

# Unified Deep Learning Framework for Joint Recognition of Modulation and FEC Encoding

Ashish Tandi\*, Aryan Jaiswal<sup>†</sup>, Nayim Ahamed<sup>‡</sup> and Swaminathan R.<sup>§</sup>

<sup>\*†‡§</sup>Department of Electrical Engineering, Indian Institute of Technology Indore,

Email: <sup>\*</sup>ee220002020@iiti.ac.in, <sup>†</sup>ee220002019@iiti.ac.in, <sup>‡</sup>phd2201202010@iiti.ac.in, <sup>§</sup>swamiramabadran@iiti.ac.in

**Abstract**—Blind recognition of modulation formats and channel encoders is essential for non-cooperative wireless communication systems, where no prior knowledge of the signal structure is available. This paper proposes a deep-learning-based joint identification framework using the feature network (FeatureNet) and one dimensional skip convolutional network (SCN1D) architectures to simultaneously classify coding schemes and modulation types under Rayleigh fading and real-time environment. The system is trained on datasets generated using convolutional, turbo, and polar encoders combined with four modulation formats, enabling 12 class combinations. Under Rayleigh fading, the FeatureNet model achieves over 96% accuracy, while the SCN1D model attains 97% classification accuracy at a signal-to-noise ratio (SNR) of 10 dB. The proposed framework maintains high recognition performance in Rayleigh fading channels, with accuracy improving steadily as the input length increases and exhibiting strong class separability at 10 dB. Hardware experiments further validate the approach, demonstrating consistently high accuracies across most classes despite waveform impairments and dense quadrature amplitude modulation (QAM) constellations. These results demonstrate the robustness and practical effectiveness of the proposed framework for blind encoder-modulation identification in realistic wireless environments.

**Index Terms**—Blind identification, channel encoders, deep learning, FeatureNet, modulation classification, Rayleigh fading, and hardware impairments.

## I. INTRODUCTION

The rapid evolution of wireless communication systems from second generation (2G) to fifth generation (5G) and beyond has led to a dramatic rise in user density, traffic load, and service diversity, thereby demanding higher spectral efficiency and intelligent physical layer processing [1]. At the transmitter, channel encoding and modulation together govern both reliability and bandwidth utilization by introducing redundancy in a structured form and mapping information onto suitable signal constellations. Modern communication standards employ a wide range of forward error correction (FEC) schemes such as convolutional and turbo codes which are valued for their strong error resilience and suitability in high mobility and fading environments [2], while adaptive digital modulation formats are used to maximize throughput under varying channel conditions. However, in many real world scenarios including military, surveillance, cognitive radio, and other uncoordinated deployments, the receiver may not have prior knowledge of either the modulation scheme or the encoder used at the transmitter. This lack of coordination requires joint blind identification of modulation and channel encoder parameters, enabling the

receiver to correctly demodulate and decode signals without relying on pilot, control overhead, or metadata. Such capability is essential for achieving reliable reception in non-cooperative, dynamic, or adversarial environments, where transmission parameters are unknown or deliberately concealed. Adaptive modulation and coding (AMC) improves spectral efficiency by dynamically selecting suitable modulation formats and coding strategies in response to varying channel conditions. However, its effectiveness relies on the receiver's ability to determine both the modulation type and channel encoding scheme without explicit signaling or overhead. Conventional approaches follow a cascaded blind identification process, where modulation is first classified and then the channel encoder is identified based on the recovered bitstream. Although this two stage approach works, it suffers from error propagation, increased latency, and reduced reliability especially under low signal-to-noise ratio (SNR) or rapidly varying channel conditions. In contrast, a joint blind identification framework aims to recognize both modulation and encoding parameters simultaneously using a single unified dataset and a common learning model. This eliminates intermediate decision dependencies, improves robustness, and accelerates receiver adaptation in non-cooperative or dynamic environments. With the advancement of deep learning, intelligent receivers can jointly classify modulation and coding, reducing latency and improving energy efficiency. This is beneficial for cognitive radio, 5G/6G systems, spectrum monitoring, and military communication and surveillance, where blind identification is essential. It also enhances spectral efficiency and real-time adaptability in complex wireless environments.

