

DualAttWaveNet: Multiscale Attention Networks for Satellite Interference Detection

Chunyu Yang, Boyu Yang, Kun Qiu, Zhe Chen, Yue Gao

School of Computer Science, Fudan University, China

{22307140114, 24110240144}@m.fudan.edu.cn, {qkun, zhechen, gao.yue}@fudan.edu.cn

Abstract—The escalating overlap between non-geostationary orbit (NGSO) and geostationary orbit (GSO) satellite frequency allocations necessitates accurate interference detection methods that address two pivotal technical gaps: 1) computationally efficient signal analysis for real-time operation, and 2) robust anomaly discrimination under varying interference patterns. Existing deep learning approaches employ encoder-decoder anomaly detectors that threshold input-output discrepancies for robustness. While the transformer-based TrID model achieves state-of-the-art performance (AUC: 0.8318, F1: 0.8321), its multi-head attention incurs prohibitive computational costs, and its decoupled training of time-frequency models overlooks cross-domain dependencies. To overcome these problems, we propose DualAttWaveNet. A bidirectional attention fusion layer dynamically correlates time-domain samples using parameter-efficient cross-attention routing. A wavelet-regularized reconstruction loss enforces multi-scale consistency. We train the model on public dataset which consists of 28 hour of satellite signals. Experiments show compared to TrID, DualAttWaveNet improves AUC by 12% and reduces inference time by 50% while maintaining F1-score.

Index Terms—interference detection, multimodal fusion, bidirectional attention, wavelet transform

I. INTRODUCTION

The accelerated deployment of low Earth orbit (LEO) satellite systems poses huge challenges for next-generation communication networks, with over 20,000 satellites projected to be launched by leading operators including SpaceX’s Starlink [1] and Starshield [2], as well as Eutelsat OneWeb [3]. These mega-constellations have become critical infrastructure to enable global connectivity, driving the commercialization of space-based communications while expanding broadband access to underserved regions [4]. However, the exponential growth in satellite numbers brings fundamental technical obstacles. Rising risk of spectrum overlap between LEO and geosynchronous orbit (GSO) satellites creates urgent demands for scalable interference management frameworks that can evolve with expanding LEO networks.

Current research in satellite interference management focuses on three main approaches: preventive measures to reduce risks before system deployment [5], [6], static mitigation methods applied after interference occurs [7], [8], and simulation-based models that predict interference using limited time or location samples [7]. While useful in controlled settings, these methods struggle in real-world environments. Preventive measures depend on assumptions that may not hold after deployment, static mitigation cannot adapt to changing interference patterns, and simulation models fail to handle unexpected environmental

shifts like solar radiation changes or atmospheric variations [9]. Additionally, traditional detection techniques relying on fixed thresholds or static signal features [7] cannot reliably identify complex, real-time interference. To address these issues, new detection solutions must achieve both fast real-time response and consistent accuracy across diverse interference conditions.

Current approaches to interference detection in satellite communications are broadly divided into traditional analytical methods and machine learning (ML)-based techniques. Conventional methods primarily rely on energy detection (ED) that computes signal energy over predefined intervals for threshold-based anomaly identification [10], or exploit spectral cyclostationary features to distinguish interference from periodic signals [11]. However, these approaches depend on manually calibrated thresholds and exhibit limited sensitivity in low signal-to-noise ratio (SNR) conditions. In contrast, ML-driven methods eliminate such constraints by automatically learning discriminative interference signatures from raw data. Classification-based solutions use deep neural networks to establish data-driven decision boundaries [12], while encoder-decoder architectures treat detection as an anomaly identification task by reconstructing interference-free waveforms from corrupted inputs [13]. Recent advancements further deploy transformer models to capture long-range spectral dependencies, improving detection accuracy for persistent anomalies [14].

