

Compact Attention-Augmented Neural Decoder for Chaos-Based Wireless Communications

Marco Siino¹, Stefano Mangione¹, *Member, IEEE*, and Ilenia Tinnirello¹

Abstract—Chaos-based modulation offers strong interference resilience but decoding remains challenging under low SNRs and fading channels. Existing deep learning receivers achieve promising results but are often too large for resource-constrained systems. We propose Ultra-CAN, an ultra-light convolutional-attention decoder for Chaos Shift Keying (CSK) that combines efficient convolutional feature extraction with a compact attention block to capture long-range dependencies. Our design achieves superior Bit Error Rate (BER) performance compared to state-of-the-art demodulators in AWGN and Rayleigh channels, with a footprint of 220.4 kB. Results demonstrate the viability of compact attention-based decoders for chaotic communications.

Index Terms—CNN, attention, chaotic maps, bit decoding, attention mechanisms, embedded communications.

I. INTRODUCTION

WIRELESS communication systems are fundamental to a wide array of applications, from everyday services to mission-critical domains such as military, secure IoT, and emergency networks. These scenarios demand **robust and efficient physical-layer techniques**. However, conventional modulation schemes often prove inadequate in the presence of strong interference, low signal-to-noise ratios (SNRs), or eavesdropping threats [1].

In this context, **chaotic communication systems**, particularly Chaos Shift Keying (CSK), have gained attention as promising alternatives due to their noise-like waveforms and inherent resilience to interference [2], [3], [4], [5], [6], [7]. Chaotic systems are deterministic nonlinear dynamical systems whose trajectories are highly sensitive to initial conditions [8]. Chaotic communication systems encode information by mapping bits onto distinct chaotic maps, such as using a Bernoulli map for ‘0’ and a logistic map for ‘1’, each initialized with a random seed to generate unique chaotic trajectories. For example, a simple discrete-time generator is the logistic map $x_{n+1} = \mu x_n(1 - x_n)$, which exhibits chaotic behavior for appropriate μ [3]. In the logistic map, x_n represents the state of the chaotic system at the n -th iteration (i.e., a random seed). The resulting waveforms are wideband, noise-like, and difficult to decode without knowledge of the transmitter’s internal state.

Received 31 October 2025; revised 18 December 2025 and 5 February 2026; accepted 1 March 2026. Date of publication 4 March 2026; date of current version 19 March 2026. This work is partially supported by the project RESTART (PE00000001) under the NRRP MUR program funded by the EU-NGEU, project Net4Future CUP: D93C22000910001. The associate editor coordinating the review of this letter and approving it for publication was C. Tsinos. (*Corresponding author: Marco Siino.*)

Marco Siino is with the Department of Electrical Electronic and Computer Engineering, University of Catania, 95131 Catania, Italy (e-mail: marco.siino@unict.it).

Stefano Mangione and Ilenia Tinnirello are with the Department of Engineering, University of Palermo, 90133 Palermo, Italy (e-mail: stefano.mangione.tlc@unipa.it; ilenia.tinnirello@unipa.it).

Digital Object Identifier 10.1109/LCOMM.2026.3670578

Despite their advantages, a significant challenge lies in the decoding process. Classical coherent receivers rely on precise synchronization to replicate the chaotic carrier, which is highly fragile under channel impairments like multipath fading [4], [9]. Non-coherent approaches (e.g., DCSK), while simpler, often suffer from lower spectral efficiency or noise degradation [10], [11]. Consequently, **Deep Learning (DL)** has emerged as a powerful data-driven alternative, capable of directly mapping distorted received signals to information bits, thereby relaxing synchronization requirements and adapting to diverse channel conditions [12], [13], [14], [15], [16], [17], [18]. At the receiver, rather than relying on precise synchronization to reconstruct the chaotic signal, a decoder—based on DL—directly interprets the distorted signal, learning to classify it back into the original transmitted bits and bypassing the fragile synchronization step of traditional coherent chaotic communications [3], [9].

