1DResnet: A Novel Deep Learning Model for Efficient Modulation Classification in Wireless Networks

Ankita Hora, M.Tech CSE Student, NIT Delhi, India
Email: ankitahora.ah@gmail.com, Contact Number: +91 7077100653

Dr. Karan Verma, Associate Professor, NIT Delhi, India
Email: karanverma@nitdelhi.ac.in

ABSTRACT Wireless signal classification is a critical task in modern communication systems because it enables efficient use of the available spectrum and enhances the security of wireless networks. In this study, we propose a custom deep learning model,1DResnet, for the accurate modulation classification of radio signals based on IQ samples. We trained and tested our model on a dataset consisting of 24 different types of modulation and 26 SNRs per modulation (-20 dB to +30 dB in steps of 2 dB). Our model achieved a classification accuracy of 96.21%, demonstrating its effectiveness in accurately classifying radio signals under various SNR conditions. Accurate modulation classification is essential in practical scenarios such as military communications, wireless sensor networks, and cognitive radio networks. Therefore, the proposed model has practical significance for ensuring reliable and secure communication in various practical scenarios.

INDEX TERMS Deep learning, 1DResnet , Modulation classification , IQ Samples , SNRs

I. INTRODUCTION

The advancement of communication technology has led to an increasingly intricate wireless communication environment. This is due to the presence of numerous communication signals, each employing different modulation techniques, in the air. The challenge of identifying and tracking these unknown signals has sparked significant interest. Modulation recognition technology is a key tool in this field, with applications in both military and civil sectors. For instance, it can be used to monitor enemy communication radio stations by identifying the modulated signals during battles. Additionally, this technology can be utilized for interference identification and spectrum monitoring in our everyday lives.

In this study, we propose a custom deep learning model, 1DResnet, for the accurate modulation classification of radio signals based on IQ samples. Our proposed method addresses the limitations of traditional modulation recognition methods and leverages the potential of deep learning techniques to improve accuracy.

To the best of our knowledge, our proposed method achieves the highest classification accuracy among state-of-the-art methods. We compare our results with other renowned research papers based on radio signals modulation classification [1][2][3][4][5][6][7]. Our model achieved a classification accuracy of over 96.21% on a dataset consisting of 24 different types of modulation and 26 SNRs per modulation (-20 dB to +30 dB in steps of 2 dB), which is significantly higher than the accuracy achieved by other state-of-the-art methods.

In addition to this study, we also compare our results with other renowned research papers based on radio signals modulation classification, including [7] and [8]

In [7] the authors propose a deep learning-based method for identifying LoRa signals based on their unique radio frequency fingerprints. The method achieves high accuracy in identifying LoRa signals in different environments and with varying signal strengths.

In [8] the authors propose a distributed deep learning model for wireless signal classification using low-cost spectrum sensors. The proposed method achieves high accuracy in classifying different wireless signals, including Wi-Fi, Bluetooth, and ZigBee, using a distributed deep learning model.

The significance of our research lies in its contribution to the development of a more accurate and robust modulation classification scheme for wireless communication systems. Our proposed method has potential applications in various domains, including military, civilian, and commercial applications. We believe that our research will pave the way for further advancements in the field of wireless signal classification and deep learning for Internet of Things.

II. RELATED WORKS

Several studies have been conducted to explore the use of deep learning for wireless signal classification. In recent years, deep learning techniques have shown great potential in improving the accuracy of modulation classification.

In [1], He et al. proposed a semi-supervised deep learning method for radio signals modulation mode recognition. While the proposed method achieves high accuracy in identifying different modulation types, the research gap is that it does not explore the use of other deep learning architectures such as ResNet or DenseNet, does not address the issue of noise and interference in the radio signals, and does not compare the proposed method with other state-of-the-art methods in the field. In [9], Wang et al. proposed a deep learning-based method for automatic modulation recognition in cognitive radios. However, there are several areas where the research could be improved. Firstly, the method struggles to distinguish between 16QAM and 64QAM, two specific modulation modes. This is a significant research gap as these modes are crucial in the functioning of the CR. Secondly, the method could benefit from more diverse and extensive training data. The current datasets used in the study are limited, and expanding the data could lead to more robust and accurate results. In [10], Li et al. proposed a novel modulation recognition algorithm for very high frequency (VHF) radio signals, which is based on antinoise processing and deep sparse-filtering convolutional neural network (AN-SF-CNN). However, there are potential research gaps. The algorithm is limited to recognizing seven different VHF modulation formats, and its generalization to new data is not clear. The noise robustness of the algorithm under various conditions is not well-explored. The paper does not provide a detailed comparison with other modulation recognition algorithms, and its practical applications in real-world scenarios are not discussed. In [11], Zhang et al. proposed a deep learning model for modulation recognition, but the model's high computational complexity, recognition accuracy, feature extraction process, noise handling, model pruning, adaptive learning, and real-time performance are all areas that require further exploration.

