Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 4, Issue 3, 2026

CNN-Based Automatic Modulation Classification of Radio Frequency Signals Under Varying SNR Conditions

[1]Eisha Tir Razia
[1]*University of Management and Technology, Lahore, Pakistan.*
[1]*eisha_tir_razia@hotmail.com*

DOI:

Keywords:
Automatic modulation classification, cognitive radio, residual CNN, random forest, support vector machine, high-SNR analysis, deep learning

Article History
Received on: 23 Feb, 2026
Accepted on: 26 March, 2026
Published on 29 March, 2026

Copyright @Author
Corresponding Author:

Abstract
This paper presents modulation classification using the RadioML 2016.10A dataset under varying SNR conditions, with a primary focus on high-SNR scenarios (≥10 dB). Automatic modulation classification (AMC) plays a vital role in modern cognitive radio and wireless communication systems by enabling dynamic spectrum access, interference mitigation, and intelligent resource allocation. A residual Convolutional Neural Network (CNN) is designed and implemented using regularization techniques and a high-SNR-focused training strategy to improve classification performance. To provide a comparative benchmark, two classical machine learning models, Random Forest (RF) and Support Vector Machine (SVM), are trained using statistical features derived from the I/Q signal representations. Experimental results show that the proposed CNN consistently achieves higher classification accuracy in high-SNR scenarios with an overall accuracy of 88%, outperforming the RF and SVM baselines, which achieve 72.29% and 70.31%, respectively. Moreover, performance analysis across different SNR levels illustrates the resilience of deep learning approaches in noisy environments compared to traditional methods. t-SNE and Grad-CAM tools visualize how the model groups data and highlight key signal features. The results validate the potential of deep residual architectures for AMC in practical communication systems and provide a reproducible framework for future research in high-SNR signal classification.

1 Introduction

The explosive growth of wireless communication technologies has placed unprecedented pressure on the radio frequency spectrum. Efficient spectrum utilization and dynamic spectrum access have therefore become critical requirements in next-generation communication systems. Automatic Modulation Classification (AMC) has emerged as one of the most important technologies in this domain. By automatically identifying the modulation scheme of a received signal without prior knowledge, AMC facilitates a wide range of applications, including cognitive radio networks, electronic warfare, spectrum monitoring, and adaptive communication systems [2, 21]. In recent years, deep learning has revolutionized AMC research by offering data-driven solutions that learn hierarchical representations directly from raw I/Q samples. Convolutional Neural Networks (CNN) have demonstrated remarkable success in extracting discriminative features from time-series and spatial data, eliminating the need for manual feature engineering. By leveraging their ability to capture local temporal dependencies in I/Q waveforms, CNN-based AMC models have consistently outperformed traditional approaches across a variety of datasets and channel conditions [4, 11, 16].

Contemporary studies have emphasized overall performance across all SNR values, often aiming to maximize accuracy in challenging low-SNR conditions. While this perspective is valuable for demonstrating model robustness, it overlooks the equally important case of high-SNR conditions, where communication links are generally stable and reliable [1, 8]. High-SNRs are critical for applications such as commercial wireless networks, mission-critical communication, and radar systems, where performance expectations are stringent and classification errors can lead to significant operational inefficiencies. Motivated by this gap, the present study adopts a targeted perspective by prioritizing performance evaluation in high-SNR scenarios [5, 7].

This study provides insights that are directly applicable to reliable communication systems, where ensuring near-perfect classification accuracy is often more important than marginal improvements at extremely low SNR levels. To this end, a residual CNN architecture trained with focus on high-SNR and regularization techniques, optimized using a cosine learning rate restart schedule to enhance stability and convergence, has been proposed. Moreover, its performance against two classical ML baselines, i.e., Random Forest and SVM, trained on statistical features extracted from I/Q samples is benchmarked. This comparative study not only highlights the advantages of deep learning in high-SNR conditions but also establishes a fair and rigorous baseline, bridging the gap between traditional AMC techniques and modern data-driven approaches. This study highlights the importance of considering SNR-specific performance in AMC research and demonstrates the potential of residual CNN architectures to serve as powerful classifiers in practical wireless communication systems.

2 Related Work

Automatic Modulation Classification (AMC) has emerged as a critical task in modern wireless communications, where accurate identification of modulation schemes enables robust spectrum monitoring, cognitive radio, and adaptive communication systems [12]. In recent years, deep learning has revolutionized AMC research, replacing traditional feature engineering with models that directly learn representations from raw in-phase and quadrature (I/Q) samples [6]. A wide body of literature has explored different neural architectures, training strategies, and evaluation protocols, yet challenges remain in ensuring robust performance across varying signal-to-noise ratios (SNRs), designing lightweight models for deployment, and providing explainability of model decisions.

