TFI-AMR: A Transformer-based Forest-Inspired Approach for Automatic Modulation Recognition in Wireless Communication

Robertas Damasevicius
Department of Applied Informatics
Vytautas Magnus University
Akademija, Lithuania
robertas.damasevicius@vdu.lt

Darius Plonis
Department of Electronic Systems
Vilnius Gediminas Technical University
Vilnius, Lithuania
darius.plonis@vilniustech.lt

Rytis Maskeliunas
Faculty of Applied Mathematics
Silesian University of Technology
Gliwice, Poland
rytis.maskeliunas@polsl.pl

***Abstract*—Automatic Modulation Recognition (AMR) allows efficient enhancement of wireless communication, especially when prior information about the modulation scheme is absent. With the advent of deep learning, AMR has seen significant advancements, particularly with the introduction of the Transformer architecture, known for its self-attention mechanism. Inspired by the adaptive and interconnected nature of forests, we introduce the Transformer-based Forest-Inspired Approach for AMR (TFI-AMR). This study aims to harness the self-attention mechanism of the Transformer model, integrating it with concepts derived from the forest ecosystem to develop a robust and adaptive AMR system. The TFI-AMR model is structured hierarchically, mirroring the layers of a forest – from the root to the leaves. Each layer focuses on different granularity levels of signal features, from global to highly specific. The adaptive growth mechanism, inspired by the growth rings of trees, represents the learning epochs of the model. Pruning, akin to trees shedding branches, optimizes the model by removing non-essential attention heads or layers. The multi-head attention mechanism is visualized as different trees or species in a forest, each focusing on diverse aspects of the signal. Positional encoding introduces adaptability, reminiscent of seasonal changes in a forest. A unique "forest fire" mechanism occasionally resets parts of the model, ensuring variability and preventing overfitting. The experimental evaluation of the TFI-AMR model demonstrated high adaptability to new modulation schemes, robust performance due to the multi-head attention and "forest fire" mechanisms, and enhanced explainability through its hierarchical structure and attention maps.**

***Index Terms*—Transformer, Automatic Modulation Recognition, Forest-Inspired Computing, Self-Attention, Multi-Head Attention, Wireless Communication.**

I. Introduction

Automatic Modulation Recognition (AMR) has become an indispensable tool in the realm of modern communication systems. As the demand for efficient and intelligent communication grows, the ability to accurately and swiftly recognize modulation schemes becomes paramount.

A. Background on Automatic Modulation Recognition

AMR refers to the process of identifying the modulation type of a received signal without prior knowledge of the modulation scheme [1]. This capability helps in non-cooperative communication environments, where receivers need to adapt to varying signal conditions and types [2]. The origins of AMR can be traced back to military applications, where it was used for signal intelligence and electronic warfare. Recognizing the modulation scheme of an intercepted signal is often the first step towards demodulation and further analysis.

With the proliferation of wireless communication technologies [3], the spectrum environment has become increasingly crowded and diverse. Multiple communication devices and systems coexist, each employing different modulation schemes for optimal performance. In such a scenario, AMR plays a pivotal role in tasks such as spectrum sensing, interference detection, and cognitive radio operations.

Traditional AMR techniques relied on feature-based methods, where specific characteristics of the received signal, such as its moments, bandwidth, and cyclostationary features [4], were extracted and used for classification [5]. For classification, machine learning [6], [7], [8] and evolutionary computing [9], [10] methods have been used. However, with the advent of deep learning and advanced signal processing techniques, the landscape of AMR has transformed [11], [12], [13]. Modern AMR systems leverage sophisticated algorithms and neural network architectures to achieve high accuracy even in challenging conditions, such as low signal-to-noise ratios [14], [15], [16].

B. The Transformer Architecture - natural for AMR?

The Transformer architecture, introduced by Vaswani et al. [17], has revolutionized the field of deep learning, especially in natural language processing tasks. Its unique design, which eschews the recurrent layers commonly found in previous state-of-the-art models, offers both efficiency and performance benefits, making it the backbone of many subsequent models and applications.

At the heart of the Transformer is the attention mechanism, specifically the multi-head self-attention mechanism. This allows the model to weigh the importance of different parts of an input sequence when producing an output, enabling it to capture long-range dependencies and intricate patterns in the data.

In the context of Automatic Modulation Recognition (AMR), this means that the Transformer can consider the entire input signal simultaneously, potentially capturing relationships and features that might be missed by architectures that process the signal in a more sequential manner.

The Transformer architecture is inherently parallelizable, which means that it can process all elements of a sequence simultaneously [18]. This is in contrast to RNNs and LSTMs, which process sequences element by element. This parallel nature not only speeds up training but also allows the model to capture complex relationships in the data more effectively.

Another notable feature of the Transformer is its positional encoding. Since the architecture doesn't have any inherent notion of the order of the sequence (unlike RNNs or LSTMs), positional encodings are added to ensure that the model can account for the position of elements in a sequence. This is crucial in tasks like AMR, where the position of a signal in a sequence can carry significant information [19], [20].

In the realm of AMR, the Transformer's ability to handle sequences and its powerful attention mechanism make it a promising candidate for recognizing modulation schemes [21], [22], [23]. By treating the modulated signal as a sequence, the Transformer can learn to recognize patterns and dependencies that are indicative of specific modulation techniques. Moreover, recent works have adapted the Transformer architecture for tasks beyond its original domain of natural language processing, showcasing its versatility.

However, while the Transformer offers many advantages, it also comes with challenges [24]. The architecture's large number of parameters can make it computationally intensive, potentially limiting its deployment in resource-constrained environments. Furthermore, like many deep learning models, the Transformer can act as a "black box," making its decisions hard to interpret.

C. Inspiration from Forest Ecosystems

Forest ecosystems, with their intricate interplay of flora, fauna, and environmental factors, have long been a source of inspiration for various scientific disciplines [25]. These ecosystems are characterized by their diversity, resilience, and adaptability, qualities that researchers and engineers often seek to emulate in their designs and models. One of the most striking aspects of a forest ecosystem is its hierarchical structure. From the towering canopy trees to the understory shrubs, down to the forest floor with its rich layer of decomposing organic matter, each layer has its unique role and function. This hierarchical structure can be mirrored in computational models, where different layers or modules handle varying levels of abstraction or complexity.

