



# An Efficient Hybrid Deep Learning Model Based on Bi-LSTM and CNN for Automatic Modulation Classification

Hesham M.Ismail<sup>a</sup> Mostfa Salah<sup>b</sup> Safwat M.ramzy<sup>b</sup>

<sup>a</sup> Faculty of Computers and Information Technology, EELU, Giza, 12611, Egypt

<sup>b</sup> Dept. of Electrical Engineering, Faculty of Engineering, Sohag University, Sohag 82524, Egypt

## Abstract

Automatic modulation classification (AMC) is vital in cognitive radio, spectrum management, and military applications. The rapid evolution of wireless networks demands reliable and effective techniques. Traditional AMC methods, like likelihood-based and feature-based algorithms, face limits in computational complexity and generalization across different channel conditions. This paper combines Bidirectional Long Short-Term Memory (Bi-LSTM) and Convolutional Neural Networks (CNN) to form a novel hybrid deep learning model. The model exploits each branch's strengths for temporal and spatial feature extraction. The Bi-LSTM branch identifies long-term temporal relationships, such as phase transitions, in signal sequences. In contrast, the CNN branch captures local spatial patterns like constellation structures. The proposed model outperforms baseline models—CLDNN, CNN-GRU, ResNet-LSTM, and VTCNN2—while maintaining lower computational complexity (410,762 parameters) on the RML2016.10b dataset. It achieves a classification accuracy of 92.32% at 18 dB SNR and performs robustly across low-to-high SNR regimes. These results reveal its potential for deployment in real-time wireless scenarios that are both noisy and dynamic.

© 2025 Published by Published by Faculty of Engineering – Sohag University. DOI: 10.21608/sej.2025.415483.1083

Keywords: Convolutional Neural Networks (CNN)- Hybrid Architectures - Bidirectional LSTM - Wireless Communication.

## 1. INTRODUCTION

The growing need for reliable and effective information transfer has led to significant advancements in wireless communications in recent years. However, identifying and categorizing signal modulation techniques is becoming more difficult as transmitted data volumes and complexities grow. Automatic Modulation Classification (AMC) is a crucial technique that enables communication systems to recognize various modulation patterns without knowing the sent data in advance. Applications such as spectrum management, cognitive radio, and improving communication network effectiveness require this capability. AMC provides vital modulation information, especially when there is no prior signal knowledge and the situation is non-cooperative. It is the first step needed for successful demodulation of received signals [1]. The problem becomes worse in military communication, where engineers work in extremely crowded electromagnetic conditions. Friendly signals must be reliably broadcast and received. Hostile signals must be identified, categorized, and blocked. These signals employ a broad range of modulation formats, from basic narrowband schemes to sophisticated wideband systems, and can span frequencies from high-frequency (HF) bands to millimeter-wave (mm-wave) bands. Thus, blind detection of the modulation format continues to be a major challenge in both commercial and military applications. To ensure safe and effective communication in these situations, complex real-time signal detection and processing techniques are needed [2].

The two primary categories of traditional AMC algorithms are likelihood-based (LB) and feature-based (FB) techniques. In general, likelihood-based techniques can be classified into the maximal likelihood (ML) theory. Tests of hybrid likelihood ratio (HLRT), generalized probability ratio (GLRT), and average likelihood ratio (ALRT). Some of them can provide a greater likelihood of successful classification, but they usually have a high computational complexity, which prevents them from being used in some real-time applications[3].

FB methods include both feature extraction and classification. The feature extraction stage is when the received signals' features are extracted. A suitable classifier uses the information gathered during the feature extraction phase to categorize the different forms of modulation[4]. The most commonly used AMC features fall into five categories: zero-crossing, statistical, transform domain, instantaneous time domain, and constellation

\* Corresponding author: heshamesmael.124@gmail.com

shape features[5]. The most commonly used classifiers are support vector machine (SVM)[6], Decision tree (DTs)[7], Naïve Bayes[8], and Nearest neighbor[9]. Both likelihood-based and feature-based techniques have serious drawbacks while being widely used. Although LB techniques are theoretically ideal in given channel conditions, they are challenging for many real-time or resource-constrained applications because of their high computational complexity and need for strong a priori knowledge. On the other hand, FB approaches mostly rely on manually created features, which might not generalize well in various channel settings or signal situations.

Deep learning (DL) is a more effective and independent option to conventional machine learning (ML) techniques because of its impressive capabilities in automatic feature extraction and decision-making. Radio modulation categorization has seen a rise in the use of DL techniques in recent years, since they have demonstrated greater performance in a variety of communication contexts.[1]. DL models that are most frequently employed include Deep Neural Network (DNN), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), and Long Short-Term Memory (LSTM). Additionally, certain hybrid DL models exist.

To capitalize on the complementary advantages of both temporal and spatial feature learning, we suggest an effective hybrid deep learning model for AMC that combines CNN and Bidirectional LSTM (Bi-LSTM). From raw I/Q samples or time-frequency representations, the CNN branch retrieves local spatial patterns (such as constellation structures or amplitude changes), whereas the Bi-LSTM branch records long-term temporal dependencies (such as phase transitions or fading dynamics) across signal sequences. Even in low-SNR and non-linear channel conditions, our model performs robust classification by hierarchically combining these features.

The main contributions of this paper can be summarized as follows:

- A novel parallel hybrid deep learning architecture that combines CNN and Bi-LSTM is proposed for efficient AMC. Unlike conventional sequential hybrid models, the proposed structure processes spatial and temporal features in parallel, enhancing feature diversity and learning efficiency.
- The model achieves superior classification accuracy, particularly under low SNR conditions, demonstrating strong robustness against channel impairments.
- The proposed architecture significantly reduces model complexity and training time compared to benchmark models such as CLDNN, GRU-CNN, ResNet-LSTM, and VTCNN2.
- A detailed class-level evaluation is conducted to analyze the model's performance for each modulation type, providing deeper insights into classification behavior across different SNR ranges

## 2. RELATED WORK

One of the earliest important studies in the field proposed a deep neural network (DNN)-based automatic modulation classification technique that used 21 extracted features [10]. The proposed DNN outperformed a shallow Artificial Neural Network (ANN)-based classifier, particularly under high Doppler fading channels. Following this, researchers focused on increasing the robustness and automation of AMC by employing deep learning architectures. CNNs became the most widely used approach due to their ability to extract hierarchical spatial features from time-frequency representations (e.g., spectrograms) and raw I/Q signals. For instance, [11] reported that CNN-based architectures achieved classification accuracies approaching 90%. Several subsequent studies have proposed lightweight and efficient CNN designs for AMC in wireless communication systems [12]–[15]. For example, the ResNet50-based model achieved 95% accuracy in identifying both modulation type and signal-to-noise ratio [13]. Other works have applied denoising techniques on cyclic spectra to improve robustness against noise, achieving reliable identification of multiple VHF modulation schemes [15].