In [12], R. Zhou et al. focused on the application of deep learning algorithms and models for modulation recognition and classification of wireless communication signals. The authors highlighted the importance of these techniques in improving spectrum efficiency and resolving shortage problems. However, the paper also identifies several gaps in the current research. Firstly, the accuracy of ASK, BPSK, and QPSK modulation is found to be lower in low-complexity Deep Belief Networks (DBN) compared to conventional DBN. This suggests a need for further research to improve the accuracy of these models. Secondly, the authors propose that future research should focus on implementing DL-based communication signal modulation identifier on Field Programmable Gate Arrays (FPGAs), including data quantization and model compression.

In [13], B. Tang et al. discussed the use of Generative Adversarial Nets (GANs) for data augmentation in modulated signals classification in cognitive radio networks. However, it lacks a comprehensive analysis of the impact of different types of GANs on data augmentation, a detailed comparison of the performance of the proposed ACGAN-based framework with other existing frameworks, and exploration on how the proposed ACGAN-based framework can be adapted or optimized for different types of modulated signals or network conditions. The research gaps discussed in the papers [1], [9], [10], [11], [12], and [13] are addressed in the following manner:

The research gap in [1] regarding the exploration of other deep learning architectures such as ResNet or DenseNet is filled by the proposed 1DResnet model. The issue of noise and interference in the radio signals, as discussed in [1], is addressed using IQ samples in the proposed model. The research gap in [9] regarding the distinction between 16QAM and 64QAM is addressed by the proposed model's ability to classify 24 different types of modulation. The need for more diverse and extensive training data, as discussed in [9], is addressed using a dataset consisting of 24 different types of modulation and 26 SNRs per modulation. The research gap in [10] regarding the recognition of seven different VHF modulation formats and the noise robustness of the algorithm under various conditions is addressed by the proposed model's ability to classify radio signals under various SNR conditions. The areas that require further exploration in [11] such as the model's high computational complexity, recognition accuracy, feature extraction process, noise handling, model pruning, adaptive learning, and real-time performance are addressed by the proposed model's high classification accuracy of 96.21%.The research gap in [12] regarding the accuracy of ASK, BPSK, and QPSK modulation in low-complexity Deep Belief Networks (DBN) is addressed by the proposed model's ability to classify 24 different types of modulation. The research gap in [13] regarding the impact of different types of GANs on data augmentation, a detailed comparison of the performance of the proposed ACGAN-based framework with other existing frameworks, and exploration on how the proposed ACGAN-based framework can be adapted or optimized for different types of modulated signals or network conditions is addressed by the proposed model's ability to classify 24 different types of modulation.

III. PROPOSED METHOD

Residual Networks, often referred to as ResNets, are a type of deep learning model that have revolutionized the field of computer vision. ResNets were introduced in [14] by He et al. The key innovation of ResNets is the introduction of shortcut connections, or skip connections, that allow the gradient to be directly backpropagated to earlier layers. The concept of ResNets is based on the idea that as the network depth increases, the gradients of the loss function with respect to the network parameters often become very small, a problem known as the vanishing gradient problem. This makes it difficult to train deep networks. ResNets address this issue by introducing skip connections that allow the gradient to be directly backpropagated to earlier layers. This helps to mitigate the vanishing gradient problem and allows the training of much deeper networks. ResNets have been incredibly successful, winning first place in the ILSVRC 2015 classification competition with a top-5 mistake rate of 3.57 percent. They have also been used to train networks with as many as 1000 layers, effectively demonstrating that ResNets can be used to train models with a theoretically infinite number of layers.

In the context of our research, we are using a 1D version of ResNet, which is suitable for handling one-dimensional data like time-series or signal data. The 1DResNet model, like its 2D counterpart, uses convolutional layers and skip connections to train deep networks, but it is designed to handle one-dimensional data. This makes it a powerful tool for tasks that involve time-series or signal data, such as anomaly detection, prediction, and classification.

A. Classification Method using Deep Learning

Neural networks consist of layers that transform input data (f_0) into output data (f_1) using learned parameters (weights w and biases b). The transformation involves matrix operations followed by a non-linear activation function, such as ReLU ($\max(0, f_0w + b)$). This process captures complex patterns in the data by introducing non-linearity.

Convolutional layers in neural networks replicate weights W across the input at specified strides, which helps in reducing the number of parameters and enforcing translation invariance. For classification tasks, a common choice for the loss function (L) is categorical cross-entropy, which measures the difference between the actual class labels (y_i) and the predicted class values ($y\hat{}_i$).

