Engineering and Technology Journal e-ISSN: 2456-3358

Volume 11 Special Issue 06 June-2026, Page No.- 14-21

DOI: 10.47191/etj/v11iSPC6.03, I.F. – 8.595

Benchmarking Deep Learning Architectures for Radio Modulation Classification Using the Radioml Dataset in Wireless Network

Dr. Neetu Sikarwar

Department of Electronice Engineering Institute of Engineeeirng , Jiwaji University Gwalior,India

ABSTRACT*:* Automatic Modulation Classification (AMC) is an essential element of the current wireless communication system, which allows the proper recognition of modulation schemes in the dynamic conditions on the channel and changing signal-to-noise ratios (SNRs). This study presents the results of DL architectural benchmarking using RadioML 2016. The architectures included in the research are CNN+BiLSTM, deep 1D ResNet, and LSTM.data set 10a. The approaches that are integrated into the proposed framework include the critical signal preprocessing steps, including the high-SNR filtering, power normalization, label encoding, and systematic split of training and testing data to get a better quality of data and better generalization of the model. The convolutional layers are used to achieve effective space features of raw I/Q samples and residual links within the ResNet framework to train the network deeper and prevent the vanishing gradient issue. Model performance is evaluated using accuracy (ACC), precision (PRE), recall (REC), F1-score (F1), and Matthews Correlation Coefficient (MCC). The experimental such as, the proposed ResNet model, which scores test ACC of 92.63%, with lower test loss (0.2903), along with higher PRE (0.9424), REC (0.9263), F1 (0.9225), and MCC (0.9213), compared with existing models, including HMF (90.70%), DenseNet (0.868), GRU (0.7891), and DRMM (73.31%) the robustness and usefulness of the ResNet model in the classification of test modulations in wireless networks.

KEYWORDS: Automatic modulation classification, RadioML 2016.10a dataset, Signal Preprocessing, Deep learning model, ResNet, Wireless Communication.

I. INTRODUCTION

The AMC plays an essential role in today's wireless communication infrastructure. It enables automatic determination of the modulation scheme of received signals, thereby improving spectral efficiency, reducing protocol overhead, and enhancing communication security[1]. AMC has important applications in drone communication, remote sensing, adaptive modulation systems, cognitive radio, and electromagnetic spectrum management. In adaptive communication systems, the modulation recognition error is minimized, enabling the receivers to pick the correct demodulation and decoding strategies to minimize the reliability and latency. Likewise, in cognitive radio networks[2], it is possible to determine which user signals are of primary importance so that it prevents interference and guarantee that the spectrum is used efficiently [3]. Real-world wireless environments can cause noise, fading, multipath propagation, and different SNR conditions to interfere with radio signals[4]. Conventional machine learning systems are very dependent on manual features and subject knowledge, which restricts their ability to scale and adapt to changes in the channel [2]. On the other hand, deep learning has been a huge hit in signal processing, parameter estimation, spectrum sensing, channel modelling, and pattern recognition generally [5], [6]. The end-to-end deep learning structures are able to automatically acquire discriminative representations directly on raw I/Q samples and do not require manual feature extraction, which can be better generalized to a wide range of wireless conditions [7].

The use of deep learning to AMC issues is on the rise, after its impressive performance in signal processing and pattern recognition applications like as parameter estimation, spectrum sensing, and channel modeling. End-to-end DL models learn input data as features automatically and without manually engineering features or defining them, compared to the traditional classifiers [8]. Deep learning models with nonlinear connections and large-scale parameters are more generalized because, as a result of big data, they are able to achieve performance stability with respect to AMC across varying channel conditions. The four most common structures used in deep learning models are CNNs, LSTMs, GCNs, and transformers. In contrast to LSTM's superiority in detecting temporal information, CNN is masterful at sensing spatial information in signals. The structure of data distribution is more susceptible to GCN [9], [10], [11]. The transformer version that incorporates an attention mechanism can adapt to cases which involve diverse parameter settings including different SNRs. However, DL models are extremely resource-intensive in terms of power consumption and computing resources. This paper has been motivated by the necessity to enhance the performance of automatic modulation classification in a situation with different and low SNR conditions with the RadioML 2016.10a dataset. Traditional methods are usually not capable of modeling the complicated spatial and time dependencies in I/Q signal data.

Thus, the work aims at exploiting the benefits of an advanced deep learning architecture suggested to bolster feature extraction, increase noise resistance, and better performance of more accurate and reliable modulation classification in a real-world wireless communication context[12]. The key findings from this study are as follows:

- Developed a robust DL framework for automatic modulation classification using the RadioML 2016.10a dataset under varying SNR conditions.
- Incorporated effective preprocessing techniques, including normalization and structured data handling, to improve model reliability and generalization.
- Implemented and comparatively analyzed CNN+BiLSTM, LSTM with Attention, and deep 1D ResNet architectures for enhanced spatial–temporal feature learning.
- Demonstrated through comprehensive evaluation (ACC, PRE, REC, F1, MCC) that the ResNet model outperforms baseline architectures in classification performance.

A. *Organization of the paper*

The structure of the paper is as follows: In Section II, classify radio modulation literature; in Section III, explore technique; in Section IV, present and analyze experimental results; and in Section V, summarize key findings and suggest directions for future research.

II. LITERATURE REVIEW

This section provides a thorough review of the existing status of automated modulation classification in wireless communication systems, with an emphasis on DL architectures for the purpose of reliably and precisely classifying radio signals. Table I supplies a summary of the relevant background research, and some reviews are based on data, results, and limitations.