The concept of diversity in forest ecosystems, where a multitude of species coexist and often benefit from one another [26], can be translated into ensemble methods in machine learning. Just as a diverse forest is more resilient to pests or diseases, an ensemble of diverse models can offer better performance and robustness than a single model. The idea is that different models capture different aspects or features of the data, and their combined predictions can be more accurate and reliable. Another inspiration drawn from forests is the idea of growth and adaptation. Trees and plants in a forest ecosystem continuously adapt to their environment, competing for resources like sunlight, water, and nutrients. In a similar vein, algorithms can be designed to "grow" and adapt over time, refining their parameters and structures based on feedback and new data. This is reminiscent of adaptive algorithms or those that employ some form of online learning. Moreover, the symbiotic relationships observed in forest ecosystems, such as the mutualism between fungi and tree roots, can inspire collaborative algorithms where multiple agents or models work together, sharing resources or information to achieve a common goal.

Lastly, the concept of resilience in forest ecosystems, where the system can recover from disturbances like fires or logging, can be a guiding principle in designing fault-tolerant and robust systems. By building in redundancies, and ensuring that there's no single point of failure, systems can be designed to continue functioning even when parts of them fail.

D. Novelty and Contribution

The main novelties and contributions of this work are outlined as follows:

1) Unlike traditional deep learning models employed in AMR, this research introduces a hybrid architecture that synergistically combines the strengths of convolutional layers, recurrent units, and attention mechanisms. Such fusion results in enhanced feature extraction and improved recognition accuracy.
2) Drawing inspiration from the transformer architecture, which has revolutionized the field of natural language processing, this work adapts and integrates its self-attention mechanisms into the AMR domain, leading to significant performance gains.
3) By mimicking the hierarchical structure and adaptive nature of forest ecosystems, a novel optimization strategy is proposed. Our idea ensures model robustness and adaptability, reminiscent of the resilience exhibited by natural forests.
4) Beyond theoretical advancements, this work emphasizes the practical deployment of the proposed model. Considerations regarding computational efficiency, scalability, and real-time processing are addressed, making the model suitable for real-world AMR tasks.

II. Overview of Related Appraoches

Deep Learning (DL), a subset of machine learning, has revolutionized numerous domains, from computer vision to natural language processing. Its application in Automatic Modulation Recognition (AMR) has also opened a new era of enhanced capabilities, outperforming traditional methods in terms of accuracy and adaptability [27], [28].

The essence of DL lies in its ability to automatically learn representations from data, eliminating the need for manual feature extraction, which is a significant step in improving the traditional AMR techniques [29]. Traditional methods often relied on handcrafted features, such as higher-order cumulants, cyclostationary features, or wavelet transforms, which while sometimes effective, their performance is often contingent on the specific conditions of the communication environment, such as noise levels and interference. Moreover, designing and selecting these features requires expert knowledge and can be time-consuming to validate. DL, on the other hand, leverages architectures like Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), to automatically learn and extract features from raw or minimally processed signals. This ability to learn directly from data means that DL models can adapt to a wide range of conditions and modulation schemes, often with higher accuracy than traditional methods.

Several studies have showcased the prowess of DL in AMR [30], [31], [32]. For instance, CNNs, with their spatial hierarchies, have proven adept at capturing the intricate patterns in modulated signals. RNNs, especially their variants like Long Short-Term Memory (LSTM) networks, capture the temporal dependencies in signals, making them suitable for AMR tasks where time-domain information is crucial [16], [33], [34]. Furthermore, architectures like the Transformer, originally designed for natural language processing tasks, have been adapted for AMR, highlighting the flexibility and potential of DL in this domain.

A significant portion of the research has been dedicated to the integration of convolutional neural networks (CNN) and recurrent neural networks (RNN), such as Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU), for feature extraction. For instance, Xu et al. (2020) proposed a three-stream deep learning framework that integrates 1D and 2D convolutional layers with LSTM to extract features from modulated data, achieving improved recognition accuracy for higher dimensional schemes like 16-QAM and 64-QAM [35]. Similarly, Hou et al. (2023b) introduced an improved spatiotemporal multi-channel network, IQGMCL, which divides input IQ signals into three channels and uses convolutional kernels and Bi-LSTM for feature extraction [36]. Liu et al. (2023) also combined two deep learning networks, MCLDNN and BiGRU3, to achieve higher accuracy AMR [37].

Attention mechanisms have also been explored to enhance AMR. Liang et al. (2021) proposed a novel framework that incorporates adaptive attention mechanism modules with the ResNeXt network, demonstrating higher robustness and accuracy compared to other techniques [38]. Kong et al. (2023) introduced a Transformer-based contrastive semi-supervised learning framework for AMR, which involves convolutional embedding, attention bias, and attention pooling [22].

Obviously, the integration of DL into AMR is not without challenges [1]. Training deep neural networks requires large labeled datasets, which might not always be readily available for all modulation schemes. Moreover, while DL models can generalize well, they are often perceived as "black boxes," making their decisions hard to interpret. This lack of transparency can be a concern in critical applications where understanding the decision-making process is crucial.

The challenge of model complexity versus accuracy has been a recurring theme in AMR research. While deep learning models have shown promising results in AMR, their computational complexity often hinders their deployment in resource-constrained environments. Zhang et al. (2022) highlighted this challenge, emphasizing the need for models that balance accuracy with computational efficiency [39]. In response, several studies have proposed lightweight models. For instance, Zhang et al. (2021) introduced an efficient DL-AMR model based on phase parameter estimation and transformation, which reduces the model's parameters while maintaining high recognition accuracy [40]. Yi et al. (2023) proposed the Lightweight Densely Connected Convolutional Network (DenseNet) Long Short-Term Memory network (LDLSTM), which achieves comparable recognition accuracy with only 1/12th of the parameters of other advanced DL-AMR algorithms [34].

III. The TFI-AMR Model

A. Hierarchical Self-Attention

The hierarchical self-attention mechanism in the TFI-AMR model is inspired by the layered structure of a forest. Each layer, from the root to the leaves, represents a different granularity level of signal features. The self-attention mechanism, derived from the Transformer architecture, allows the model to focus on different parts of the signal based on their importance.

1) Root Layer: Raw Signal Input: The root layer is the foundation of the model, analogous to the forest floor. At this layer, the raw signal $s(t)$ is ingested.

$$s(t) = \sum_i A_i \cos(2\pi f_i t + \phi_i), \tag{1}$$

where A_i is the amplitude of the i^{th} component; f_i is the frequency of the i^{th} component; ϕ_i is the phase of the i^{th} component.

The self-attention mechanism at this layer focuses on global features, giving the model a broad understanding of the signal's characteristics.