Several techniques have been investigated for the blind identification of channel encoders, as reported in [3]. In [4], an algebraic reconstruction strategy based on Gauss–Jordan elimination was introduced for LDPC, convolutional, and turbo codes; however, it was limited to noise-free conditions. To address noisy environments, rank-deficiency-based methods for code classification and parameter estimation were explored in [5]. To improve reliability at low SNR, likelihood-ratio-based techniques were proposed in [6], [7], while [8] employed an expectation-maximization (EM)-based multi-receiver likelihood method, achieving near-perfect identification accuracy at an SNR of 3 dB. More recent studies have adopted deep learning (DL) techniques, particularly convolutional neural network (CNN)- and recurrent neural network (RNN)-based classifiers, for blind FEC recognition, demonstrating improved robustness

to channel impairments [9]–[11]. A DL-based framework for identifying LDPC, convolutional, turbo, and polar codes under multipath fading was presented in [12]. More recently, [13] proposed an aerial-computing-assisted CNN9 framework for joint modulation and encoder recognition. Although it achieved good accuracy across different SNRs, its large parameter count and computational complexity limit lightweight real-time deployment.

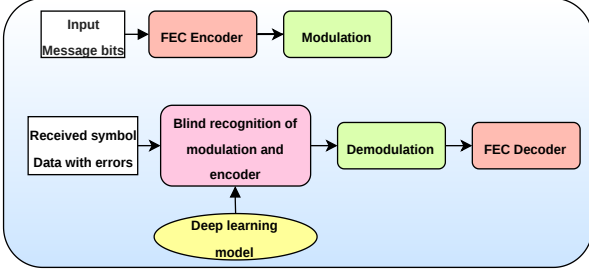


Fig. 1: Joint encoder and modulation recognition system model for feature extraction and classification.

#### A. Motivations and Contributions

Modern wireless communication systems, including cognitive radio, military communications, and AMC-based receivers, often operate without reliable side information regarding the modulation and FEC coding schemes. This necessitates blind joint identification at the receiver. However, existing studies largely treat modulation and FEC recognition separately, with limited deep learning solutions available for joint classification. Motivated by this gap and the limitations of existing approaches under low-SNR conditions, this work investigates the joint classification of 12 encoder–modulation combinations formed by three FEC encoder families (convolutional, turbo, and polar) and four modulation schemes (BPSK, 8-PSK, 16-QAM, and 64-QAM). Two deep learning architectures, FeatureNet and SCNID, are proposed and validated using both simulated Rayleigh fading data and real RF signals collected from a GNU Radio–USRP B210 testbed, with performance compared against CNN- and RNN-based baselines.

### II. FRAMEWORK FOR BLIND RECOGNITION SYSTEM

#### A. System Model

At the transmitter, a binary message sequence  $m \in \text{GF}(2)^k$  of length  $k$  bits is encoded into a codeword  $c \in \text{GF}(2)^n$  of length  $n$  bits using one of several FEC coding schemes such as convolutional, turbo, or polar coding, where  $n > k$  as shown in Fig. 1, and the code rate is given by  $R = \frac{k}{n}$ . The encoded bitstream  $c = [c_1, c_2, \dots, c_n]$  is subsequently modulated using a  $M$ -ary digital modulation scheme  $\mathcal{M} \in \{\text{BPSK}, \text{8-PSK}, \text{16-QAM}, \text{64-QAM}\}$ , mapping groups of bits into complex symbols.

The baseband signal is transmitted through a Rayleigh fading channel with additive white Gaussian noise (AWGN), representing realistic multipath propagation conditions. The

instantaneous complex channel coefficient is denoted by  $h$ , modeled as a circularly symmetric complex Gaussian random variable with zero mean and unit variance, i.e.,  $h \sim \mathcal{CN}(0, 1)$ . Accordingly, the received signal can be written as

$$y = h \cdot s + n, \quad (1)$$

where  $n \sim \mathcal{CN}(0, N_0)$  represents complex Gaussian noise and  $s$  denotes the modulated symbols.