RNNs are a family of deep learning models that utilize a hidden state to capture temporal dependencies over time in order to handle sequential data. Because RNNs use feedback connections instead of standard feedforward neural networks, they can remember past inputs and simulate time-varying patterns, which makes them especially well-suited for signal processing applications like (AMC). It was first proposed by Pollack[16]. To solve the issue of vanishing gradients, the Long Short-Term Memory (LSTM) network was created[17]. The authors in [18] suggest an innovative and successful RNN-based AMC technique. The suggested approach outperforms the CNN-based approach at high SNR regimes and increases recognition accuracy from 80% to 91%. Using the LSTM, the classification accuracy is about 90%[19]. For Automatic Modulation Classification (AMC), a novel data-driven approach based on long short-term memory (LSTM) is suggested [20]. Analysis reveals that under a range of SNR settings, from 0dB to 20dB, the suggested model produces an average classification accuracy of about 90%.

A hybrid deep learning model is an architecture that overcomes the drawbacks of individual models by combining two or more neural network frameworks to capitalize on their corresponding strengths. Investigation of signal feature extraction by including recurrent, convolutional, and other types of residual layers into a deep neural network architecture. It was determined that the optimum classification architecture was a Convolutional Long Short-term Deep Neural Network (CLDNN), which increases accuracy by about 13.5%[21]. CNN and Gated Recurrent Units (GRU) are combined in the hybrid network-based AMC approach[22]. This method automatically

extracts and classifies data with different dimensionalities using several structures. On various datasets, the overall recognition accuracy was 60.64% and 73.2%, respectively.

### 3. SIGNAL MODEL

The goal of AMC is to identify and classify the modulation-specific types of incoming radio signals by analyzing and modeling their signal characteristics. The received signal ( $r(t)$ ) in a typical wireless communication system can be represented as the convolution of the sent signal ( $x(t)$ ) with the channel impulse response ( $h(t)$ ), distorted by additive noise ( $n(t)$ ) which is handled as AWGN as expressed in [2]. The mathematical expression for this relationship is

$$R(t)=x(t)*h(t)+n(t) \quad (1)$$

### 4. PROPOSED METHODOLOGY

Based on the motivations and contributions discussed in the introduction. The proposed model shown in Fig. 1 adopts a parallel hybrid architecture that integrates two specialized feature extraction branches: a spatial feature branch based on CNN and a temporal feature branch based on (Bi-LSTM network. Unlike conventional sequential hybrid models, in which the output of one block is fed into the next (e.g., CNN  $\rightarrow$  LSTM or LSTM  $\rightarrow$  CNN), the parallel design processes the raw I/Q signal input simultaneously in both branches. This arrangement enables each branch to independently learn complementary feature representations — the CNN focusing on localized spatial correlations, and the Bi-LSTM capturing bidirectional temporal dependencies. The absence of cascading dependency reduces the risk of information loss and ensures that both spatial and temporal patterns are preserved in their most expressive form. After independent feature extraction, the two representations are concatenated into a unified feature space and passed to a fully connected classification layer. The parallel approach offers several advantages over sequential designs, including improved generalization to varying channel conditions, richer combined feature spaces, and greater architectural flexibility for branch-specific optimization. The following subsections describe each branch in detail.

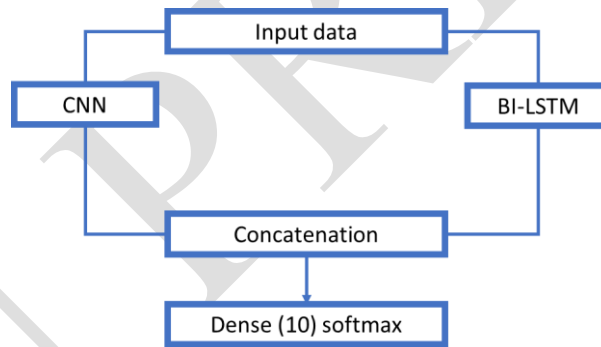


Fig. 1. The proposed model architecture

#### 4.1. The CNN Architecture Design

The spatial feature extraction branch uses a four-hierarchical, two-dimensional (CNN) to analyze I/Q input samples with the complex signal components represented by the shapes (2, 128, 1). The architecture is made up of First Feature Extraction: ReLU activation and a 128-filter convolutional layer with 1x8 kernels for feature standardization and dimensionality reduction, uses batch normalization and max pooling (2x2 window). Batch normalization is used in the second 128-filter convolutional layer (1x8 kernels) of intermediate processing. A transition layer employing 64 filters (1x3 kernels) and max pooling (1x2 window) comes next. Final 64-filter convolutional layer (1x8 kernels) for feature refinement. Regularization using batch normalization and dropout ( $p=0.3$ ). To gradually change the input, the network uses: To capture both wide and tiny signal features, the kernel sizes were carefully chosen (8 samples first, followed by 3 samples). Reducing temporal dimensions while maintaining spatial links using pooling operations, Layers for batch normalization, and maintaining training at all depths. Dropout layers prevent overfitting in deeper layers. For real-time processing scenarios, this architecture was tuned to extract discriminative spatial information while preserving computational efficiency.

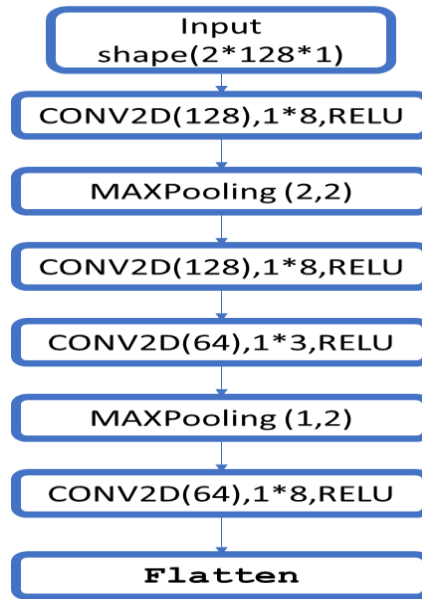


Fig. 2. CNN architecture design

#### 4.2 .BI- LSTM Architecture Design

Bi-LSTM is an advanced variant of the RNN designed to overcome the limitations of standard RNNs in capturing long-term dependencies within sequential data. The Long Short-Term Memory (LSTM) architecture introduces memory cells and gating mechanisms (input, forget, and output gates) that regulate the flow of information, enabling the network to selectively retain relevant information and discard irrelevant patterns over extended time spans.