One of the most influential works in this field has been proposed by O'Shea et al. (2016), who introduced convolutional neural networks (CNNs) for AMC trained directly on raw I/Q samples using the RadioML dataset [15]. This study established a strong baseline by demonstrating that end-to-end learning outperformed classical approaches relying on handcrafted features. The study provided analysis of model robustness at very low SNRs or insights into model interpretability. Building on this foundation, Liu, Yang, and El Gamal (2017) systematically compared deep architectures for AMC, including CNNs, recurrent neural networks (RNNs), and hybrid models. Their findings highlighted the importance of network depth, kernel size, and recurrent layers, showing that different architectures offer complementary strengths in recognizing complex modulation patterns [13]. However, this study, like earlier works, often restricted evaluation to limited SNR ranges and did not always benchmark against strong classical baselines. Subsequent research has also explored the efficiency of deep learning models for AMC.

Ramjee et al. (2019) investigated the problem of deploying deep models in real-time environments, where inference speed and model size are critical constraints. By applying model pruning and architectural optimizations, they demonstrated fast AMC solutions that maintained competitive accuracy. Nevertheless, speed-oriented approaches often compromise performance in noisy, low-SNR scenarios, leaving a gap between efficiency and robustness [18]. More recently, Mohsen et al. (2024) revisited CNN-based AMC with modern training recipes, including data augmentation and advanced regularization, to further improve classification accuracy. Their work emphasized the need for reproducible pipelines and highlighted ongoing inconsistencies across the literature, such as varying evaluation protocols and differences in how training/test splits are defined [14].

In addition to CNN-based methods, contemporary studies have also incorporated residual connections, deeper one-dimensional convolutional architectures, and SNR-aware training strategies. Learning approaches, which begin with high-SNR data before progressively including lower-SNR signals, have shown promise in improving model convergence and robustness. At the same time, explainability techniques such as t-SNE visualizations of learned feature spaces and adaptations of Grad-CAM for one-dimensional data have provided new insights into model decision-making processes [22]. While many recent works focus exclusively on deep learning, classical approaches such as Random Forests and Support Vector Machines (SVMs) remain competitive, especially at certain SNR levels, and continue to serve as important baselines for fair comparison. Against this backdrop, the present work builds on and extends prior studies by combining a residual one-dimensional CNN architecture with an SNR-aware training strategy focused on high-SNR scenarios. A comparison of different studies with the present study is demonstrated in Table 1.

Table 1. *Comparison with some other studies*

Study	Dataset	Method Used	Limitations	Improvements in this Study
O'Shea et al. (2016) [15]	RadioML-2016.10a	CNN (2 conv + 2 dense)	Shallow model, low-SNR accuracy, limited interpretability	ResNet1D with residual blocks, high-SNR focused learning, ~ 88% accuracy at SNR≥10

Liu et al. (2017) [13]	RadioML-2016.10a	CNN, CLDNN, DenseNet (~ 86.6% at high SNR)	Complex models, no hybrid baselines or visualizations	Competitive accuracy with lighter ResNet1D, plus RF/SVM baselines
Ramjee et al. (2019) [18]	RadioML-2016.10b	CNN, ResNet, LSTM variants	Focus on speed and reduced sampling, limited interpretability	Maintains high-SNR accuracy with ResNet1D and adds explainability
Mohsen et al. (2024) [14]	RadioML-2016.10a	Tuned CNNs (~ 94.4% in some ranges)	High accuracy but no baseline comparisons	Achieved ~ 88% at high SNR⩾10, baseline compar-isons
Some Other CNN Works [5] [7] [8]	RadioML variants	Standard CNNs, spectro-gram-based	>90% accuracy at high SNR but lack baselines and visualizations	Holistic evaluation: CNN + RF + SVM, baselines and visualizations

3 Dataset and Preprocessing

The RadioML 2016.10A dataset has been used in this research study, which contains 220,000 signal samples spanning 11 modulation types (such as QPSK, QAM16, AM-DSB, and WBFM) and SNR values ranging from −20 to +18 dB as shown in Fig. 1. Some example waveforms can be seen in Fig. 2. Each sample consists of 128 complex I/Q time steps, which are reshaped into a (128, 2) format to separate the in-phase (I) and quadrature (Q) channels. To focus on reliable communication scenarios, only samples with SNR ⩾10 dB are retained for CNN training, and labels are one-hot encoded before performing a stratified train-test split to preserve class balance. For the classical machine learning baselines (Random Forest and SVM), additional handcrafted statistical features, including mean, variance, skewness, kurtosis, signal power, and peak amplitudes, are extracted from the I/Q samples to facilitate model training.

