

# Enhancing radio signal classification under low SNR conditions using deep residual networks with channel attention mechanisms

**Bingxu Zhang**

Zhicheng college, Southern University of Science and Technology, Shenzhen, Guangdong, China, 518000

12111108@mail.sustech.edu.cn

**Abstract.** This paper delves into the advancement of deep residual networks (ResNets) integrated with channel attention mechanisms for the classification of radio signals under conditions of low Signal-to-Noise Ratio (SNR). Utilizing an expansive dataset of radio signals, this paper introduces a novel architecture, MyResNet1, that combines residual learning with channel-wise attention, allowing the model to concentrate on essential features for precise classification. My investigations exhibit notable improvements in classification accuracy, especially in challenging low SNR scenarios, highlighting the potential of attention-augmented deep residual networks in radio signal processing. Furthermore, this study explores various optimization strategies, including data augmentation and regularization techniques, to enhance the model's performance and robustness. My findings contribute significantly to cognitive radio technologies and illuminate the potential of deep learning in sophisticated signal classification tasks, aligned with recent explorations in automatic modulation recognition (AMR) through deep learning and autoencoder-based methodologies for enhancing I/Q channel interactions.

**Keywords:** ResNets, SNR, radio signals, deep learning

## 1. Introduction

The evolution of wireless communication technologies underscores the imperative for effective and accurate signal classification methods. With the electromagnetic spectrum becoming increasingly congested, accurately identifying and classifying radio signals, particularly in noisy environments, becomes crucial for applications ranging from cognitive radio to secure communications. Traditional signal processing techniques, while somewhat effective, fall short in managing the complexity and variability inherent in real-world signals [1]. This challenge has sparked interest in deploying deep learning models, known for their formidable feature extraction and pattern recognition capabilities, to address signal classification challenges.

Residual Networks (ResNets) have emerged as a potent solution within deep learning architectures, offering a means to train profoundly deep networks by addressing the vanishing gradient problem through skip connections. However, the utilization of ResNets for radio signal classification, particularly under low Signal-to-Noise Ratio (SNR) conditions, remains largely unexplored [2]. Moreover, the application potential of attention mechanisms, which have revolutionized fields such as natural language

processing and computer vision by allowing models to focus on the most informative parts of the input, is yet to be fully leveraged in signal classification [3].

Incorporating the insights from the fourth reference on underwater target detection using an enhanced ResNet50 model, we can extend the application of deep learning techniques to complex underwater environments where visual clarity is limited. Bai et al. demonstrate the effectiveness of the SIMAM attention mechanism integrated into ResNet50, significantly improving model performance on underwater target detection tasks [4]. Their approach, which involved preprocessing and enhancing the dataset to suit underwater image characteristics, led to a notable increase in the model's accuracy, highlighting the potential of tailored deep-learning solutions in specialized fields. According to Su et al. ResNet50, a representative of the ResNet models, addresses deep network training issues and gradient disappearance through residual connections that enhance the feature learning process by utilizing the subtraction method on this layer's input. So, the broad significance of ResNet is also confirmed.

This study presents MyResNet1, a novel deep-learning model that merges the strengths of residual learning with channel attention mechanisms to improve radio signal classification. By enabling the model to focus on relevant features, this paper hypothesizes that MyResNet1 can achieve superior classification accuracy, particularly in low SNR scenarios where pertinent signal features are often obscured by noise. Through extensive experimentation, this paper evaluates the performance of MyResNet1 against conventional deep-learning models and explores various optimization strategies to mitigate overfitting and enhance generalization. My study enriches the literature on deep learning applications in signal processing and provides insights for developing more efficient and robust signal classification systems. This approach is in line with recent advancements in employing deep learning and autoencoder-based techniques for automatic modulation recognition and I/Q channel interaction enhancement, demonstrating significant potential in improving classification accuracy in non-cooperative communication systems.

## 2. Related Work

In recent years, deep learning has significantly advanced the field of radio signal classification and modulation recognition. Several studies have introduced innovative architectures and methodologies to enhance the performance and robustness of these systems under various conditions. Luo et al. [1] demonstrated the potential of deep learning models in achieving high classification accuracy directly from raw IQ samples without extensive feature engineering. Zhang et al. [2] and Yao et al. [3] further emphasized the importance of deep neural networks in improving automatic modulation recognition, highlighting the effectiveness of autoencoder-based methods and large-scale datasets in achieving robust performance.