M. M. Tahir et al. (2026) propose that an innovative AMC is crucial to contemporary wireless communication systems, especially in areas where interference resistance is paramount, such cognitive radio, the Internet of Things (IoT), and defense (as seen in the RadioML 2018 dataset). In every experiment, HFDNN proved to be more effective than the more conventional deep learning models. At -20 dB, it attains an ACC of 66.72% [13]. K. A. Mohamed et al. (2025) wireless communication systems' anticipated automated modulation categorization. A CNN model is trained to distinguish. Deep learning in intelligent receiver design offers a scalable, hardware-efficient solution for real-time modulation classification, and it was evaluated on the RFSC-classification dataset, achieving 91.64% ACC [14].

Gurucharan et al. (2025) proposed an image dataset of 24,460 images using eight classes of digital modulation techniques—2PSK, QPSK, 8PSK, 16PSK, 8QAM, 16QAM, 32QAM is used to test the model created by TM entire number of images into 85% for training and 15% for testing. According to the success rate statistics, 3,417 out of 3,669 images passed the generated model's ACC test with 94.19% [15]. B. E. Altuntas et al. (2024) address radio networks is AMR. A novel AMR method based on DL is being developed, which is being utilized in cognitive radio networks, which are part of current communication systems. This network design uses a number of separate convolutional blocks to learn the spatio-temporal signal correlations with the GELU activation function all at once. The suggested strategy achieves an overall 6-modulation classification rate of 80% at 20 dB SNR in the simulations run with the provided dataset [16].

M. Abdollahi et al. (2023) proposed digital modulation of non-line-of-sight signals using the commonly utilized QRSS, QPSK, and QAM techniques in contemporary telecommunications. In addition, for a variety of identification and classification tasks, they employ three well-known DL methods: CNN, DNN, and LSTM [17]. M. B. Mahieddine et al. (2022) work makes use of radio technology, a novel approach to radio design that seeks to improve cognitive radio's spectrum sensing by strengthening modulation recognition. The CNN-inception method was created for modulation categorization using Convolutional Neural Networks. A complex time-domain (I/Q) is the first input, and the fast Fourier transform of the signal is the second. The findings demonstrate that a maximum ACC of 93.22% may be achieved using the 10a dataset [18].

A. *Research Gap*

Despite the development of various DL models such as CNN, LSTM, GRU, DenseNet, and combined models for AMC tasks, current researches are still struggling with their inability to address low-SNR wireless scenarios, ensure robustness in spatial-temporal feature extraction processes, and remain computationally effective. Majority of earlier developed methods deal mostly with specific cases of modulation schemes and transformations of images rather than applying both deep residual learning and sequential feature modelling techniques. Moreover, a number of models demonstrated insufficient ACC during testing in noisier channels, and they lacked rigorous experiments on balanced real-life benchmark databases like RadioML 2016.10a. Consequently, a need exists for developing an efficient and robust deep learning model to increase modulation ACC and noise resistance while processing raw I/Q signal data.

TABLE I. SUMMARY OF THE RELATED WORK ON RADIO MODULATION CLASSIFICATION USING MACHINE LEARNING

Author (Year)	Dataset	Technique	Key Findings	Challenges	Future Work
M. M. Tahir et al. (2026)	RadioML 2018	Hybrid Deep Feedforward	Achieved 66.72% accuracy at −20 dB; outperformed	Low accuracy at extremely noisy SNR levels	Improve robustness under severe noise and

		Neural Network (HFDNN)	traditional DL models in low SNR		optimize model efficiency
K. A. Mohamed et al. (2025)	RFSC Classification Dataset	CNN with RTL/Verilog Hardware Implementation	Achieved 91.64% accuracy; enabled real-time hardware inference	Hardware complexity and resource constraints	Optimize hardware efficiency and extend to more modulation schemes
Gurucharan et al. (2025)	Custom Image Dataset (24,460 images)	Deep Learning (Image-based AMC Model)	Achieved 94.19% accuracy across 8 digital modulation schemes	Limited to image-transformed signals	Extend to raw I/Q datasets and real-time systems
B. E. Altuntas et al. (2024)	Generated Dataset	DL with Multi-Convolution Blocks + GELU	Achieved 80% accuracy at 20 dB SNR (6 modulations)	Performance drop at low SNR	Improve low-SNR robustness and scalability
M. Abdollahi et al. (2023)	Custom Dataset (BPSK, QPSK, QAM)	DNN, CNN, LSTM	Demonstrated DL effectiveness for NLOS digital signals	Limited modulation types	Expand to multi-class and real-world datasets
M. B. Mahieddine et al. (2022)	MIGOU-MOD & RadioML2016.10a	CNN-Inception Model	Achieved 93.22% accuracy; evaluated I/Q & FFT inputs	Increased computational complexity	Develop lightweight architectures for deployment

III. METHODOLOGY

The methodology begins with loading the DeepSig RadioML 2016.10a dataset and parsing the data to extract I/Q signal samples and corresponding labels. The whole process shows in Fig. 1. The data is prepared for model training during the signal preprocessing step by using high SNR filtering, power normalization, and label encoding. Separate training and testing sets are then created from the processed dataset. Next, three DL models CNN+BiLSTM, LSTM, and ResNet, are implemented and trained for automatic modulation classification. The ResNet model outperforms the others after training when measured using efficiency measures like F1, REC, ACC, and PRE.