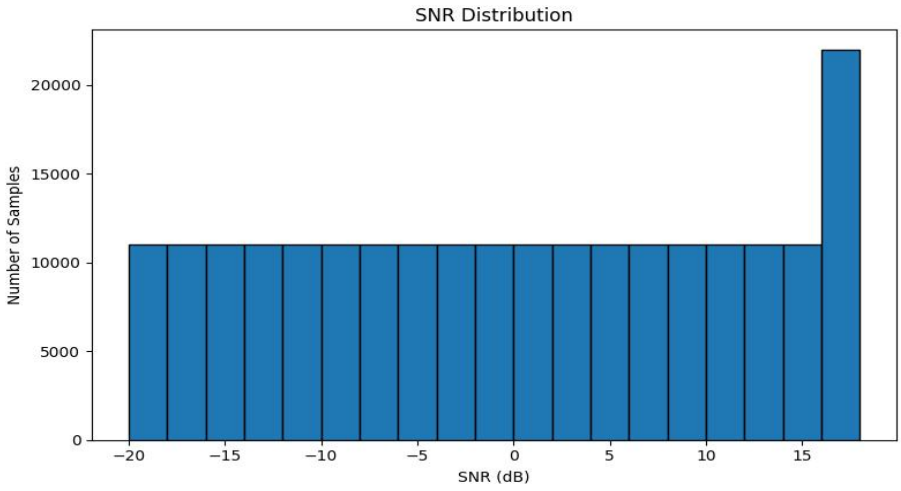


Fig. 1. SNR distribution histogram (Number of samples vs. SNR).

IEER
Institute for Excellence in Education & Research

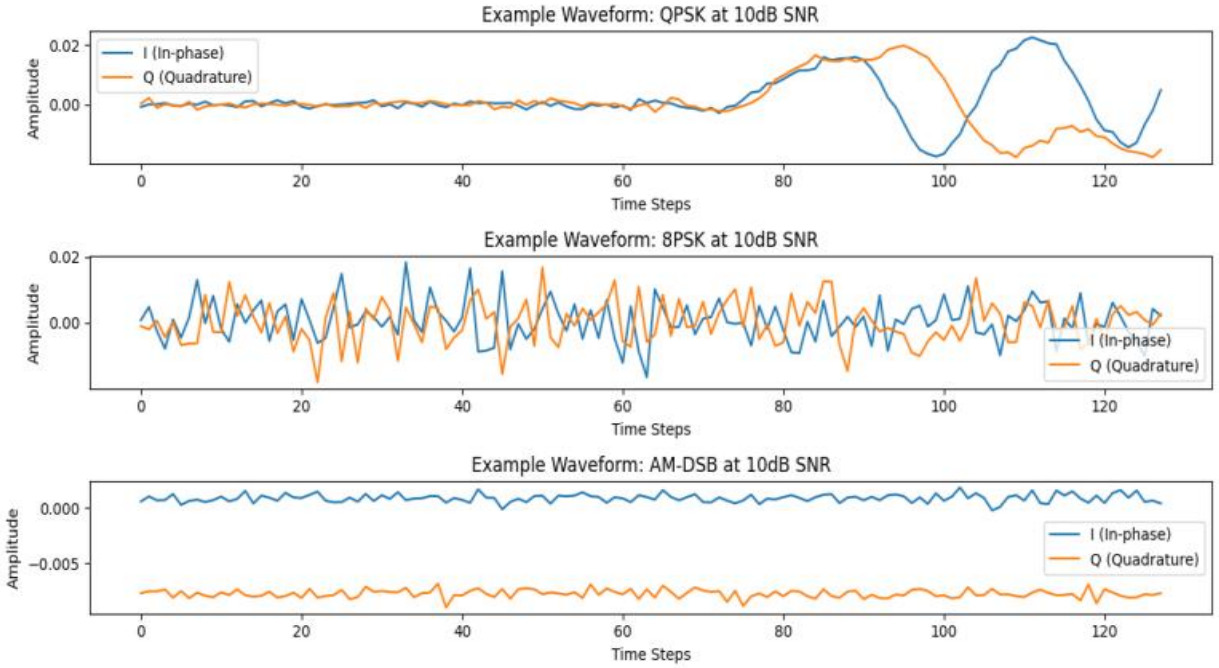


Fig. 2. Example waveforms (Amplitude vs. Time steps).

4 Methodology

4.1 Residual CNN

The core deep learning model employed in this work is a one-dimensional Residual Convolutional Neural Network (Residual CNN), specifically designed to handle raw I/Q sequences efficiently [9, 10]. The architecture integrates stacked residual blocks with progressively decreasing filter sizes, 256 filters with a kernel size of 7, followed by 128 filters with kernel sizes of 5 and 3 allowing the network to capture both wide and fine temporal dependencies.

Regularization is achieved through the application of dropout layers and batch normalization, which help mitigate overfitting and improve generalization. A Global Average Pooling layer is used to condense the learned features into a compact representation, followed by a dense softmax classifier to predict one of the 11 modulation classes. To further enhance training, a cosine learning-rate restart scheduler has been employed in conjunction with early stopping and model check pointing, ensuring optimal convergence while avoiding overtraining.