Nevertheless, existing DL-based CSK decoders, such as those based on CNNs or LSTMs, often face a critical trade-off: they either exhibit poor Bit Error Rate (BER) under harsh channel conditions or require a large number of parameters, rendering them unsuitable for resource-constrained devices. While transformer-based architectures and attention mechanisms excel at capturing long-range dependencies [19], their substantial computational and memory requirements hinder practical deployment. Therefore, model compression techniques, and in particular pruning, become essential to obtain compact yet effective decoders.

To bridge this gap, we propose **Ultra-CAN**, an **ultra-lightweight CNN-attention decoder** for CSK. Ultra-CAN synergistically combines efficient convolutional layers for local feature extraction with a compact attention block designed to capture long-range dependencies with minimal overhead. **Our main contributions are:** 1) The **design and implementation** of a novel hybrid neural architecture that integrates 1D convolutional layers with a lightweight attention mechanism specifically tailored for chaotic signal decoding; 2) The **application** of an iterative cosine similarity-based pruning strategy to systematically reduce model redundancy without compromising decoding accuracy, resulting in a model variant requiring only ~ 200 kB of memory; 3) A comprehensive **comparative analysis** of the proposed decoder against three state-of-the-art deep learning baselines and a traditional DCSK receiver under both AWGN and Rayleigh fading channel conditions, demonstrating superior BER performance (particularly at 0 dB SNR); and 4) An architectural analysis highlighting the trade-offs between model size and BER performance, showcasing a variant requiring only ~ 200 kB of memory.

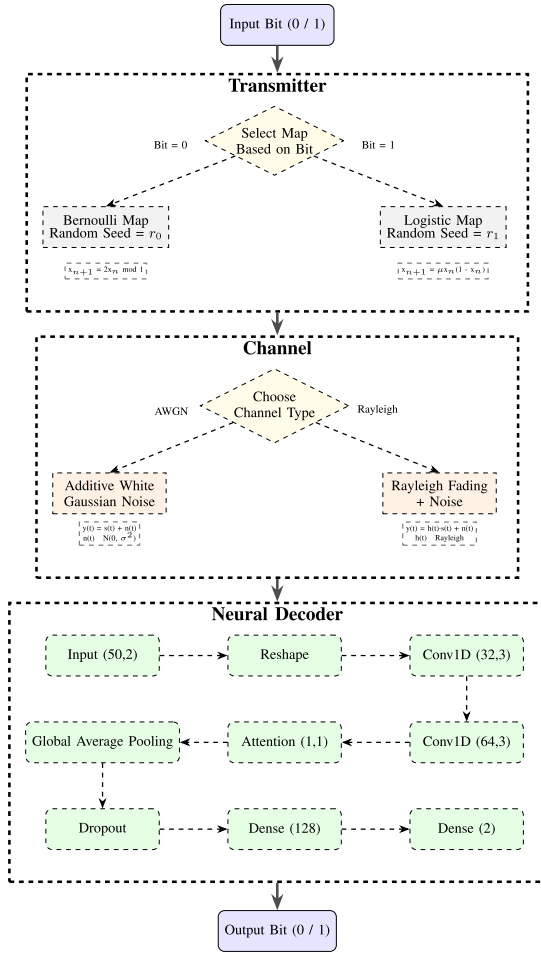


Fig. 1. Block diagram of the proposed DL-based communication system.

All code, experiments, and results are available on GitHub.¹

II. PROPOSED METHOD

Our receiver, **Ultra-CAN**, combines convolutional layers with a lightweight attention mechanism to enable robust bit recovery while keeping a compact footprint. The architecture is shown in Figure 1. The input of our model is a real-valued chaotic sequence $r[n]$ of length $N_{\text{seq}} = 100$ (i.e., 50 for real part and 50 for imaginary part of the signal).