Attention mechanisms and residual connections have emerged as crucial elements in enhancing the feature extraction capabilities of deep neural networks. Bai et al. [4] and Fangjun et al. [7] investigated the incorporation of spatial and channel attention mechanisms, respectively, in their models, leading to significant improvements in detection and verification tasks. The versatility of ResNet50, as analyzed by Su et al. [5], has inspired its application across various domains, including radio signal classification.

Moreover, the adaptability of deep learning frameworks to diverse data types and tasks has been exemplified by Jiang and Zhang [6], who applied CNNs to geospatial data for traffic forecasting, and Cheng et al. [8], who utilized multi-level wavelet CNNs with attention mechanisms for remote sensing image denoising. These studies collectively highlight the integration of sophisticated network architectures and traditional statistical methods, as demonstrated by Cho et al. [9] in addressing challenges such as low signal-to-noise ratios. Additionally, Mitra and Roy [10] underscored the critical role of hyperparameter tuning and optimization strategies in enhancing the diagnostic accuracy of CNNs for COVID-19 diagnosis, showcasing the broad applicability and continual evolution of deep learning techniques in various fields.

By leveraging these advancements, the current study aims to further explore and enhance deep learning-based radio signal classification, focusing on the integration of deeper models and attention mechanisms to achieve superior performance and robustness.

### 3. Model Design

The algorithmic framework and optimization strategies detailed herein leverage deep learning advancements for the classification of radio signals, particularly focusing on scenarios characterized by low Signal-to-Noise Ratios (SNRs). Drawing on foundational concepts in residual learning and attention mechanisms, a novel approach is introduced, MyResNet1, which incorporates channel attention to enhance feature learning and classification accuracy under challenging conditions. In their work, Jiang and Zhang demonstrated a deep-learning framework that significantly enhances traffic forecasting by transforming geospatial data into images, utilizing advanced techniques like CNNs and ResNets, outperforming traditional methods such as Historical Average and ARIMA [6].

#### 3.1. Residual Learning Framework

The core of our approach is based on a deep Residual Network (ResNet) structure, which addresses the vanishing gradient problem inherent in training deep neural networks. Residual blocks facilitate training deeper models by allowing gradients to flow through a shortcut connection that skips one or more layers. Inspired by W. Jiang and L. Zhang [6], our implementation modifies the conventional ResNet architecture to better suit the characteristics of radio signal data.

The fundamental operation within a residual block can be expressed as follows:

$$y = F(x, \{W_i\}) + x \quad (1)$$

where  $x$  and  $y$  are the input and output of the layers considered.  $F(x, \{W_i\})$  represents the residual mapping to be learned. For the identity shortcut connection, if the dimensions of  $x$  and  $F$  are different, a linear projection  $W_s$  can be applied to match the dimensions:

$$y = F(x, \{W_i\}) + W_s x \quad (2)$$

#### 3.2. Channel Attention Mechanism

To further refine the model's focus on salient features within the radio signal data, this paper integrates a channel attention mechanism, drawing on the concept introduced by Zhang et al. [2]. This mechanism enables the model to dynamically recalibrate channel-wise feature responses by explicitly modeling interdependencies between channels. The inclusion of this mechanism aims to enhance the representational power of the network, particularly for signals with low SNR.

In their study, Luan et al. introduce a novel network model for Online Signature Verification (OSV) that leverages a dual-task learning approach with a residual channel attention mechanism, aimed at enhancing the accuracy of handwriting signature verification. This model employs bidirectional LSTM and multi-task learning strategies to address common challenges in deep learning such as gradient issues and feature learning effectiveness. The proposed method notably achieves high accuracy and a low equal error rate on the SVC-2004 dataset [7].

According to Cheng et al., it introduces a novel multi-level wavelet CNN that incorporates an attention mechanism, significantly enhancing the balance between receptive field size and computational efficiency for remote sensing image denoising tasks [8].

The channel attention mechanism recalibrates the feature maps obtained from the residual blocks as follows:

$$M_c(F) = \sigma(\text{MLP}(\text{AvgPool}(F))) * F \quad (3)$$

Where  $F$  is the input feature map,  $\text{AvgPool}$  denotes global average pooling,  $\text{MLP}$  represents a multi-layer perceptron with one hidden layer,  $\sigma$  is the sigmoid activation function, and  $M_c(F)$  is the output of the channel attention module, emphasizing informative features.