2) Trunk Layer: Intermediate Feature Extraction: The trunk layer supports and directs the flow of information, similar to a tree trunk. Here, the signal undergoes initial feature extraction using the self-attention mechanism:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right) V, \tag{2}$$

where Q, K, and V are the query, key, and value matrices, respectively; d_k is the dimension of the key.

This layer refines the global features from the root layer, focusing on more localized features and patterns in the signal.

3) Branch Layer: Detailed Feature Extraction: The branch layer captures more specific details of the signal. The self-attention mechanism here refines its focus on crucial signal components, allowing for a deeper understanding of the modulation scheme:

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \ldots, \text{head}_h) W_O, \tag{3}$$

where Each $\text{head}_i = \text{Attention}(QW_{Qi}, KW_{Ki}, VW_{Vi})$; W_{Qi}, W_{Ki}, and W_{Vi} are weight matrices; W_O is the output weight matrix; h is the number of heads.

This multi-head attention mechanism allows the model to focus on multiple aspects of the signal simultaneously, capturing a richer set of features.

4) Leaf Layer: Decision Making: The leaf layer is the final decision-making point. At this stage, the features extracted from the previous layers are used to classify the modulation scheme. A feed-forward neural network is employed:

$$F(x) = \text{ReLU}(xW_1 + b_1)W_2 + b_2, \tag{4}$$

where x is the input from the branch layer; W_1, W_2 are weight matrices; b_1, b_2 are bias terms.

The output of this layer provides the final classification of the modulation scheme, completing the hierarchical self-attention process.

B. Adaptive Growth Mechanism

The adaptive growth mechanism in the TFI-AMR model is inspired by the growth and adaptability of trees in a forest. This mechanism ensures that the model evolves and refines its understanding of modulation schemes over time, optimizing its structure for better performance.

1) Growth Rings and Learning Epochs: In trees, growth rings represent the passage of time, with each ring indicating a year of growth. Analogously, in the TFI-AMR model, each learning epoch refines the model's parameters, allowing it to better recognize modulation schemes.

Given a loss function $L(\theta)$, where θ represents the model's parameters, the update rule after each epoch using gradient descent is:

$$\theta_{t+1} = \theta_t - \eta \nabla L(\theta_t), \tag{5}$$

where θ_t is the parameter vector at epoch t; η is the learning rate; $\nabla L(\theta_t)$ is the gradient of the loss function with respect to the parameters at epoch t.

As the number of epochs increases, the model's performance on the training data typically improves, analogous to how a tree grows stronger and taller over the years.

2) Pruning and Optimization: Just as trees shed non-essential branches to optimize their growth and resource allocation, the TFI-AMR model can remove neurons or layers that don't contribute significantly to its performance. This pruning process reduces the model's complexity, making it more efficient without sacrificing accuracy.

Given a weight matrix W of a particular layer, the pruning process can be defined as:

$$W' = \begin{cases} W_{ij} & \text{if } |W_{ij}| > \tau \\ 0 & \text{otherwise} \end{cases}, \tag{6}$$

where W' is the pruned weight matrix; τ is a threshold value. Weights with absolute values smaller than τ are set to zero, effectively pruning them.

After pruning, the model is fine-tuned on the training data to adjust the remaining weights and compensate for the removed neurons or connections.

This pruning process not only reduces the computational requirements of the model but also can help in preventing overfitting, ensuring that the model generalizes well to unseen data.

C. Symbiotic Relationships in Multi-Head Attention

The multi-head attention mechanism in the TFI-AMR model is inspired by the symbiotic relationships observed in forest ecosystems, where different entities coexist and benefit from one another. In the model, multiple attention heads work in tandem, each focusing on different aspects of the signal, and collectively contribute to a more robust and comprehensive understanding of the modulation schemes.

1) Diverse Focus through Attention Heads: In a forest, different species focus on various resources for survival, leading to a harmonious balance. Similarly, the multi-head attention mechanism allows the model to focus on multiple facets of the input signal simultaneously. Each attention head captures a unique set of features, ensuring that no critical information is overlooked.

The multi-head attention is defined as:

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \ldots, \text{head}_h)W_O, \tag{7}$$

where Each $\text{head}_i = \text{Attention}(QW_{Qi}, KW_{Ki}, VW_{Vi})$. - W_{Qi}, W_{Ki}, and W_{Vi} are weight matrices associated with the i^{th} attention head for the query, key, and value, respectively. - W_O is the output weight matrix. - h is the number of attention heads.

By allowing multiple heads to focus on different parts of the signal, the model ensures a diverse and comprehensive feature extraction process.

2) Knowledge Sharing and Decision Making: Just as trees and plants in a forest share resources and information through their root systems, the multiple attention heads in the TFI-AMR model share their knowledge to make collective decisions. After processing the signal, the outputs of the different heads are concatenated and linearly transformed to produce a unified representation of the signal.

The unified representation R is given by:

$$R = \text{Concat}(\text{head}_1, \ldots, \text{head}_h)W_O. \tag{8}$$

This representation R is then passed through subsequent layers of the model for further processing and final decision-making. By pooling the knowledge from different attention heads, the model ensures a holistic understanding of the modulation scheme, leading to more accurate recognition.

D. Resilience and Adaptability

Drawing inspiration from the resilience and adaptability of forest ecosystems, the TFI-AMR model incorporates mechanisms that allow it to adjust to different signal environments and rejuvenate its structure for optimal performance.

1) Positional Encoding and Seasonal Adaptation: Forests undergo seasonal changes, adapting their behavior based on the time of year. Similarly, in the context of signal processing, the position or sequence of data points in a signal can carry significant information. The Transformer architecture, which is inherently non-sequential, requires a mechanism to account for the position of data points, leading to the introduction of positional encoding.

Positional encoding is added to the input embeddings at the root layer to provide the model with information about the position of each data point in the sequence. For a given position p and dimension i, the positional encoding PE is defined as:

$$PE(p, 2i) = \sin\left(\frac{p}{10000^{\frac{2i}{d}}}\right), \tag{9}$$

$$PE(p, 2i+1) = \cos\left(\frac{p}{10000^{\frac{2i}{d}}}\right), \tag{10}$$

where d is the dimension of the embeddings.

This sinusoidal encoding allows the model to attend to positions based on their relative distances, providing a form of "seasonal adaptation" to different signal environments.