A. Convolutional Feature Extraction

The input is reshaped into $X_{\text{input}} \in \mathbb{R}^{N_{\text{seq}} \times 1}$ and processed by two Conv1D layers. The received sequence $r[n]$ is reshaped into $X_{\text{input}} \in \mathbb{R}^{N_{\text{seq}} \times 1}$ before being fed into the convolutional layers, ensuring proper formatting for neural network processing while maintaining the temporal characteristics of the chaotic communication signal. The first has $F_1 = 32$ filters of size $K_1 = 3$:

$$H^{(1)}[i, f] = \text{ReLU} \left(W_f^{(1)} * X_{\text{input}} \right) [i] + B_f^{(1)} \quad (1)$$

The equation describes the first convolutional layer, where $*$ is the convolution operator, and $W_f^{(1)}$ and $B_f^{(1)}$ are the weights and bias for filter f , respectively.

The second layer uses $F_2 = 64$ filters, $K_2 = 3$:

$$H^{(2)}[i, f] = \text{ReLU} \left(W_f^{(2)} * H^{(1)} \right) [i] + B_f^{(2)} \quad (2)$$

The first convolutional layer with $F_1 = 32$ filters of size $K_1 = 3$ produces output $H^{(1)} \in \mathbb{R}^{98 \times 32}$, while the second layer with $F_2 = 64$ filters of size $K_2 = 3$ yields $H^{(2)} \in \mathbb{R}^{96 \times 64}$. The progressive reduction in sequence length results from valid convolution operations.

B. Lightweight Convolutional Attention

To capture long-range dependencies at low cost, we apply a Conv1D layer with $K_A = 1$ and $F_A = 1$ to $H^{(2)}$, yielding attention weights:

$$\alpha[i] = \sigma \left(\sum_{h=0}^{F_2-1} W_{0,h}^{(A)} H^{(2)}[i, h] + B^{(A)} \right) \quad (3)$$

The attended features are:

$$H^{\text{att}}[i, f] = H^{(2)}[i, f] \cdot \alpha[i] \quad (4)$$

C. Pooling and Classification

Global average pooling condenses H^{att} into

$$V_f = \frac{1}{N_{\text{seq}}} \sum_{i=0}^{N_{\text{seq}}-1} H^{\text{att}}[i, f] \quad (5)$$

After dropout, two dense layers follow: first 128 ReLU units,

$$O_j^{(1)} = \text{ReLU} \left(\sum_f W_{f,j}^{(D1)} V_f + B_j^{(D1)} \right) \quad (6)$$

then a 2-unit Softmax output:

$$P_k = \frac{e^{O_k^{(2)}}}{\sum_{j=0}^1 e^{O_j^{(2)}}}, \quad (7)$$

$$O_k^{(2)} = \sum_j W_{j,k}^{(D2)} O_j^{(1)} + B_k^{(D2)} \quad (8)$$

D. Cosine Similarity Pruning

To achieve an ultra-light footprint while preserving performance, we employ iterative cosine similarity pruning. This automated procedure identifies and removes redundant convolutional filters with high feature similarity, progressively reducing model complexity while monitoring BER degradation. The algorithm iteratively prunes the most similar convolutional filter pairs (i.e., W_i and W_j), fine-tunes the network, and validates performance, ensuring the final compressed model maintains reliable decoding capability. We prune redundant convolutional filters based on cosine similarity. For filters W_i, W_j :

$$\begin{aligned} \text{sim}(W_i, W_j) &= \frac{W_i \cdot W_j}{\|W_i\| \|W_j\|} \\ &= \frac{\sum_{k,d} W_i[k, d] W_j[k, d]}{\sqrt{\sum_{k,d} W_i[k, d]^2} \sqrt{\sum_{k,d} W_j[k, d]^2}} \end{aligned} \quad (9)$$

Highly similar filters ($\text{sim} > \tau$) are pruned iteratively until BER degradation exceeds tolerance.