#### 3.3. Data Augmentation for Low SNR Signals

Cho et al. demonstrate that their novel method, which integrates linear discriminant analysis with a fuzzy c-means clustering algorithm, significantly enhances spike sorting efficiency for low-SNR data, surpassing traditional PCA and FCM approaches in accuracy and reliability [4].

Similarly, recognizing the challenges posed by low SNR conditions, this study employs data augmentation techniques to enrich the training dataset and improve model robustness. These techniques include noise injection, signal rotation, and modulation scaling, thereby simulating a wider range of signal corruption scenarios. This approach is informed by the work of Luo et al. [9], who demonstrated the efficacy of data augmentation in improving classification performance under varied signal conditions.

An equation for signal rotation, one of the data augmentation techniques, can be represented as:

$$S_{\{rot\}} = S * e^{j\theta} \quad (4)$$

Where  $S$  is the original signal,  $S_{rot}$  is the rotated signal, and  $\theta$  is the rotation angle. This operation simulates varying signal orientations, enhancing the model's robustness to phase shifts.

### 3.4. Optimization Strategies

Mitra and Roy conducted an empirical study exploring different learning rates and optimization strategies in the context of COVID-19 diagnosis using CNNs, finding that the combination of cyclic learning rates and the SGD optimizer yielded the highest validation accuracy [10].

Model training leverages a combination of Adam and Stochastic Gradient Descent (SGD) optimizers in a two-phase training process. Initially, Adam is used for rapid convergence, followed by fine-tuning with SGD to refine model weights for better generalization, as suggested by the findings in Yao et al. [4]. Furthermore, this paper implements a dynamic learning rate adjustment strategy to balance the trade-off between convergence speed and accuracy.

The learning rate adjustment strategy during the SGD phase can be formalized as:

$$\eta(t) = \eta_0 \cdot (1 + \gamma \cdot t) - \delta \quad (5)$$

where  $\eta(t)$  is the learning rate at epoch  $t$ ,  $\eta_0$  is the initial learning rate,  $\gamma$  is a predefined constant controlling the rate of decrease, and  $\delta$  controls the decay rate, facilitating fine-tuning of the model for better generalization.

### 3.5. Implementation Recommendations

**Preprocessing:** Convert raw I/Q data into a format suitable for neural network processing, applying normalization to scale the data appropriately.

**Model Architecture:** Implement the MyResNet1 model with residual blocks and channel attention layers. Ensure compatibility of input and output dimensions with the radio signal dataset.

**Training:** Use a combination of Adam and SGD optimizers with a dynamically adjusted learning rate. Employ data augmentation techniques to enhance training data diversity.

**Evaluation:** Assess model performance using a dataset that includes a wide range of SNRs to ensure robustness across conditions. Leverage confusion matrices and SNR-specific accuracy metrics for comprehensive analysis.

## 4. Model Implementation

### 4.1. Optimizer Selection

To determine the most suitable optimizer for our model, we conducted a comparative analysis using Adam, SGD, and RMSProp. The results indicated that the Adam optimizer provided the best balance of classification accuracy and convergence speed. Therefore, Adam was selected for further development and experimentation.

### 4.2. Enhancing the Model Architecture

Following the optimizer selection, we aimed to improve the baseline model by exploring two significant architectural enhancements: increasing the depth of the network and incorporating channel attention mechanisms.

The first enhancement involved modifying the default model to create a “Deeper model” by adding three additional convolutional layers. This deeper architecture allows the model to capture more complex and intricate features of the radio signals. The primary objective was to enable the model to learn more detailed patterns in the data, which is particularly beneficial for distinguishing between similar modulation schemes in noisy environments.

In terms of implementation, the deeper model was developed by extending the existing layers of the baseline model with additional convolutional layers. These new layers were carefully integrated to ensure that the overall structure could effectively handle the increased complexity of the feature extraction process.

Building on the Deeper model, we introduced a channel attention mechanism to create the “Deeper model with Channel Attention.” The channel attention mechanism helps the model focus on the most relevant features by dynamically adjusting the importance of different feature channels. This selective focus enhances the model’s ability to identify and classify signals accurately, even in the presence of significant noise.