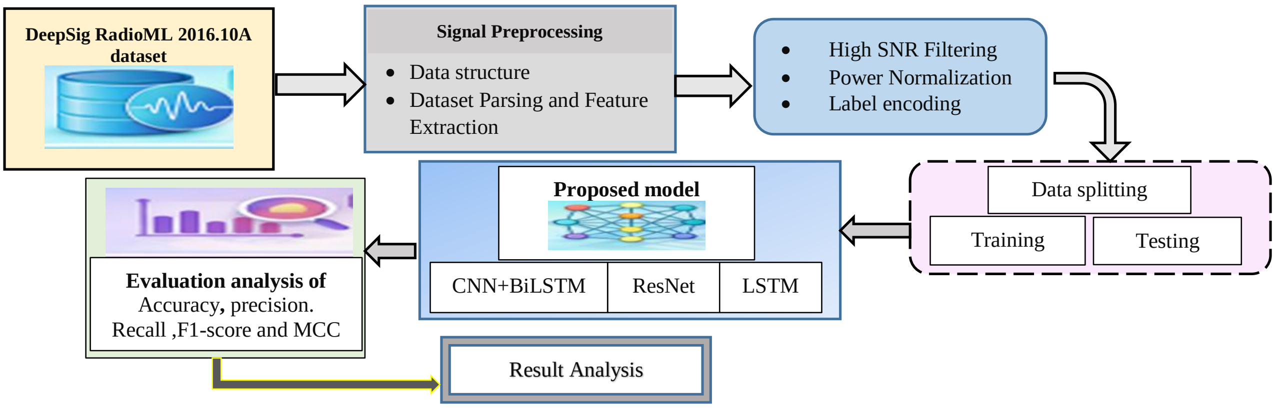


Fig. 1. Flowchart for Modulation classification using RadioML dataset.

Each step of proposed flowchart of methodology is briefly explained below:

A. *Data Collection*

RadioML 2016.10a (DeepSig) is a named radio signal collection from Kaggle that is used for automatic modulation classification. The 220,000 signal samples it includes are all of shape (2,128), and they represent the in-phase (I) and quadrature (Q) components, respectively. 11 different formats of modulation are included in the dataset. When loaded from the pickled file it is stored as a Python dictionary where each key represents a (modulation type, SNR) pair and information of modulation are describe in Table II.

TABLE II. INFORMATION ON MODULATION DATASET

Attribute	Description
Modulation Classes	BPSK, QPSK, 8PSK, CPFSK, GFSK, PAM4, QAM16, QAM64, AM-DSB, AM-SSB, WBFM
SNR Range	−20 dB to +18 dB
Number of SNR Levels	20 SNR values
Samples per SNR per Class	1000 signal samples
Total Samples	11 classes × 20 SNR levels × 1000 samples = 220,000 signal samples

The distribution of modulation types in the dataset, including all class as shown in Fig. 2. Each class contains approximately 20,000 samples, indicating a balanced dataset suitable for reliable training. The balanced distribution of datasets for all classes is confirmed by the modulation type vs SNR heatmap presented in Fig. 3, which shows that BPSK, AM-DSB, AM-SSB, CPFSK, GFSK, PAM4, QAM16, QAM64, QPSK, and WBFM all keep consistent sample counts of 1000 throughout SNR levels ranging from −20 dB to 18 dB.

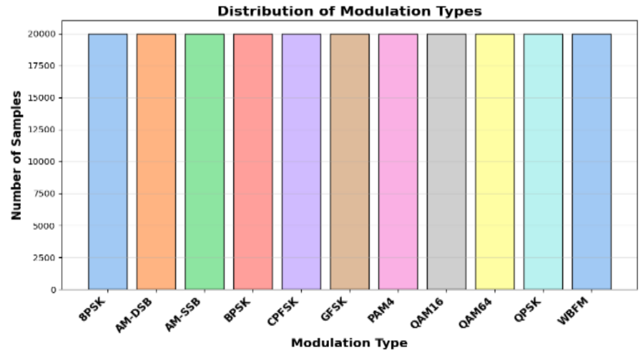


Fig. 2. Distribution of Modulation Types.

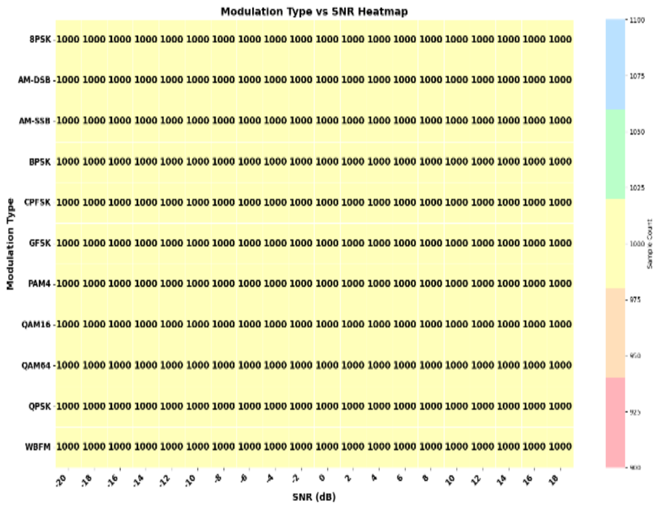


Fig. 3. Correlation Heatmap for Modulation type vs SNR.