Unlike conventional RNNs or unidirectional LSTMs, which process sequences in a single temporal direction (forward or backward), a Bi-LSTM processes the input sequence in both forward and backward directions simultaneously. This dual processing is achieved through two separate LSTM layers:

- The forward layer captures dependencies from past to future.
- The backward layer captures dependencies from future to past.

Each layer maintains its own hidden states and memory cells, and their outputs are typically concatenated to form a comprehensive representation that encodes contextual information from both temporal directions. This bidirectional structure significantly enhances the model's ability to capture complex patterns and contextual dependencies, making it highly effective for tasks involving sequential data, such as speech recognition, natural language processing, and signal classification.

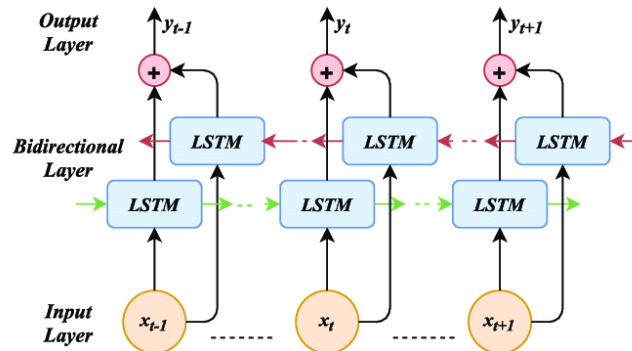


Fig. 3. Bidirectional processing

As shown in Fig.4, A two-layer bidirectional LSTM (Bi-LSTM) architecture is used by the temporal feature extraction branch to thoroughly examine sequential signal patterns from the  $(2, 128)$  input I/Q sequences. With `return sequences=True`, the 64-unit Bi-LSTM with ReLU activation in the first layer maintains temporal

dimensions and fully captures both forward and backward signal dependencies. A second 64-unit Bi-LSTM layer (`return_sequences=False`) processes this bidirectional context and combines the temporal characteristics while preserving the advantages of ReLU activation, which mitigates vanishing gradients. When combined, this design captures long-range dependencies that are essential for reliable modulation classification in difficult channel conditions and efficiently represents time-varying signal properties, like as phase transitions and frequency shifts. Because the Bi-LSTM is bidirectional, it can analyze each time step in chronological and reverse chronological settings. It provides a refined temporal representation that can be fused with spatial information for superior temporal analysis.

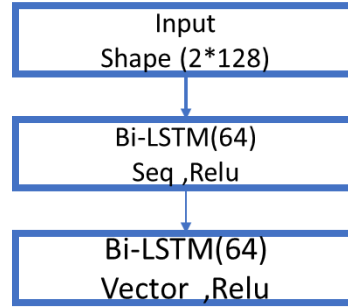


Fig. 4. BI- LSTM architecture design

## 5. EXPERIMENTAL RESULTS AND EVALUATION

### 5.1. Dataset and Simulation Environment

RML2016.10b is a popular wireless signal detection dataset for machine learning, especially modulation classification. It will be used for training and testing for the proposed model. The dataset was synthetically generated using a software-defined radio (SDR) simulation framework (GNU), which created realistic radio signals with varying modulations and noise conditions. To simulate actual situations, the center frequency, symbol rate, and other signal parameters were changed. Multiple SNR levels were simulated by adding AWGN (Additive White Gaussian Noise) in 2 dB steps, ranging from -20 dB to +18 dB. All SNR levels were used to generate each modulation in order to ensure robustness testing. These datasets were first made available[23]. These are split into WBFM, AM-SSB, and AM-DSB for analog modulation and 8PSK, BPSK, SPFSK, GFSK, PAM4, QAM16, QAM64, and QPSK for digital modulation. Python 3.8 was used for all simulations and model training, together with the TensorFlow and Keras deep learning packages. After normalizing all input signals before training, the dataset was divided into 80% for training and 20% for testing. Key attributes of the dataset used in our tests are compiled in Table 1. With a learning rate of 0.0001, the model was trained using the Adam optimizer. Because the classification job involved multiple classes, categorical cross-entropy was the loss function employed. A batch size of 64 was used for training across 30 epochs, and the accuracy metric was used to assess the model's performance.

TABLE 1. SUMMARY OF DATASET

| Attribute        | Description                                                     |
|------------------|-----------------------------------------------------------------|
| Name             | RML2016.0b                                                      |
| Total samples    | 1200000                                                         |
| Sampling rate    | 1 MSps                                                          |
| Snr range        | From -20 to +18                                                 |
| Modulation types | QPSK, PAM 4, 8PSK, BPSK, CPFSK, BFSK QAM64, QAM16 AM-DSB, WB-FM |
| Signal duration  | 128 samples per signal (I/Q format)                             |
| Data format      | Complex baseband I/Q samples                                    |

## 5.2. Evaluation Metrics

The use of suitable assessment criteria is essential for evaluating the effectiveness of the suggested hybrid deep learning model for AMC. These measurements aid in assessing the model's ability to successfully and accurately distinguish between various modulation types under a range of signal situations. We can learn more about the model's advantages, disadvantages, and effectiveness for practical application by examining its predictions with statistical and graphical tools. This section explains the primary metrics we used in our studies and how they fit into the process of evaluating performance.

### 5.2.1. Classification accuracy

The suggested Bi-LSTM and CNN-based hybrid model's classification accuracy is shown in Fig. 5 across a broad range of Signal-to-Noise Ratio (SNR) levels, from -20 dB to +18 dB. As can be seen, the model performs well in a range of noise scenarios. Due to the high noise level, which makes signal features harder to discern, the classification accuracy stays very poor at very low SNR levels (e.g., -20 dB to -10 dB). Nonetheless, the accuracy of the model greatly improves as the SNR rises. Around 8 dB, there is a clear improvement that lasts until it saturates at +6 dB, where accuracy exceeds 90%. This pattern demonstrates the model's capacity for generalization and high classification accuracy in settings with moderate to high SNR, which is essential for real-world wireless communication situations. The hybrid architecture is effective in learning discriminative features across various modulation types, even under difficult noise situations, as evidenced by the stable performance at high SNR levels. The model performs well, reaching a peak accuracy of 92.32% at 18 dB and increasing to 87.51% at 0 dB. The model exhibits stability and dependability in low-noise situations, maintaining a continuously high accuracy above 91% between 6 dB and 18 dB.