Despite advancements, critical challenges persist in existing interference detection methods. First, threshold-dependent traditional methods - exemplified by energy detection - exhibit degraded reliability in low-SNR regimes, where static thresholds fail to dynamically adapt to noise fluctuations or interference intensity variations, leading to frequent false negatives in rapidly evolving orbital conditions [14]. Second, while attention-driven architectures achieve state-of-the-art detection sensitivity, their computational overhead from multi-head attention mechanisms imposes huge latency during both training and inference, making them unsuitable for resource-constrained satellite edge devices. Third, contemporary deep learning models often process time-domain and frequency-domain signal representations in isolation by training separate networks for each modality. While this decoupled approach achieves moderate performance (e.g., 0.8318 AUC for time-domain models and 0.7106 AUC for frequency-domain models alone), it fundamentally limits detection capability. In this paper we demonstrate that unified architectures jointly learning from both domains can elevate AUC to 0.9327—exposing a 12-22%

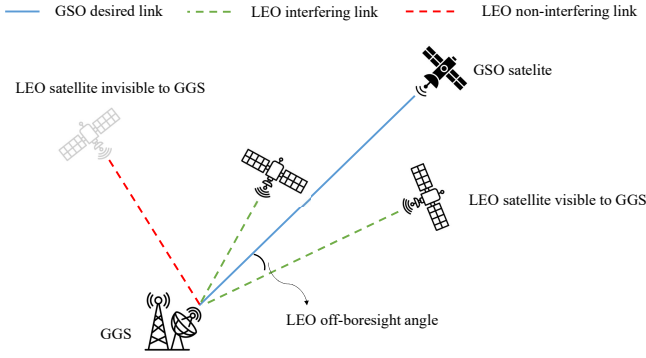


Fig. 1. Interference scenario between GSO and LEO satellite systems.

performance gap compared to isolated modality training.

To overcome these challenges, we propose DualAttWaveNet, a unified model integrating cross-domain signal fusion for real-time interference detection. Our primary contributions are threefold:

- 1) The proposed architecture jointly processes time-domain IQ samples and their frequency-domain representations through a bidirectional attention mechanism. This design eliminates the need for explicit multi-head computation while enabling adaptive correlation learning between spectral and temporal features.
- 2) To enhance robustness against minor signal fluctuations, we propose a wavelet-constrained reconstruction loss. This is achieved by applying discrete wavelet transform (DWT) decomposition to both raw and reconstructed signals, enforcing cross-band consistency through multi-scale subspace constraints.
- 3) Comprehensive evaluation on a public satellite communication dataset (48 hours duration) demonstrates DualAttWaveNet's superiority: 12% higher AUC and 50% faster inference compared to state-of-the-art baselines, while maintaining competitive F1-score.

The rest of the paper is organized as follows. Section II introduces background and motivations. Section III details the architecture of DualAttWaveNet. Section IV conducts experiments and analysis. Finally, Section V concludes the paper.

II. BACKGROUND

A. Interference Scenario

Satellite communication systems using the same frequency band often interfere with each other. As shown in Fig. 1, we focus on two types of satellites. GSO Satellite flies in a fixed position 36,000 km above Earth. It serves as the main signal source for a ground station (GGS). LEO Satellites move rapidly at 500-2,000 km altitudes. Their signals can interfere with GSO signals due to spectrum overlap. The composite received waveform of GGS contains both desired carrier signals and interference.

The received GSO carrier power is determined through classical satellite link relationships:

$$C = \frac{\text{EIRP}_{\text{gso}} \cdot G_{r, \text{gso}}(\theta_0)}{L_{\text{FS, gso}} \cdot L_{\text{add}}} \quad (1)$$

where EIRP is GSO satellite equivalent isotropic radiated power, $G_{r, \text{gso}}(\theta_0)$ denotes the maximum receive antenna gain at boresight angle θ_0 , $L_{\text{FS, gso}}$ represents free-space path loss, and L_{add} accounts for aggregate atmospheric impairments.

Interference contributions from K LEO satellites are modeled as individual components:

$$I_k = \frac{\text{EIRP}_k \cdot G_{r, k}(\theta_k) \cdot B_{\text{adj}, k}}{L_{\text{FS, k}} \cdot L_{\text{add}}} \quad (2)$$

The angular gain term $G_{r, k}(\theta_k)$ reflects spatial relationships caused by LEO orbital motion, while $B_{\text{adj}, k} \in [0, 1]$ is the spectral overlap between GSO and LEO transmissions.