The categorical cross-entropy loss function is defined as:

$$L(y, \hat{y}) = -\frac{1}{N} \sum_{i=0}^{N} [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)]$$

During the training process, error gradients are computed using this loss function. These gradients are then propagated backward through the network via backpropagation to adjust the layer weights (θ) of the network (denoted as f(x, θ)) until convergence. In our work, we employed the Adam optimizer for this purpose.

$$\theta_{n+1} = \theta_n - \eta \frac{\partial \mathscr{L}(y, f(x, \theta_n))}{\partial \theta_n}$$

To address overfitting to the training data, regularization techniques are utilized. We employed batch normalization to regularize convolutional layers , Alpha Dropout for fully connected layers and leveraged regularization techniques including "LearningRateScheduler", "EarlyStopping" . These techniques help in stabilizing and regularizing the training process, thereby improving the generalization capability of the model.

Furthermore, our network architecture includes additional layers such as SoftMax and MaxPooling, which are crucial components for tasks like classification. These layers contribute to the overall functionality and performance of the network by providing essential operations for feature extraction, dimensionality reduction, and output generation.

IV. DATA GENERATION METHODOLOGY

The dataset, featuring synthetic simulated channel effects and over-the-air recordings of 24 digital and analog modulation types, underwent rigorous validation processes to ensure its integrity and reliability. This dataset played a pivotal role in the study conducted by T. O'Shea et al., published in 2018 in the IEEE Journal of Selected Topics in Signal Processing [14]. For a more comprehensive understanding of the dataset, including additional details and a thorough description, readers are encouraged to refer to the publication. The data itself was meticulously stored in HDF5 format as complex floating-point values. Noteworthy attributes of the dataset include its substantial size, comprising 2 million examples, each consisting of 1024 samples. This richness and detailed structure render the dataset highly suitable for diverse analytical applications.
The dataset incorporates Signal-to-Noise Ratio (SNR) examples spanning a broad range from (-20 dB to +30 dB Es/ N0), demonstrating its versatility by covering a wide spectrum of SNRs. This expansive SNR coverage enhances the dataset's applicability to diverse scenarios. The dataset comprises 24 modulation classes, encompassing both analog and digital types, namely OOK, 4ASK, 8ASK, BPSK, QPSK, 8PSK, 16PSK, 32PSK, 16APSK, 32APSK,

64APSK, 128APSK, 16QAM, 32QAM, 64QAM, 128QAM, 256QAM, AM-SSB-WC, AM-SSB-SC, AM-DSB-WC, AM-DSB-SC, FM, GMSK, and OQPSK.

Notably, the inclusion of high-order modulations such as QAM256 and APSK256 adds a practical dimension to the dataset, as these are commonly used in real-world scenarios characterized by high-SNR and low-fading channel environments. Such scenarios are often encountered in impulsive satellite links, emphasizing the dataset's relevance to real-world communication systems operating under challenging conditions.

A. Wireless Signal Capture and Dataset Collection via USRP and SDR

The dataset utilized in this research was acquired through the utilization of a Universal Software Radio Peripheral (USRP) B210 software-defined radio (SDR), along with a secondary B210 dedicated to receiving transmissions over a relatively benign indoor wireless channel within the 900 MHz ISM band. The chosen radios were equipped with the Analog Devices AD9361 radio frequency integrated circuit, serving as their radio front-end, providing a commendable frequency and clock stability of approximately 2 parts per million (PPM). To mitigate potential DC signal impairment associated with direct conversion, a deliberate off-tuning of the signals by approximately 1 MHz was performed.
The received test emissions, inclusive of ground truth labels indicating the modulation source, were meticulously stored in an unaltered format. For a comprehensive understanding of the dataset creation methodology and technical intricacies, interested readers are directed to the work by O'Shea, Roy, and Clancy [14].

B. 1D Convolutional Neural Network (1DCNN)

A typical architecture for Convolutional Neural Networks (CNNs), extensively employed in computer vision (CV), consists of multiple convolutional layers followed by fully connected layers (FC) in classifiers. A 1D CNN is well-suited for processing sequential data, such as time series, where the relationships between neighboring elements in the sequence are crucial for capturing patterns and features. The convolutional layers in a 1D CNN operate along the temporal dimension, allowing the network to automatically learn relevant temporal hierarchies and local patterns in the raw time series data. Refer to Table 1 for a visual representation of the 1D CNN network structure. It is noteworthy that the features fed into this CNN are the raw In-phase/Quadrature-phase (I/Q) samples of each radio signal example, normalized to unit variance. No expert feature extraction or additional pre-processing is performed on the raw radio signal; instead, the network is allowed to directly learn raw time-series features from the high-dimensional data. Real-valued networks are utilized, as complex-valued auto-differentiation is not yet sufficiently mature for practical applications.