At the receiver, the noisy signal  $y$  is demodulated to recover the distorted symbol sequence. In the proposed joint recognition framework, the received I/Q samples are directly processed by a DL-based classifier that performs simultaneous identification of both the modulation and channel coding schemes without any prior knowledge of the transmitted configuration. The model learns to discriminate among 12 possible signal classes, each corresponding to a unique combination of modulation and FEC encoding type. A large-scale synthetic dataset comprising 12 modulation-coding classes was generated in MATLAB across multiple SNR levels. The transmitted signals were encoded using different FEC schemes, modulated, and passed through a Rayleigh fading channel to emulate realistic wireless conditions. At the receiver, the proposed blind recognition network jointly identifies the modulation format and channel encoder, enabling adaptive demodulation and decoding. This unified framework enhances physical-layer signal characterization and supports robust intelligent receiver operation in adaptive, cognitive, and military communication systems.

#### B. FEC Encoding and Modulation

In the proposed system, three widely used FEC encoders are considered: convolutional, turbo, and polar codes. Convolutional codes introduce redundancy through generator polynomials to improve error correction capability. Turbo codes employ parallel concatenated recursive systematic convolutional (RSC) encoders separated by an interleaver and achieve near-capacity performance through iterative decoding. Polar codes utilize channel polarization to transmit information over reliable channels while assigning frozen bits to unreliable ones, making them well suited for modern standards such as 5G. In this work, a rate-1/2 convolutional encoder, a rate-1/3 turbo encoder, and a rate-1/2 polar encoder are used. The corresponding encoder parameters are summarized in Table I. For modulation, BPSK, 8-PSK, 16-QAM, and 64-QAM schemes are considered. Following the MATLAB modulation framework, the input bitstream is grouped according to the bits-per-symbol requirement of each scheme and mapped to Gray-coded constellation points, ensuring consistent and robust dataset generation.

#### C. Dataset Generation

The dataset for 12 classes was generated in MATLAB considering Rayleigh fading channels to simulate system generated noise and multipath effects, with a signal length of 1024 I/Q samples. As shown in Fig. 2, the real-time dataset was also generated using USRP B210 SDR hardware to introduce realistic radio frequency impairments under identical signal parameters. The system parameters are provided in Table II. For simulations, the SNR was varied from  $-25$  dB to  $20$  dB in  $5$

TABLE I: Channel Encoder Parameters

| Encoder Type            | Code Rate | Constraint Length / Block Length | Generator / Notes                       |
|-------------------------|-----------|----------------------------------|-----------------------------------------|
| Convolutional (133,171) | 1/2       | 7 (constraint length)            | Generator polynomials $[171, 133]_8$    |
| Turbo                   | 1/3       | Input block length variable      | Two parallel RSC encoders + interleaver |
| Polar                   | 1/2       | 128 bits                         | Channel polarization with frozen bits   |



Fig. 2: GNU Radio-based hardware implementation setup.

TABLE II: RF Parameter used in USRP Hardware

| Systems parameters | Values  |
|--------------------|---------|
| Center frequency   | 2.4 GHz |
| Bandwidth          | 1 MHz   |
| Transmitter gain   | 200 dB  |
| Receiver gain      | 40 dB   |
| Frame bits         | 1024    |

TABLE III: Features of Joint Encoder-Modulation Schemes

| Encoder | Mod.         | Learnable RF Fingerprints                                  |
|---------|--------------|------------------------------------------------------------|
| Conv.   | BPSK/PSK/QAM | Short-range trellis correlations and periodic transitions. |
| Turbo   | PSK/QAM      | Interleaver-induced long-range dependencies.               |
| Polar   | BPSK/QAM     | Frozen-bit structure and unequal bit reliability.          |
| All     | 16/64-QAM    | Dense constellations with encoder-dependent redundancy.    |

dB steps, yielding 120,000 training samples. The testing dataset was generated independently using a 80%-20% train-test split.