¹<https://github.com/marco-siino/commlett>

III. EXPERIMENTAL EVALUATION

For the experimental evaluation, a comprehensive training dataset was generated offline to facilitate the training of both our proposed model and benchmark architectures. The training dataset comprised a total of 105,000 symbols for training. A total of 5,000 symbols per each SNR point were generated. For the final evaluation and BER computation, a separate testing methodology was adopted. For each SNR point, testing continued until at least 100 bit errors were accumulated, ensuring statistically significant BER estimates with narrow confidence intervals for reliable performance comparison. To ensure computational efficiency and avoid excessively long test runs for very low BERs, a maximum ceiling of 10 million samples was generated per SNR point. The resulting BER performance curves, derived from this rigorous testing, are presented and thoroughly discussed in the subsequent results section. For a robust comparison with the recent literature, two training dataset versions were created. The first spanned SNRs from 0–20 dB, exposing models to both noisy and clean conditions. The second targeted 11–15 dB, following recent findings [18] that identify this range as most relevant for practical systems. During the training, a validation set - with the same number and distribution of samples of the training set - was used to monitor generalization and apply **early stopping**, halting training if the validation loss failed to improve for 10 consecutive epochs.

The generation process involved simulating chaotic map transmissions (i.e., logistic and Bernoulli maps) across various SNR levels. In our proposed scheme, the random seed for chaotic map generation is **regenerated for each transmitted symbol** to ensure waveform diversity and enhance security. The experimental evaluation employed two fundamental channel models to assess decoder performance under different propagation conditions. For the **AWGN channel**, complex-valued additive white Gaussian noise was added to the transmitted signal, with noise power calculated based on the target SNR and the average signal power. The noise was generated independently for both real and imaginary components. For the **Rayleigh fading channel**, an uncorrelated fading model was implemented where each subcarrier experienced independent Rayleigh fading modeled as complex Gaussian coefficients. The faded signal was power-normalized before AWGN addition to maintain consistent average signal power across channel conditions, ensuring fair SNR comparison between AWGN and fading scenarios.

To thoroughly assess the efficacy and superior performance of our architecture, its performance was compared against three DL-based demodulators from the state-of-the-art literature, along with traditional non-deep DCSK approach. These benchmark models were chosen for their relevance in the context of DL for communication systems and their diverse architectural approaches. The first benchmark is the **MC-DLCSK** (Multi-Carrier Deep Learning Chaotic Shift Keying) [18] demodulator. For this benchmark, we adopted the standard configuration with $M = 4$ subcarriers and a spreading factor $\beta = 50$ to match the proposed input dimensions. It is worth noting that our reproduction of the MC-DLCSK baseline achieves slightly superior BER performance compared to the

original results reported in [18]. This improvement is attributed to our use of the Adam optimizer (as opposed to SGD used in [18]) and a larger training dataset (1.05×10^5 samples vs. 4×10^4 in [18]), which allowed the benchmark model to reach better convergence. The second benchmark is an **LSTM-OFDM-DCSK** [20], a recurrent neural network architecture known for its strong temporal sequence processing capabilities. Finally, we included the **TDNN-LSTM** (Time Delay Neural Network - Long Short-Term Memory) model [21], which combines the strengths of TDNNs in capturing local temporal dependencies with LSTMs' ability to model long-range contexts. All benchmark models were trained and evaluated under the identical hardware and dataset conditions outlined previously to ensure a fair and consistent comparison against our proposed compact solution. However, we differentiate some settings according to the authors recommendations (i.e., number of epochs, batch size, number of parameters).

A. Experimental Setup

All computational experiments and model training were conducted on a single NVIDIA Tesla T4 GPU utilizing the TensorFlow framework, ensuring a standardized environment. Memory allocation comprised 13.28 GB of total RAM, complemented by approximately 75 GB of available disk space for dataset storage and model checkpoints. The software stack predominantly featured Python 3, with deep learning models developed using the TensorFlow 2 and Keras API, alongside standard scientific libraries such as NumPy and Matplotlib for data handling and visualization. This standardized and robust environment ensured the reproducibility and integrity of our experimental results. We used Adam optimizer and an early stopping training strategy up to a maximum of 30 epochs.

IV. RESULTS

We evaluate Ultra-CAN against state-of-the-art chaotic demodulators under both AWGN and Rayleigh fading channels. As shown in Figures 2 and 3, the proposed Ultra-CAN decoder consistently outperforms all benchmark methods across the majority of the evaluated SNR range.