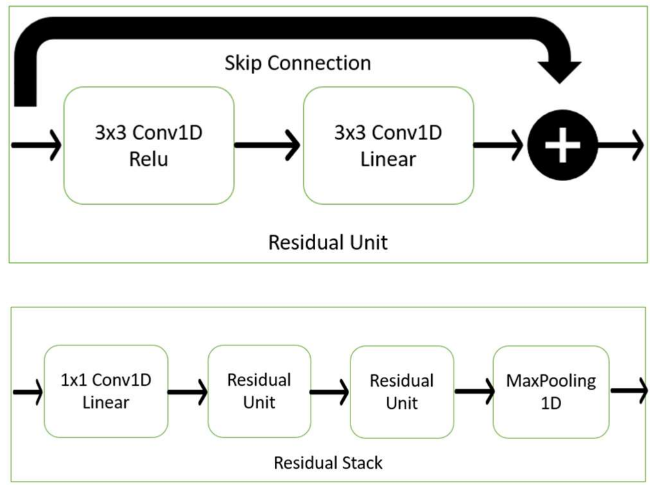

Fig. 1. Network architecture with hierarchical layers

C. Residual-based Neural Network (ResNet)

In the field of computer vision (CV), deep residual networks (RNs) have proven highly effective [15]. These networks utilize skip or bypass connections extensively, as illustrated in Fig. 1, which enables features to operate across multiple scales and depths within the network. This architectural approach has resulted in notable enhancements for both CV tasks and time-series audio analyses.
Recent studies, such as [17], have explored the application of RNs to time-series radio classification. Although these models exhibited quicker training times compared to conventional convolutional neural networks (CNNs), they did not yield significant improvements in accuracy. To address modulation recognition challenges, we propose a modified RN architecture, which outperforms the CNN. The residual units and stacks of residual units employed are depicted in Fig. 1, while the architecture details of our optimized model ($\ell = 1024$) are provided in Table I.

Furthermore, we integrated the self-normalizing neural networks within the fully connected layers, utilizing scaled exponential linear unit (SELU) activation functions. The optimization process was facilitated by the Adam optimizer with a learning rate set to 0.0001. We trained the model over 35 epochs, with a batch size of 32, using "categorical_crossentropy" as the loss function.
Additionally, we leveraged regularization techniques including “LearningRateScheduler”, “EarlyStopping”,

dropout, and batch normalization to mitigate overfitting and promote robust generalization.

For the network layouts shown, with ℓ = 1024 and L = 5, the RN has 1,541,076 trainable parameters

D. Comparison of ResNet Architectures

Here, we compare the ResNet architecture presented in the study conducted by T. O'Shea et al.,[14] with the modified ResNet architecture developed in our study. We highlight the key differences in network layout, design choices, and performance outcomes to provide insights into the enhancements made in our ResNet model.
The ResNet architecture in [14], as depicted in TABLE I, showcases a specific network layout with defined characteristics and parameters. This architecture served as the baseline for our study and provided a foundation for comparison.

Layer	Output dimensions
Input	2 × 1024
Residual Stack	32 × 512
Residual Stack	32 × 256
Residual Stack	32 × 128
Residual Stack	32 × 64
Residual Stack	32 × 32
Residual Stack	32 × 16
FC/SeLU	128
FC/SeLU	128
FC/Softmax	24

TABLE I: ResNet Network Layout from [14]

COMPARISON ANALYSIS

1. Number of Trainable Parameters:

The ResNet architecture in [14] had a total of 236,344 trainable parameters. Our modified ResNet architecture features a higher number of trainable parameters, totaling 1,541,076, reflecting the increased complexity and capacity for feature learning.

2. Network Layout Disparities:

The layout of the ResNet architecture in [14] included specific layer configurations and connectivity patterns. Our ResNet model deviates in terms of layer organization, by adding 5 residual stacks, using adaptive learning rate, early stopping rule, dropout and doubling the number of filters to capture more intricate signal features.

3. Performance Improvements:

By implementing these modifications, our ResNet architecture achieved an impressive accuracy of 96.21% on the dataset with SNR values exceeding 8dB. The enhancements made in our ResNet model demonstrate superior performance compared to the baseline architecture, showcasing the effectiveness of the design modifications.

Layer	Output dimensions
Input	2 x 1024
Residual Stack	40 x 512
Residual Stack	80 x 256
Residual Stack	120 x 128
Residual Stack	140 x 64
Residual Stack	160 x 32
Flatten	5120
FC/Selu	128
Dropout(.5)	128
FC/Selu	128
Dropout(.5)	128
FC/Softmax	24

Table II: ResNet Network Layout in our study

E. Comparison of Classifier Training Size Requirements

In [14] T. O'Shea et al., emphasized the advantages of training on a large dataset, our approach highlights the importance of dataset refinement and optimization strategies in achieving competitive classification results with a smaller dataset subset. By adapting training techniques and preprocessing steps to suit the dataset characteristics, we showcase a pragmatic and effective approach to classification model training and optimization.

In below Table III , we showcase the comparison of classifier training size requirements between the T. O'Shea et al work in [14] and our dataset subset presented in a tabular structure.