### III. DL ARCHITECTURES

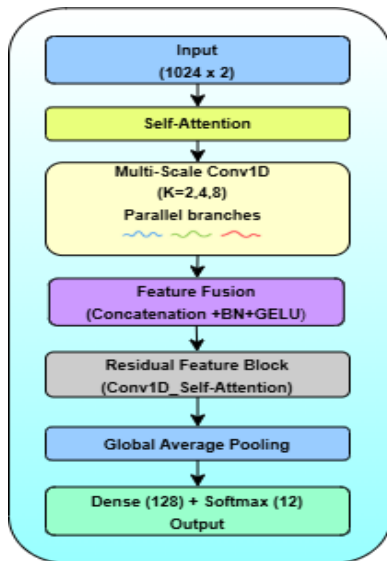


Fig. 3: FeatureNet with self-attention and multi-scale learning.

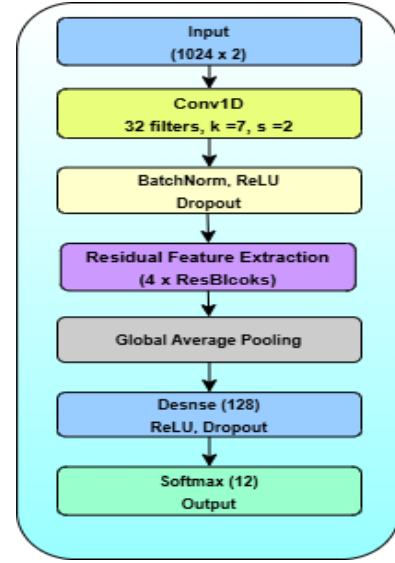


Fig. 4: Lightweight SCN1D architecture with residual feature extraction.

FeatureNet is a lightweight hierarchical 1D CNN designed for joint modulation and channel encoder recognition from raw complex baseband I/Q samples of size  $1024 \times 2$ . As shown in Fig. 3, the network first employs a self-attention layer to emphasize informative temporal regions in the received waveform. A multi-scale Conv1D module consisting of three parallel convolution branches with kernel sizes  $k = \{2, 4, 8\}$  extracts complementary short-, medium-, and long-range temporal features, as represented by the colored waveform branches in the architecture. The resulting feature maps are fused through concatenation, batch normalization (BN), and GELU activation to generate discriminative representations. A residual feature learning block further refines the extracted features, followed by global average pooling and a dense classification layer with softmax activation to perform joint modulation and channel encoder recognition.

As shown in Fig. 4, SCN1D is a compact residual 1D CNN designed for joint modulation and channel encoder recognition from raw complex I/Q samples of size  $1024 \times 2$ . The network begins with a Conv1D layer followed by batch normalization, ReLU activation, and dropout for initial feature extraction.

A stack of residual blocks then learns hierarchical temporal representations while preserving efficient gradient propagation through shortcut connections. The extracted features are aggregated using global average pooling and subsequently processed by a dense layer with dropout for high-level feature integration. Finally, a softmax classifier maps the learned representation to the 12 modulation-coding categories.

Further Table III summarizes, the relationship between joint encoder-modulation parameters and the corresponding learned features at the DL input.

#### IV. SIMULATION RESULTS AND DISCUSSIONS

We implement the model in Keras using cross-entropy loss and the Adam optimizer ( $L_r = 0.0001$ ). Training is performed in an Anaconda environment on a system with 16 GB RAM, a Core i9 CPU, and an NVIDIA RTX 3060 (6 GB VRAM). The dataset is split into 80% training and 20% testing, with training conducted for 75 epochs. Performance is evaluated using classification accuracy across all 12 classes, defined as

$$\text{Accuracy (\%)} = \frac{\text{Total Correct Predictions}}{\text{Total Samples}} \times 100. \quad (2)$$