Fig. 2 shows that Ultra-CAN consistently achieves the lowest BER across the entire SNR range in AWGN, outperforming MC-DLCSK, TDNN-LSTM, and LSTM-OFDM-DCSK. The improvement is especially evident at low SNRs, where conventional approaches fail to maintain reliable performance. Also in the easier high-SNR regime (11–15 dB) - not shown in the Figure - where benchmarks like TDNN-LSTM tend to saturate, Ultra-CAN maintains a lower error floor, demonstrating high precision in clean channel conditions. In Rayleigh fading (Fig. 3), Ultra-CAN again demonstrates superior robustness and performance up to 5 dB SNR, after which MC-DLCSK slightly outperforms it. However Ultra-CAN remains highly competitive and strictly outperforms the LSTM-based and traditional benchmarks, providing the best trade-off given its reduced computational footprint. In the Rayleigh channel the performance of all the models are more modest compared to the AWGN channel. Fig. 4 compares Ultra-CAN variants in AWGN. The default configuration offers the best BER in

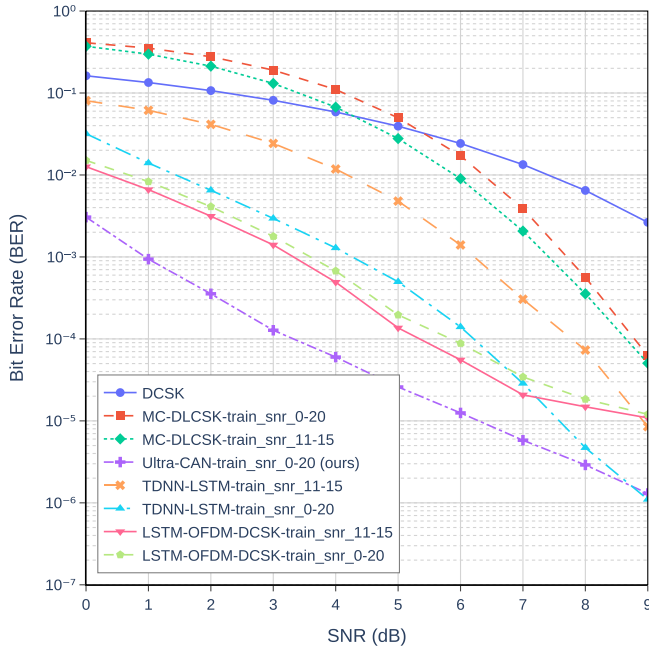


Fig. 2. BER comparison in AWGN. ultra-CAN consistently achieves the lowest BER across all SNRs compared to the benchmarks: MC-DLCSK [18], TDNN-LSTM [21] and LSTM-OFDM-DCSK [20].

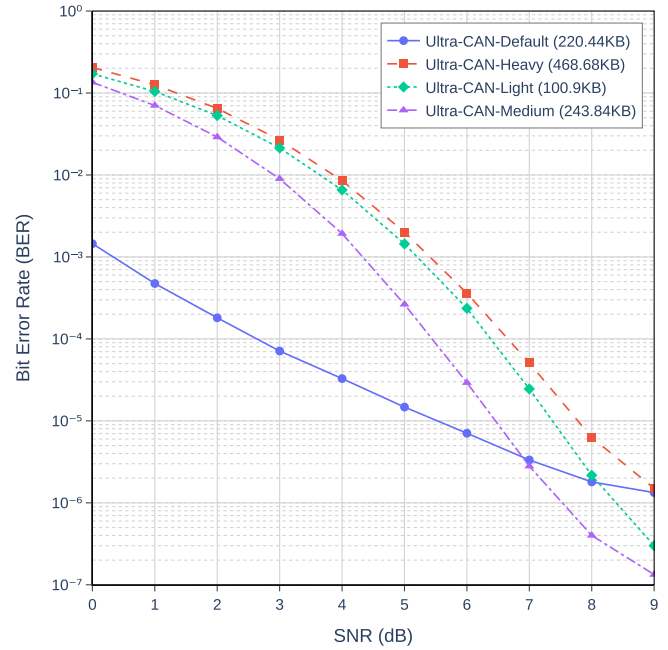


Fig. 4. BER of Ultra-CAN variants in AWGN. Default achieves best accuracy, light version is most compact.

