dB * 0.3 + accuracy at 10 dB * 0.2 + accuracy at 20 dB * 0.1)

Table 1: Performance comparison data

Model	-10 dB Accuracy	0 dB Accuracy	10 dB Accuracy	20 dB Accuracy	Comprehensive Score
Basic CLDNN	4.2%	49.4%	82.4%	78.8%	40.9
Enhanced CLDNN	8.3%	48.3%	81.5%	83.5%	42.5

As shown in Table 1, the enhanced CLDNN model shows improved accuracy in low SNR conditions (e.g., -10 dB from 4.2% to 8.3%) and a higher comprehensive score (42.5), indicating better generalizability.

6.1.5. Conclusion of enhancing generalizability

This study optimized the CLDNN model for radio modulation signal recognition using the DeepSig Dataset. Through enhanced data preprocessing (normalization), multi-SNR training (-10 dB to 20 dB), and network adjustments (additional CNN layers, extended LSTM, dual-head attention), the model's weighted accuracy improved by 3.9% (40.9 to 42.5), with a significant gain at -10 dB (4.2% to 8.3%). These enhancements boost robustness and generalizability, making the CLDNN effective for real-world applications like spectrum monitoring and defense communications.

6.2. Design and comparison of GRU-optimized model

6.2.1. Principle and advantages of GRU

In the field of modulation signal recognition, optimizing deep learning models is crucial for improving recognition accuracy and processing efficiency. This section explores the application of Gated Recurrent Units (GRU) in optimizing the CLDNN model, analyzing its working principle and advantages, and comparing its performance with the original CLDNN model.

GRU is an improved type of recurrent neural network (RNN) unit designed to solve the gradient vanishing and exploding problems in traditional RNNs when handling long sequences. By introducing gating mechanisms, GRU can effectively control the flow and memory of information. Compared to Long Short-Term Memory (LSTM), GRU has a simpler structure with fewer parameters but often performs similarly or even better in many tasks. Inspired by the efficient modulation classification using deep LSTM-aided convolutional networks proposed by Chen et al., the application of GRU in this study leverages its efficiency to enhance the CLDNN model for modulation signal recognition tasks [14].

A. Principle of GRU:

GRU units contain two key gating structures: the update gate and the reset gate.

- Update Gate: Determines the fusion weight of the current input and historical information, controlling how much new information updates the hidden state.
- Reset Gate: Decides whether to ignore historical information, allowing the model to "forget" irrelevant data when necessary.
- The hidden state update equations for GRU are as follows:

$$z_t = \sigma(Wzxt + Uzht - 1 + bz) \quad (1)$$

$$rt = \sigma(Wrxt + Urht - 1 + br) \quad (2)$$

$$\tilde{h}_t = \tanh(W_h x_t + U_h(r_t \odot h_{t-1}) + b_h) \quad (3)$$

$$h_t = (1 - zt) \odot h_{t-1} + zt \odot \tilde{h}_t \quad (4)$$

where:

x_t : Input at time t

h_{t-1} : Hidden state at time $t - 1$

σ : Sigmoid activation function

$\odot$: Element -wise multiplication

zt : Update gate output

rt : Reset gate output

$\tilde{h}_t$: Candidate hidden state

h_t : Current hidden state

A. Advantages of GRU

Compared to LSTM, GRU offers several notable advantages that enhance its suitability for modulation signal recognition tasks: it features fewer parameters with only two gating units (update and reset gates) as opposed to LSTM's three (input, forget, and output gates), resulting in reduced complexity and higher training efficiency; its simplified structure also leads to greater computational efficiency, particularly when processing large datasets; and research indicates that GRU delivers comparable or even superior performance to LSTM across various tasks, especially in resource-constrained environments [15]. These characteristics position GRU as an ideal alternative to LSTM in scenarios demanding rapid training and real-time inference.

6.2.2. Performance evaluation

To evaluate the performance of GRU in the CLDNN model, this research replaced the LSTM layer in the original CLDNN with a GRU layer and compared it with the original model. The evaluation focuses on processing speed and performance under low SNR conditions.

A. Processing Speed Analysis:

Processing speed is a key indicator for measuring the practicality of the model, especially in real-time signal processing. We compared the training time to reach optimal parameters for both GRU and LSTM in CLDNN.

Table 2: Training time comparison

Model	Training Time (seconds)
CLDNN (LSTM)	2891
CLDNN (GRU)	1177

From Table 2, it can be seen that the training time for the GRU model is reduced by approximately 59.3% compared to the LSTM model. This improvement is mainly due to the fewer parameters and lower computational complexity of GRU, significantly speeding up the training process.

B. Performance under Low SNR Conditions

Under low signal-to-noise ratio (SNR) conditions, signal recognition becomes more challenging, and the model's robustness is particularly critical. This research evaluated the classification accuracy of GRU and LSTM models at 10 dB SNR and analyzed their performance.

C. Accuracy Comparison (10 dB SNR)

In evaluating the performance of CLDNN models, the study found that LSTM and GRU exhibit comparable overall accuracy rates of 82.4% and 82.2%, respectively; however, their performance varies significantly across specific modulation types: for simple modulations, both GRU and LSTM achieve near-100% accuracy, performing similarly; in high-order modulations such as 64QAM, GRU's accuracy drops to only 40.1%, substantially lower than LSTM by approximately 20%-30%; and for analog modulations, GRU performs comparably to LSTM on AM-DSB-SC (around 90%) but achieves only 44.0% on AM-SSB-WC, indicating poorer performance.