4.2 Baseline Models

To provide a fair benchmark against the proposed CNN, two classical machine learning models have been implemented: Random Forest (RF) and Support Vector Machine (SVM). The Random Forest model is constructed with 200 decision trees and trained using flattened I/Q samples, capturing a variety of non-linear relationships through ensemble averaging. In parallel, the SVM model is trained with a radial basis function (RBF) kernel, known for its effectiveness in handling non-linear classification tasks. Its hyperparameters are carefully tuned, with the penalty parameter set to C=10 and gamma set to 'scale', ensuring robust decision boundaries. Both models relied on handcrafted statistical features extracted from the I/Q signals, offering a traditional contrast to the end-to-end feature learning capabilities of CNNs [3, 20].

4.3 Evaluation Strategy

The evaluation framework is designed to comprehensively assess the performance and robustness of all models. Standard classification metrics, including accuracy, precision, recall, and F1-score, have been calculated to quantify overall performance.

Additionally, per-SNR accuracy is measured by evaluating each model's performance across individual SNR levels, offering insights into robustness under varying noise conditions. Beyond numerical evaluation, interpretability and visualization techniques have been employed [17, 19]. Confusion matrix has been generated to identify common misclassifications among modulation types.

Using Grad-CAM heatmaps revealed the temporal I/Q regions most influential in CNN decision-making, and t-SNE plots visualized feature-space separation, thereby validating the representational strength of deep learning approaches compared to classical baselines. An overall system flow of this research project can be seen in Fig. 3.

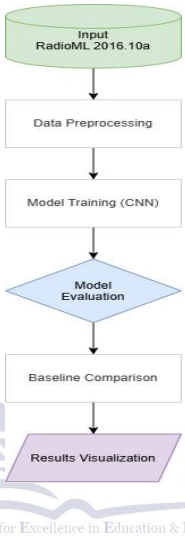


Fig. 3. System flow diagram

5 Results and Discussion

The experimental evaluation highlights the effectiveness of the proposed residual CNN compared to classical machine learning baselines (RF and SVM), especially under high-SNR conditions.

5.1 Per-SNR Performance

To further analyze robustness, per-SNR accuracy trends have been examined. The CNN exhibited graceful degradation as the SNR decreased. At extremely low SNR values (≤ -10 dB), accuracy hovered near random chance, reflecting the difficulty of the task. However, performance steadily improved with increasing SNR, surpassing 70% accuracy at 2 dB, and reaching close to 90% accuracy above 10 dB. This steady performance curve (Fig. 4) highlights the CNN's resilience in noisy environments, even when conditions are not ideal.

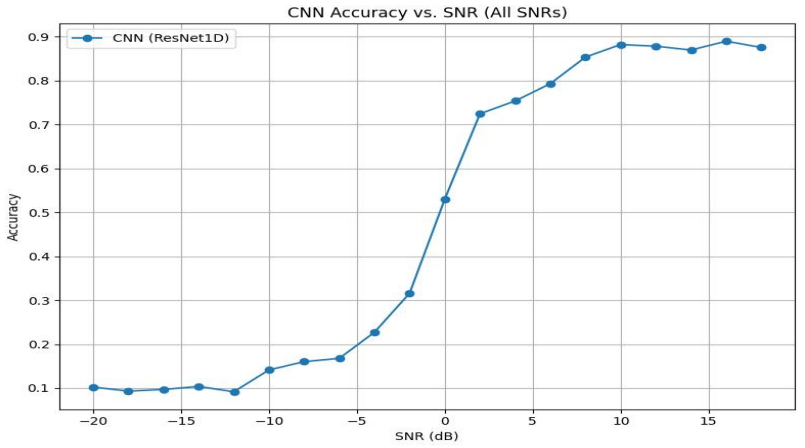


Fig. 4. Evaluating CNN across all SNR values.

5.2 High-SNR Classification

When restricted to SNR levels ⩾10 dB, the CNN clearly demonstrated superior classification accuracy compared to Random Forest and SVM. The CNN achieved an overall accuracy of 88%, substantially higher than the 72.29% obtained by Random Forest and the 70.31% from SVM. The CNN has been particularly effective for digital modulation schemes such as BPSK, QPSK, 8PSK, CPFSK, GFSK, and PAM4, achieving near-perfect classification. On the other hand, classical models performed moderately but consistently lagged behind the CNN. While RF and SVM classified low-order modulations and CPFSK/GFSK with reasonable accuracy, they struggled with higher-order constellations such as QAM16 and QAM64. The CNN also showed some weakness in QAM16 and QAM64 recognition, but its performance is still higher than RF and SVM, which misclassified many of these cases. Analog modulations such as WBFM posed challenges across all models. Nonetheless, the CNN outperformed both classical baselines, establishing it as the stronger candidate for high-SNR AMC tasks (Fig. 5).