The performance of the proposed FeatureNet and SCN1D models was evaluated under Rayleigh fading channel across an SNR range of  $-20$  dB to  $25$  dB, and compared with RNN and CNN-based architectures [13]. As shown in Table IV, FeatureNet achieves the highest F1-score of  $0.963$  with only  $182.7k$  parameters, which is significantly fewer than those of the other models, including the existing CNN9 [13]. Furthermore, FeatureNet requires  $1.74$  ms per sample, which is slightly slower than SCN1D ( $1.13$  ms) due to its multiscale convolution and channel-wise concatenation. In contrast, SCN1D's streamlined residual design enables faster inference despite a higher parameter count. The RNN model exhibits the lowest overall performance (F1-score of  $0.873$ ) and the highest inference time ( $3.59$  ms), primarily due to its sequential processing nature, which limits its ability to efficiently capture discriminative features compared to residual architectures.

We study the ablation of the core FeatureNet components in Table V. The baseline CNN achieves an accuracy of  $84.7\%$ , which improves to  $89.1\%$  ( $+4.4\%$ ) with the multiscale convolution block. Adding the self-attention module further increases the accuracy to  $92.4\%$  by capturing long-range dependencies. Residual connections provide a substantial improvement to  $95.6\%$ , while the complete FeatureNet with GELU activation attains the highest accuracy of  $96.3\%$ , confirming the cumulative benefit of all components.

Fig. 5 shows the accuracy performance of different models under Rayleigh fading conditions, where fading-induced amplitude and phase distortions degrade all models, particularly at low SNR. Although CNN9 [13] peaks at mid-high SNRs and SCN1D maintains stable performance across the SNR range, FeatureNet rapidly climbs from  $64.40\%$  at  $-20$  dB to approximately  $96\%$  by  $10$ – $15$  dB and remains stable thereafter. Despite its sensitivity at extremely low SNRs, FeatureNet

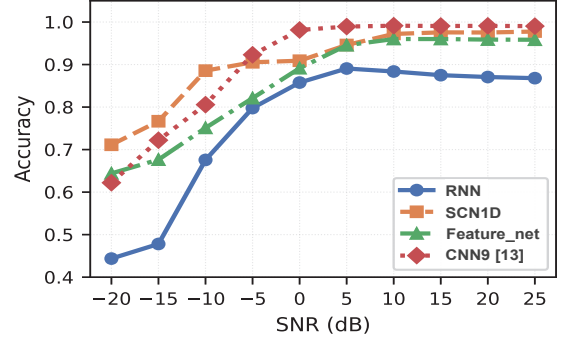


Fig. 5: Comparison of deep learning model accuracy versus SNR.

achieves strong high-SNR performance with far fewer parameters compared to SCN1D and CNN9 [13] as shown in Table IV, making it the most efficient and accurate model in practical, moderate-to-high SNR conditions. In contrast, the RNN exhibits the lowest accuracy across the SNR range and saturates below  $90\%$  even at high SNR due to its sequential processing and limited robustness under deep fading. The strong performance is attributed to distinct I/Q-domain features arising from constellation geometry, symbol dynamics, and coding-induced redundancy. These characteristics are effectively captured by the multiscale Conv1D branches and the attention mechanism, improving class separability.

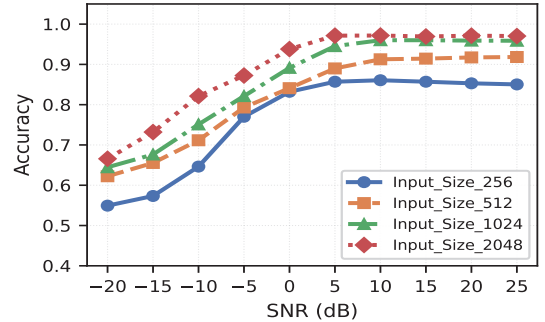


Fig. 6: Classification accuracy across SNR for different input lengths.

