V3

1st Chunyu Yang
School of Computer Science
Fudan University
Shanghai, China
22307140114@m.fudan.edu.cn

***Abstract*—This document is a model and instructions for LaTeX. This and the IEEEtran.cls file define the components of your paper [title, text, heads, etc.]. *CRITICAL: Do Not Use Symbols, Special Characters, Footnotes, or Math in Paper Title or Abstract.**

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

I. Introduction

The rapid increase of non-geostationary orbit (NGSO) satellite constellations, fueled by demand for low-latency broadband and advancements in cost-effective launch technologies, has intensified orbital congestion in low-Earth orbit (LEO). With thousands of satellites now competing for limited spectrum resources, the risk of cross-system interference between NGSO and geostationary orbit (GSO) networks has reached critical levels. Spectral overlap between these systems disrupts service reliability, as even minor deviations in orbital positioning or beam alignment can generate harmful interference-to-noise ratios (I/N) [1]. Mitigating these effects requires real-time, adaptive detection mechanisms capable of discerning interference within dynamic, multi-operator environments. However, there has been an absence of scalable solutions for modern satellite coexistence scenarios.

Standardized frameworks define permissible interference thresholds and aggregation methodologies. For instance, the International Telecommunication Union (ITU) mandates a -10 dB I/N ceiling over 99.9% of operational intervals to protect GSO receiver integrity [2]. Similarly, EPFD aggregation techniques measure the total power flux density from distributed LEO transmitters to ensure that the aggregate interference at GSO terminals remains within acceptable limits [3].

Traditional interference detection methods, such as energy detection (ED), operate by measuring signal energy over a predefined time window and comparing it to a static threshold [4]. While computationally efficient, ED often struggles to differentiate weak interference from background noise in low signal-to-noise (SNR) scenarios, leading to missed detections or unsustainable false-alarm rates. Enhanced techniques like cyclostationary feature detection improve robustness by exploiting periodic signal properties [5], yet they incur significant computational costs when applied to LEO systems. Two-step hybrid approaches integrating pilot cancellation and hardware-based filtering have demonstrated limited success but lack the adaptability to support evolving satellite architectures [6].

Recent advances in machine learning (ML) seek to mitigate the shortcomings of traditional detection frameworks. Initial efforts utilized deep neural networks, such as LSTM-based architectures [7], to identify interference by processing in-phase/quadrature (IQ) signal components or time-domain amplitude signal. Contemporary methodologies view the detection task as an anomaly discrimination problem, training models to isolate interference from nominal transmissions by learning their intrinsic signatures. Autoencoder-based strategies [8], for instance, reconstruct idealized interference-free signals from raw inputs; significant energy deviations between the input and reconstructed waveforms are flagged as potential interference. Transformers have further pushed detection boundaries, attaining state-of-the-art area-under-curve (AUC) scores by modeling long-range spectral and temporal correlations in signal [9]. Nevertheless, these architectures impose substantial computational burdens during training, making

them impractical for latency-sensitive satellite operations.

To overcome these barriers, we propose DualAttWaveNet, a multimodal architecture that unifies time-frequency signal representations, achieving an AUC of 0.9327 while reducing inference latency by 3.8 times compared to transformer baselines. The core contributions include:

1) Bidirectional Cross-Domain Attention: A lightweight mutual attention mechanism that dynamically fuses temporal waveforms and spectral features, learning interdependencies between domains without concatenation bottlenecks.
2) Multi-Scale Wavelet Regularization: A loss function enforcing reconstruction fidelity across four predefined wavelet scales, aligning interference patterns with broader spectral trends for robust detection.
3) Our detector is validated using the system model in [9], which synthesizes realistic GSO/LEO coexistence scenarios. This ensures evaluation under interference conditions mirroring actual orbital dynamics and spectral constraints.

II. System Model

A. Interference Scenario Configuration

An interference scenario between geostationary orbit (GSO) and low Earth orbit (LEO) satellite systems is considered, as conceptually illustrated in Figure 1. The GSO satellite operates as the primary communication node, while multiple LEO satellites from non-geostationary systems introduce unintended interference components observed at a GSO ground station (GGS). The composite received waveform contains both desired carrier signals and interference.

B. Link Budget Analysis

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

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

where $G_{\mathrm{r,\,gso}}(\theta_0)$ denotes the maximum receive antenna gain at boresight angle θ_0, $L_{\mathrm{FS,\,gso}}$ represents free-space path loss, and L_{add} accounts for aggregate atmospheric impairments.