The signal received by physical layer at the GGS has three components:

$$\begin{aligned} y(t) = & x(t)\sqrt{\text{CNR}} \text{ (Desired GSO)} \\ & + \sum_{k=1}^K i_k(t)e^{j2\pi\Delta f_k t}\sqrt{\text{INR}_k} \text{ (LEO interference)} \\ & + \zeta(t) \text{ (Thermal noise)} \end{aligned} \quad (3)$$

where $\Delta f_k = f_{c, k} - f_{c, \text{gso}}$ captures carrier frequency offsets from Doppler effects. The exponential terms induce time-varying phase rotations proportional to relative satellite motion. Here, CNR (Carrier-to-Noise Ratio) and INR_k (Interference-to-Noise Ratio) respectively characterize the desired signal quality and interference intensity relative to the noise floor,

Dual signal representations are derived for machine learning processing:

- Time-domain: y^A captures instantaneous amplitude variations through uniform sampling
- Frequency-domain: Welch's power spectral density estimation generates logarithmic magnitude spectra via overlapping windowed transforms: $y^F = 10 \log_{10}(\phi(y(t)))$

B. Deep Learning Approaches for Interference Detection

Recent advancements in machine learning have reshaped interference detection paradigms through a transition from traditional signal processing to anomaly discrimination. A popular strategy employs encoder-decoder architectures by training models to reconstruct idealized interference-free waveforms from raw inputs. This approach builds on the fundamental principle that interference anomalies create deviations from primary signals. Unlike regression-based methods that face challenges in modeling heterogeneous interference sources, reconstruction frameworks directly isolate deviations through input/output pattern comparisons.

Early implementations by Pellaco et al. [12] developed a long short-term memory autoencoder (LSTMAE) for detecting multi-scale anomalies in non-geostationary satellite signals, but its recursive architecture limits computational parallelism,

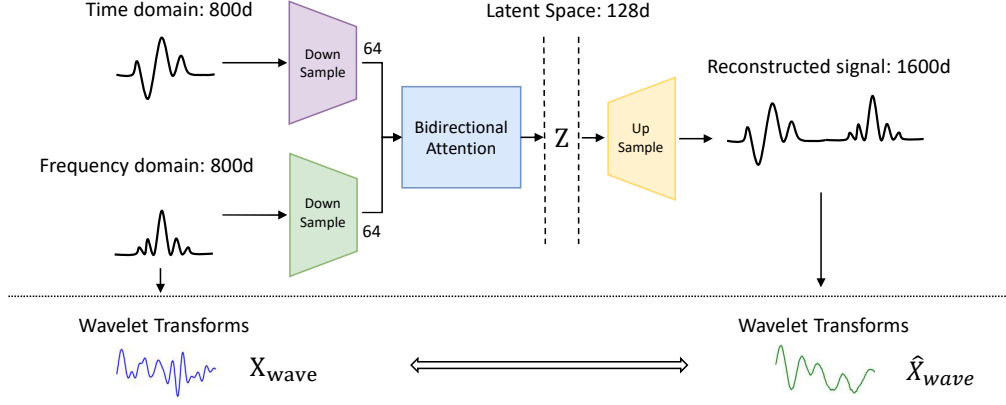


Fig. 2. Summary of our approach. Architecture of the cross-modal signal enhancement model showing dual-branch processing: time domain (800d) and frequency domain (800d) inputs are downsampled to 64d, fused through bidirectional attention into a 128d latent space (\mathcal{Z}), then upsampled to produce a reconstructed 1600d signal. Wavelet transforms (X_{wave} and \hat{X}_{wave}) are applied to the original and reconstructed signals.

leading to inference delays incompatible with latency-sensitive operations. Subsequent work by Saifaldawla et al. [13] addressed temporal dependency constraints through convolutional autoencoders (CAE), enabling direct in-phase/quadrature (IQ) sample processing for terrestrial interference detection.

Recent studies [14] integrate transformer architectures into detection pipelines to resolve long-range spectral dependencies. Through self-attention mechanisms, these models achieve superior performance in capturing global interference signatures that conventional convolutional filters or recurrent units lack. Further innovations introduced dual-domain analysis by expanding input representations to joint time-frequency spaces.

Specifically, each input sample consisting of 800 sequential time-domain sampling points is first encoded through stacked transformer layers into a 128-dimensional latent representation. The decoder then inversely maps this latent space back to a reconstructed signal through learned upsampling operations. During training, a naïve mean squared error (MSE) loss between the original and reconstructed sequences drives the model to emulate interference-free waveform patterns. At inference, anomalous samples that deviate from interference-free signals will have a larger difference when reconstructed. Comparing this error with a predefined threshold will determine the presence of interference.