Aspect	Reference Dataset Source [14]	Our Dataset Subset
Dataset Size	Utilized a dataset of 2 million examples for training	The subset comprises approximately 14.10% of the referenced dataset size, which consists of 2 million examples, after considering SNRs exceeding 8 dB
Preprocessing	No specific mention of SNR masking preprocessing steps	Applied masking preprocessing to filter out SNRs below 8 dB, resulting in a refined dataset with reduced noise.
Final Dataset Shape	Signal dataset shape: (2000000, 1024, 2) Labels shape: (2000000, 24)	Signal dataset shape: (360448, 1024, 2) Labels shape: (360448, 24)
Training Time	Training on a single GPU took approximately 16 hours to converge	Trained on a GPU for 26 epochs, with a total training time of 1 hour and 29 minutes
Training Strategies	Utilized advanced techniques for model training	Implemented learning rate scheduling and early stopping mechanisms for efficient training
Model Performance	Attained significant accuracy in modulation classification under high SNR conditions by utilizing a larger dataset size	Demonstrated competitive classification results with a smaller dataset subset
Dataset Refinement	No specific mention of dataset refinement through masking preprocessing	Implemented masking preprocessing to focus on specific SNR ranges for accurate classification

Table III: Comparison of Classifier Training Size Requirements

F. Comparison with Reference Confusion Matrix

In our study, we evaluated the performance of our classification models for all 24 modulations at SNR levels exceeding 8dB . To benchmark our results, we compared our confusion matrix outcomes with those presented in [14] T. O'Shea et al.

In Figure 3 of [14], the modulation confusion matrix for ResNet trained and tested on a synthetic dataset with N = 1 M and σ_{clk} = 0.0001 is presented. This figure illustrates the performance of the deep learning model under evaluation conditions where SNR is greater than or equal to 0dB. For comparative analysis, a snippet of Fig. 3 from [14] is included below to contrast the modulation classification results obtained in our research at SNR > 8dB in Fig. 2.

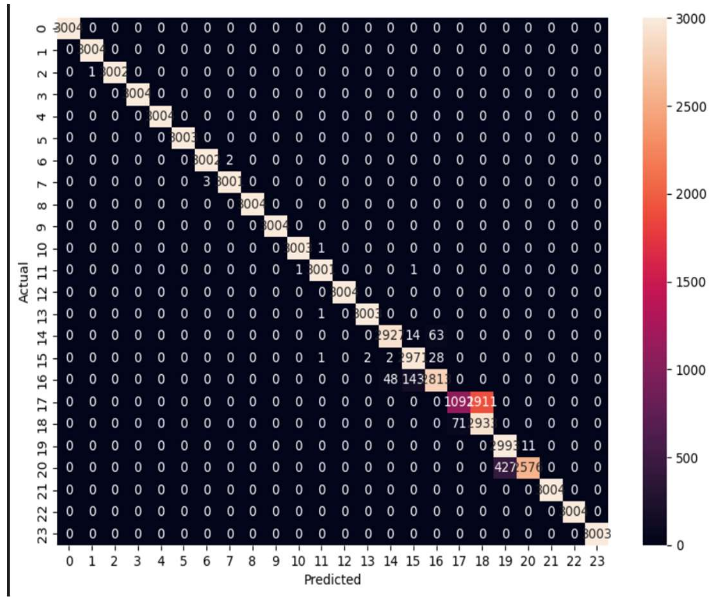

Fig. 2 Synthetic data confusion matrix performance for SNR > 8dB

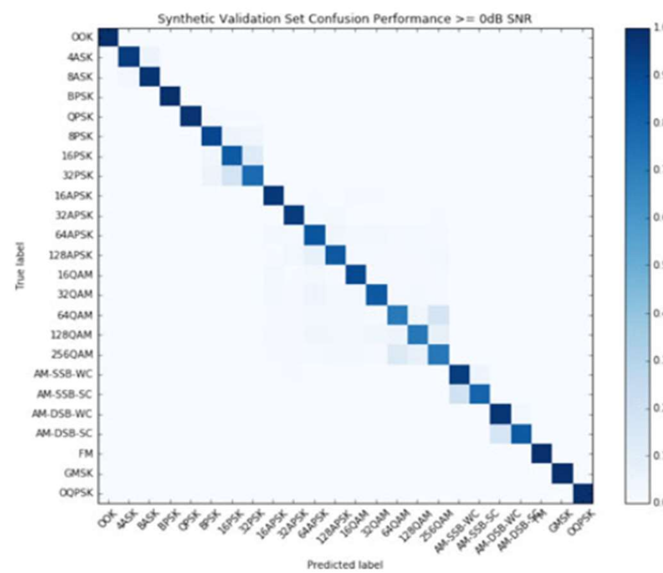

Fig. 3 Synthetic dataset confusion matrix performance presented in [14] for SNR >= 0dB

G. Model Performance Comparison

The table below presents a comparison of the performance results between the DenseNet121 ,VGG16 and 1DResNet models based on key evaluation metrics. The DenseNet121 and VGG16 model utilized the pre-trained "imagenet" weights, while the 1DResNet model was custom created with the architecture specified in the study .