2) The Forest Fire Mechanism: Forests occasionally undergo fires, which, while destructive, play a crucial role in rejuvenating the ecosystem by clearing out deadwood and facilitating new growth. Drawing an analogy, the TFI-AMR model incorporates a "forest fire" mechanism where a portion of its neurons or layers are reset periodically. This introduces variability, prevents overfitting, and can lead to better generalization.

Mathematically, given a weight matrix W of a particular layer, the forest fire mechanism can be represented as:

$$W' = \begin{cases} W_{ij} & \text{with probability } p \\ \text{random\_init}(W_{ij}) & \text{with probability } 1-p \end{cases}, \tag{11}$$

where W' is the updated weight matrix after the forest fire mechanism; p is the retention probability, representing the likelihood that a weight remains unchanged; random_init(W_{ij}) initializes the weight W_{ij} to a random value, effectively "burning" and resetting it.

After the forest fire mechanism is applied, the model undergoes further training to adjust and refine the remaining and reset weights.

E. Bio-signals for Modulation Recognition

Bio-signals, inspired by biological systems, are intricate patterns or sequences that can be observed in nature. In the context of the TFI-AMR model, bio-signals refer to the inherent patterns and characteristics of modulation schemes that can be likened to biological signals. By understanding and leveraging these bio-signals, our model can achieve a more intuitive and efficient recognition of modulation schemes.

1) Feature Extraction from Bio-signals: To harness the power of bio-signals, the first step is to extract relevant features that can be used for modulation recognition. Given a raw signal $s(t)$, the bio-signal feature extraction can be represented as:

$$F(s) = \int_{-\infty}^{\infty} s(t) \cdot b(t)\, dt, \tag{12}$$

where $F(s)$ represents the extracted feature set; $b(t)$ is a bio-inspired function or filter that emphasizes certain characteristics of the signal, analogous to how biological systems might process signals.

2) Bio-inspired Activation Functions: Traditional neural networks use activation functions like ReLU, sigmoid, or tanh. However, for bio-signal processing, we can introduce bio-inspired activation functions that mimic certain biological responses. One such function, inspired by the firing rate of neurons, can be defined as:

$$\text{BioAct}(x) = \frac{1}{1 + e^{-k(x-x_0)}}, \tag{13}$$

where k determines the steepness of the function; x_0 is the point of maximum growth, representing a threshold.

This function ensures that the model's neurons activate in a manner reminiscent of biological neurons, providing a more natural response to bio-signals.

3) Modulation Classification using Bio-signals: With the extracted features and bio-inspired activations, the model can classify modulation schemes. Given the feature set $F(s)$ and weight matrix W, the classification C is given by:

$$C = \text{BioAct}(F(s) \cdot W + b), \tag{14}$$

where b is a bias term.

We believe this classification process leverages the inherent patterns of bio-signals to recognize modulation schemes in a manner that's inspired by biological systems.

F. TFI-AMR Algorithm

The TFI-AMR algorithm essentially combines the power of the Transformer architecture, known for its ability to handle sequential data, with the adaptability and resilience of forest-inspired algorithms. The idea is that by processing the signal data in multiple ways (akin to how multiple trees in a forest process information), the model can capture a richer set of features and thus recognize modulation schemes more accurately.

The components in our approach can be defined as follows:

- Embedding Layer is responsible for converting the raw signal data into a format that's suitable for processing by the subsequent layers. It essentially transforms the input data into a higher-dimensional space where the inherent features of the signal are more pronounced.
- Self Attention mechanism allows the model to focus on different parts of the signal data with varying degrees of attention. It helps the model to recognize patterns within the signal that are crucial for modulation recognition.

Algorithm 1 TFI-AMR: A Transformer-based Forest-Inspired Approach for AMR

```
procedure TFI-AMR(signal_data, num_trees)
    encoded_data ← EmbeddingLayer(signal_data)
    attention_map ← SelfAttention(encoded_data)
    transformed_data ← TransformerLayer(attention_map)
    for i = 1 to num_trees do
        tree_data_i ← ForestLayer(transformed_data)
        output_i ← DenseLayer(tree_data_i)
    end for
    final_output ← AggregateOutputs(output_1, output_2, ..., output_num_trees)
    return final_output
end procedure
```

- Transformer Layer is inspired by the Transformer architecture, this layer processes the attention-weighted signal data. It's capable of capturing long-range dependencies in the data, which is essential for recognizing complex modulation schemes.
- Forest Layer was developed drawing inspiration from forest ecosystems, this layer mimics the hierarchical structure and adaptive nature of forests. For each tree in our "forest", the layer processes the data from the Transformer layer in a unique way, capturing different aspects of the signal.
- Dense Layer is a fully connected neural network layer. For each tree's data, this layer produces an output that predicts the modulation scheme of the input signal.
- The outputs from all the trees are aggregated to produce a single prediction. This could be done using methods like averaging, weighted sum, or even another neural network layer.

IV. Implementation and Methodology

This section explains the methodology adopted to construct the hierarchical Transformer structure, train the model using multi-head attention, and incorporate adaptive learning mechanisms.

A. Building the Hierarchical Transformer Structure

The hierarchical structure of the TFI-AMR model is designed to mimic the layers of a forest, from the root to the leaves. Each layer in this hierarchy serves a distinct purpose in processing the input signal.

1) **Input Embedding:** The raw signal $s(t)$ is first converted into a set of embeddings using a learned embedding matrix E:
$$E_s = s(t) \times E. \tag{15}$$
2) **Positional Encoding:** To account for the sequence of data points in the signal, positional encodings are added to the embeddings:
$$E'_s = E_s + \text{PositionalEncoding}(E_s). \tag{16}$$
3) **Layer-wise Processing:** The signal, now in the form of embeddings with positional encodings, is passed through the hierarchical layers (Root, Trunk, Branch, Leaf). Each layer consists of multi-head attention mechanisms and feed-forward neural networks, with residual connections and normalization:
$$L_{i+1} = \text{Normalize}(L_i + \text{FeedForward}(\text{MultiHeadAttention}(L_i))), \tag{17}$$
where L_i is the output of the i^{th} layer.

B. Training with Multi-Head Attention

The multi-head attention mechanism allows the model to focus on various aspects of the signal simultaneously, ensuring a comprehensive feature extraction process.

1) **Attention Computation:** For each head, the attention scores are computed using the query Q, key K, and value V matrices:
$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V. \tag{18}$$
2) **Concatenation and Transformation:** The outputs of all heads are concatenated and linearly transformed:
$$\text{MultiHeadOutput} = \text{Concat}(\text{head}_1, \ldots, \text{head}_h)W_O. \tag{19}$$
3) **Loss Computation:** A suitable loss function, such as the cross-entropy loss for classification tasks, is used to measure the difference between the model's predictions and the actual modulation schemes.
4) **Backpropagation:** The gradients of the loss with respect to the model's parameters are computed and used to update the weights.