However, transformer-based detectors exhibit quadratic computational complexity growth with sequence length, posing challenges for wideband satellite implementations. Current approaches yet often employ separated processing for temporal and spectral inputs, neglecting cross-domain correlations that could improve detection robustness through synergistic feature learning.

III. PROPOSED DEEP LEARNING MODEL

A. Overview Design

We propose DualAttWaveNet, an autoencoder that takes both time and frequency domain signal as input, and try to reconstruct both of them. The model (show in Fig. 2) consists

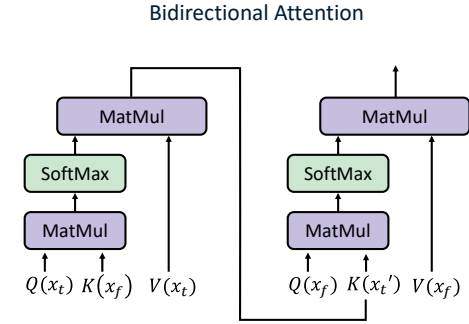


Fig. 3. DualAttnWaveNet's bidirectional attention mechanism. Both attention phase share the same parameter Q, K and V . To decrease channel numbers, they are implemented as 1d-Convolution.

of separate encoder and decoder for both domains, and a fusion module that utilizes bidirectional attention before concatenating. The loss function is regularized with wavelet transform to enforce reconstruction fidelity across multiple scales. The input signals from both time and frequency domain are of shape $B \times 800$, where B is the batch size. After going through separate convolution modules acting as downsampling layers, the signals are transformed into $B \times 128$ latent representations. Then these latent representations go through the bidirectional attention module and are concatenated before going through the upsampling layers. The final output is of shape $B \times 1600$. The model is trained with a combination of MSE loss and wavelet loss, described in Section III-C.

B. Bidirectional Attention

We propose a parameter-efficient mutual attention mechanism for cross-modal feature fusion. Unlike conventional multi-head attention in [15], our design employs single-head dot-product attention with spatial reduction to minimize computational overhead while maintaining inter-domain alignment capacity, as show in Fig. 3.

Given input feature maps $\mathbf{X} \in \mathbb{R}^{B \times C \times L}$ (temporal domain) and $\mathbf{Y} \in \mathbb{R}^{B \times C \times L}$ (spectral domain), the mutual attention operator computes:

$$\begin{aligned} \text{MutualAttn}(\mathbf{X}, \mathbf{Y}) &= \mathbf{X} + \gamma \cdot \text{AttentionGate}(\mathbf{X}, \mathbf{Y}) \\ \text{AttentionGate}(\mathbf{X}, \mathbf{Y}) &= \mathbf{V}_y \cdot \text{Softmax} \left(\frac{\mathbf{Q}_x \mathbf{K}_y^\top}{\sqrt{d}} \right) \end{aligned} \quad (4)$$

where γ is a learnable scalar initialized to 0, $\mathbf{Q}_x = \mathcal{W}_Q(\mathbf{X}) \in \mathbb{R}^{B \times L \times \frac{C}{8}}$, $\mathbf{K}_y = \mathcal{W}_K(\mathbf{Y}) \in \mathbb{R}^{B \times \frac{C}{8} \times L}$, and $\mathbf{V}_y = \mathcal{W}_V(\mathbf{Y}) \in \mathbb{R}^{B \times C \times L}$. The projection matrices $\mathcal{W}_{Q,K,V}$ implement 1D convolutions with kernel size 1, reducing channel dimensions by $8\times$ for queries/keys to optimize memory footprint.

The compatibility scores between temporal (\mathbf{X}) and spectral (\mathbf{Y}) features are computed through matrix multiplication of the reduced-channel representations. This produces an $L \times L$ affinity matrix that represents position-wise cross-domain correlations before applying row-wise softmax normalization.

To preserve original feature stability during early training, the residual connection is initially dampened ($\gamma = 0$) and progressively strengthened through learning. The symmetric architecture applies identical attention operations in both temporal \rightarrow spectral and spectral \rightarrow temporal directions, in sequential form:

$$\begin{aligned} \hat{\mathbf{X}} &= \text{MutualAttn}(\mathbf{X}, \mathbf{Y}) \\ \hat{\mathbf{Y}} &= \text{MutualAttn}(\mathbf{Y}, \mathbf{X}) \end{aligned} \quad (5)$$

This bidirectional design shares parameters across both call of attention, enables joint refinement of both modalities without major computational overhead.