6.2.3. Conclusion of optimizing speed

As an efficient recurrent neural network unit, GRU significantly reduces training time by approximately 59.3% when replacing LSTM in the CLDNN model, achieving faster convergence without compromising core performance. Experimental results show that under 10 dB SNR conditions, the classification accuracy only slightly decreases by 0.2% from 82.4% to 82.2%, maintaining strong recognition capabilities for simple modulations like BPSK and QPSK. While GRU demonstrates comparable performance for AM-DSB-SC ($\approx 90\%$) and minor declines in high-order modulations such as 64QAM (40.1% vs LSTM's 60-70%), its parameter reduction advantage (2 gating mechanisms vs LSTM's 3) makes it particularly suitable for real-time applications requiring fast inference. Future research should explore GRU integration with advanced architectures like attention mechanisms or residual networks to address accuracy gaps in complex modulations while preserving computational efficiency [16].

7. Conclusion

In conclusion, this paper highlights the effectiveness of the CLDNN neural network in the automatic recognition of radio modulation signals, achieving an impressive 82.4% accuracy at 10 dB SNR, surpassing the performance of CNN and LSTM models. Through optimization efforts, including the integration of GRU, the weighted accuracy increased by 4%, while training time was significantly reduced by 59.3%. These improvements position CLDNN as a highly efficient and viable solution for real-time applications in radio signal processing. The model's enhanced performance and reduced computational demands make it particularly promising for practical deployment in scenarios requiring fast and accurate modulation recognition, such as wireless communication systems and spectrum monitoring. Despite challenges in complex real-world environments and resource-constrained devices, the optimized CLDNN demonstrates substantial potential for advancing signal recognition technology.

References

- [1] S. Hu, Y. Pei, P. P. Liang, and Y.-C. Liang, "Deep learning based automatic modulation recognition: Models, datasets, and challenges," *Digital Signal Processing*, vol. 129, p. 103685, 2022.
- [2] J. Jiang, Z. Wang, H. Zhao, and S. Qiu, "Modulation recognition method of satellite communication based on CLDNN model," in *Proc. IEEE Int. Conf. Signal Process., Commun. Comput. (ICSPCC)*, pp. 1-5, 2021.
- [3] S. Wei, Q. Qu, H. Su, M. Wang, J. Shi, and K. Hao, "Intra-pulse modulation radar signal recognition based on CLDN network," *IET Radar, Sonar & Navigation*, vol. 14, no. 5, pp. 803-810, 2020.
- [4] S. Li, Y. Zhang, X. Wang, and Z. Xu, "Deep learning-based automatic modulation classification using robust CNN architecture for cognitive radio networks," *Sensors (Basel)*, vol. 23, no. 23, p. 9467, 2023.
- [5] Y. Wang, M. Liu, J. Yang, and G. Gui, "A survey of applications of deep learning in radio signal modulation recognition," *Applied Sciences*, vol. 12, no. 23, p. 12052, 2022.
- [6] H. Zhang et al., "A novel automatic modulation classification scheme based on multi-scale networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 8, no. 1, pp. 97-110, 2022.
- [7] J. Jiang, Z. Wang, H. Zhao, and S. Qiu, "Modulation recognition method of satellite communication based on CLDNN model," in *Proc. IEEE Int. Conf. Signal Process., Commun. Comput. (ICSPCC)*, pp. 1-5, 2021.

- [8] Y. Wang, M. Liu, J. Yang, and G. Gui, "A survey of applications of deep learning in radio signal modulation recognition," *Applied Sciences*, vol. 12, no. 23, p. 12052, 2022.
- [9] S. Wei, Q. Qu, H. Su, M. Wang, J. Shi, and K. Hao, "Intra-pulse modulation radar signal recognition based on CLDN network," *IET Radar, Sonar & Navigation*, vol. 14, no. 5, pp. 803-810, 2020.
- [10] X. Li et al., "Research on modulation signal recognition based on CLDNN network," *Electronics*, vol. 11, no. 9, p. 1379, 2022.
- [11] M. A. M. Hasan, S. M. S. Reza, and M. M. Rahman, "A comparative study between CNN, LSTM, and CLDNN models in the context of radio modulation classification," in *2020 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, pp. 1-6, 2020.
- [12] B. Zou, X. Zeng, and F. Wang, "Research on modulation signal recognition based on CLDNN network," *Electronics*, vol. 11, no. 9, p. 1379, 2022.
- [13] Z. He and X. Zeng, "Research on communication signal modulation recognition based on a CCLDNN," *Electronics*, vol. 13, no. 9, p. 1604, 2024.
- [14] Z. Chen, Y. Wang, and X. Wang, "A novel efficient automatic modulation classification algorithm using deep LSTM aided convolutional networks," *Acta Automatica Sinica*, vol. 47, no. 2, pp. 377-386, 2021.
- [15] T. J. O'Shea, T. Roy, and T. C. Clancy, "Deep learning for wireless communications," in *Machine Learning for Future Wireless Communications*, F. L. Luo, Ed., Wiley, pp. 217-235, 2020.
- [16] Y. Qu, "Enhancing automatic modulation recognition through robust global feature extraction," *arXiv preprint arXiv:2401.01056*, 2024.