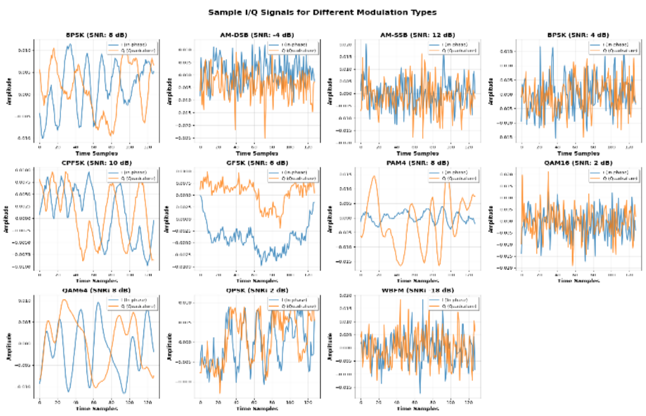


Fig. 4. Sample I/Q signal for different modulation types.

Fig. 4 displays the sampled in-phase (I) and quadrature (Q) signals for several modulation schemes, such as BPSK, AM-DSB, AM-SSB, BFSK, GFSK, QAM16, QAM64, QPSK, and WBFM. The signals were measured at varying SNR levels ranging from 2 dB to 12 dB, and noise was added to the mix.

B. Data Preprocessing

Data preprocessing refines the raw I/Q signals by improving data structure, signal clarity, and label organization before model training. It ensures consistent and reliable input for accurate classification. Further pre-processing steps are given in next steps.

- **Data structure:** Complex baseband I/Q samples make up the dataset. The form of each signal instance is (2, 128), which represents the in-phase and quadrature components, respectively.
- **Dataset Parsing and Feature Extraction:** The dataset is read from its dictionary structure and reorganized into structured arrays containing I/Q signal samples, modulation labels, and SNR values.

C. High SNR Filtering

High SNR filtering is used to keep signal samples that have high SNR values to minimize the effect of severe noise during training. Once the data is transposed to an input shape of (220000, 128, 2), a High SNR filtering method used to enhance the quality of data. A SNR threshold of 10 dB or more is chosen to filter out samples with a very high noise level. Out of the 220,000 signal instances, 55,000 of them meet this requirement, which makes the retention rate 25% of the total. The dataset that has been retained has an SNR of 10 dB to 18 dB. This is done by concentrating on larger SNR signals, making the I/Q patterns easier and less distorted.

D. Power Normalization

The normalization of power is the process which scales every RF signal to unit average power by dividing the signal by its RMS value. This eliminates amplitude variations due to channel effects, differences in transmission power or hardware variations. Mathematically, every signal x is normalized as shown in Equation (1):

$$x_{norm} = \frac{x}{\sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2 + \epsilon}} \tag{1}$$

N is the number of samples and ϵ is a small constant that prevents division by zero. Low amplitude magnitude is shown by the dataset's −0.000381 mean and 0.005838 standard deviation prior to normalization. After applying RMS-based normalization, the mean becomes −0.068759 and the standard deviation increases to 0.997632, which approximates unit variance.

E. Label encoding

Label encoding converts categorical modulation labels into numerical form using, and then transforms them into one-hot encoded vectors for multi-class classification. Since modulation types are categorical, direct integer labels may introduce artificial ordinal relationships. One-hot encoding represents each class as a binary vector of length 11, such as BPSK-[0,0,0,1,0,…,0], ensuring independent class representation. This enables the SoftMax layer to generate class-wise probability distributions and supports correct loss computation using categorical cross-entropy.

F. Data Splitting

A stratified technique is used to maintain an equal distribution of classes by splitting the dataset into 80% training and 20% testing sets. This ensures balanced representation of all modulation types and enables reliable evaluation on unseen data.

G. Proposes Models

It provides the theoretical explanation of the DL A model are CNN+BiLSTM, ResNet and LSTM created to capture both temporal and spatial relationships in radio signals. The architecture consists of the following modules in this study in this section.

1) CNN+BiLSTM

The CNN+BiLSTM model begins by taking the input signal and passing it through three Conv1D layers with 64, 128, and 256 filters to extract important spatial features. The CNN output are then reshaped and passed through a BiLSTM layer with 128 units [19], producing output shapes of (None, 126, 64), (None, 61, 128), and (None, 28, 256), respectively, each followed by BatchNormalization and MaxPooling for feature extraction and dimensionality reduction. The extracted features are then passed to two Bidirectional LSTM layers, where the first BiLSTM outputs (14, 256), capturing forward and backward temporal dependencies. Finally, Dense layers (256 units) and a SoftMax output layer (11 units) The model contains a total of 719,947 parameters, enabling effective spatial and temporal feature learning for modulation classification.

2) ResNet

A ResNet (Residual Network) model is evaluated for signal the proposed ResNet architecture for signal classification. The model uses Conv1D layers with BatchNormalization and MaxPooling to extract features from shapes that are input with a shape of (128, 2). Global Average Pooling and fully connected dense layers with dropout are followed by stacked residual blocks with increasing filters (64 → 128 → 256), which enable deep feature learning. The final Softmax layer classifies the 11 modulation types as Equation (2):

$$H(x) = F(x) + x \tag{2}$$

The residual learning principle which helps reduce the vanishing gradient problem and improves deep network training attention, Q (Query) is compared with K (Key) using QK^T to measure similarity, scaled by, $\sqrt{d_k}$ and passed through SoftMax to generate importance weights. The attention mechanism is expressed as Equation (3):

$$Attention(Q, K, V) = SoftMax\left(\frac{QK^T}{\sqrt{d_k}}\right) \tag{3}$$

Categorical Cross-Entropy loss is used to train the model, as shown in Equation (4):

$$L = -\sum y_i \log(\hat{y}_i) \tag{4}$$

with the AdamOptimizer (LearningRate 0.001), BatchSize 128, and Early Stopping. This model achieved efficient and reliable modulation classification performance

3) LSTM

The LSTM-based classification model starts with an input signal of shape (128, 2), which first passes through two Conv1D layers (64 and 128 filters) followed by BatchNormalization and Max-Pooling to extract Spatial Features. After that, the features are passed into two stacked Bidirectional LSTM layers, which provide rich sequential representations by capturing the signal's forward and backward temporal relationships. The model contains 521,164 total parameters, enabling effective spatial and temporal feature learning.