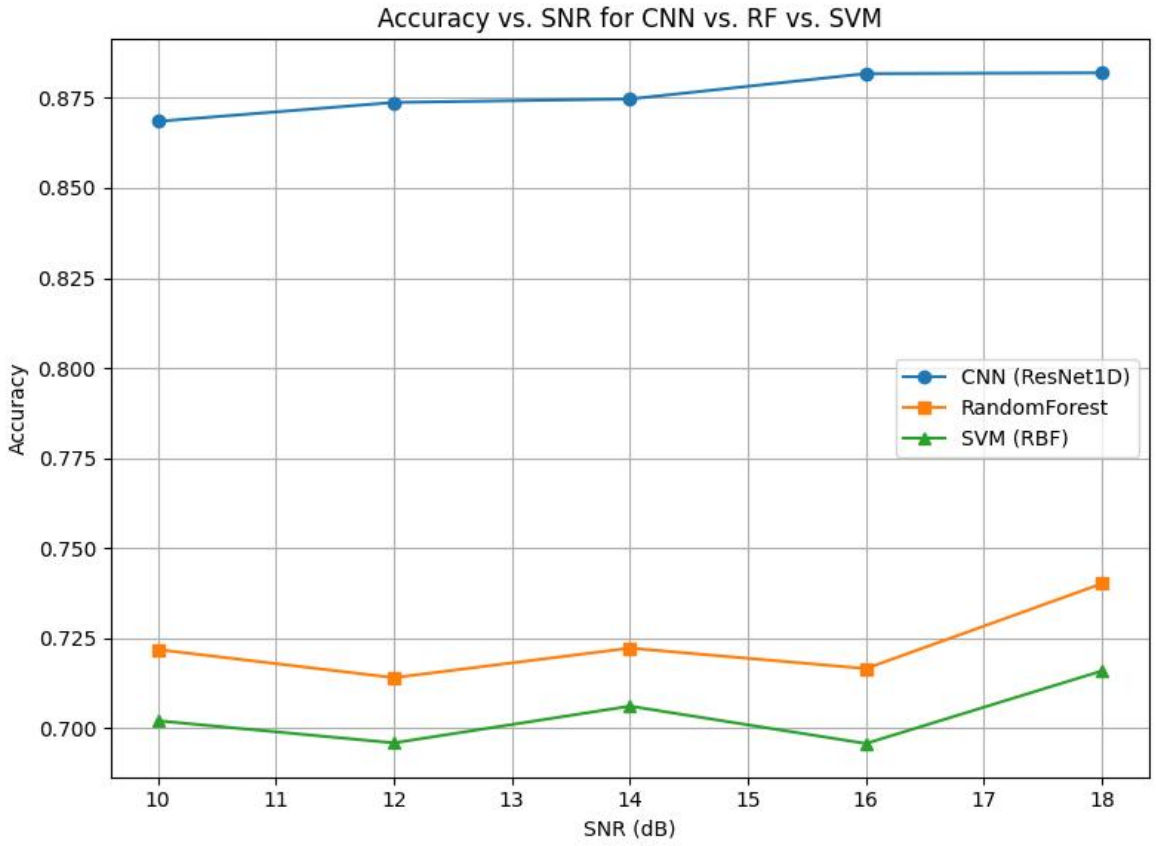


Fig. 5. Evaluating CNN vs. RF vs. SVM per higher SNR.

5.3 Interpretability and Visualization

Model interpretability and feature-space analysis provided deeper insights into classifier behavior. Confusion matrix showed that the majority of misclassifications occurred between visually and structurally similar modulation schemes, such as QAM16 vs. QAM64 (Fig. 6). For analog modulations like WBFM and AM-DSB, errors stemmed from overlapping spectral patterns under noisy conditions. Grad-CAM analysis confirmed that the CNN did not make decisions arbitrarily but instead learned to focus on distinctive I/Q temporal regions that carried modulation-specific signatures (Fig. 7).

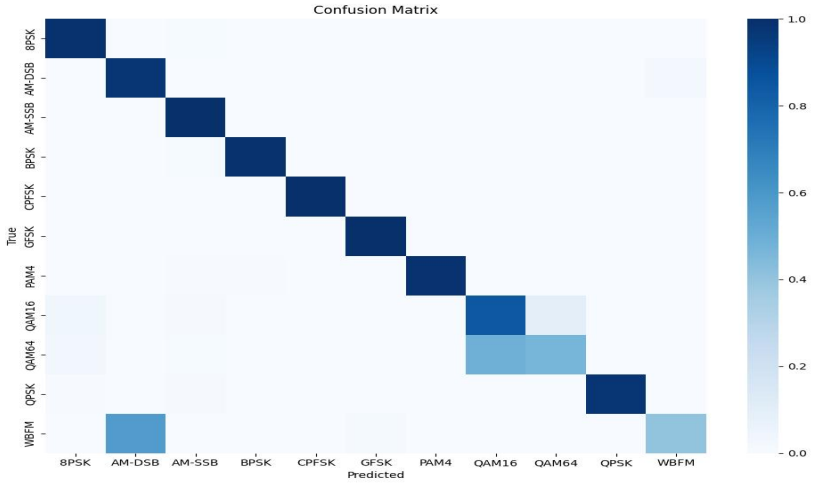


Fig. 6. Confusion matrix (normalized, for high SNR ⩾10 dB).