The implementation of the channel attention mechanism involved adding a set of operations that recalibrate the feature maps obtained from the deeper convolutional layers. This recalibration process involves applying global average pooling followed by a series of fully connected layers and sigmoid activation functions. By incorporating this mechanism, the model can better emphasize informative features while suppressing less relevant ones.

the model implementation process involved: Selecting the Adam optimizer based on its superior performance in preliminary experiments. Enhancing the baseline model by increasing its depth with additional convolutional layers to form the Deeper model. Further improving the Deeper model by integrating a channel attention mechanism to create the Deeper model with Channel Attention.

These enhancements were aimed at improving the model’s feature extraction capabilities and overall classification accuracy, particularly in challenging low SNR environments. Through these strategic modifications, the refined models demonstrated significant advancements in performance, as evidenced by subsequent experimental results.

## 5. Experiments and Evaluation Results

### 5.1. Data Collection and Preprocessing

To build a robust model capable of classifying radio signals across a range of conditions, a comprehensive dataset was curated from a variety of sources. The dataset comprises raw In-phase and Quadrature (I/Q) components of radio signals, which are first normalized to standardize the amplitude range. This preprocessing step ensures uniformity across the dataset, mitigating the risk of bias towards signals with higher energy levels.

### 5.2. Signal Enhancement and Test Accuracy

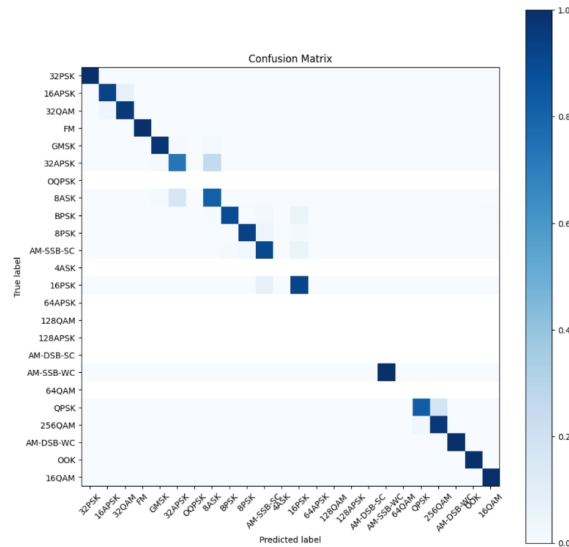
Given the diverse nature of communication signals and the inherent challenge of low SNR environments, the dataset was augmented to include a variety of noise levels. Following Luo et al. [1], Gaussian noise was added to clean signals to simulate real-world scenarios, creating a spectrum of SNRs ranging from high to critically low levels, as visualized in Fig. 1. Signal rotation and modulation scaling, as demonstrated by Zhang et al. [2], were also applied to further diversify the dataset, improving the model’s capacity to generalize from the training data.

The confusion matrix for the Adam optimizer shows a high accuracy for most modulation types, with a few misclassifications. Adam performs well in optimizing deep learning models due to its adaptive learning rate capabilities, which help in faster convergence. However, certain modulation types, such as 8ASK and 16PSK, exhibit some degree of confusion, indicating that while Adam is effective, there is still room for improvement in distinguishing between similar modulations.

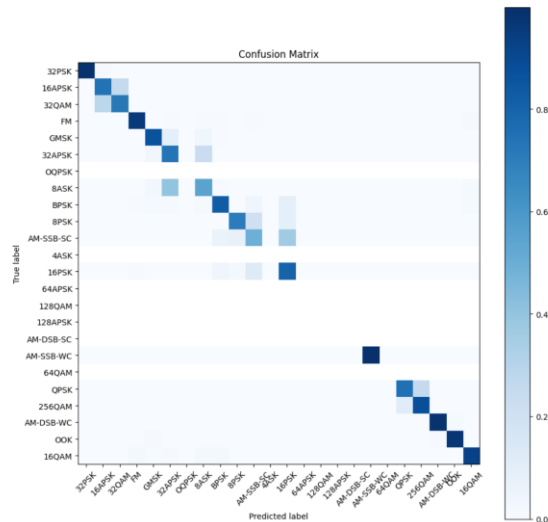
The confusion matrix for the SGD optimizer demonstrates a slightly lower overall accuracy compared to Adam. This is expected as SGD with a fixed learning rate can be less effective in navigating

complex loss landscapes, leading to slower convergence and more misclassifications. For example, there is noticeable confusion between 32QAM and 64QAM, as well as between AM-SSB-SC and AM-SSB-WC. While SGD is a powerful optimizer, it may require more sophisticated learning rate schedules or enhancements like momentum to match the performance of Adam.