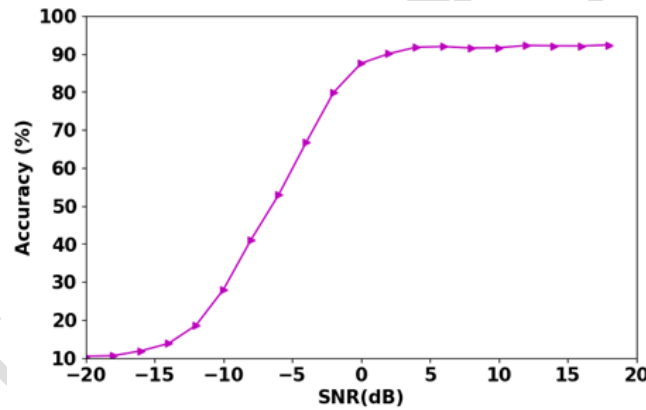


Fig. 5. Classification accuracy of the proposed model

### 5.2.2. confusion matrix

In classification problems, the confusion matrix is a popular evaluation tool that offers a thorough analysis of the model's prediction accuracy across many classes. In contrast to accuracy alone, which provides a broad performance metric, the confusion matrix shows how well the classifier differentiates between different kinds of modulation.

Fig.6 displays the confusion matrices at SNR levels of -2 dB, 0 dB, 10 dB, and 18 dB, which demonstrate the robust classification capabilities of the suggested model under various channel conditions. For digital modulations with distinct constellation patterns (CPFSK: 98%, GFSK: 96%, PAM4: 99%), the model performs well at -2 dB. However, analog signals (AM-DSB: 64%, WBFM: 57%) are more difficult to handle because of noise susceptibility. For QPSK (93% vs. 75% at -2 dB) and QAM16 (92% vs. 86%), classification accuracy greatly improves as SNR rises to 0 dB, with misclassifications mostly happening between spectrally similar modulations

such as QAM16/QAM64. For most digital modulations, high SNR situations (10–18 dB) exhibit near-perfect accuracy (BPSK/CPFSK/GFSK: 99–100%, PAM4/QPSK: 99%), while results at 18 dB show remarkable robustness (99% for 8PSK/BPSK).