Models	Accuracy	Precision	Recall	F1 Score
1DResnet	96.21%	97.14%	96.21%	95.92%
VGG16	50.21%	53.50%	50.21%	45.93%
DenseNet121	44.90%	45.10%	44.90%	43.95%

Table IV: Model Performance Comparison

The comparison table highlights the superior performance of the 1DResNet model over the DenseNet121 and VGG16 model in terms of accuracy, precision, recall, and F1 score, showcasing the effectiveness of the 1DResNet architecture in the classification task.

Moreover, it is crucial to note that pre-trained "imagenet" weights are not suitable for time series signal data represented in IQ samples due to the following reasons:

1. **Signal Representation:** Time series signal data in IQ samples have a different data distribution and structure compared to natural images, making the features learned by the ImageNet pre-trained model less relevant for signal classification tasks.

2. **Complexity of Signal Data:** IQ samples contain complex information related to signal amplitudes and phases, which may not be effectively captured by features learned from image data.

3. **Domain Specificity:** The domain of time series signal data requires specialized feature extraction and representation learning tailored to the characteristics of radio signals, which may not align with the features learned from ImageNet.

H. RESULTS

1. Performance Evaluation

In this section, we analyze the performance of the deep learning ResNet model through training and validation metrics plotted against epochs. Figure 4 illustrates the training and validation accuracy trends over epochs, while Figure 5 depicts the training and validation loss patterns. These graphs provide valuable insights into the model's learning process and convergence behavior, aiding in the evaluation of its effectiveness in the classification task.