H. Model Evaluation

The performance of the presented models has been assessed using several assessment indicators. The fundamental indicator is ACC, which may be measured as the proportion of properly recognized signals across all modulation categories. Below, it presents the metrics used in dueling analysis, which include the F1, the Confusion Matrix, REC, f1, and MCC, as well as changes in SNR circumstances (Equations 5 to 9):

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \tag{5}$$

$$Precision = \frac{TP}{TP+FP} \tag{6}$$

$$Recall = \frac{TP}{TP+FN} \tag{7}$$

$$F1 - Score = 2 * \frac{(Precision*Recall)}{Precision+Recall} \tag{8}$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \tag{9}$$

ROC curve analysis allows calculation of the area under the curve to determine a model's performance with the given threshold values.

IV. RESULT ANALYSIS AND DISCUSSION

Experiment result of proposed models In the current study, the DeepSig RadioML 2016.10A is used where 11 types of modulation are tested on SNR value between - 20 dB and 18 dB. It is implemented on the Windows 10 system with the Intel i5 9th generation processor, 8 GB of RAM, and NVIDIA GTX 1650 graphic cards and developed in Python in the Jupyter Notebook and Visual Studio code using NumPy, Pandas, TensorFlow/Keras, Scikit-learn, and Matplotlib packages. It has been experimentally demonstrated as shown in Table III that the ResNet-like CNN has a test ACC of 92.63%, test loss of 0.2903, PRE of 0.9424 and REC of 0.9263 as well as F1 and MCC of 0.9213, which is better than CNN+BiLSTM and Attention-based LSTM with test ACC of 90.15% and 90.18% respectively and hence more capable to the field of reliable automatic modulation.

TABLE III. EXPERIMENT RESULTS OF PROPOSED MODELS FOR WIRELESS CLASSIFICATION NETWORK

Evaluation Metrics	CNN+Bi-LSTM	ResNet	LSTM
Test ACC	90.15%	92.63%	90.18%
Test Loss	0.3784	0.2903	0.3136
Precision	0.9190	0.9424	0.9185
Recall	0.9015	0.9263	0.9018
F1-Score	0.8968	0.9225	0.8977
MCC	0.8942	0.9213	0.8944

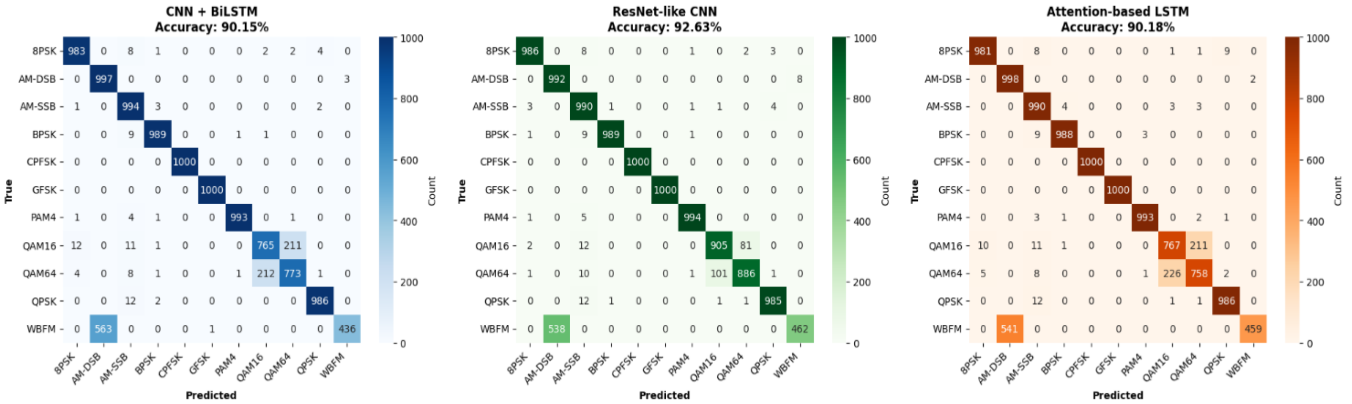


Fig. 5. Confusion matrix of Proposed Model.

Fig. 5 shows the confusion matrices of CNN+BiLSTM, ResNet-like CNN and Attention-based LSTM models in modulation classification. The total accuracies are 90.15%, 92.63% and 90.18% respectively. ResNet-like CNN performs the correct classification on 990 AM-SSB and 1000 CPFSK samples. The strong performance in the sense of class-wise prediction in BPSK, QPSK, QAM, and other related signal types is reflected in the diagonal dominance.