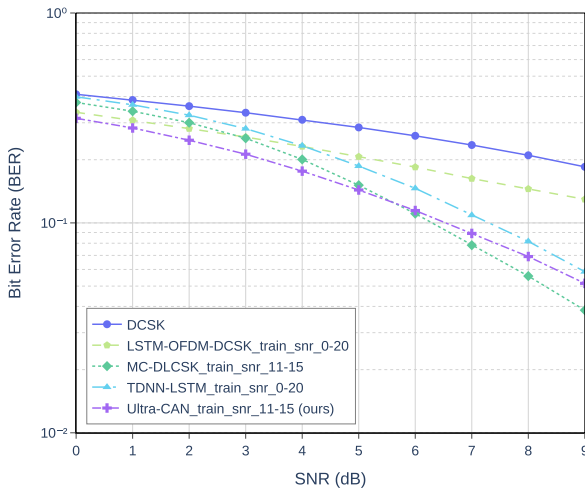


Fig. 3. BER comparison in rayleigh fading. Ultra-CAN achieves lowest BER up to 5 dB, compared against benchmarks MC-DLCSK [18], TDNN-LSTM [21], and LSTM-OFDM-DCSK [20].

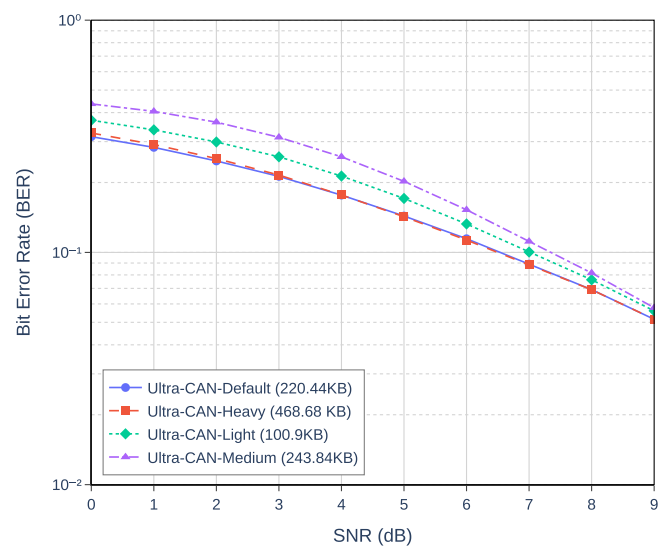


Fig. 5. BER of Ultra-CAN variants in Rayleigh fading. Default again dominates, light version remains competitive.

the critical low-SNR regime (0–7 dB). However, at higher SNRs (>7 dB), the light and medium variants exhibit superior performance, likely due to the regularization effect of pruning which enhances generalization when channel noise is low. This highlights a trade-off: the default model provides maximum robustness against heavy noise, while the light variant offers extreme efficiency and excellent clarity in better channel conditions. Fig. 5 confirms that across channel types, Ultra-CAN-Default provides the most reliable balance, while Ultra-CAN-Light offers the smallest footprint with acceptable BER.

Overall, Ultra-CAN is the only architecture among those tested that explicitly integrates attention, which proves to be

effective in extracting salient features and achieving strong BER performance with reduced model size. Ultra-CAN still maintains superiority at low SNRs (up to 5 dB) due to its attention mechanism, which helps in focusing on less corrupted parts of the signal. Finally, across all examined channel scenarios and the entire 0–20 dB SNR range, **all variants** of our DL model, including the most compact **Ultra-CAN-Light** configuration, consistently and significantly surpass the BER performance of the conventional **DCSK** system. This substantial and consistent gain confirms the inherent efficacy of the AI-based approach for robust decoding of chaotic signals compared to traditional correlator receivers.