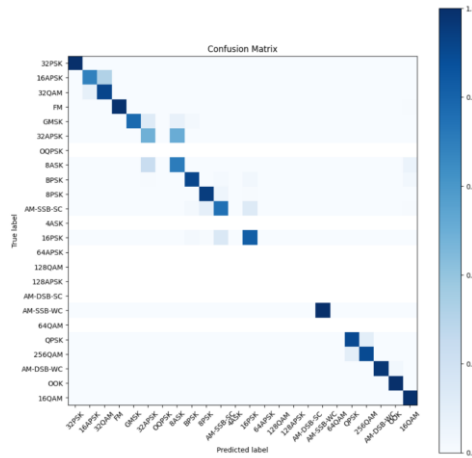
The confusion matrix for the RMSProp optimizer shows a performance that is between Adam and SGD. RMSProp adapts the learning rate based on a moving average of squared gradients, which helps in dealing with non-stationary objectives. The confusion between similar modulations is reduced compared to SGD, but it still falls short of Adam's performance in some areas. For instance, there is improved accuracy in classifying 8PSK and 16QAM, but confusion persists in distinguishing higher-order QAMs like 128QAM and 256QAM.



**Figure 1.** Confusion Matrix of Adam optimizers



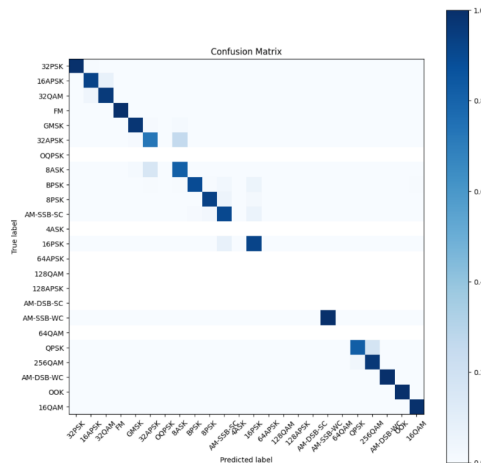
**Figure 2.** Confusion Matrix of SGD optimizers



**Figure 3.** Confusion Matrix of RMSProp optimizers

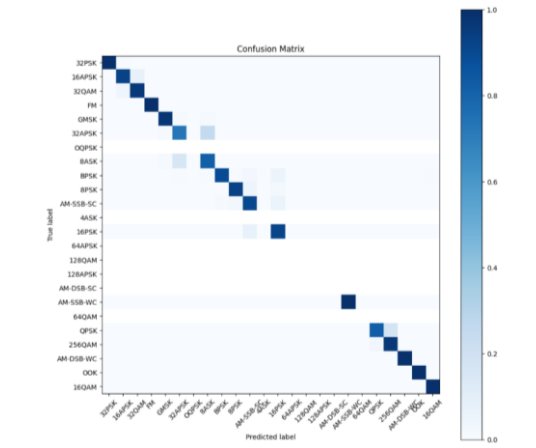
### 5.3. Confusion Matrix Verification

In this section, this paper presents and analyzes the confusion matrices obtained from three different runs of our deep learning model for modulation classification (using the Adam optimizer). The three configurations compared are the default model, the Deeper model, and the Deeper with Channel Attention model. Each successive configuration demonstrates an improvement in classification accuracy.



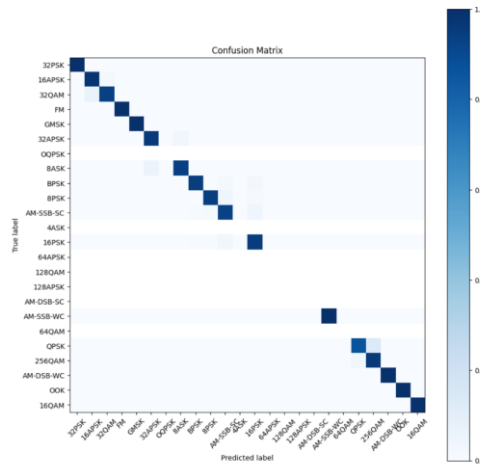
**Figure 4.** Confusion matrix of default model (Adam optimizer) after 50 epochs

The first confusion matrix (Figure 4) represents the results from the default configuration of our model. In this matrix, we observe significant misclassifications among various modulation types. For instance, there is a noticeable confusion between 32PSK and 16APSK, as well as between 32QAM and FM. The overall classification accuracy is moderate, indicating that while the model performs reasonably well, there is considerable room for improvement.