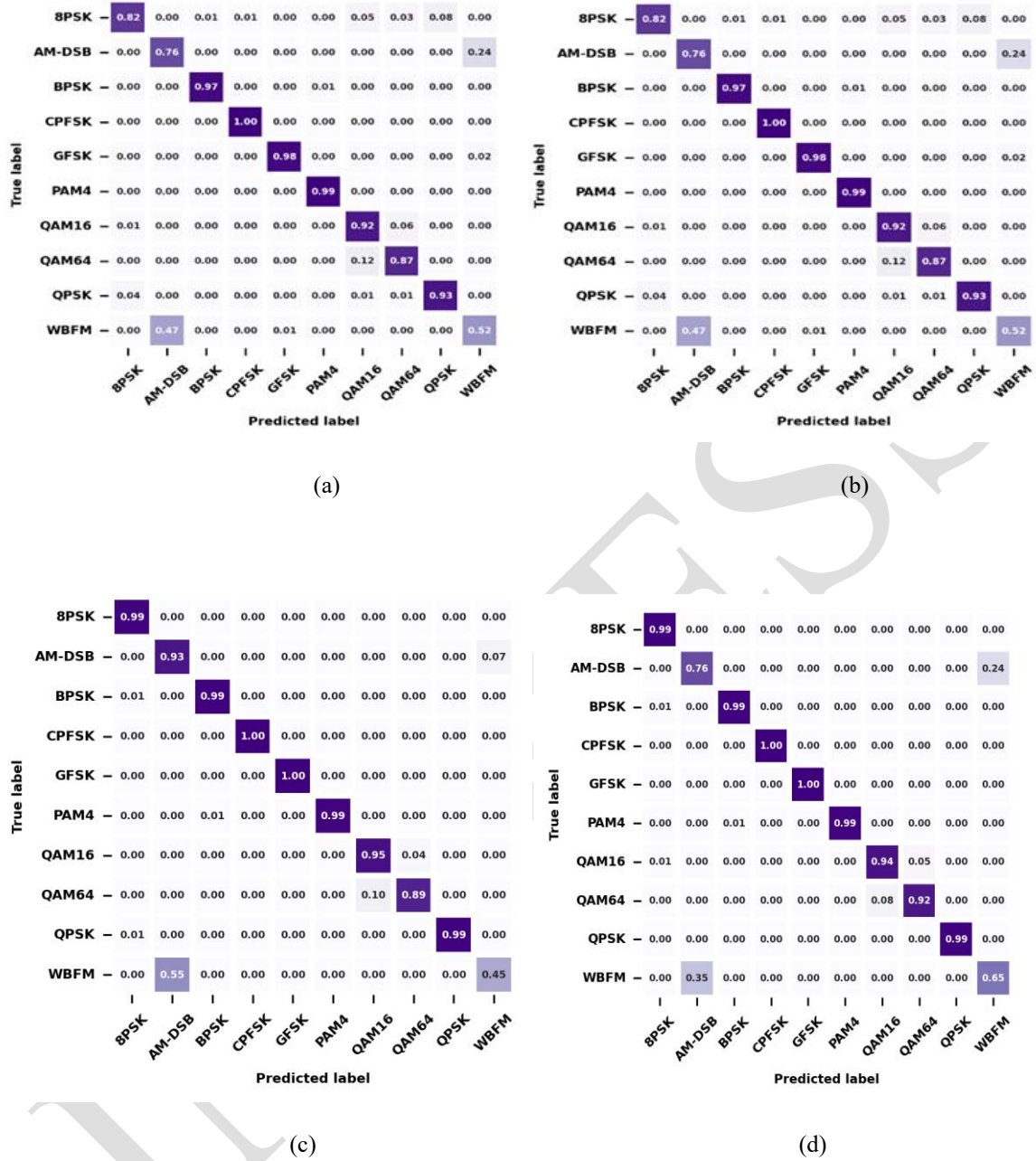


Fig. 6. Confusion matrix (a) at -2 dB, (b) at 0 dB, (c) at 10 dB, (d) at 18 dB

### 5.2.3. Modulation Accuracy Analysis Across SNR Levels

To better understand the model's classification performance, we separately assessed the accuracy of each type of modulation over a broad range of SNR values, from -20 dB to 20 dB. This per-class analysis makes it possible to comprehend each modulation scheme's response to various noise situations in greater depth. According to the results, every type of modulation has a distinct accuracy curve that reflects the properties of the signal and the model's capacity to distinguish between them at different SNR levels. Early convergence to high accuracy is demonstrated by modulations like BPSK, QPSK, and GFSK, although other modulations progressively get better as SNR rises. These curves offer a straightforward and understandable representation of the model's robustness and sensitivity to each modulation technique separately. This method of analysis facilitates the assessment of the system's performance both overall and class-wise, which is useful for real-world applications where some modulation patterns could be more important or common than others. Accuracy of per-class classification for various modulation types at varied SNR levels (ranging from -20 dB to +20 dB) is shown in Fig.7. The suggested

hybrid Bi-LSTM and CNN model's ability to differentiate between each modulation class separately is shown in this picture. It is clear that higher SNR values greatly increase classification accuracy.

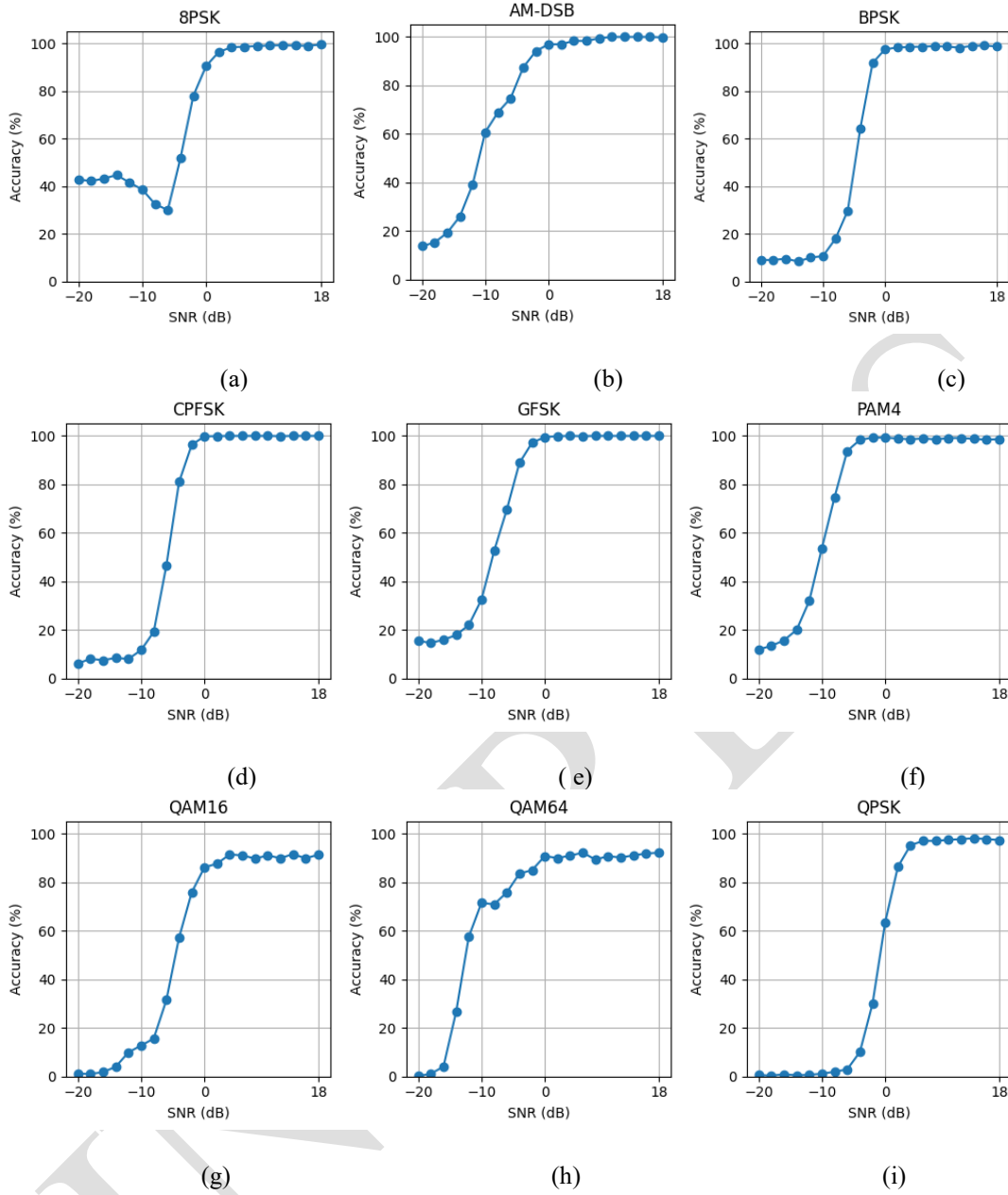


Fig. 7. Classification accuracy for different modulation types over varying SNR (a) 8PSK (b) AM-DSB (c) BPSK (d) CPFSK (e) GFSK (f) PAM4 (g) QAM16 (h) QAM64 (i) QPSK

### 5.3. Comparison with Baseline Models

The performance of the suggested hybrid deep learning model must be compared to a number of well-known baseline models in order to show how successful it is. Four competitive and modern hybrid architectures—CLDNN, GRU-CNN, ResNet-LSTM, and VTCNN—have been chosen for comparison in this work. These models are frequently used in related research and have demonstrated encouraging outcomes in the field of automated modulation classification (AMC). To guarantee a fair and consistent comparison, all models were trained and assessed using the RadioML2016.10b dataset under the same experimental conditions.