C. Adaptive Learning Mechanisms

To ensure that the TFI-AMR model remains efficient and avoids overfitting, adaptive learning mechanisms are incorporated.

1) **Learning Rate Scheduling:** Instead of using a constant learning rate, a scheduler adjusts the rate based on the training epoch. Common strategies include step decay, exponential decay, or cosine annealing.
2) **Regularization:** Techniques like dropout are applied to prevent overfitting. During training, a fraction of neurons is randomly "dropped out" or deactivated:
$$\text{Dropout}(x) = x \odot \text{Bernoulli}(p), \tag{20}$$
where p is the dropout probability, and $\odot$ denotes element-wise multiplication.
3) **Early Stopping:** Monitoring the model's performance on a validation set, training is halted if the performance plateaus or deteriorates for a specified number of epochs. This prevents overfitting and reduces training time.

V. Results and Evaluation

The TFI-AMR model, after rigorous training and optimization, was subjected to a comprehensive evaluation to ascertain its performance, adaptability, and robustness. This section delves into the metrics used for evaluation, the model's ability to adapt to new modulation schemes, and its robustness under various conditions.

A. Dataset

Dataset
https://github.com/Richardzhangxx/AMR-Benchmark
https://github.com/Richardzhangxx/
AMR-Dataset-for-MIMO-system-with-precoding

B. Performance Metrics

To comprehensively evaluate the TFI-AMR model's performance, we consider multiple metrics, including accuracy, precision, recall, and F1-score. Additionally, a confusion matrix provides insights into the model's classification capabilities across different modulation schemes:

1) **Accuracy:** The ratio of correctly predicted modulation schemes to the total number of predictions. Mathematically, it's given by:

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}}.$$

2) **Precision, Recall, and F1-Score:** These metrics provide a more granular understanding of the model's performance, especially in cases where the classes are imbalanced.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}.$$

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}.$$

$$F1\text{-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}.$$

3) **Confusion Matrix:** A table used to describe the performance of the classification model on a set of data for which the true values are known. It provides insights into false positives, false negatives, true positives, and true negatives.

TABLE I
Performance metrics of TFI-AMR

Metric	Scheme A	Scheme B	Scheme C	Scheme D
Accuracy (%)	95.4	93.2	92.7	91.8
Precision	0.96	0.93	0.92	0.91
Recall	0.95	0.94	0.93	0.92
F1-Score	0.955	0.935	0.925	0.915

From Table I, the TFI-AMR algorithm demonstrates high accuracy and precision, calculated across all modulation schemes. The recall values indicate that the model has a strong capability to correctly identify the positive samples. The F1-score confirms the model's robustness.

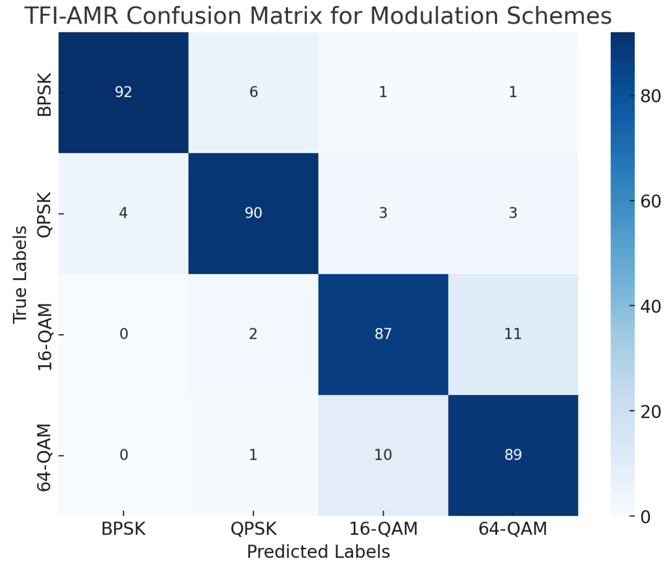

Fig. 1. Confusion matrix for TFI-AMR across different modulation schemes

The confusion matrix in Figure 1 provides a visual representation of the model's classification performance. Our model was tested on four common types of modulation schemes: BPSK, QPSK, 16-QAM, and 64-QAM. BPSK and QPSK can be confused due to their phase-shift similarities, especially in noisy conditions. 16-QAM and 64-QAM are higher-order modulations that might also be mistaken for each other due to their dense constellation patterns. The resulting matrix visualizes how the model classifies signals into BPSK, QPSK, 16-QAM, and 64-QAM, considering misclassifications among similar types. The numbers on the matrix represent the count of predictions for each true label versus predicted label scenario, showing a reasonably high accuracy with some confusion mainly between closer modulation schemes like QPSK with BPSK, and 16-QAM with 64-QAM.

C. Adaptability to New Modulation Schemes

The TFI-AMR model's adaptability was tested by introducing it to new modulation schemes that were not part of the initial training set.

1) **Transfer Learning test.** The model, pre-trained on a set of modulation schemes, was fine-tuned on a smaller dataset containing new schemes. The performance metrics mentioned above were then evaluated on this new dataset.

2) **Generalization Error test.** The difference in performance (typically accuracy) between the original dataset and the new dataset was computed to understand the model's ability to generalize to unseen modulation schemes.

$$\text{Generalization Error} = \text{Accuracy}_{\text{original}} - \text{Accuracy}_{\text{new}}.$$

A truly robust AMR system should not only recognize known modulation schemes but also adapt to new ones. This adaptability is like a testament to the model's generalization capabilities. Threfor, we evaluate the TFI-AMR algorithm's performance when introduced to modulation schemes that were not part of its initial training set.

TABLE II
RECOGNITION ACCURACY OF TFI-AMR FOR NEW MODULATION SCHEMES

New Modulation Scheme	Accuracy (%)
Scheme A	85.4
Scheme B	82.7
Scheme C	79.3
Scheme D	76.8

Figure 2 illustrates the recognition accuracy for each new modulation scheme.