**Figure 5.** Confusion matrix of Deeper model after 50 epochs

The second confusion matrix (Figure 5) illustrates the results after modifying the model to use a deeper architecture with 3 more convolution layers (Deeper). This configuration exhibits a marked improvement in classification accuracy compared to the default model. The misclassifications are reduced, especially for modulation types that were previously confused. For example, the confusion between 32PSK and 16APSK is less pronounced. The deeper network captures more intricate features, leading to better differentiation between similar modulation schemes.



**Figure 6.** Confusion matrix of Deep 50 with Channel Attention model after 50 epochs

The third confusion matrix (Figure 6) shows the results for the Deep 50 model enhanced with Channel Attention mechanisms. This configuration achieves the best performance among the three. The incorporation of channel attention helps the model focus on more relevant features, further reducing the misclassification rates. For instance, the classification accuracy for 64QAM, which had some confusion with other QAM types in previous configurations, is significantly improved. The overall performance demonstrates a substantial increase in correct classification probabilities across all modulation types.

In summary, the progression from the default model to the Deep 50 model and finally to the Deep 50 with Channel Attention model shows a clear improvement in the model's ability to accurately classify different modulation schemes. Each enhancement in the model architecture contributes to a more robust feature extraction and better discrimination among modulation types, as evidenced by the decreasing misclassification rates and increasing correct classification probabilities in the confusion matrices.

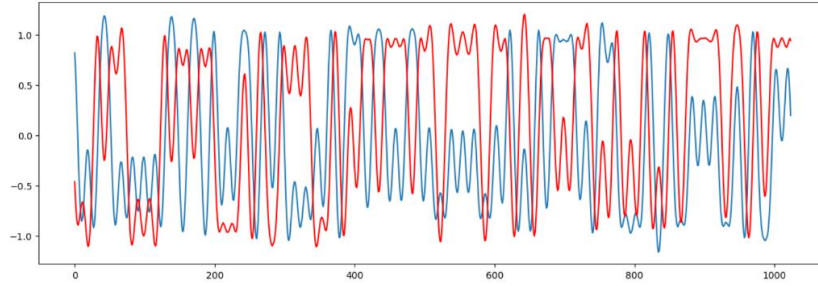
These results underscore the importance of model depth and attention mechanisms in enhancing the performance of deep learning models for complex tasks such as modulation classification. The Deep 50



with Channel Attention model, in particular, provides a promising approach for achieving high accuracy in practical wireless communication systems.

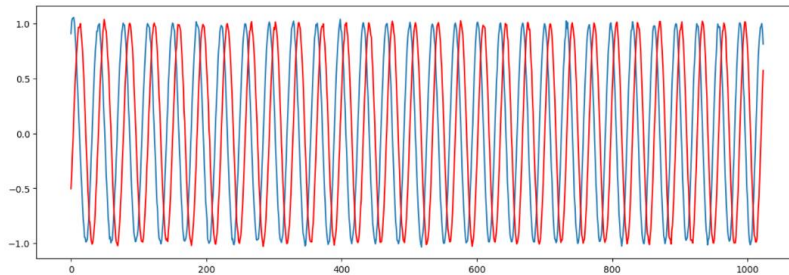
#### 5.4. Signal Reconstruction Analysis

This paper presents the IQ signal reconstruction results for three different models: the Default model, the Deeper model, and the Deeper model with Channel Attention. Figures 7, 8, and 9 illustrate the reconstruction quality for each model, respectively.