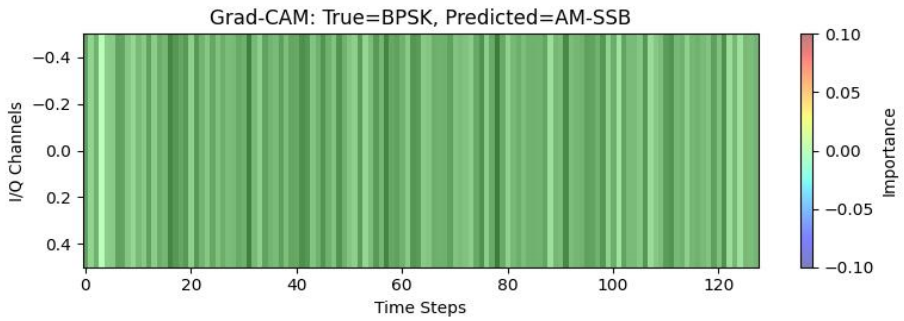


Fig. 7. Grad-CAM visualization (I/Q Channels vs. Time Steps).

This demonstrates that deep learning is capturing meaningful discriminative features rather than overfitting noise. Similarly, t-SNE visualizations of the CNN's feature space revealed well-separated clusters corresponding to modulation types, underscoring the model's strong ability to learn compact and discriminative embeddings (Fig. 8). In contrast, feature spaces derived from handcrafted features used by RF and SVM are less structured, explaining their weaker classification performance. The CNN consistently outperformed classical baselines in both per-SNR evaluations and high-SNR conditions. Its robustness, interpretability, and feature-space discriminability highlight its suitability for practical AMC applications in modern wireless systems.

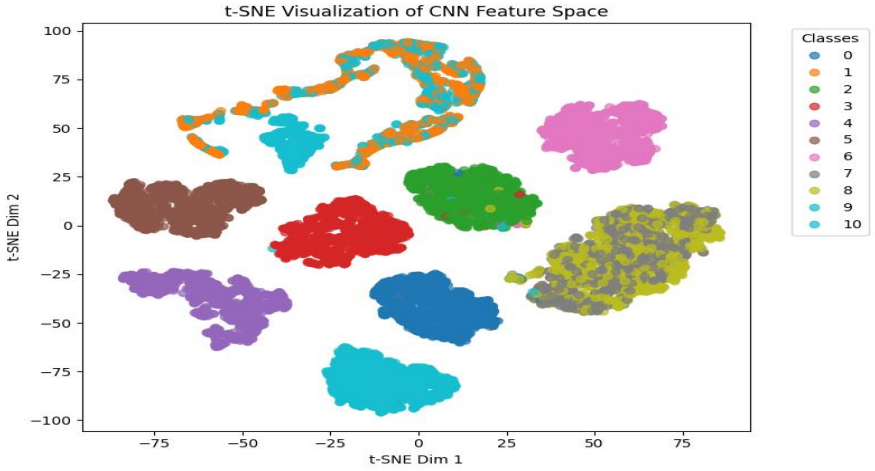


Fig. 8. t-SNE visualization of CNN feature space.

The results confirm that the deep residual CNNs provide strong discriminative power for modulation classification, particularly in high-SNR regimes. Classical ML models, while computationally cheaper, fail to generalize well across modulation families. The interpretability analysis further establishes the CNN as not only more accurate but also more insightful in capturing signal characteristics. Waveform visualizations of some correct vs. incorrect predictions and a few examples of average waveforms per class at 10 dB are shown in Fig. 9 and Fig. 10.