#### 5.3.1 Convolutional LSTM dense neural network (CLDNN)

Recently, CLDNN was introduced, utilizing the complementary advantage of CNNs, LSTMs, and DNNs to combine the architectures of LSTM and CNN into a deep neural network.[19] The convolutional layers with pooling in CLDNN designs perform feature extraction and dimensionality reduction, converting the lengthy input sequences into significantly shorter representations of high-level features. To learn long-term temporal coherence of various modulation types, the shorter sequences are sequentially fed into later LSTM layers.[24] Fig.8 illustrates the proposed CLDNN architecture. The input signal first goes through three stacked convolutional layers, which extract local patterns and spatial information that are essential for hierarchically differentiating modulation schemes. An important benefit of using an LSTM layer instead of a pure CNN architecture is that it can capture the dynamic evolution of wireless signals over time by modeling temporal relationships. Finally, the processed spatiotemporal representations are classified via dense and output layers.

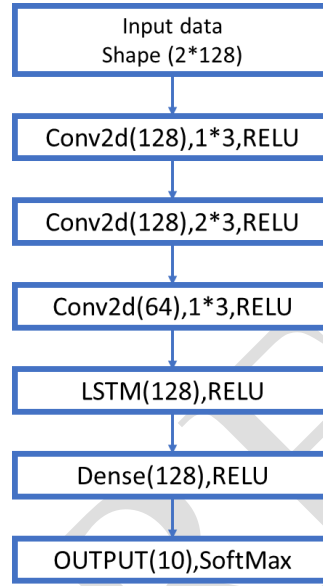


Fig. 8. CLDNN architectures

### 5.3.2 Convolutional Neural Networks and Gated Recurrent Units (CNN-GRU)

A GRU is a more advanced version of an LSTM, consisting of just two gates. It uses a reset gate in place of an output gate and merges the input and forget gates into a single update gate. Additionally, it mixes the hidden and cell states, requiring fewer training parameters[22]. As seen in Fig. 9, the model proposed uses a hierarchical deep learning framework that is tailored for interpreting temporal signals. Three stacked Conv2D layers make up the first part of the architecture, which uses localized filter operations to extract spatially invariant characteristics from the input data. The Time Distributed Flatten layer, which prepares the spatial features for recurrent processing while maintaining the temporal sequence structure, is then applied to these convolutional outputs. The gating methods of subsequent bidirectional GRU layers allow for the selective retention of significant signal properties over time steps, thereby capturing long-range temporal dependencies in the data. Once the features have been analyzed, a Dense layer with softmax activation maps them to the target classes, producing probabilistic modulation classifications. Specifically for time-varying wireless signals under low-SNR conditions, this hybrid design offers better performance than standalone convolutional or recurrent architectures in modulation classification tasks by skillfully fusing the temporal modeling strengths of GRUs with the spatial feature extraction capabilities of CNNs.

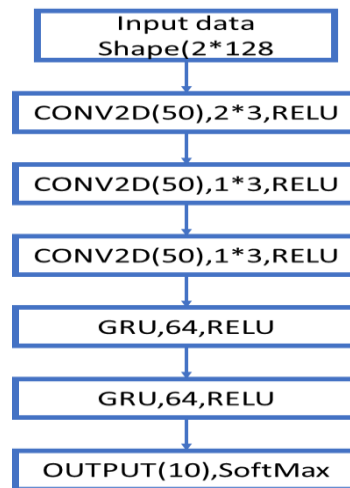


Fig. 9. CNN-GRU

### 5.3.3. Vector Temporal Convolutional Neural Network (VTCNN)

A deep learning architecture called the Vector Temporal Convolutional Neural Network (VTCNN) uses temporal convolutional layers to identify sequential correlations in time-series data and is intended for automatic modulation classification. Unlike recurrent neural networks, VTCNN processes input data efficiently and in parallel by modeling temporal interactions using convolutional operations. This method enhances the model's capacity to differentiate between various modulation patterns by efficiently extracting hierarchical temporal information from multivariate inputs. The VTCNN2 model builds on the original VTCNN and incorporates architectural improvements to improve training efficiency and classification accuracy, such as additional convolutional layers, different kernel sizes to capture features at different temporal scales, and optimized pooling strategies. VTCNN2 relies only on convolutional mechanisms without recurrent layers, which allows it to achieve faster convergence and robustness in complex or noisy signal environments, making it a good choice for real-time communication systems. The first layer of the convolutional neural network architecture shown in Fig. 10 is an input layer made to take in raw signal data in its original format. This is fed into two convolutional layers (Conv Layer 1 and Conv Layer 2) that learn filter banks and gradually extract hierarchical spatial features, each of which captures progressively more intricate patterns. As a link between the classifier components and the spatial feature extractors, the network then uses a flatten layer to convert the two-dimensional feature maps into a one-dimensional vector representation.

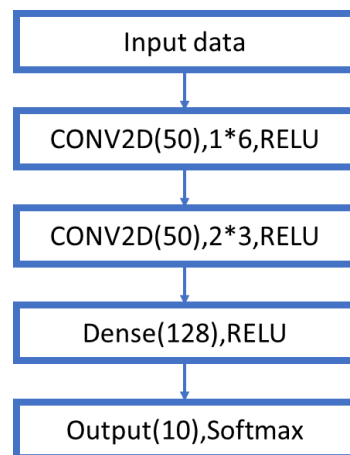


Fig. 10. VTCNN

### 5.3.4 ResNet and LSTM

The approach was established in [25]. To efficiently analyze time-series signal data for modulation classification, the suggested deep learning framework integrates recurrent layers with residual convolutional networks, as shown in Fig. 11. Two residual blocks that allow training of deeper networks through skip connections while maintaining gradient flow come after a CNN block that extracts preliminary features from the input tensor ( $2 \times 128 \times 1$ ). The Time Distributed Flatten layer prepares the spatial information for sequence processing by transforming it while preserving its temporal ordering. After that, the system uses a dual-phase LSTM structure, with the first LSTM layer (64 units) processing the entire sequence and the second LSTM generating the final context vector (64 units). Both short-term and long-term relationships in communication signals are uniquely captured by this hierarchical temporal framework. In the end, the network produces probabilities for each of the 11 modulation classes through a thick layer with softmax activation. In difficult RF situations with time-varying signal characteristics, this hybrid architecture is especially well-suited since it combines the complementary advantages of residual CNNs for reliable spatial feature extraction and LSTMs for complex temporal modeling.