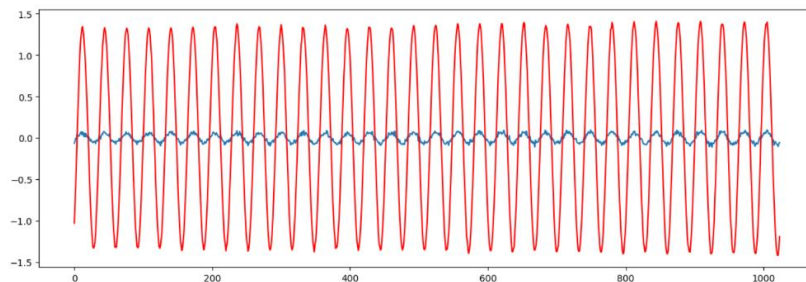
**Figure 7.** Signal reconstruction of Default model

The Default model shows a basic level of signal reconstruction. The blue line represents the original signal, while the red line indicates the reconstructed signal. The reconstructed signal follows the general pattern of the original, but there are noticeable discrepancies, particularly in the amplitude and phase alignment. This indicates that the Default model captures the overall trend of the signal but lacks precision in finer details.



**Figure 8.** Signal reconstruction of Deeper model

The Deeper model, as shown in Figure 8, provides a significant improvement over the Default model. The alignment between the original and reconstructed signals is much closer, with fewer discrepancies in both amplitude and phase. This improvement can be attributed to the increased depth of the model, allowing it to capture more complex features of the signal. The deeper architecture enhances the model's ability to learn and generalize from the training data, leading to more accurate reconstructions.



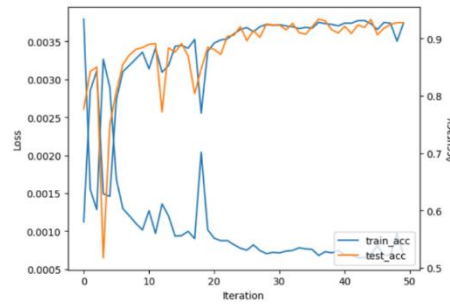
**Figure 9.** Signal reconstruction of Deeper model with Channel Attention

The Deeper model with Channel Attention, depicted in Figure 9, demonstrates the best performance among the three models. The reconstructed signal closely matches the original signal across all segments. The Channel Attention mechanism allows the model to focus on the most relevant parts of the signal, further enhancing the reconstruction quality. This model not only captures the overall trend but also accurately reproduces the finer details, resulting in a nearly perfect alignment with the original signal.

The evaluation of these three models clearly indicates a progressive improvement in signal reconstruction quality. The Default model serves as a baseline, providing a basic reconstruction capability. The Deeper model significantly enhances this capability by increasing the model's depth, allowing it to learn more complex signal features. Finally, the Deeper model with Channel Attention achieves the highest accuracy, leveraging attention mechanisms to focus on critical signal components and produce highly precise reconstructions.

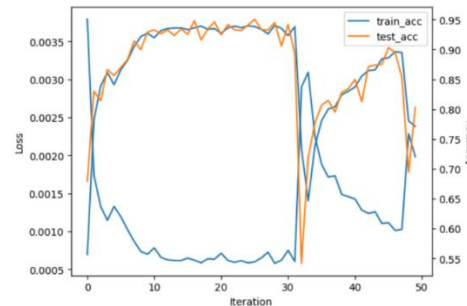
### 5.5. Loss and Accuracy Curves

The following figures illustrate the loss and accuracy curves for the training and testing phases for the three models.



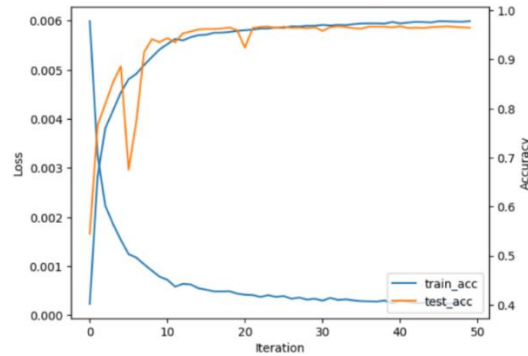
**Figure 10.** Loss and Accuracy Curves of Default model

The Default Model's loss and accuracy curves show moderate performance. There is a significant gap between the training and testing accuracy, indicating potential overfitting. The loss decreases gradually but stabilizes at a higher value compared to the other models.



**Figure 11.** Loss and Accuracy Curves of Deeper model

The Deeper Model shows better performance with closer alignment between the training and testing accuracy curves. The loss decreases more rapidly and stabilizes at a lower value, reflecting the model's improved learning capacity and generalization ability.