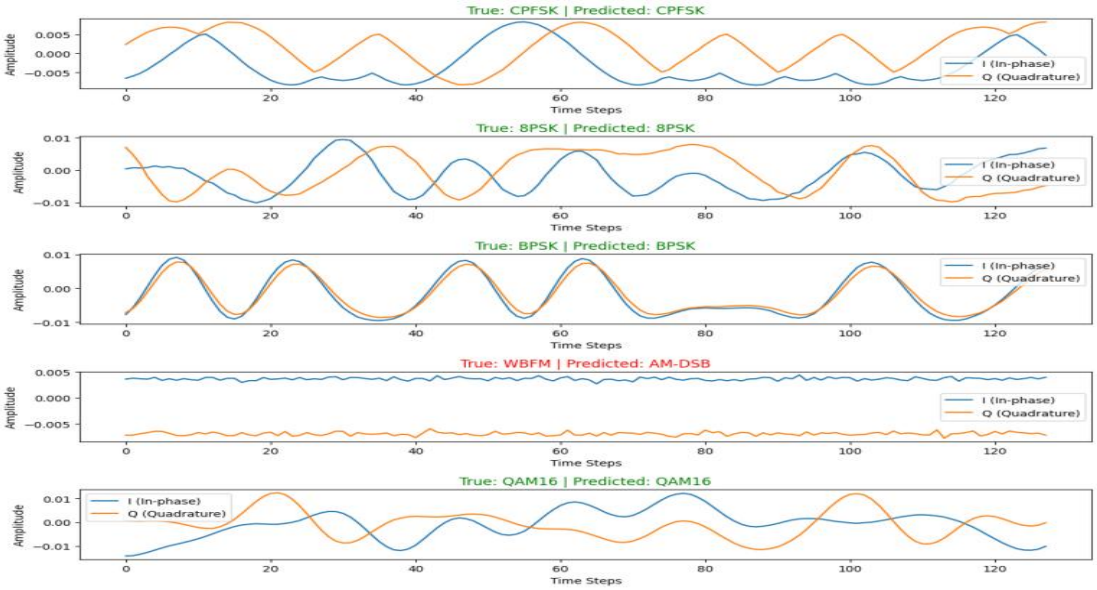


Fig. 9. Waveform visualizations of correct vs. incorrect predictions.

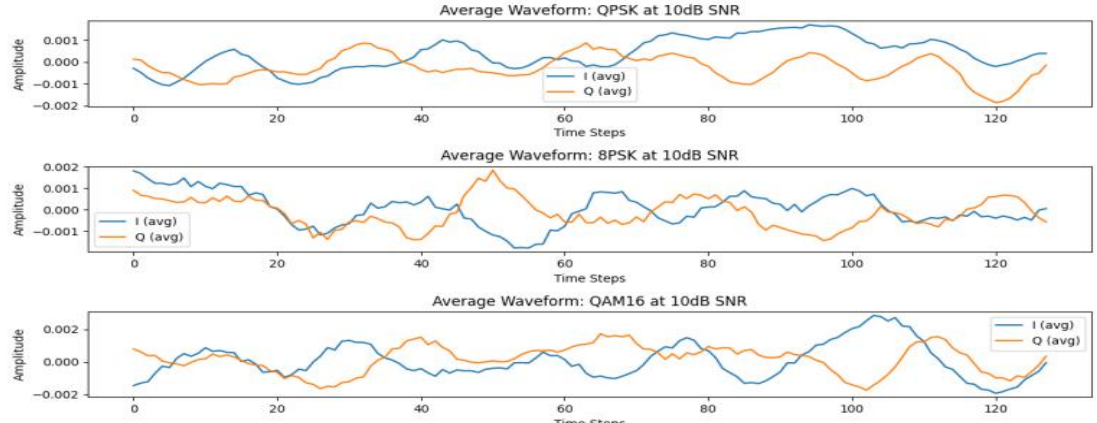


Fig. 10. Average waveforms per class at 10 dB.

6 Conclusion

This study has presented a comprehensive investigation into the use of Convolutional Neural Network (CNN) for automatic modulation classification (AMC), under different SNR conditions that are crucial for reliable communication systems. By leveraging a residual CNN architecture optimized with high-SNR-focused learning, dropout regularization, batch normalization, and cosine learning-rate restarts, it has been demonstrated that deep learning models can significantly outperform traditional machine learning approaches such as Random Forest (RF) and Support Vector Machines (SVM). The proposed CNN consistently achieved superior accuracy across a wide range of modulation types, excelling especially in digital modulation schemes while showing resilience in noisy environments. One of the key contributions of this work is the establishment of a reproducible benchmark pipeline for AMC research using the widely adopted RadioML 2016.10A dataset. This not only facilitates fair comparisons across models but also integrates interpretability tools, including Grad-CAM visualizations and t-SNE feature space mapping, to provide deeper insights into model behavior.

These analyses revealed that the CNN is capable of learning discriminative temporal and spectral patterns directly from raw I/Q sequences, leading to meaningful feature representations that classical models relying on handcrafted features fail to capture. The interpretability aspect also strengthens confidence in deploying such deep learning solutions in real-world cognitive radio and spectrum monitoring systems. The comparative evaluation further highlights the limitations of RF and SVM, which, despite their effectiveness in controlled or high-SNR scenarios, fail to maintain robustness under noise or when distinguishing between closely related modulation schemes. In contrast, the CNN maintained competitive accuracy even under challenging low-SNR conditions, thereby underscoring the scalability and adaptability of deep learning methods in practical wireless communication environments. Looking ahead, this research opens several promising directions.

Future work can focus on hybrid deep architectures that combine convolutional, recurrent, and attention-based mechanisms to capture both local and long-range dependencies in I/Q signals. Transfer learning and domain adaptation techniques also hold potential for bridging the gap between synthetic datasets such as RadioML and real-world radio environments, ensuring that trained models generalize effectively in the presence of hardware impairments, fading channels, and interference.

Lightweight CNN architectures can be explored for deployment on resource-constrained edge devices, enabling real-time AMC in Internet of Things (IoT) and tactical communication scenarios. In conclusion, the study demonstrates that deep residual CNNs are not only accurate but also robust and interpretable solutions for AMC tasks, setting a strong foundation for future research in intelligent wireless communication systems. By combining methodological rigor, reproducibility, and interpretability, this work advances the field towards

practical deployment of deep learning-based AMC solutions.