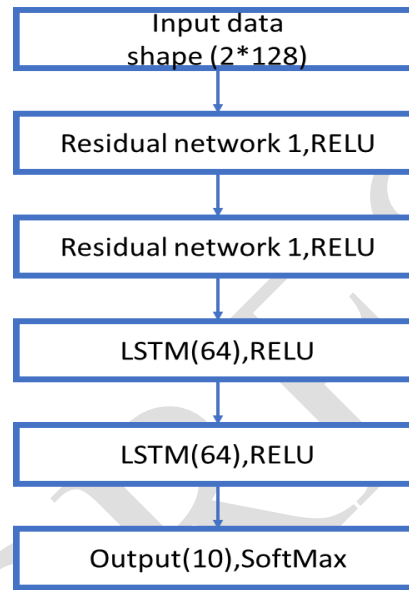


Fig. 11. Resnet-LSTM

## 5.4 PERFORMANCE ANALYSIS

The performance of the suggested hybrid deep learning model is thoroughly analyzed in this part in comparison to the chosen baseline models to fully evaluate its efficacy. The assessment centers on how well the model can identify different types of modulation in a range of channel conditions. By examining the outcomes over a broad range of signal-to-noise ratio (SNR) values, we hope to show how reliable and broadly applicable the suggested design is in real-world wireless communication situations.

### 5.4.1 Classification Accuracy

Fig. 12 presents a comparison of classification accuracy across a range of SNR values for five models: the proposed Bi-LSTM-CNN and four baseline hybrid models—CLDNN, CNN-GRU, ResNet-LSTM, and VTCNN2. According to the Fig, the suggested approach continuously performs better than the baseline models, particularly in low and medium SNR scenarios (such as between -6 and 8 dB), where noise-induced signal distortion is more pronounced. The accuracy of all models stabilizes at high SNR levels (over 10 dB), however, the suggested model continues to retain the best accuracy, nearing 92.5%, while ResNet-LSTM performs at about 92% and the other models at about 86%. By zooming in on the high-SNR zone (0–20 dB), the inset subplot makes it easier to see the small variations in model performance in low-noise settings.

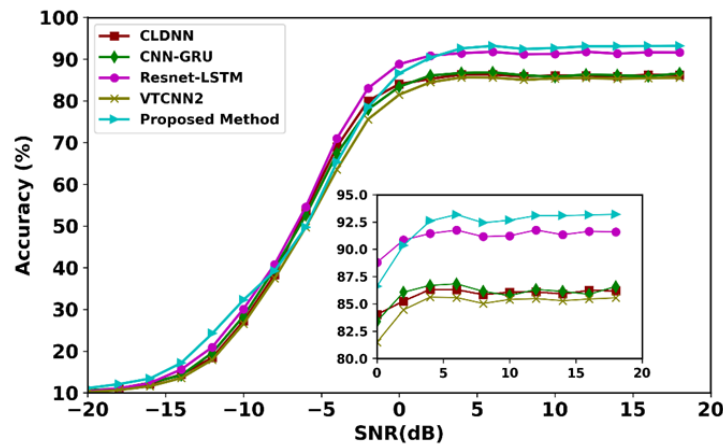


Fig.12. Classification accuracy for the suggested model and baseline models

TABLE 2. COMPARISON OF THE SUGGESTED MODEL'S CLASSIFICATION ACCURACY WITH BASELINE MODELS IN RELATION TO SNR

| SNR( dB) | CNN-GRU % | VTCNN2 % | CLDNN % | Resnet-LSTM % | Proposed model % |
|----------|-----------|----------|---------|---------------|------------------|
| -2       | 77.9      | 75.5     | 82      | 85.4          | 87.5             |
| 0        | 83.4      | 81.4     | 87      | 90.1          | 90               |
| 10       | 85.7      | 85.4     | 89      | 91.2          | 91.6             |
| 18       | 86.5      | 85.4     | 90      | 92            | 92.3             |

#### 5.4.2. Model complexity

Model complexity is a crucial indicator for assessing whether deep learning models can be implemented in practical applications, particularly on devices with limited resources. The total number of parameters for each model under evaluation is shown in Table 2

The ResNet-LSTM model contains 770,378 parameters, whereas the VTCNN2 model has over 5.7 million, making it the most complex model. The high number of parameters in these models is largely due to the use of deep convolutional or residual structures. With an architecture size of roughly 624,514 parameters each, the CNN-GRU and CLDNN models exhibit a moderate degree of complexity as a result of the integration of recurrent and convolutional layers. The suggested BiLSTM-CNN model, on the other hand, has the fewest number of parameters (410,762), indicating its portability and potential for effective deployment in low-power or real-time systems. These findings highlight how well the suggested model balances classification performance and architectural simplicity, making it a solid contender for real-world implementation.

TABLE 3. MODEL COMPLEXITY AND PERFORMANCE COMPARISON

| Models                     | CNN-GRU   | VTCNN2    | CLDNN   | Resnet-LSTM | Proposed model |
|----------------------------|-----------|-----------|---------|-------------|----------------|
| Total parameters           | 2,014,090 | 5,698,714 | 624,514 | 770,378     | 410,762        |
| Epoch                      | 15        | 20        | 20      | 20          | 30             |
| Training time (sec) /epoch | 380       | 200       | 235     | 517         | 145            |
| Predication time           |           |           |         |             |                |
| Time( $\mu$ s) /sample     | 220       | 183       | 175     | 180         | 173            |

An important consideration when assessing the effectiveness of deep learning models is training time, particularly when deploying them in real-time or resource-constrained settings. The total amount of time needed for each model's training under identical hardware and condition settings is shown in Table 2. According to the results, the suggested BiLSTM-CNN model successfully strikes a balance between training efficiency and accuracy. Even though models like ResNet-LSTM and VTCNN2 have comparatively good accuracy, their training periods were noticeably higher. The suggested model, on the other hand, showed improved classification performance with quicker convergence and less training time.

## 6. CONCLUSION

The drawbacks of traditional approaches are addressed by this study's effective hybrid deep learning model for AMC, which combines CNN and Bi-LSTM architectures. The model is very good at extracting both temporal and spatial information from raw I/Q signals. It is robust and has high accuracy (above 90% over 6 dB SNR), especially under difficult low-SNR conditions. Comparative evaluation against cutting-edge baselines validates its better scalability, performance, and computational efficiency. Hardware optimization for edge-device deployment and adaptive architectures for new modulation methods may be the subjects of future research. Providing a viable alternative for next-generation wireless systems, the suggested framework closes the gap between theoretical developments in AMC and useful real-world applications.