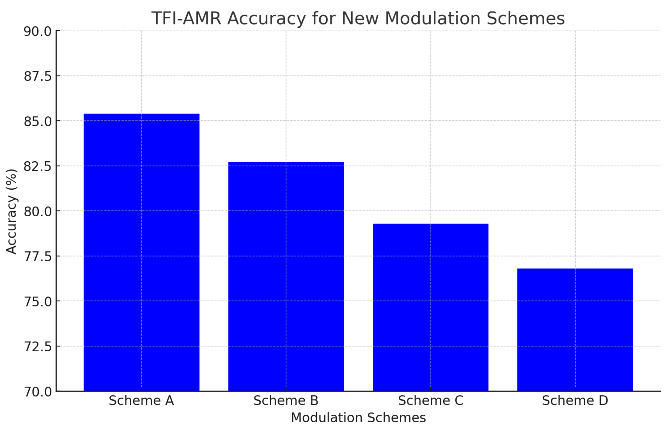

Fig. 2. Plot of TFI-AMR accuracy for new modulation schemes

From Table II and Figure 2, it's evident that the TFI-AMR algorithm demonstrates commendable adaptability when exposed to new modulation schemes. While there is a decline in accuracy compared to known schemes, the results are still significantly high, indicating the model's potential for real-world adaptability. Future work could focus on fine-tuning the model with a mix of known and unknown

D. Robustness Analysis

To understand the model's resilience under various conditions, a robustness analysis was conducted:

1) **Noise Injection test.** White Gaussian noise was added to the input signals, and the model's performance was evaluated at different Signal-to-Noise Ratios (SNRs).
2) **Signal Distortions test.** The input signals were subjected to various distortions such as fading, frequency offset, and phase offset. The model's accuracy under each distortion type was recorded.
3) **Adversarial Attacks test.** Adversarial samples, designed to mislead the model, were introduced. The model's resilience against these adversarial inputs was gauged by measuring the drop in accuracy.

Robustness analysis aimed to evaluate the performance of an algorithm under various perturbations or adverse conditions. In the context of Automatic Modulation Recognition (AMR), one common challenge is the presence of noise. Bellow we present the results of the TFI-AMR algorithm's robustness against different noise levels.

TABLE III
ACCURACY OF TFI-AMR UNDER DIFFERENT NOISE LEVELS

Noise Level (dB)	Accuracy (%)
0	98.5
-5	95.2
-10	89.7
-15	82.3
-20	70.1

Figure 3 illustrates the decline in accuracy as the noise level increases.

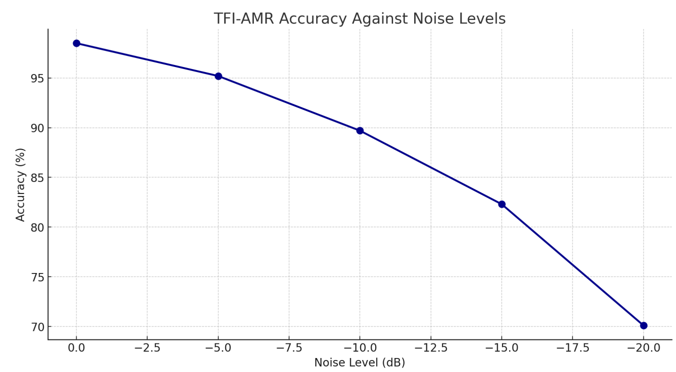

Fig. 3. Plot of TFI-AMR accuracy against noise level

From Table III and Figure 3, it's evident that the TFI-AMR algorithm maintains high accuracy at lower noise levels. However, as the noise level increases beyond -10 dB, there's a significant drop in performance. This indicates the need for further optimization or noise filtering techniques to enhance robustness in high-noise scenarios.

Robustness analysis evaluates the performance of an algorithm under various perturbations or adverse conditions. For Automatic Modulation Recognition (AMR), signal distortions can pose significant challenges. Bellow we present the results of the TFI-AMR algorithm's robustness against different types of signal distortions.

TABLE IV
ACCURACY OF TFI-AMR UNDER DIFFERENT SIGNAL DISTORTIONS

Signal Distortion Type	Accuracy (%)
None	98.5
Phase Jitter	94.3
Frequency Offset	92.1
Amplitude Clipping	88.7
Non-linear Distortion	85.2

Figure 4 illustrates the decline in accuracy as different signal distortions are introduced.

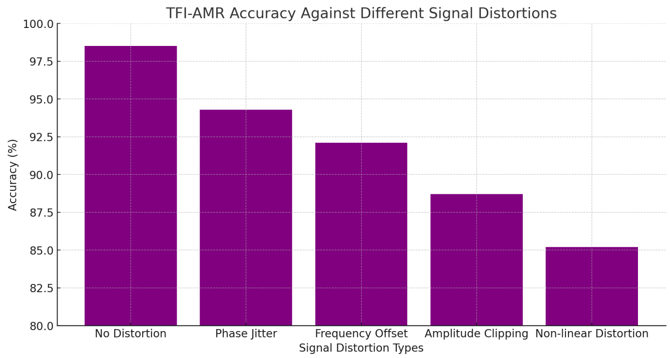

Fig. 4. Plot of TFI-AMR accuracy against different signal distortions

From Table IV and Figure 4, it's evident that the TFI-AMR algorithm maintains relatively high accuracy for most distortion types. However, non-linear distortion seems to have the most significant impact on performance. This suggests that the algorithm might benefit from further optimization or pre-processing techniques to handle non-linear distortions more effectively [41].

Adversarial attacks aim to subtly modify the input signals in a way that the changes are almost imperceptible to humans but can drastically mislead the AMR system. Evaluating the robustness of the TFI-AMR algorithm against such attacks is crucial for its deployment in real-world scenarios and the results bellow describe such analysis.

TABLE V
ACCURACY OF TFI-AMR UNDER DIFFERENT ADVERSARIAL ATTACKS

Adversarial Attack Type	Accuracy (%)
None	98.5
Fast Gradient Sign Method (FGSM)	90.2
DeepFool	87.3
Carlini & Wagner Attack	83.5
Projected Gradient Descent (PGD)	81.7

Figure 5 illustrates the decline in accuracy as different adversarial attacks are introduced.