Fig. 6 compares the FeatureNet accuracy for different input lengths ( $N = 256$ – $2048$ ) across SNR. Increasing the input length consistently improves performance, since longer I/Q sequences provide richer temporal context and enable the network to capture encoder-induced correlations and modulation-specific features more reliably. At low SNR (e.g.,  $-20$  dB), the accuracy ranges from approximately  $55\%$  ( $N = 256$ ) to  $67\%$  ( $N = 2048$ ). At higher SNR, performance improves significantly, where the  $256$ -sample input saturates near  $85\%$  while the  $2048$ -sample input reaches approximately  $96\%$ – $97\%$ . The observed performance gap is mainly due to channel-induced amplitude and phase variations, which distort the signal structure and suppress discriminative redundancy patterns, particularly for shorter input lengths.

In the real-time channel, the confusion matrix heatmap for the proposed FeatureNet model, shown in Fig. 7, demonstrates consistently strong performance across most coding-modulation

TABLE IV: Performance Comparison of Different Models

| Model      | Avg Acc.     | Max Acc.     | Min Acc.     | Prec.        | Recall       | F1           | Params    | Time (ms) |
|------------|--------------|--------------|--------------|--------------|--------------|--------------|-----------|-----------|
| RNN        | 0.864        | 0.984        | 0.475        | 0.897        | 0.864        | 0.873        | 60,812    | 3.59      |
| SCN1D      | 0.944        | 0.999        | 0.748        | 0.946        | 0.944        | 0.944        | 676,972   | 1.13      |
| FeatureNet | <b>0.963</b> | <b>0.999</b> | <b>0.837</b> | <b>0.963</b> | <b>0.963</b> | <b>0.963</b> | 182,700   | 1.74      |
| CNN9 [13]  | 0.924        | 0.999        | 0.671        | 0.932        | 0.924        | 0.924        | 1,312,244 | 1.79      |

TABLE V: Ablation Study of Core FeatureNet Components

| Model Variant                | Accuracy (%) | Observation                                            |
|------------------------------|--------------|--------------------------------------------------------|
| Baseline CNN (single Conv1D) | 84.7         | Limited temporal feature capture.                      |
| + Multiscale Conv Block      | 89.1         | Captures short- and mid-range dependencies.            |
| + Self-Attention             | 92.4         | Improves long-range dependency modeling.               |
| + Residual connection        | 95.6         | Enhances feature diversity and robustness.             |
| + GELU (Full FeatureNet)     | 96.3         | Smooth nonlinearity improves convergence and accuracy. |

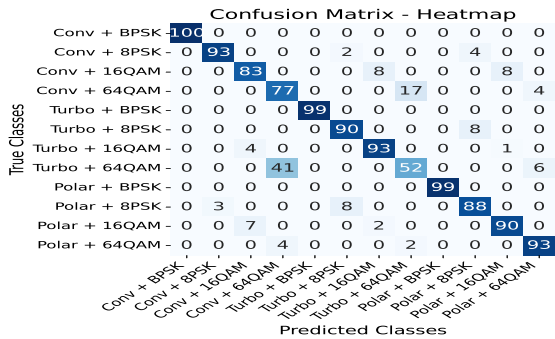


Fig. 7: Confusion matrix of FeatureNet using real-time hardware data

classes, with diagonal accuracies predominantly in the 90–100% range. However, the FeatureNet heatmap also reveals a notable 48% misclassification for one turbo + 64-QAM class. This performance degradation is primarily attributed to the dense constellation structure of 64-QAM, which is more susceptible to channel impairments and hardware-induced distortions under real-time conditions. Consequently, the extracted features exhibit increased similarity to neighboring coding-modulation combinations, leading to higher classification ambiguity.

## V. CONCLUSION

This work presented a unified DL-based framework comprising FeatureNet and SCN1D models for the joint blind identification of channel coding schemes and modulation formats across Rayleigh fading and real-time hardware environments. By generating diverse datasets and employing convolutional, turbo, and polar encoding structures paired with practical modulation formats, the proposed approach demonstrates strong adaptability to varying channel impairments. Compared with the existing baselines such as CNN9 and RNN, the proposed models achieve superior recognition performance and robustness under fading, while requiring significantly fewer parameters than CNN9 and substantially lower inference latency than RNN. FeatureNet consistently extract discriminative RF features and maintain reliable recognition capability in real-time hardware evaluation, despite RF impairments and dense signal constellations. The hardware experiments further confirm

the feasibility of deploying the proposed framework in practical non-cooperative communication scenarios.