Acknowledgments

I gratefully acknowledge the DeepSig platform for providing open access to the dataset that made this study possible. The data was retrieved from https://www.deepsig.ai/datasets/.

References

A. Kumar, K. K. Srinivas, and S. Majhi. Automatic modulation classification for adaptive OFDM systems using convolutional neural networks with residual learning. IEEE Access, 11:61013–61024, 2023.

A. Kumar, U. Satija, et al. Residual stack-aided hybrid CNN-LSTM-based automatic modulation classification for orthogonal time-frequency space system. IEEE Communications Letters, 27(12):3255–3259, 2023.

A. P. Hermawan, R. R. Ginanjar, D.-S. Kim, and J.-M. Lee. CNN-based automatic modulation classification for beyond 5g communications. IEEE Communications Letters, 24(5):1038–1041, 2020.

B. Bhargava, A. Deshmukh, M. V. Rupa, R. P. Sirigina, S. K. Vankayala, and A. Narasimhadhan. Deep learning approach for wireless signal and modulation classification. In 2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall), pp. 1–6. IEEE, 2021.

H. S. Hussein, M. H. E. Ali, M. Ismeil, M. N. Shaaban, M. L. Mohamed, and H. A. Atallah. Automatic modulation classification: Convolutional deep learning neural networks approaches. IEEE Access, 11:98695–98705, 2023.

J. Chen, H. Cui, S. Miao, C. Wu, H. Zheng, S. Zheng, L. Huang, and Q. Xuan. Fem: Feature extraction and mapping for radio modulation classification. Physical Communication, 45:101279, 2021.

J. R. Satapathy, R. M. Bodade, and J. Ayeelyan. CNN based modulation classifier and radio fingerprinting for electronic warfare systems. In 2024 4th International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), pp. 1880–1885. IEEE, 2024.

K. Jung, J. Woo, and S. Mukhopadhyay. On-chip acceleration of rf signal modulation classification with short-time fourier transform and convolutional neural network. IEEE Access, 11:144051–144063, 2023.

M. Leblebici, A. C¸ alhan, and M. Cicio ˘glu. CNN-based automatic modulation recognition for index modulation systems. Expert Systems with Applications, 240:122665, 2024.

P. Osinowo. CNN-Based Signal Modulation Classifier for Shared Spectrum in Cognitive Radio Networks. PhD thesis, University of Strathclyde, 2024.

R. R. Yakkati, R. R. Yakkati, R. K. Tripathy, and L. R. Cenkeramaddi. Radio frequency spectrum sensing by automatic modulation classification in cognitive radio system using multiscale deep CNN. IEEE sensors journal, 22(1):926–938, 2021.

S. Mohsen, A. M. Ali, and A. Emam. Automatic modulation recognition using CNN deep learning models. Multimedia Tools and Applications, 83(3):7035–7056, 2024.

S. Rajesh, S. Geetha, S. Babu Sudarson, and S. Ramesh. Deep learning based real time radio signal modulation classification and visualization. International Journal of Engineering and Manufacturing, 13(5):30–37, 2023.

S. Ramjee, S. Ju, D. Yang, X. Liu, A. Gamal, and Y. Eldar. Fast deep learning for automatic modulation classification. IEEE Transactions on Cognitive Communications and Networking, Aug. 2021.

S. T. H. Chahil, M. Zakwan, K. Khan, and A. Fazil. Performance anal-ysis of different signal representations and optimizers for CNN based automatic modulation classification. Wireless Personal Communications, 139(4):2503–2528, 2024.

S.-H. Kim, J.-W. Kim, W.-P. Nwadiugwu, and D.-S. Kim. Deep learning-based robust automatic modulation classification for cognitive radio networks. IEEE access, 9:92386–92393, 2021.

T. Huynh-The, Q.-V. Pham, T.-V. Nguyen, T. T. Nguyen, R. Ruby, M. Zeng, and D.-S. Kim. Automatic modulation classification: A deep architecture survey. IEEE Access, 9:142950–142971, 2021.

T. O'Shea, J. Corgan, and T. C. Clancy. Convolutional radio modulation recognition networks. In International Conference on Engineering Applications of Neural Networks, 2016.

T. Wang, G. Yang, P. Chen, Z. Xu, M. Jiang, and Q. Ye. A survey of applications of deep learning in radio signal modulation recognition. Applied Sciences, 12(23):12052, 2022.

T. Xu and Y. Ma. Signal automatic modulation classification and recognition in view of deep learning. IEEE Access, 11:114623–114637, 2023.

X. Liu, C. J. Li, C. T. Jin, and P. H. Leong. Wireless signal representation techniques for automatic modulation classification. IEEE Access, 10:84166–84187, 2022.

X. Liu, D. Yang, and A. E. Gamal. Deep neural network architectures for modulation classification. In 2017 51st Asilomar Conference on Signals, Systems, and Computers, pp. 915–919, 2017. doi: 10.1109/ACSSC.2017.8335483

IEER
Science in Education & Research