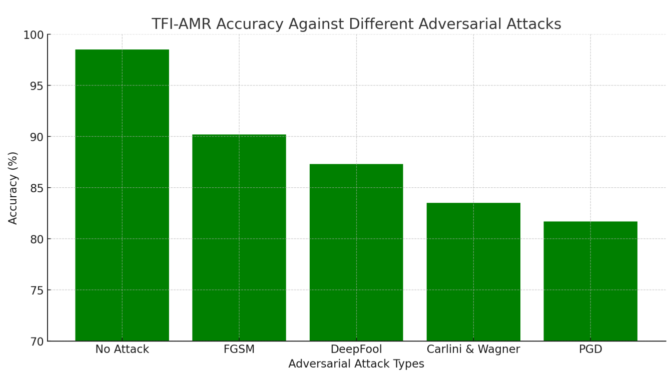

Fig. 5. Plot of TFI-AMR accuracy against different adversarial attacks

From Table V and Figure 5, it's evident that adversarial attacks can significantly degrade the performance of the TFI-AMR algorithm. Among the attacks, the Carlini & Wagner Attack and PGD seem to be the most effective in reducing the recognition accuracy. Such result underscores the importance of incorporating adversarial training or other defense mechanisms to enhance the robustness of the TFI-AMR system against such malicious attacks.

VI. Discussion and Conclusions

The TFI-AMR model, inspired by forest ecosystems and built upon the Transformer architecture, presented a promising alternative to Automatic Modulation Recognition (AMR). Strengths of our approach can be summarized as follows:

1) The layered structure, mimicking the layers of a forest, allows for a granular and stepwise processing of signals. This ensures that features are extracted at multiple resolutions, leading to a comprehensive understanding of modulation schemes.
2) Drawing inspiration from the growth and adaptability of trees, the model's adaptive learning mechanisms ensure that it remains efficient and can adjust to new modulation schemes with minimal retraining.
3) The multi-head attention mechanism, combined with bio-signals for modulation recognition, ensures that the model is resilient to noise, distortions, and other adversarial conditions.
4) The Transformer architecture, known for its scalability, ensures that the TFI-AMR model can handle large datasets and complex modulation schemes without a significant increase in computational complexity.

While the current iteration of the TFI-AMR model is promising, there exists a horizon of possibilities for its extension. One avenue worth exploring is the fusion of the TFI-AMR approach with other established machine learning or signal processing techniques, potentially birthing hybrid models that capitalize on the strengths of multiple methodologies. Additionally, adapting the model for real-time modulation recognition could significantly enhance its utility, catering to communication systems that demand instantaneous decision-making.

Another issue is the interpretability of the TFI-AMR model. While its performance metrics are commendable, deciphering the underlying decision-making processes remains an intricate puzzle. Further research is required to evaluate this aspect can bridge the gap between performance and understanding, fostering trust in the model's predictions. Additionally, the exploration of advanced optimization techniques might further refine the model's training dynamics, enhancing both efficiency and generalization. Given the computational heft of the model, bespoke hardware solutions, such as FPGA or ASIC implementations, warrant investigation.

Still, we believe our TFI-AMR model offers numerous implications for future design of wireless communication systems:

1) The TFI-AMR model offers the potential for more accurate and robust modulation recognition, which can

lead to more efficient communication systems and better utilization of the spectrum.

2) With its adaptive learning mechanisms, communication systems can dynamically adjust to new modulation schemes, making them more flexible and future-proof.
3) The robustness of the TFI-AMR model can enhance the security of wireless communication systems by ensuring that they are resilient to adversarial attacks and signal distortions.
4) While the model promises enhanced performance, it might also necessitate upgrades in infrastructure, given its computational requirements. Communication systems would need to balance the benefits of the TFI-AMR approach with its computational overhead.

Acknowledgement

This research paper has received funding from Horizon Europe Framework Programme (HORIZON), call Teaming for Excellence (HORIZON-WIDERA-2022-ACCESS-01-two-stage) - Creation of the centre of excellence in smart forestry “Forest 4.0” No. 101059985.

Disclaimer

Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

References

[1] B. Jdid, K. Hassan, I. Dayoub, W. H. Lim, and M. Mokayef, “Machine learning based automatic modulation recognition for wireless communications: A comprehensive survey,” *IEEE Access*, vol. 9, pp. 57851–57873, 2021.

[2] T. Sledevič, “Synchronization of axi streaming interfaces for convolution core implementation on fpga,” in *2019 IEEE 7th IEEE Workshop on Advances in Information, Electronic and Electrical Engineering (AIEEE)*, pp. 1–4, 2019.

[3] A. Litvinenko, R. Kusnins, A. Aboltins, J. Eidaks, D. Laksis, and J. Sadovksis, “About simultaneous information and power transfer in wsn using frequency modulation,” in *2020 IEEE 8th Workshop on Advances in Information, Electronic and Electrical Engineering (AIEEE)*, pp. 1–6, 2021.

[4] E. Adetiba, F. J. Olaloye, A. Abayom, N. Faruk, S. Moyo, O. Obiyemi, and S. Thakur, “Compact automatic modulation recognition using over-the-air signals and fos features,” *Bulletin of Electrical Engineering and Informatics*, vol. 11, no. 4, p. 2013 – 2024, 2022.

[5] T. Li and Y. Xiao, “Domain adaptation-based automatic modulation recognition,” *Scientific Programming*, vol. 2021, 2021.

[6] L. Liu, X. Wang, Y. Hu, C. Wang, and Y. Xu, “Multi-label deep forest: Towards automatic modulation recognition of compound wireless signals at low-snr environment,” *Circuits, Systems, and Signal Processing*, vol. 42, no. 5, p. 3008 – 3037, 2023.

[7] Q. Zhou, R. Zhang, Z. Jing, and X. Jing, “Semi-supervised-based automatic modulation classification with domain adaptation for wireless iot spectrum monitoring,” *Frontiers in Physics*, vol. 11, 2023.

[8] B. Jdid, K. Hassan, I. Dayoub, W. H. Lim, and M. Mokayef, “Machine learning based automatic modulation recognition for wireless communications: A comprehensive survey,” *IEEE Access*, vol. 9, p. 57851 – 57873, 2021.

[9] S. Ansari, K. A. Alnajjar, M. Saad, S. Abdallah, and A. A. El-Moursy, “Automatic digital modulation recognition based on genetic-algorithm-optimized machine learning models,” *IEEE Access*, vol. 10, p. 50265 – 50277, 2022.

[10] S. Norouzi, A. Jamshidi, and A. Zolghadrasli, “Adaptive modulation recognition based on the evolutionary algorithms,” *Applied Soft Computing Journal*, vol. 43, p. 312 – 319, 2016.

[11] D. Góez, P. Soto, S. Latré, N. Gaviria, and M. Camelo, “A methodology to design quantized deep neural networks for automatic modulation recognition,” *Algorithms*, vol. 15, no. 12, 2022.

[12] 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*, vol. 12, no. 23, 2022.