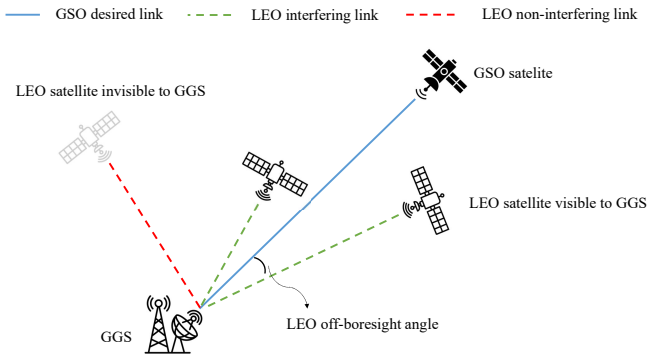

Figure 1. Interference scenario between GSO and LEO satellite systems.

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

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

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

The composite signal quality is characterized through the carrier-to-interference-plus-noise ratio:

$$\mathrm{CINR} = \frac{C}{\sum_{k=1}^{K} I_k + k_{\mathrm{B}} T B} \tag{3}$$

where k_{B} denotes the Boltzmann constant, T the system noise temperature, and B the operational bandwidth.

C. Signal Composition

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

$$\begin{aligned} y(t) = {} & x(t)\sqrt{\mathrm{CNR}}\,(\text{Desired GSO}) \\ & + \sum_{k=1}^{K} i_k(t) e^{j2\pi\Delta f_k t}\sqrt{\mathrm{INR}_k}\,(\text{LEO interference}) \\ & + \zeta(t)\,(\text{Thermal noise}) \end{aligned} \tag{4}$$

where $\Delta f_k = f_{\mathrm{c},k} - f_{\mathrm{c,gso}}$ captures carrier frequency offsets from Doppler effects and orbital dynamics. The exponential terms induce time-varying phase rotations proportional to relative satellite motion.

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: $y^F = 10\log_{10}(\phi(y(t)))$

III. Proposed Deep Learning Model

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

A. Bidirectional Attention

We propose a parameter-efficient mutual attention mechanism for cross-modal feature fusion. Unlike conventional multi-head attention in [10], our design employs single-head dot-product attention with spatial reduction to minimize computational overhead while maintaining inter-domain alignment capacity.

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:

$$\begin{aligned}\text{MutualAttn}(\mathbf{X},\mathbf{Y}) &= \mathbf{X} + \gamma \cdot \text{AttentionGate}(\mathbf{X},\mathbf{Y}) \\ \text{AttentionGate}(\mathbf{X},\mathbf{Y}) &= \mathbf{V}_y \cdot \text{Softmax}\left(\frac{\mathbf{Q}_x\mathbf{K}_y^\top}{\sqrt{d}}\right)\end{aligned} \tag{5}$$

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× 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→spectral and spectral→temporal directions, in sequential form:

$$\begin{aligned}\widehat{\mathbf{X}} &= \text{MutualAttn}(\mathbf{X},\mathbf{Y}) \\ \widehat{\mathbf{Y}} &= \text{MutualAttn}(\mathbf{Y},\mathbf{X})\end{aligned} \tag{6}$$

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

B. Wavelet-Domain Spectral Regularization

Our temporal-spectral analysis employs parameterized Morlet wavelets with learned scale distributions. Given input sequence $x \in \mathbb{R}^{B\times C\times L}$, we construct a filter bank $\mathcal{F} \in \mathbb{R}^{S\times 1\times K}$ through:

$$\mathcal{W}_s(\tau) = \text{Norm}\left(\cos(\frac{1.75\tau}{2}) \odot \mathcal{G}(\tau, s)\right) \tag{7}$$

where $\mathcal{G}(\tau, s) = \exp(-\tau^2/(2s^2))$ is the Gaussian window, $s \in \mathbb{S}$ denotes wavelet scales, and $K = 4s_{\max}$ defines the kernel size from maximum scale $s_{\max}$. Each filter is L_2-normalized to preserve energy consistency across scales.