TABLE I
MODEL EFFICIENCY COMPARISON

Model	Attention	Size (kB)	Params	MFLOPs
LSTM-ODFM-DCSK	NO	162.9	3.2k	0.21
MC-DLCSK	NO	364.3	21k	2.66
TDNN-LSTM	NO	719.9	55.4k	5.63
Ultra-CAN-Default	YES	220.4	14.9k	1.65
Ultra-CAN-Light	YES	100.9	4.7k	0.63
Ultra-CAN-Medium	YES	243.8	16.1k	3.98
Ultra-CAN-Heavy	YES	468.7	36.1k	5.98

A. Model Efficiency Comparison

Table I summarizes efficiency metrics. While LSTM-ODFM-DCSK has the lowest parameter count, Ultra-CAN variants achieve the best trade-off between compactness and accuracy. Notably, Ultra-CAN-Light requires only 4.7k parameters and 0.63 MFLOPs, yet outperforms heavier models in low-SNR regimes, providing an ultra-light solution suitable for resource-constrained edge devices while still maintaining competitive BER in low-SNR scenarios. The optimal BER performance demonstrated in Fig. 2 and Fig. 3 is achieved using the **Ultra-CAN-Default** configuration, which employs $F_1 = 32$ and $F_2 = 64$ convolutional filters, resulting in 14.9k parameters. This balanced configuration (220.44 kB) is selected to maximize decoding accuracy. To address the critical **trade-off between BER performance and computational complexity**, we leverage different architectural variants and structural pruning. This confirms the viability of compact, attention-augmented decoders for resource-constrained chaotic communication systems.

V. CONCLUSION

This letter presented **Ultra-CAN**, an ultra-light CNN-attention decoder for CSK. By combining compact convolutional layers with a lightweight attention block, Ultra-CAN achieves robust bit recovery while maintaining a model size of ~ 200 kB. **Simulation results confirm the suitability of the proposed method for both static and multipath environments:** experiments show superior BER performance at low SNRs in AWGN and demonstrate robust decoding capabilities in **Rayleigh fading channels**, outperforming state-of-the-art DL-based benchmarks in severe noise conditions. The use of cosine similarity-based pruning further reduces redundancy without degrading accuracy.

Our study highlights the trade-off between complexity and performance in chaotic decoding, showing that compact attention mechanisms can be employed for efficient and reliable communications. Future work will extend Ultra-CAN to more complex chaotic maps, resilience against jamming, and adaptive channel estimation.

AUTHOR CONTRIBUTIONS

Marco Siino: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Resources, Software, Visualization, Writing—Original Draft, Writing—Review and Editing. Stefano Mangione: Investigation, Validation, Writing—Review and Editing. Ilenia Tinnirello: Investigation, Supervision, Writing—Review and Editing.