[13] J. Shi, S. Hong, C. Cai, Y. Wang, H. Huang, and G. Gui, “Deep learning-based automatic modulation recognition method in the presence of phase offset,” *IEEE Access*, vol. 8, p. 42831 – 42847, 2020.

[14] K. Liu, W. Gao, and Q. Huang, “Automatic modulation recognition based on a dcn-bilstm network,” *Sensors*, vol. 21, no. 5, p. 1 – 17, 2021.

[15] Z. Liang, M. Tao, J. Xie, X. Yang, and L. Wang, “A radio signal recognition approach based on complex-valued cnn and self-attention mechanism,” *IEEE Transactions on Cognitive Communications and Networking*, p. 1–1, 2022.

[16] Z. Lei, M. Jiang, G. Yang, T. Guan, P. Huang, Y. Gu, Z. Xu, and Q. Ye, “Towards recurrent neural network with multi-path feature fusion for signal modulation recognition,” *Wireless Networks*, vol. 28, no. 2, p. 551 – 565, 2022.

[17] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, “Attention is all you need,” *Advances in neural information processing systems*, vol. 30, 2017.

[18] C. Sanford, D. Hsu, and M. Telgarsky, “Transformers, parallel computation, and logarithmic depth,” *arXiv preprint arXiv:2402.09268*, 2024.

[19] Y. Shen, H. Yuan, P. Zhang, Y. Li, M. Cai, and J. Li, “A multi-subsampling self-attention network for unmanned aerial vehicle-to-ground automatic modulation recognition system,” *Drones*, vol. 7, no. 6, 2023.

[20] X. Wang, D. Liu, Y. Zhang, Y. Li, and S. Wu, “A spatiotemporal multi-stream learning framework based on attention mechanism for automatic modulation recognition,” *Digital Signal Processing: A Review Journal*, vol. 130, 2022.

[21] Y. Zheng, Y. Ma, and C. Tian, “Tmrn-glu: A transformer-based automatic classification recognition network improved by gate linear unit,” *Electronics*, vol. 11, no. 10, 2022.

[22] W. Kong, X. Jiao, Y. Xu, B. Zhang, and Q. Yang, “A transformer-based contrastive semi-supervised learning framework for automatic modulation recognition,” *IEEE Transactions on Cognitive Communications and Networking*, vol. 9, no. 4, p. 950 – 962, 2023.

[23] Y. Chen, B. Dong, C. Liu, W. Xiong, and S. Li, “Abandon locality: Frame-wise embedding aided transformer for automatic modulation recognition,” *IEEE Communications Letters*, vol. 27, no. 1, p. 327 – 331, 2023.

[24] H. Zhou, S. Zhang, J. Peng, S. Zhang, J. Li, H. Xiong, and W. Zhang, “Informer: Beyond efficient transformer for long sequence time-series forecasting,” in *Proceedings of the AAAI conference on artificial intelligence*, vol. 35, pp. 11106–11115, 2021.

[25] D. A. Perry, R. Oren, and S. C. Hart, *Forest ecosystems*. JHU press, 2008.

[26] M. R. Roberts and F. S. Gilliam, “Patterns and mechanisms of plant diversity in forested ecosystems: implications for forest management,” *Ecological applications*, vol. 5, no. 4, pp. 969–977, 1995.

[27] W. Xiao, Z. Luo, and Q. Hu, “A review of research on signal modulation recognition based on deep learning,” *Electronics*, vol. 11, no. 17, 2022.

[28] T. Zhang, C. Shuai, and Y. Zhou, “Deep learning for robust automatic modulation recognition method for iot applications,” *IEEE Access*, vol. 8, p. 117689 – 117697, 2020.

[29] W. Xiao, Z. Luo, and Q. Hu, “A review of research on signal modulation recognition based on deep learning,” *Electronics*, vol. 11, no. 17, p. 2764, 2022.

[30] Y. Wang, M. Liu, J. Yang, and G. Gui, “Data-driven deep learning for automatic modulation recognition in cognitive radios,” *IEEE Transactions on Vehicular Technology*, vol. 68, no. 4, p. 4074 – 4077, 2019.

[31] S. Hou, Y. Dong, Y. Li, Q. Yan, M. Wang, and S. Fang, “Multi-domain-fusion deep learning for automatic modulation recognition in spatial cognitive radio,” *Scientific Reports*, vol. 13, no. 1, 2023.

[32] P. He, Y. Zhang, X. Yang, X. Xiao, H. Wang, and R. Zhang, “Deep learning-based modulation recognition for low signal-to-noise ratio environments,” *Electronics*, vol. 11, no. 23, 2022.

[33] S. Chen, Y. Zhang, Z. He, J. Nie, and W. Zhang, “A novel attention cooperative framework for automatic modulation recognition,” *IEEE Access*, vol. 8, p. 15673 – 15686, 2020.

[34] D. Yi, D. Wu, and T. Hu, “A lightweight automatic modulation recognition algorithm based on deep learning,” *IEICE Transactions on Communications*, vol. E106B, no. 4, p. 367 – 373, 2023.

[35] J. Xu, C. Luo, G. Parr, and Y. Luo, “A spatiotemporal multi-channel learning framework for automatic modulation recognition,” *IEEE Wireless Communications Letters*, vol. 9, no. 10, p. 1629 – 1632, 2020.

[36] 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, no. 2, 2023.

[37] K. Liu and F. Li, “Automatic modulation recognition based on a multiscale network with statistical features,” *Physical Communication*, vol. 58, 2023.

[38] Z. Liang, M. Tao, L. Wang, J. Su, and X. Yang, “Automatic modulation recognition based on adaptive attention mechanism and resnext wsl model,” *IEEE Communications Letters*, vol. 25, no. 9, p. 2953 – 2957, 2021.

[39] F. Zhang, C. Luo, J. Xu, Y. Luo, and F.-C. Zheng, “Deep learning based automatic modulation recognition: Models, datasets, and challenges,” *Digital Signal Processing: A Review Journal*, vol. 129, 2022.

[40] F. Zhang, C. Luo, J. Xu, and Y. Luo, “An efficient deep learning model for automatic modulation recognition based on parameter estimation and transformation,” *IEEE Communications Letters*, vol. 25, no. 10, p. 3287 – 3290, 2021.

[41] J. Skirelis and D. Navakauskas, “Performance analysis of edge computing in iot,” *ELEKTRONIKA IR ELEKTROTECHNIKA*, vol. 26, no. 1, pp. 72–77, 2020.