The discrete wavelet transform is implemented as depth-wise separable convolution:

$$\begin{aligned}\mathbf{X}_{\text{wave}} &= \text{Conv1D}(\mathbf{X}, \mathcal{F}) \\ &= \bigcup_{s\in\mathbb{S}} \mathbf{X} * \mathcal{W}_s \in \mathbb{R}^{B\times C\times S\times L}\end{aligned} \tag{8}$$

where $*$ denotes cross-correlation with reflection padding for boundary alignment. The transform preserves temporal resolution through same-stride (stride = 1) convolution across S parallel scale dimensions. See Figure 2 for an illustration of the wavelet kernels and their input-output relationships.

We optimize both signal-space and wavelet-domain reconstructions through:

$$\mathcal{L} = \lambda_1\|\widehat{\mathbf{X}} - \mathbf{X}\|_2^2 + \lambda_2\|\widehat{\mathbf{X}}_{\text{wave}} - \mathbf{X}_{\text{wave}}\|_2^2 \tag{9}$$

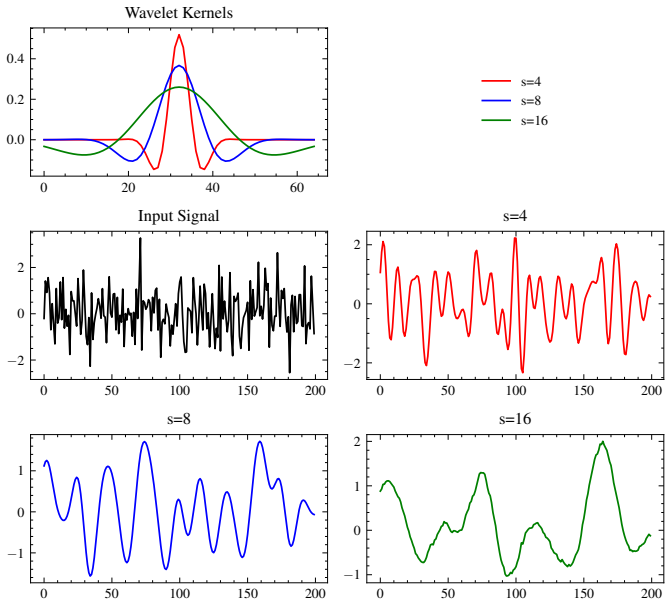

Figure 2. (Top) Normalized Morlet-style wavelet kernels at scales [4, 8, 16]. (Bottom) Input–output relationships: Gaussian white noise input (top–left) and its filtered components using the respective wavelets (red: $s = 4$, blue: $s = 8$, green: $s = 16$). Responses exhibit progressive high–frequency suppression and temporal smoothing with increasing scale. The visualizer implements bank normalization and preserves temporal resolution via same–padding convolutions (using a β–spline differentiable implementation).

where $\lambda_1 = 1$, $\lambda_2 = 0.5$ balance reconstruction error and wavelet transform loss.

IV. Experiments

A. Dataset Generation

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 [9]. 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 (INR $<$ Γ_{th}) below the system protection threshold, and class 1 indicates substantial interference (INR $\geq$ Γ_{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. Evaluation Results

DualAttWaveNet achieves state-of-the-art interference detection performance across metrics, as shown in Table I. With an AUC of 0.9327, it outperforms ConvAE (0.9175), ConvAE+Attention (0.8719), and TransAE (0.6812), as quantified by ROC analysis (Figure 3). Notably, the model surpasses the previous best method, TrID [9], which achieved AUCs of 0.8318 (time domain) and 0.7106 (spectral domain), highlighting superior cross-modal harmonization through dual-attention synergy.

When conducting confusion matrix analysis, we set the detection threshold to $\mu + 1 \cdot \sigma$, where μ and σ are mean and standard deviation of reconstruction error on validation set. Reconstruction error over this threshold is considered to be in class 1. Figure 4 shows that DualAttWaveNet has balanced precision-recall tradeoff (F1: 83.66%). Baselines exhibit critical limitations: ConvAE+Attention suffers a 4.56% AUC drop relative to ConvAE, shows the insufficiency of naive attention for cross-domain alignment. TransAE' s poor convergence (AUC: 0.6812) suggests transformer architectures is hard to converge to optimal point during training.