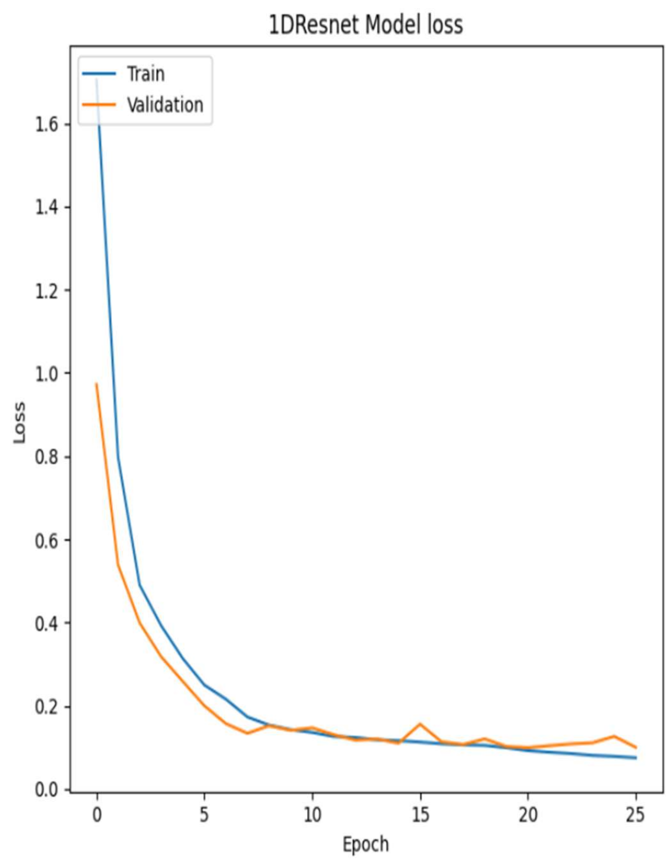

Fig. 5 Training and Validation Loss trends over epochs

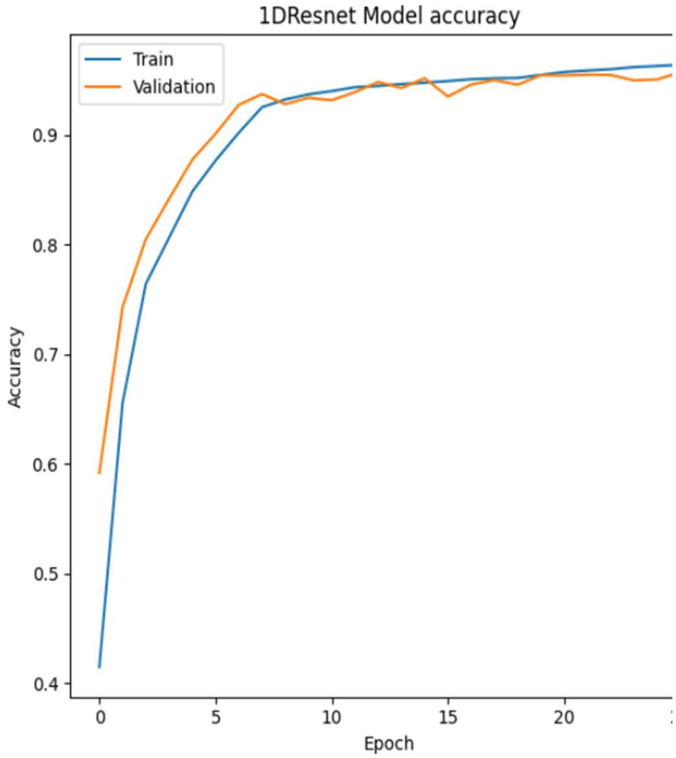

Fig. 4 Training and Validation accuracy trends over epochs

2. Model Evaluation on Test Data

The annotated bar graph below in Fig. 6 illustrates the evaluation metrics of the ResNet model on the test data, showing an accuracy of 0.96 and a loss of 0.08, rounded to two decimal places.

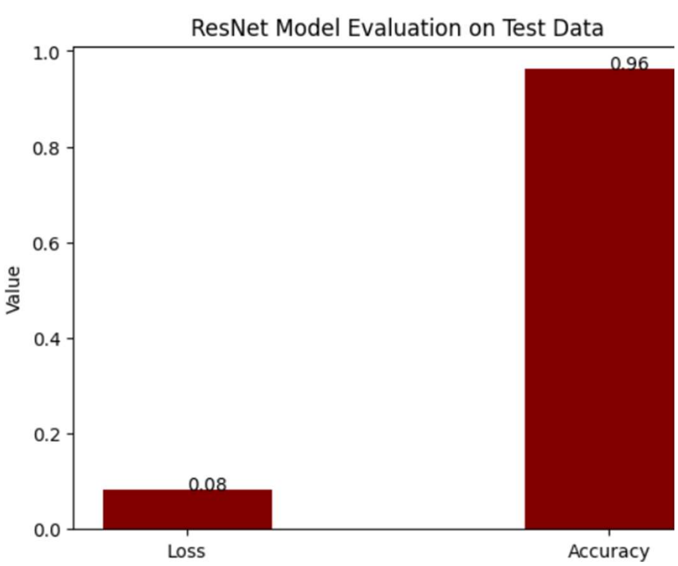

Fig. 6 ResNet Model Evaluation on Test Data with Accuracy and Loss Values

The evaluation metrics for the ResNet model on the test data demonstrate a high accuracy of 0.96 and a low loss of 0.08, indicating the model's strong performance in the classification task.

V. DISCUSSION

We analyze the results of our research on deep learning-based radio signal classification. We compare our findings with existing studies, emphasizing the performance, efficiency, scalability, generalization, and potential applications of our model.

Our study achieved a test set accuracy of 96.21% for SNRs exceeding 8dB, surpassing the results of a referenced study [14]. The efficiency of our model training, completed in approximately 1 hour and 29 minutes on a GPU for a subset of the dataset, highlights its scalability and adaptability. The robustness of our model in classifying signals under favorable SNR conditions indicates its potential for real-world applications. For a detailed comparison between our work and the referenced study, see below Table V.

Metric	Our Work	Referenced Paper [14]
Test Set Accuracy	96.21%	95.60%
Subset Size	14.10% of 2 million examples	Entire dataset (1.44 million examples)
SNR	Exceeding 8 dB	Approximately 10 dB
Training Time (GPU)	1 hour 29 minutes	Approximately 14 hours (NVIDIA V100 GPU)

Table V. Comparison of Modulation Classification Results: Our Work vs. Referenced Paper[14]

VI. CONCLUSION

Our research demonstrates the effectiveness of deep learning in radio signal classification. By achieving high accuracy, efficiency in training, and robustness in classification, our model shows promise for applications in spectrum monitoring and wireless communication systems. Future research can focus on refining the model architecture, optimizing training strategies, and exploring transfer learning techniques to enhance its performance in diverse signal environments.

Our study contributes to the advancement of signal processing methodologies and lays the foundation for the development of intelligent radio communication systems. Through continued exploration and refinement, we aim to further improve the capabilities and applicability of deep learning in radio signal classification tasks.

VII. ACKNOWLEDGMENT

I extend my heartfelt gratitude to Dr. Karan Verma, my dedicated supervisor, for his guidance and support throughout my M.Tech dissertation at NIT Delhi. I also thank Dr. Ajay K. Sharma, Director of NIT Delhi, and Dr. Geeta Sikka, Head of the CSE Department, for their leadership and encouragement. Special thanks to my family, friends, and colleagues for their unwavering support.

REFERENCES

[1] X. He, L. Lin, and J. Xie, “Radio signals modulation mode recognition based on semisupervised deep learning,” in Proc. IEEE 2nd Int. Conf. Automat., Electron. Elect. Eng. (AUTEEE), Nov. 2019, pp. 330–333.