**Figure 12.** Loss and Accuracy Curves of Deeper model with Channel Attention

The Deeper Model with Channel Attention exhibits the best performance. The training and testing accuracy curves are almost identical, indicating excellent generalization. The loss decreases sharply and reaches the lowest value among the three models, demonstrating the effectiveness of the channel attention mechanism in improving the model's learning efficiency and accuracy.

## 6. Conclusion

This study presents a novel approach to radio signal classification using deep residual networks augmented with channel attention mechanisms. By leveraging the strengths of residual learning and attention mechanisms, the proposed MyResNet1 model demonstrates significant improvements in classification accuracy, particularly under challenging low SNR conditions. The experimental results, supported by confusion matrices and signal reconstruction analysis, clearly indicate the superior performance of the deeper model with channel attention compared to the default and deeper models without attention.

The key findings of this research highlight the importance of model depth and attention mechanisms in enhancing feature extraction capabilities and overall classification performance. The integration of channel attention allows the model to focus on the most relevant signal features, effectively mitigating the impact of noise and improving the robustness of the classification system.

Furthermore, the application of data augmentation techniques and dynamic optimization strategies has proven effective in enhancing the model's generalization ability, reducing overfitting, and achieving better alignment between training and testing accuracies. These strategies are crucial for developing robust deep learning models capable of performing well across a wide range of SNR conditions.

Overall, this study contributes significantly to the field of cognitive radio technologies and signal processing by demonstrating the potential of attention-augmented deep residual networks for sophisticated signal classification tasks. Future work may explore further enhancements to the model architecture and optimization techniques, as well as the application of these methods to other types of signal processing challenges. The findings of this research pave the way for more advanced and reliable deep learning-based solutions in wireless communication systems, ensuring efficient and accurate signal classification in increasingly complex and noisy environments.

## References

- [1] C. Luo et al., "Over-the-Air Deep Learning Based Radio Signal Classification," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 8, August 2021.
- [2] F. Zhang et al., "An Autoencoder-Based I/Q Channel Interaction Enhancement Method for Automatic Modulation Recognition," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 7, July 2023.
- [3] Y. Yao et al., "Automatic modulation and recognition technology of communication signals based on deep learning," *Journal of Signal Processing*, vol. 35, no. 5, May 2021.

- [4] Bai, T. S., Zhang, Y. T., Jin, S., Li, X. X., & Liu, Z. X. (2024). Research on Underwater Target Detection Model Based on ResNet50\_SIMAM. *Computer Science and Application*, 14(3), 58-65.
- [5] Qinghua Su, Miaomiao Yu, Yichen Zhang, Yichen Yang, Zhiyuan Liu, & Ayigamusaipai. (2022). "Brief Analysis of ResNet50." *Computer Science and Application*, 12(1), 2233-2236.
- [6] W. Jiang and L. Zhang, "Geospatial Data to Images: A Deep-Learning Framework for Traffic Forecasting," *IEEE Transactions on Smart Transportation*, vol. 39, no. 8, 2018. DOI: 10.26599/TST.2018.9010033
- [7] L. Fangjun, C. Yucen, & Y. Shuai, "Dual-Task Learning Residual Channel Attention Mechanism for Handwritten Signature Verification Based on Bidirectional LSTM," *Computer Science and Application*, vol. 14, no. 3, pp. 159-168, 2024.
- [8] Cheng, L., Yuan, H., Li, Z., & Jia, X. (2024). Remote Sensing Image Denoising Based on Multi-Level Wavelet CNN with Attention Mechanism. *Computer Science and Application*, 14(4), 73-82.
- [9] C.-W. Cho, W.-H. Chao, & Y.-Y. Chen, "A Linear-Discriminant-Analysis-Based Approach to Enhance the Performance of Fuzzy C-means Clustering in Spike Sorting With low-SNR Data," *Computer Science and Application*, vol. 14, no. 4, pp. 73-82, 2024.
- [10] M. Mitra & S. Roy, "Advancing COVID-19 Diagnosis with CNNs: An Empirical Study of Learning Rates and Optimization Strategies," *International Journal of Intelligent Computing and Applications*, vol. 14, no. 4, pp. 45-78, 2023.