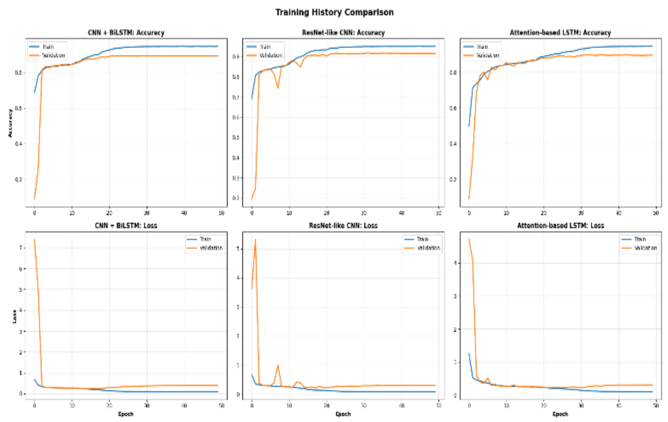
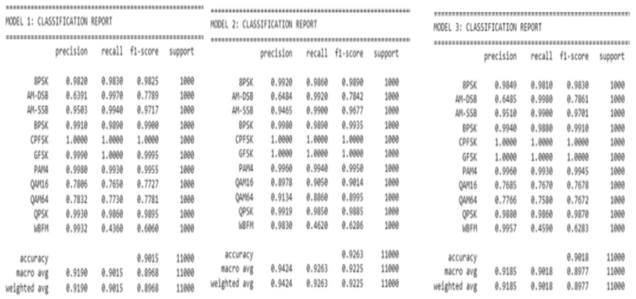


Fig. 6. Training of Accuracy and Loss graph in CNN+Bi-LSTM, ResNet and LSTM model for wireless network.

ACC and Loss curve of proposed models as shown in Fig. 6. Training ACC converges above 0.90 within 10 epochs, reaching nearly 0.93. Validation loss decreases rapidly below 0.2, while training loss stabilizes around 0.1, indicating consistent convergence and minimal overfitting behavior.



MODEL 1: CLASSIFICATION REPORT

	precision	recall	f1-score	support
8PSK	0.9820	0.9830	0.9825	1000
AM-DSB	0.6391	0.9970	0.7789	1000
AM-SSB	0.9503	0.9940	0.9717	1000
BPSK	0.9910	0.9890	0.9900	1000
CPFSK	1.0000	1.0000	1.0000	1000
GFSK	0.9990	1.0000	0.9995	1000
PAM4	0.9980	0.9930	0.9955	1000
QAM16	0.7806	0.7650	0.7727	1000
QAM64	0.7832	0.7730	0.7781	1000
QPSK	0.9930	0.9860	0.9895	1000
WBFM	0.9932	0.4360	0.6060	1000
accuracy			0.9015	11000
macro avg	0.9190	0.9015	0.8968	11000
weighted avg	0.9190	0.9015	0.8968	11000

MODEL 2: CLASSIFICATION REPORT

	precision	recall	f1-score	support
8PSK	0.9920	0.9860	0.9890	1000
AM-DSB	0.6484	0.9920	0.7842	1000
AM-SSB	0.9465	0.9900	0.9677	1000
BPSK	0.9980	0.9890	0.9935	1000
CPFSK	1.0000	1.0000	1.0000	1000
GFSK	1.0000	1.0000	1.0000	1000
PAM4	0.9960	0.9940	0.9950	1000
QAM16	0.8978	0.9050	0.9014	1000
QAM64	0.9134	0.8860	0.8995	1000
QPSK	0.9919	0.9850	0.9885	1000
WBFM	0.9830	0.4620	0.6286	1000
accuracy			0.9263	11000
macro avg	0.9424	0.9263	0.9225	11000
weighted avg	0.9424	0.9263	0.9225	11000

MODEL 3: CLASSIFICATION REPORT

	precision	recall	f1-score	support
8PSK	0.9849	0.9810	0.9830	1000
AM-DSB	0.6485	0.9980	0.7861	1000
AM-SSB	0.9510	0.9900	0.9701	1000
BPSK	0.9940	0.9880	0.9910	1000
CPFSK	1.0000	1.0000	1.0000	1000
GFSK	1.0000	1.0000	1.0000	1000
PAM4	0.9960	0.9930	0.9945	1000
QAM16	0.7685	0.7670	0.7678	1000
QAM64	0.7766	0.7580	0.7672	1000
QPSK	0.9880	0.9860	0.9870	1000
WBFM	0.9957	0.4590	0.6283	1000
accuracy			0.9018	11000
macro avg	0.9185	0.9018	0.8977	11000
weighted avg	0.9185	0.9018	0.8977	11000

Fig. 7. Classification report of CNN+Bi-LSTM, ResNet and LSTM model for wireless network.

Fig. 7 classification reports of proposed model for modulation recognition. ResNet-like CNN achieves the highest ACC of 0.9263 with weighted F1 of 0.9225, while CNN+BiLSTM and LSTM obtain 0.9095 and 0.9018, respectively. For BPSK, PRE reaches 0.9940 in LSTM, and CPFSK attains REC of 1.0000 in ResNet, indicating strong class-level performance.

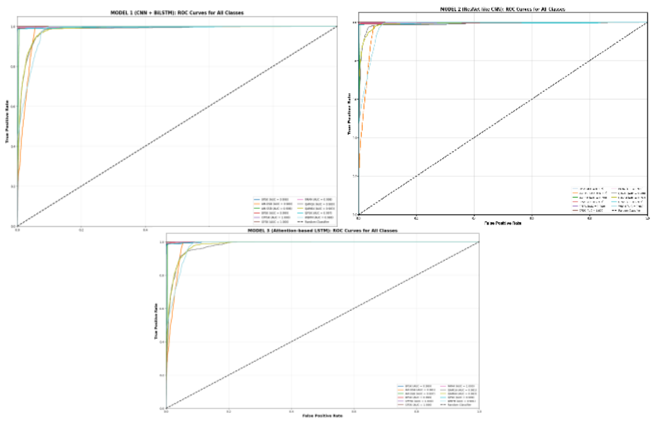


Fig. 8. ROC Curve for CNN+Bi-LSTM, ResNet and LSTM model for wireless Network Classification.