C. Wavelet-Domain Spectral Regularization

Traditional MSE loss is not enough to capture the complex structure of the signal. We propose a wavelet loss term to enforce the reconstructed output to be consistent across all scales. Given an input tensor sequence $\mathbf{x} \in \mathbb{R}^{B \times C \times L}$ where B , C , and L denote batch size, channels, and temporal length respectively, we construct a learnable filter bank $\mathcal{F} \in \mathbb{R}^{S \times 1 \times K}$ through the following parameterization:

$$\mathcal{W}_s(\tau) = \text{Norm} \left(\cos \left(\frac{t}{s} \tau \right) \odot \mathcal{G}(\tau, s) \right) \quad (6)$$

where $\mathcal{G}(\tau, s) = \exp(-\tau^2/(2s^2))$ represents the Gaussian envelope function, $s \in \mathbb{S}$ enumerates the wavelet scale parameters, and $K = 4s_{\max}$ determines the kernel size from the maximum scale s_{\max} to ensure sufficient spatial support. Each filter undergoes L_2 -normalization to preserve energy consistency across multiple scales.

The discrete wavelet transform is implemented as asymmetric depth-wise separable convolution with spectral preservation properties:

$$\begin{aligned} \mathbf{X}_{\text{wave}} &= \text{Conv1D}(\mathbf{X}, \mathcal{F}) \\ &= \bigcup_{s \in \mathbb{S}} \mathbf{X} * \mathcal{W}_s \in \mathbb{R}^{B \times C \times S \times L} \end{aligned} \quad (7)$$

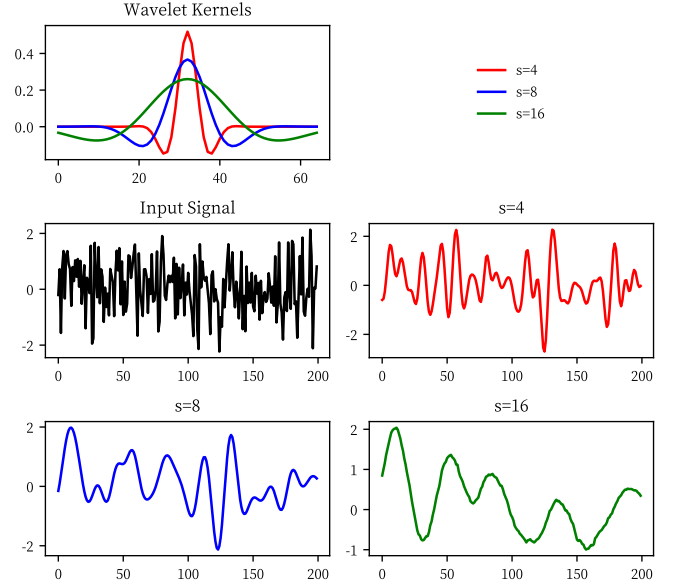


Fig. 4. (Top) Normalized Morlet wavelet kernels spanning scales $s \in \{4, 8, 16\}$, designed to capture variable temporal-spectral features. (Bottom) Example analysis of a synthetic signal: Gaussian white noise (top-left) decomposed into distinct frequency components (red: $s = 4$, high-frequency; blue: $s = 8$, mid-frequency; green: $s = 16$, low-frequency). Larger scales broaden temporal support for low-frequency emphasis, while smaller scales localize high-frequency transients.

Reflection padding strategy mitigates boundary artifacts while maintaining temporal resolution through stride-1 convolution across S parallel scale dimensions. Fig. 4 illustrates how wavelet kernels at different scales filter input signals.

Our joint optimization framework combines both signal-space and wavelet-domain consistency through an adaptive loss weighting mechanism:

$$\mathcal{L} = \lambda_1 \|\hat{\mathbf{X}} - \mathbf{X}\|_2^2 + \lambda_2 \sum_{s \in \mathbb{S}} \|\hat{\mathbf{X}}_{\text{wave}}^{(s)} - \mathbf{X}_{\text{wave}}^{(s)}\|_2^2 \quad (8)$$

where λ_1 and λ_2 represent empirically optimized weighting and the second term denotes the norm over wavelet subbands. This dual-domain regularization ensures simultaneous preservation of temporal features and spectral distribution.