C. Ablation Study

To validate the necessity of DualAttWaveNet' s architectural innovations, we conducted component-wise ablation experiments (II). The full model achieves peak performance (Accuracy: 83.67%, F1: 83.66%, AUC: 0.9327), while ablating the mutual attention module reduces AUC to 0.9294 (-0.33%), indicating its critical role in aligning temporal-spectral features. Removing the wavelet coherence loss further degrades AUC to 0.9283 (-0.44%), confirming its utility in preserving time-frequency structure. The vanilla implementation (no mutual attention

Table I
PERFORMANCE COMPARISON OF DUALATTNWAVENET AGAINST BASELINE MODELS

Model	Accuracy (%) ↑	F1 Score ↑	AUC↑	FLOPS (G)↓
DualAttnWaveNet	0.8367	0.8366	0.9327	0.00
LinearAE	0.7987	0.7966	0.9176	0.00
CNNAE	0.7996	0.7975	0.9175	0.00
TrID	0.8318	0.8321	0.8318	0.00
Transformer AE	0.5678	0.5634	0.6812	0.00

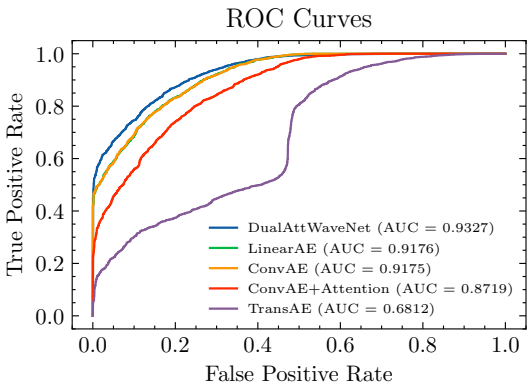

Figure 3. ROC curves for DualAttnWaveNet and baseline models. DualAttnWaveNet achieves the highest AUC score.

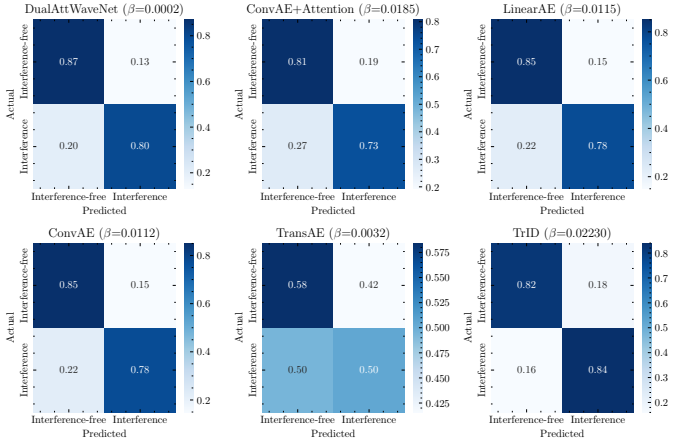

Figure 4. Confusion matrix for DualAttnWaveNet and baselines on the test set

or wavelet loss) yields the lowest metrics (AUC: 0.9175, Accuracy: 79.95%), underscoring that naively combining branches without alignment mechanisms underperforms all proposed variants by significant margins.

REFERENCES

[1] *Radio Regulations*, International Telecommunication Union (ITU) Std., 2020.

[2] *Protection Criteria Related to the Operation of Data Relay Satellite Systems*, International Telecommunication Union (ITU) Recommendation ITU-R SA.1155-2, 2017.

[3] *Methodologies for Calculating Aggregate Downlink Equivalent Power Flux-Density Produced by Multiple Non-Geostationary Fixed-Satellite Service Systems into a Geostationary Fixed-Satellite Service Network*, International Telecommunication Union (ITU-R) Report ITU-R S.1588-0, 2002.

[4] S. M. Kay, *Fundamentals of Statistical Processing, Volume 2: Detection Theory*. Pearson Education, 2009.

[5] 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.

[6] T. Wang, W. Li, and Y. Li, "Co-frequency interference analysis between large-scale ngso constellations and gso systems," in *Proc. 2020 International Conference on Wireless Communications and Signal Processing (WCSP)*, 2020, pp. 679–684.

[7] 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.

[8] 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.

[9] A. Saifaldawla, F. Ortiz, E. Lagunas, A. B. M. Adam, and S. Chatzinotas, "Genai-based models for ngso satellites interference detection," *IEEE Transactions on Machine Learning in Communications and Networking*, vol. 2, pp. 904–924, 2024.

[10] 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.

Table II
Ablation Study of DualAttnWaveNet Components

Model Variant	**Accuracy (%) ↑**	**F1 Score ↑**	**AUC ↑**
DualAttnWaveNet (Full)	0.8367	0.8366	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