## VI. ACKNOWLEDGMENT

This work was supported by the EMR-II scheme of CSIR-HRDG (22/0881/23/EMR-II), the TTDF scheme of DoT under 6G project (ID: 209), and by the 5G Use Case Lab at IIT Indore, established with funding from DoT.

## REFERENCES

- [1] M. Rowshan, M. Qiu, Y. Xie, X. Gu and J. Yuan, “Channel coding toward 6G: Technical overview and outlook,” *IEEE open journal of the Commun. Society*, vol. 5, pp. 2585-2685, April. 2024.
- [2] M. Eroz, Feng-Wen Sun and Lin-Nan Lee, “An innovative low-density parity-check code design with near-Shannon-limit performance and simple implementation,” *IEEE Trans. on Commun.*, vol. 54, no. 1, pp. 13-17, Jan. 2006.
- [3] R. Swaminathan, A. S. Madhukumar, G. Wang and T. Shang Kee, “Blind reconstruction of Reed-Solomon encoder and interleavers over noisy environment,” *IEEE Trans. Broadcast.*, vol. 64, no. 4, pp. 830-845, Dec. 2018.
- [4] A. Refaey, R. Niati, X. Wang and J. Y-Chouinard, “Blind detection approach for LDPC, convolutional, and turbo codes in non-noisy environment,” *IEEE Conf. on Commun. and Netw. Security.*, 2014, pp. 502–503.
- [5] Swaminathan R, A.S. Madhukumar, W. Guohua, and T. S. Kee, “Blind recognition of LDPC code parameters over erroneous channel conditions,” *IET Signal Processing*, vol. 13, no. 1, pp. 86-95, Feb. 2019.
- [6] T. Xia and H.-C. Wu, “Novel blind identification of LDPC codes using average LLR of syndrome a posteriori probability,” in *Proc. 12th Int. Conf. ITS Telecommun.*, Taipei, Taiwan, 2012.
- [7] L. Longqing, H. Zhiping and Z. Jing, “Blind identification of LDPC codes based on a candidate set,” *IEEE Commun. Lett.*, vol. 25, no. 9, pp. 2786-2789, Sept. 2021.
- [8] Y. Liu, F. Wang, J. Zhang, B. Ai and Z. Zhong, “Blind identification of LDPC codes in multipath fading channel via expectation maximization,” *IEEE Global Commun. Conf.(GLOBECOM)*, Abu Dhabi, United Arab Emirates, 2018, pp. 1-6.
- [9] F. Mei, H. Chen, and Y. Lei, “Blind recognition of forward error correction codes based on recurrent neural network,” *Sensors*, vol. 21, no. 11: 3884, June 2021.
- [10] S. Li, J. Zhou, Z. Huang, and X. Hu, “Recognition of error correcting codes based on CNN with block mechanism and embedding,” *Digit. Signal Process.*, vol. 111, no. 102986, April 2021.
- [11] J. Wang, C. Tang, H. Huang, H. Wang, and J. Li, “Blind identification of convolutional codes based on deep learning,” *Digit. Signal Process.*, vol. 115, no. 103086, Aug. 2021.
- [12] S. Dehdashtian, M. Hashemi and S. Salehkalebar, “Deep-learning-based blind recognition of channel code parameters over candidate sets under AWGN and multi-path fading conditions,” *IEEE Wireless Commun. Lett.*, vol. 10, no. 5, pp. 1041-1045, May 2021.
- [13] B. Wang, Z. Jiang, H. Zhou, and S. Zheng, “Blind recognition of non-cooperative signals via aerial computing,” *IEEE Trans. Netw. Sci. Eng.*, vol. 11, no. 6, pp. 5244–5252, Nov.–Dec. 2024.