D. Threshold Selection

The final output of the model is a reconstructed signal. To determine whether the input signal contains interference, we compare the reconstructed signal with the original signal. If the difference is larger than a predefined threshold, we classify the input signal as interference. The threshold is determined by the following formula:

$$\Gamma_{\text{th}} = \mu + \text{std}(\mathbf{L}) \quad (9)$$

where μ is the mean of the loss in validation set. $\text{std}(\mathbf{L})$ is the standard deviation of the loss. This strategy assumes that the loss of the model on the validation set is normally distributed, and the threshold is set to be one standard deviation above

the mean. Any loss above this threshold is considered outliers, hence containing interference.

IV. EXPERIMENTS

A. Data Configuration

The synthetic dataset is generated through a 48-hour MATLAB simulation sampling Ku-band (10.7-12.7 GHz) interference scenarios at 10-second intervals, producing 17,281 temporal snapshots following [14]. Each instance contains synchronized time-domain and frequency-domain representations: an 800-point waveform captures signal amplitudes, while an 800-bin spectral magnitude is derived via FFT processing.

Binary classification labels are assigned through link budget analysis, where class 0 denotes non-interference scenarios ($\text{INR} < \Gamma_{\text{th}}$) below the system protection threshold, and class 1 indicates substantial interference ($\text{INR} \geq \Gamma_{\text{th}}$) exceeding operational limits. We normalize the input signals in both domains separately to zero mean and unit variance.

The dataset is partitioned under anomaly detection constraints, with training (11,509 samples) and validation (1,302 samples) sets containing exclusively non-interference data (class 0). The test set comprises balanced proportions of 2,235 class 0 and 2,235 class 1 instances. The simulation incorporates time-varying link losses with 0-9 dB range, extreme interference cases reaching peak aggregate INR of 32.47 dB, with background CNR fluctuations between 6.40-15.40 dB.

B. Baseline Models

We use the following models as baselines to benchmark our framework: **LinearAE** with fully-connected encoder-decoder architectures as a simplistic reconstruction baseline; **CNNAE** utilizing 1D convolutional layers for encoder and MLP for decoder; **CNNAE+Attention** augmenting CNNAE with temporal self-attention modules; domain-specific **TrID** embedding spectral correlation priors; and **Transformer AE** constructed with stacked multi-head attention layers for global context modeling. All baselines adhere to the standard autoencoder paradigm that reconstructs inputs from bottleneck embeddings, implemented with identical training protocols and hyperparameter tuning strategies for fair comparison.

C. Evaluation Results

As shown in Table I, DualAttnWaveNet establishes state-of-the-art performance across all key metrics, showcasing its superior design synergy. It achieves the highest AUC score (0.9327), surpassing the best baseline (LinearAE/CNNAE: 0.9176) by 1.6% and demonstrating a 36.9% absolute improvement over the cumbersome Transformer AE (0.6812). Simultaneously, our framework completes inference in 0.5409s – 46% of TrID’s 1.0156s – while maintaining competitive F1 score parity (0.8351) with comparably lightweight models like CNNAE (0.8169). Notably, the F1 score shows only 0.3% performance decay from the task-specialized TrID (0.8321), effectively balancing precision-recall tradeoffs through integrated dual-attention mechanisms.

TABLE I
PERFORMANCE COMPARISON OF DUALATTNWAUNET AGAINST
BASELINE MODELS

Model	Accuracy (%) \uparrow	F1 Score \uparrow	AUC \uparrow	Time(s) \downarrow
DualAttnWaveNet	0.8351	0.8351	0.9327	0.5409
LinearAE	0.8149	0.8149	0.9176	0.0966
CNNAE	0.8170	0.8169	0.9175	0.1654
CNNAE+Attention	0.7695	0.7691	0.8718	0.5969
TrID	0.8318	0.8321	0.8318	1.0156
Transformer AE	0.6812	0.5921	0.6812	1.8354

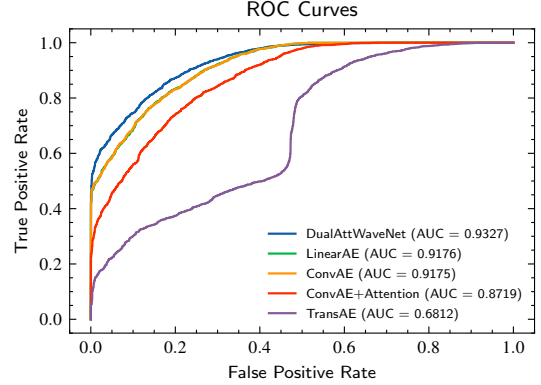


Fig. 5. ROC curves for DualAttnWaveNet and baseline models. DualAttnWaveNet achieves the highest AUC score.