## References

- [1] 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," *Appl. Sci.*, vol. 12, no. 23, 2022, doi: 10.3390/app122312052.
- [2] O. A. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, "Survey of automatic modulation classification techniques: Classical approaches and new trends," *IET Commun.*, vol. 1, no. 2, pp. 137–156, 2007, doi: 10.1049/iet-com:20050176.
- [3] J. L. Xu, W. Su, and M. Zhou, "Likelihood-ratio approaches to automatic modulation classification," *IEEE Trans. Syst. Man Cybern., Part C Appl. Rev.*, vol. 41, no. 4, pp. 455–469, 2011, doi: 10.1109/TSMCC.2010.2076347.
- [4] M. A. Abdel-Moneim, W. El-Shafai, N. Abdel-Salam, E. S. M. El-Rabaie, and F. E. A. El-Samie, "A survey of traditional and advanced automatic modulation classification techniques, challenges, and some novel trends," *Int. J. Commun. Syst.*, vol. 34, no. 10, pp. 1–36, 2021, doi: 10.1002/dac.4762.
- [5] A. Hazza, M. Shoaib, S. A. Alshebeili, and A. Fahad, "An overview of feature-based methods for digital modulation classification," in *Proc. 1st Int. Conf. Commun. Signal Process. Their Appl. (ICCSA)*, 2013, pp. 1–8, doi: 10.1109/ICCSA.2013.6487244.
- [6] Y. Wang, "Signal classification based on spectral correlation analysis and SVM," in *Proc. Int. Conf. Adv. Inf. Netw. Appl.*, 2008, pp. 883–887, doi: 10.1109/AINA.2008.27.
- [7] M. V. Subbarao and P. Samundiswary, "Automatic modulation recognition in cognitive radio receivers using multi-order cumulants and decision trees," *J. Commun. Eng.*, no. Nov. 2020.
- [8] M. L. D. Wong and S. K. Ting, "Naïve Bayes classification of adaptive broadband wireless modulation schemes with higher-order cumulants," in *Proc. Int. Conf. Mach. Learn.*, 2008, pp. 1–4.
- [9] W. Shen, "Automatic digital modulation recognition based on locality preserved projection," in *Proc. Int. Conf. Wireless Commun. Signal Netw.*, 2014, pp. 348–352, doi: 10.1109/WCSN.2014.78.
- [10] B. Kim, J. Kim, H. Chae, D. Yoon, and J. W. Choi, "Deep neural network-based automatic modulation classification technique," in *Proc. Int. Conf. Inf. Commun. Technol.*, 2016, pp. 579–582.
- [11] W. Yongshi, G. Jie, L. Hao, L. Li, and W. Houjun, "CNN-based modulation classification in the complicated communication channel," in *Proc. Int. Conf. Wireless Commun.*, 2017, pp. 1–5.
- [12] Z. Wang, K. Gong, and W. Wang, "A lightweight CNN architecture for automatic modulation classification," *IEEE Access*, 2021.
- [13] X. Tian and C. Chen, "Modulation pattern recognition based on ResNet-50 neural network," in *Proc. 2nd IEEE Int. Conf. Inf. Commun. Signal Process. (ICICSP)*, 2019, pp. 34–38, doi: 10.1109/ICICSP48821.2019.8958555.
- [14] O. F. Abd-Elaziz, M. Abdalla, and R. A. Elsayed, "Deep learning-based automatic modulation classification using robust CNN architecture for cognitive radio networks," *J. Commun. Netw.*, pp. 1–19, 2023.
- [15] R. Li, L. Li, S. Yang, and S. Li, "Robust automated VHF modulation recognition based on deep convolutional neural networks," *IEEE Commun. Lett.*, vol. 22, no. 5, pp. 946–949, 2018, doi: 10.1109/LCOMM.2018.2809732.
- [16] P. J. Werbos, "Backpropagation through time: What it does and how to do it," *Proc. IEEE*, vol. 78, no. 10, pp. 1550–1560, 1990, doi: 10.1109/5.58337.
- [17] A. Sherstinsky, "Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network," *Phys. D Nonlinear Phenom.*, vol. 404, p. 132306, 2020, doi: 10.1016/j.physd.2019.132306.
- [18] D. Hong, Z. Zhang, and X. Xu, "Automatic modulation classification using recurrent neural networks," in *Proc. 3rd IEEE Int. Conf. Comput. Commun. (ICCC)*, 2017, pp. 695–700, doi: 10.1109/CompComm.2017.8322633.
- [19] N. E. West and T. O'Shea, "Deep architectures for modulation recognition," in *Proc. IEEE Int. Symp. Dyn. Spectr. Access Networks (DySPAN)*, 2017, doi: 10.1109/DySPAN.2017.7920754.
- [20] S. Rajendran, W. Meert, D. Giustiniano, V. Lenders, and S. Pollin, "Deep learning models for wireless signal classification with distributed low-cost spectrum sensors," *IEEE Trans. Cogn. Commun. Netw.*, vol. 4, no. 3, pp. 433–445, 2018.

- [21] X. Liu, D. Yang, and A. El Gamal, "Deep neural network architectures for modulation classification," in Proc. 51st Asilomar Conf. Signals, Syst. Comput. (ACSSC), 2017, pp. 915–919, doi: 10.1109/ACSSC.2017.8335483.
- [22] X. Hao, Y. Luo, Q. Ye, Q. He, G. Yang, and C.-C. Chen, "Automatic modulation recognition method based on a hybrid model of convolutional neural networks and gated recurrent units," *Sensors Mater.*, vol. 33, 2021.
- [23] T. J. O'Shea, J. Corgan, and T. C. Clancy, "Convolutional radio modulation recognition networks," *Commun. Comput. Inf. Sci.*, vol. 629, pp. 213–226, 2016, doi: 10.1007/978-3-319-44188-7\_16.
- [24] A. Emam, M. Shalaby, M. A. Aboelazm, H. E. A. Bakr, and H. A. A. Mansour, "A comparative study between CNN, LSTM, and CLDNN models in the context of radio modulation classification," in Proc. 12th Int. Conf. Electr. Eng. (ICEENG), 2020, pp. 190–195, doi: 10.1109/ICEENG45378.2020.9171706.
- [25] M. M. Elsayagheer and S. M. Ramzy, "A hybrid model for automatic modulation classification based on residual neural networks and long short-term memory," *Alex. Eng. J.*, vol. 67, pp. 117–128, 2023, doi: 10.1016/j.aej.2022.08.019