ROC curves of all modulation classes proposed models are shown in Fig. 8. The use of ResNet-like CNN depicts better results as the AUC values tend to approach 0.991.00. Both CNN+BiLSTM and Attention-based LSTM attain a higher AUC of more than 0.97 and the curves are clustered of strong classification.

A. *Model complexity comparison of proposed model*

The ResNet model exhibits the highest complexity, CNN+BiLSTM shows moderate complexity, and the Attention-based LSTM provides a lightweight yet efficient architecture for modulation classification, as shown in Table IV.

TABLE IV. MODEL COMPLEXITY COMPARISON OF PROPOSED DEEP LEARNING ARCHITECTURES

Model	Total Parameters	Trainable Parameters	Parameters (Millions)
CNN + BiLSTM	719,947	719,051	0.719947
ResNet	1,226,571	1,222,859	1.226571
Attention-based LSTM	521,164	520,780	0.521164

B. *Comparative Analysis*

A comprehensive performance comparison of different DL architectures for AMC is presented in Table V. The highest ACC of 92.63 is attained in the proposed ResNet-based model that has the best residual learning and feature extraction potential. The Hierarchical Multi-Feature Fusion (HMF) model achieves 90.70% ACC which is an indication of successful hierarchical integration of features. GRU model records the highest levels of 86.8% ACC whereas the DenseNet records 78.91% which is moderate levels of classification. Conversely, the DeepResidual Neural Network with Masked Modeling (DRMM) achieves the ACC of 73.31% that proves its better residual learning process and strong classification functionality in modulation classes.

TABLE V. MODEL PERFORMANCE COMPARISON OF DIFFERENT EVALUATION METRICS ON RADIOML DATASET

Model	Accuracy
HMF[20]	90.70
DenseNet[21]	0.7891
GRU[22]	0.868
DRMM[23]	73.31
ResNet	92.63%

The proposed framework improves automatic modulation the proposed architectures of the entire process of automatic modulation classification using the RadioML 2016.10a dataset under different SNR conditions. The framework improves the temporal and spatial feature extraction and utilizes the residual learning to train further. The approach is justified by experimental results in which ResNet gets the most efficient and reliable performance in terms of classification.

C. *Justification and Novelty*

This paper's originality comes in the implementation and comparative analysis of various architectures for automated modulation categorization using RadioML 2016.10a dataset with various SNR values. The framework enhances spatial and temporal feature extraction while leveraging residual learning for deeper training. Experimental results justify the approach, with ResNet achieving the most efficient and reliable classification performance.

V. CONCLUSION AND FUTURE SCOPE

The efficiency of sophisticated DL models for automated modulation categorization in wireless communication systems was examined in this work, utilizing the RadioML 2016.10a dataset. A systematic framework created to incorporate signal preprocessing actions like normalization, label encoding, and suitable dataset partitioning to improve steady learning and generalization. ResNet designs were analyzed to have the capability to extract meaningful spatial and temporal information out of unprocessed I/Q signals. The proposed ResNet model demonstrated significantly better performance with an experiment result of 92.63% test ACC. The spatial connections allowed further training of the network and better learning of features without vanishing gradient problems that caused increased robustness in the presence of a noisy channel. The suggested framework can be elaborated on in future as the work, the proposed structure may be further expanded to real-time hardware execution, be optimized to be more lightweight, and include more advanced attention or transformer-based mechanisms to enhance the capability of the proposed structure to adapt to various and large-scale communication datasets.

REFERENCES

1. S. K. Davuluri, V. Challagulla, V. Mudapaka, and U. Konka, "Telcoformrix: An AI-Augmented Framework for Declarative and Scalable Provisioning of Real-Time Communication Infrastructure (Work in Progress)," in *2025 IEEE International Conference and Expo on Real Time Communications at IIT (RTC)*, Chicago, IL, USA: IEEE, 2025, pp. 1–4, October. doi: 10.1109/RTC66985.2025.11211725.
2. S. Singh, *Enabling Sustainable 5G Networks Through Energy-Efficient Open Radio Access Networks (ORAN) Ecosystems*. 2026. doi: 10.1007/978-981-96-8793-0_11.
3. Q. Zheng *et al.*, "MobileRaT: A Lightweight Radio Transformer Method for Automatic Modulation Classification in Drone Communication Systems," *Drones*, vol. 7, no. 10, 2023, doi: 10.3390/drones7100596.
4. V. K. Sharma, "STRATEGIC FRAMEWORK FOR AI-ENHANCED PORTFOLIOS IN WIRELESS ENGINEERING: A LITERATURE REVIEW," *Int. J. Core Eng. Manag.*, vol. 8, no. 03, pp. 77–83, 2025.