REFERENCES

- [1] 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, p. 4762, Jul. 2021.
- [2] X. Cai, C. Huang, P. Chen, E. Basar, and C. Yuen, "Design of non-coherent RIS-empowered DCSK with two-level nested index modulation," *IEEE Trans. Wireless Commun.*, vol. 24, no. 4, pp. 3044–3058, Apr. 2025.
- [3] G. Kaddoum, "Wireless chaos-based communication systems: A comprehensive survey," *IEEE Access*, vol. 4, pp. 2621–2648, 2016.
- [4] X. Cai, W. Xu, L. Wang, and G. Kaddoum, "Joint energy and correlation detection assisted non-coherent OFDM-DCSK system for underwater acoustic communications," *IEEE Trans. Commun.*, vol. 70, no. 6, pp. 3742–3759, Jun. 2022.
- [5] X. Cai, C. Yuen, C. Huang, W. Xu, and L. Wang, "Toward chaotic secure communications: An RIS enabled M-ary differential chaos shift keying system with block interleaving," *IEEE Trans. Commun.*, vol. 71, no. 6, pp. 3541–3558, Jun. 2023.
- [6] X. Cai, D. Wang, W. Xu, L. Wang, and F. C. M. Lau, "Design of carrier index keying-aided M-ary differential chaos cyclic shift keying for D2D communications," *IEEE Trans. Commun.*, vol. 71, no. 6, pp. 3233–3250, Jun. 2023.
- [7] M. Mobini, G. Kaddoum, and M. Herceg, "Design of a SIMO deep learning-based chaos shift keying (DLCSK) communication system," *Sensors*, vol. 22, no. 1, p. 333, Jan. 2022.
- [8] L. M. Pecora and T. L. Carroll, "Synchronization in chaotic systems," *Phys. Rev. Lett.*, vol. 64, no. 8, pp. 821–824, 1990.
- [9] B. Jovic, *Synchronization Techniques for Chaotic Communication Systems*. Cham, Switzerland: Springer, 2011.
- [10] C. Bai, X.-H. Zhao, H.-P. Ren, G. Kolumbán, and C. Grebogi, "Double-stream differential chaos shift keying communications exploiting chaotic shape forming filter and sequence mapping," *IEEE Trans. Wireless Commun.*, vol. 21, no. 7, pp. 4954–4972, Jul. 2022.
- [11] G. Narang, M. Aggarwal, H. Kaushal, A. Kumar, and S. Ahuja, "Performance analysis of differential chaos shift keying in free space optical communication with diversity techniques," *IEEE Access*, vol. 11, pp. 54438–54447, 2023.
- [12] Y. Fang, Y. Pan, H. Ma, D. Ma, and M. Guizani, "A novel DCSK-based linear frequency modulation waveform design for joint radar and communication systems," *IEEE Trans. Green Commun. Netw.*, vol. 9, no. 1, pp. 354–366, Mar. 2025.
- [13] H. Yang, W. K. S. Tang, G. Chen, and G.-P. Jiang, "Multi-carrier chaos shift keying: System design and performance analysis," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 64, no. 8, pp. 2182–2194, Aug. 2017.
- [14] J. Zou, Y. Tao, Y. Fang, H. Ma, and Z. Yang, "Receiver design for ICI-CSK system: A new perspective based on GRU neural network," *IEEE Commun. Lett.*, vol. 27, no. 11, pp. 2983–2987, Nov. 2023.
- [15] Y. Tao, Y. Fang, P. Chen, H. Ma, and M. Guizani, "Matrix reconstruction algorithm-assisted multi-carrier DCSK scheme: An effective solution for frequency-selective fading channel," *IEEE Wireless Commun. Lett.*, vol. 12, no. 11, pp. 1941–1945, Nov. 2023.
- [16] M. Mobini, M. Herceg, and G. Kaddoum, "Design of an M-ary DLCSK communication system using deep transfer learning," *IEEE Open J. Commun. Soc.*, vol. 4, pp. 2318–2342, 2023.
- [17] F. J. Escribano, G. Kaddoum, A. Wagemakers, and P. Giard, "Design of a new differential chaos-shift-keying system for continuous mobility," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 2066–2078, May 2016.
- [18] M. Mobini, G. Kaddoum, and M. Herceg, "A multi-carrier deep learning CSK communication system," *IEEE Commun. Lett.*, vol. 29, no. 6, pp. 1191–1195, Jun. 2025.
- [19] A. Vaswani et al., "Attention is all you need," in *Proc. Adv. Neural Inform. Process. Syst. (NIPS)*, 2017, pp. 5998–6008.
- [20] L. Zhang, H. Zhang, Y. Jiang, and Z. Wu, "Intelligent and reliable deep learning LSTM neural networks-based OFDM-DCSK demodulation design," *IEEE Trans. Veh. Technol.*, vol. 69, no. 12, pp. 16163–16167, Dec. 2020.
- [21] H. Zhang, L. Zhang, Y. Jiang, and Z. Wu, "Reliable and secure deep learning-based OFDM-DCSK transceiver design without delivery of reference chaotic sequences," *IEEE Trans. Veh. Technol.*, vol. 71, no. 8, pp. 8059–8074, Aug. 2022.