Baseline analyses reveal critical architectural limitations: The Transformer AE’s stacked multi-head attention layers incur huge latency (1.8354s) while suffering catastrophic performance collapse (-22.5% F1 vs ours), exposing fundamental overparameterization issues. Although domain-specific TrID achieves near-parity accuracy (0.8318), its AUC score plunges to 0.8318 – diverging 10.8% from our method – implying overfitting to spectral correlation priors. The temporal attention failure in CNNAE+Attention has computation time 3.6 \times over vanilla CNNAE (0.5969s vs 0.1654s) while degrading AUC by 5.2%, contrasting sharply with our efficiently optimized dual-attention paradigm. LinearAE has the highest throughput (0.0966s) but suffers from suboptimal performance.

As visualized in Fig. 5, the ROC curves quantify the discriminative power hierarchy among competing models. DualAttnWaveNet’s left-skewed trajectory (AUC = 0.9327) dominates the upper-left quadrant, achieving 83.5% true positive rate at less than 10% false positives. The performance collapse in ConvAE+Attention (AUC = 0.8719) is shown as rightward curve drift, exposing the detrimental impact of ill-configured attention modules on feature selectivity. TransAE’s trajectory (AUC = 0.6812) hovers near the diagonal chance line, showing Transformer AE’s failure to learn meaningful decision boundaries despite its 3.6 \times higher computational cost versus our method. Fig. 6 reveals class-specific recognition patterns through confusion matrices across six models.

The temporal localization parameter t mentioned in Eq. 6 governs the precision of interference feature extraction. The model achieves peak performance (0.9327 AUC) at $t=2$. When

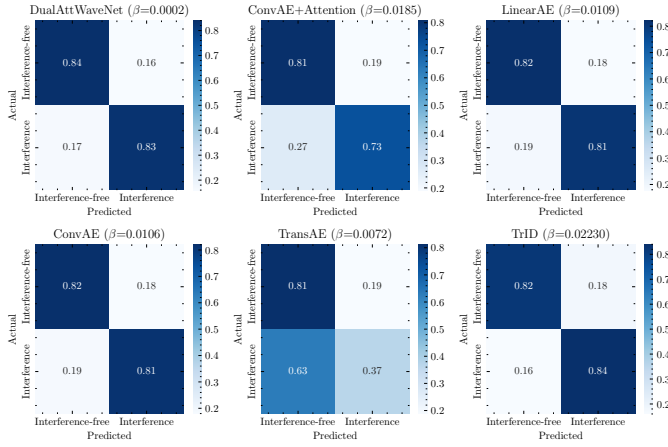


Fig. 6. Confusion matrix for DualAttnWaveNet and other baseline models.

TABLE II
CHOOSING TEMPORAL LOCALIZATION PARAMETER t

t Value	Test AUC
1	0.9257
2	0.9327
4	0.9221
8	0.9046
16	0.8874

$t = 1$ introduces over-localization performance drops by 0.7%. Conversely, larger t values gradually blur critical temporal details: $t = 16$ cause 4.5% AUC degradation. These findings confirm $t = 2$ is the best choice for our model.

D. Ablation Study

As shown in Table III, the full DualAttnWaveNet achieves the best performance with 83.51% accuracy, 83.51% F1 score, and 0.9327 AUC. Removing the mutual attention mechanism ("w/o Mutual Attention") causes consistent performance drops (0.6% accuracy, 0.6% F1, 0.3% AUC). Disabling the wavelet loss ("w/o Wavelet Loss") further degrades accuracy by 1.0%, proving the necessity of joint time-frequency optimization. The vanilla implementation exhibits significant performance limitations (79.95% accuracy, 0.9175 AUC), demonstrating a 3.6% accuracy gap compared to the full model, which highlights the collective contribution of our dual-attention design and spectral-aware constraints to robust classification capability.