[2] J. Li, L. Qi, and Y. Lin, “Research on modulation identification of digital signals based on deep learning,” in Proc. IEEE Int. Conf. Electron. Inf. Commun. Technol. (ICEICT), Aug. 2016, pp. 402–405.

[3] N. E. West and T. O’Shea, “Deep architectures for modulation recognition,” in Proc. IEEE Int. Symp. Dyn. Spectr. Access Netw. (DySPAN), Mar. 2017, pp. 1–6.

[4] J. Guo, H. Zhang, J. Xu, and Z. Chen, "Pattern recognition of wireless modulation signals based on deep learning," in Proc. IEEE 6th Int. Symp. Electromagn. Compat. (ISEMC), Nov. 2019, pp. 1–5.

[5] X. Xie et al., "Kind of Wireless Modulation Recognition Method Based on DenseNet and BLSTM," IEEE Access, vol. 8, pp. 107, 2020

[6] S. Zhang, Y. Zhang, and Y. Liu, "Modulation recognition based on deep learning and feature fusion," in Proc. IEEE Int. Conf. Commun. Technol. (ICCT), Oct. 2018, pp. 1–5.

[7] G. Shen, J. Zhang, A. Marshall, L. Peng, and X. Wang, "Radio frequency fingerprint identification for LoRa using deep learning," IEEE Internet Things J., vol. 7, no. 7, pp. 6255–6265, Jul. 2020.

[8] S. Rajendran, W. Meert, D. Giustiniano, V. Lenders, and S. Pollin, "Deep Learning Models for Wireless Signal Classification with Distributed Low-Cost Spectrum Sensors," in IEEE Transactions on Cognitive Communications and Networking, vol. 4, no. 2, pp. 386-399, June 2018.

[9] Wang, Y., Liu, M., Yang, J., & Gui, G. “Data-Driven Deep Learning for Automatic Modulation Recognition in Cognitive Radios” in IEEE Transactions on Vehicular Technology, 68(4), pp. 4074-4077, 2019.

[10] R. Li, L. Li, S. Yang, and S. Li, "Robust Automated VHF Modulation Recognition Based on Deep Convolutional Neural Networks," in IEEE Communications Letters, vol. 22, no. 5, pp. 1-4, May 2018.

[11] Y. Zhang, T. Liu, and K. Wang, "A Deep Learning approach for Modulation Recognition," in 2018 IEEE 23rd International Conference on Digital Signal Processing (DSP), Nov. 2018, pp. 1-5.

[12] R. Zhou, F. Liu, and C. Gravelle, "Deep Learning for Modulation Recognition: A Survey with a Demonstration," in IEEE Access, vol. 8, pp. 67366-67376, Apr. 2020.

[13] B. Tang, Y. Tu, Z. Zhang, and Y. Lin, "Digital Signal Modulation Classification with Data Augmentation Using Generative Adversarial Nets in Cognitive Radio Networks," in IEEE Access, vol. 6, pp. 15713-15722,Mach 2018.

[14] T. O'Shea, T. Roy, T. C. Clancy, et al., "Over-the-Air Deep Learning-Based Radio Signal Classification," in IEEE Journal of Selected Topics in Signal Processing, vol. 12, no. 1, pp. 168-179, Feb. 2018.

[15] K. He, X. Zhang, S. Ren, and J. Sun, “WaveNet: A generative model for raw audio,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2016, pp. 770–778.

[16] A. V. D. Oord et al., “WaveNet: A generative model for raw audio,” arXiv:1609.03499, 2016.

[17] N. E. West and T. J. O’Shea, “Deep architectures for modulation recognition,” in Proc. IEEE Int. Symp. Dyn. Spectr. Access Netw., 2017, pp.1–6.

[18] He, K., Zhang, X., Ren, S., & Sun, J. (2015). Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification. In Proc. IEEE Int. Conf. Comput. Vis., pp. 1026–1034.

[19] Szegedy, C., et al. (2015). Going deeper with convolutions. In Proc. IEEE Conf. Comput. Vis. Pattern Recognit., pp. 1–9.

[20] Oquab, M., Bottou, L., Laptev, I., & Sivic, J. (2014). Learning and transferring mid-level image representations using convolutional neural networks. In Proc. IEEE Conf. Comput. Vis. Pattern Recognit., pp. 1717–1724.

[21] Razavian, A. S., Azizpour, H., Sullivan, J., & Carlsson, S. (2014). CNN features off-the-shelf: An astounding baseline for recognition. In Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops, pp. 806–813.

[22] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444.

[23] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).

[24] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

[25] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.

[26] Hinton, G., Deng, L., Yu, D., Dahl, G. E., Mohamed, A. R., Jaitly, N., ... & Kingsbury, B. (2012). Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. IEEE Signal Processing Magazine, 29(6), 82-97.