5. Vaidehi Shah, "Next-Gen Emergency Communication Using Low-Power Wide-Area and Software-Defined WANS," *Int. J. Adv. Res. Sci. Commun. Technol.*, vol. 2, no. 1, pp. 600–609, Sep. 2022, doi: 10.48175/IJARSCT-8349M.
6. H. P. Cyril, "EVENT-DRIVEN PROVISIONING ARCHITECTURES FOR MODERN TELECOM NETWORKS: OVERCOMING LEGACY LIMITATIONS AND ENABLING AUTONOMOUS 6G OPERATIONS," *Int. J. Adv. Res. Comput. Sci.*, vol. 16, no. 6, pp. 75–82, 2025.
7. S. Chen, S. Zheng, L. Yang, and X. Yang, "Deep Learning for Large-Scale Real-World ACARS and ADS-B Radio Signal Classification," *IEEE Access*, vol. 7, pp. 89256–89264, 2019, doi: 10.1109/ACCESS.2019.2925569.
8. L. Huang, Y. Zhang, W. Pan, J. Chen, L. P. Qian, and Y. Wu, "Visualizing Deep Learning-Based Radio Modulation Classifier," *IEEE Trans. Cogn. Commun. Netw.*, vol. 7, no. 1, pp. 47–58, 2021, doi: 10.1109/TCCN.2020.3048113.
9. S. Singh, "Advancing Wireless Communications with Open Radio Access Network," in *2025 6th International Conference on Data Intelligence and Cognitive Informatics (ICDICI)*, Tirunelveli, India: IEEE, Jul. 2025, pp. 682–687. doi: 10.1109/ICDICI66477.2025.11135206.
10. S. K. Chintagunta and S. Amrale, "A Deep Learning Framework for Adaptive E- Learning : Integrating Learning Style Detection in Web-Based Platforms," *Int. J. Adv. Res. Sci. Commun. Technol.*, pp. 716–727, 2024, doi: 10.48175/IJARSCT-19397.
11. L. Huang, W. Pan, Y. Zhang, L. Qian, N. Gao, and Y. Wu, "Data Augmentation for Deep Learning-Based Radio Modulation Classification," *IEEE Access*, vol. 8, pp. 1498–1506, 2020, doi: 10.1109/ACCESS.2019.2960775.
12. S. Singh, S. A. Pahune, P. Chatterjee, and R. Sura, "Advanced Machine Learning Methods for Churn Prediction and Classification in Telecom Sector," in *2025 IEEE 6th India Council International Subsections Conference (INDISCON)*, Rourkela, India: IEEE, 2025, pp. 1–7, December. doi: 10.1109/INDISCON66021.2025.11252233.
13. M. M. Tahir *et al.*, "HFDNN : A Hybrid Fusion Deep Neural Network for Robust Automatic Modulation Classification in Adverse Wireless Environments," *IEEE Access*, vol. 14, no. February, 2026.
14. K. A. Mohamed *et al.*, "FPGA-Optimized CNN Architecture for Automatic Modulation Classification," in *2025 37th International Conference on Microelectronics (ICM)*, 2025, pp. 1–6. doi: 10.1109/ICM66518.2025.11321328.
15. J. Gurucharan, M. Padmapriya, M. R. Grace, S. B. Krishna, P. Ganesan, and V. Elamaran, "Deep Learning on Modulation Classification using Google's Teachable Machine," in *2025 1st International Conference on Radio Frequency Communication and Networks (RFCoN)*, 2025, pp. 1–6. doi: 10.1109/RFCoN62306.2025.11085309.
16. [16] B. E. Altuntas, O. Aksu, M. Y. Celik, and M. H. Durak, "Deep Learning Based Automatic Modulation Recognition Using GELU Activation Function," in *2024 4th International Conference on Emerging Smart Technologies and Applications (eSmarTA)*, IEEE, Aug. 2024, pp. 1–4. doi: 10.1109/eSmarTA62850.2024.10638998.
17. M. Abdollahi, R. Sabzalizadeh, S. Javadinia, S. Mashhadi, S. S. Mehrizi, and A. Baniasadi, "Automatic Modulation Classification for NLOS 5G Signals with Deep Learning Approaches," in *2023 10th International Conference on Wireless Networks and Mobile Communications (WINCOM)*, 2023, pp. 1–6. doi: 10.1109/WINCOM59760.2023.10322928.
18. M. B. Mahieddine, N. Mellah, A. Bassou, and M. Khelifi, "Implementation of CNN-Inception Deep Learning for Cognitive Radio Based on Modulation Classifications," in *2022 2nd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)*, IEEE, Mar. 2022, pp. 1–5. doi: 10.1109/IRASET52964.2022.9738405.
19. T. Bayan, D. Das, and N. Deb, "Deep Hybrid CNN–BiLSTM–Attention Model for EEG Classification Using Wavelet Features," *Ann. Neurosci.*, Jan. 2026, doi: 10.1177/09727531251396337.
20. N. Wang, Y. Liu, L. Ma, Y. Yang, and H. Wang, "Multidimensional CNN-LSTM Network for Automatic Modulation Classification," *Electronics*, vol. 10, no. 14, 2021, doi: 10.3390/electronics10141649.
21. S. Hou, Y. Fan, B. Han, Y. Li, and S. Fang, "Signal Modulation Recognition Algorithm Based on Improved Spatiotemporal Multi-Channel Network," *Electronics*, vol. 12, p. 422, 2023, doi: 10.3390/electronics12020422.
22. S. Ansari *et al.*, "Attention-Enhanced Hybrid Automatic Modulation Classification for Advanced Wireless Communication Systems: A Deep Learning-Transformer Framework," *IEEE Access*, vol. 13, no. June, pp. 105463–105491, 2025, doi: 10.1109/ACCESS.2025.3580574.
23. Y. Peng *et al.*, "Automatic Modulation Classification Using Deep Residual Neural Network with Masked Modeling for Wireless Communications," *Drones*, vol. 7, no. 6, 2023, doi: 10.3390/drones7060390.