V. CONCLUSION

In this paper, we present DualAttWaveNet, a computationally-efficient multimodal framework for interference detection in GSO/LEO coexistence systems. We used a bidirectional attention mechanism to fuse time and frequency domain signals. This design integrates the capability of attention module and at the same time uses less computational resources. Additionally, we introduced a wavelet loss besides traditional MSE loss to enforce the reconstructed output to be consistent across all scales. Extensive evaluations in synthetic Ku-band scenarios show DualAttWaveNet achieves 0.9327 AUC with

TABLE III
ABLATION STUDY OF DUALATTNWAVERNENET COMPONENTS

Model Variant	Accuracy (%) \uparrow	F1 Score \uparrow	AUC \uparrow
DualAttnWaveNet (Full)	0.8351	0.8351	0.9327
w/o Mutual Attention	0.8289	0.8288	0.9294
w/o Wavelet Loss	0.8273	0.8273	0.9283
Vanilla Implementation	0.7995	0.7975	0.9175

83.51% accuracy, surpassing state-of-the-art baselines TrID. The model is trained and tested on an edge NVIDIA 3050Ti GPU, demonstrating 50% faster inference time compared to TrID while maintaining competitive F1 score. Future work will optimize spectral downsampling strategies for different spectrum overlaps scenarios.

REFERENCES

- [1] "Starlink," <https://www.starlink.com/>, accessed: March 23, 2025.
- [2] "SpaceX - starshield," <https://www.spacex.com/starshield/>, accessed: March 23, 2025.
- [3] "Eutelsat oneweb," <https://oneweb.net/>, accessed: March 23, 2025.
- [4] Y. M. Reddy, V. H. Raj, H. P. Thethi, S. Gupta, P. Maan, and R. H. Ghani, "Low earth orbit (leo) satellite networks: A new era in global communication," in *2023 10th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON)*, vol. 10, 2023, pp. 1563–1568.
- [5] S. K. Sharma, S. Chatzinotas, and B. Ottersten, "In-line interference mitigation techniques for spectral coexistence of geo and ngeo satellites," *International Journal of Satellite Communications and Networking*, vol. 34, no. 1, pp. 11–39, 2016. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/sat.1090>
- [6] R. Li, P. Gu, and C. Hua, "Optimal beam power control for co-existing multibeam geo and leo satellite system," in *2019 11th International Conference on Wireless Communications and Signal Processing (WCSP)*, 2019, pp. 1–6.
- [7] T. Wang, W. Li, and Y. Li, "Co-frequency interference analysis between large-scale ngeo constellations and gso systems," in *Proc. 2020 International Conference on Wireless Communications and Signal Processing (WCSP)*, 2020, pp. 679–684.
- [8] C. Zhang, J. Jin, H. Zhang, and T. Li, "Spectral coexistence between leo and geo satellites by optimizing direction normal of phased array antennas," *China Communications*, vol. 15, no. 6, pp. 18–27, 2018.
- [9] G. Facsko, G. Koban, N. Biro, and M. Lkhagvadorj, "Space Weather Effects on Critical Infrastructure," Jul. 2023.
- [10] S. M. Kay, *Fundamentals of Statistical Processing, Volume 2: Detection Theory*. Pearson Education, 2009.
- [11] F. Dimc, G. Baldini, and S. Kandeepan, "Experimental detection of mobile satellite transmissions with cyclostationary features," *International Journal of Satellite Communications and Networking*, vol. 33, no. 2, pp. 163–183, 2015.
- [12] L. Pellaco, N. Singh, and J. Jaldén, "Spectrum prediction and interference detection for satellite communications," in *Proc. 37th International Communications Satellite Systems Conference (ICSSC-2019)*, 2019, pp. 1–18.
- [13] A. Saifaldawla, F. Ortiz, E. Lagunas, and S. Chatzinotas, "Convolutional autoencoders for non-geostationary satellite interference detection," in *Proc. 2024 IEEE International Conference on Communications Workshops (ICC Workshops)*, Denver, CO, USA, 2024, pp. 1334–1339.
- [14] A. Saifaldawla, F. Ortiz, E. Lagunas, A. B. M. Adam, and S. Chatzinotas, "Genai-based models for ngeo satellites interference detection," *IEEE Transactions on Machine Learning in Communications and Networking*, vol. 2, pp. 904–924, 2024.
- [15] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. ukasz Kaiser, and I. Polosukhin, "Attention is All you Need," in *Advances in Neural Information Processing Systems*, vol. 30. Curran Associates, Inc., 2017.