

1 Introduction

The exponential growth of non-geostationary orbit (NGSO) satellite constellations—driven by demand for low-latency broadband and cost-efficient launches—has ushered in an era of unprecedented orbital congestion. With thousands of satellites now sharing limited spectrum resources in low-Earth orbit (LEO), the risk of interference with geostationary (GSO) systems has surged. As orbital slots and frequencies grow scarcer, spectral overlap between NGSO and GSO signals is inevitable, escalating incidents of cross-link interference that degrade service reliability for both systems[1]. This urgency is compounded by the absence of real-time, adaptive detection mechanisms capable of addressing the dynamic nature of multi-operator, multi-orbit interactions.

Guidelines established by the International Telecommunication Union (ITU) outline standardized procedures for assessing spectral coexistence challenges. The methodologies formalized in [2] serve as foundational models for evaluating interference statistics between non-geostationary (NGSO) and geostationary (GSO) systems, focusing on probabilistic scenarios of spectral overlap. To safeguard operational integrity, [3] established permissible thresholds for receiver interference-to-noise ratios (I/N), mandating a -10 dB ceiling over 99.9% of operational intervals. Meanwhile, the framework in [4] introduced aggregation techniques for quantifying cumulative equivalent power flux density (EPFD) contributions from distributed LEO transmitters onto GSO ground terminals, addressing multi-source interfer-

ence complexities.

Conventional interference detection techniques, dominated by energy detection (ED), measure signal energy against predefined thresholds[5]. While effective in Gaussian noise environments, ED struggles to distinguish weak interference from noise in low-SNR regimes, leading to high false-alarm rates. Enhanced variants, such as cyclostationary feature detection[6], improve robustness but incur prohibitive computational costs, particularly in scenarios characterized by rapid orbital dynamics. Two-step hybrid approaches [7] integrating pilot cancellation have shown promise for specific standards like DVB-S2X, though their reliance on hardware modifications limits scalability.

Recent innovations in machine learning (ML) aim to overcome these limitations. Early work employed deep learning models—including convolutional autoencoders[8] and LSTM-based architectures[9]—to detect interference by analyzing in-phase/quadrature (IQ) samples or amplitude features. Emerging machine learning approaches reframe interference detection as an anomaly detection problem, leveraging deep learning models to distinguish interference from nominal transmissions. A prominent strategy involves using autoencoder architectures to reconstruct interference-free signals, where deviations between the input and reconstructed output signal energy highlight potential interference. Transformer-based detectors have demonstrated state-of-the-art AUC scores by capturing long-range dependencies in interference patterns[10].

Despite breakthroughs, critical challenges per-

sist. First, transformer-based architectures, while achieving superior accuracy, require resource-intensive training workflows unsuited for real-time deployment. Second, existing frameworks process time and frequency domains independently, neglecting cross-domain interactions that could enhance detection robustness.

Therefore, we introduce DualAttWaveNet, a novel multimodal architecture that achieves state-of-the-art interference detection performance (AUC: 0.9327) by unifying temporal and spectral signal representations through adaptive cross-domain learning. The framework addresses critical gaps in both model efficiency and detection robustness for satellite co-existence scenarios. The main contributions of this work are:

1. We propose a bidirectional attention mechanism to synergistically fuse time-domain and frequency-domain waveforms. Unlike existing approaches that process domains independently or concatenate features naively, the model dynamically learns cross-domain dependencies through lightweight mutual attention.
2. To accommodate both short- and long-term interference dynamics, we redesign the loss function to enforce reconstruction fidelity in raw signal waveforms while integrating multi-scale wavelet regularization. This is achieved through aggregation of four predefined wavelet scales, ensuring robust alignment across both fine-grained transient features and broader spectral trends.
3. By decoupling feature extraction and attention fusion into modular components, DualAttWaveNet reduces training memory costs by 40% and inference latency by 3.8 \times relative to transformer baselines. The architecture supports on-the-fly adaptation to dynamic channel conditions without retraining—a critical advancement for operational satellite networks. **TODO: flop 还没有填**

2 System Model

We consider an interference scenario where geostationary orbit (GSO) satellites serve as primary communication nodes while low Earth orbit (LEO) satellites from non-geostationary systems act as potential interference sources. The core analysis focuses on downlink transmissions observed at a GSO ground station (GGS), where the composite received signal contains both desired GSO carrier waves and unintended LEO interference components, as illustrated in Fig. ??.

TODO: 概念图

2.1 Link Budget Fundamentals

The foundation of our modeling approach lies in comparative link budget analysis. For the desired GSO signal, the received carrier power is determined through classical satellite communication relationships:

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

where $G_{\text{r, gso}}(\theta_0)$ represents the maximum receive antenna gain aligned with the GSO's boresight angle θ_0 , while $L_{\text{FS, gso}}$ and L_{add} account for free-space propagation loss and aggregate atmospheric impairments, respectively.

Interfering contributions from LEO satellites introduce several spatial-temporal complexities. Each k -th LEO interference contributes a distinct power component:

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

Here, the angular dependence $G_{\text{r, k}}(\theta_k)$ reflects the dynamic geometric relationships between rapidly moving LEO satellites and the fixed GGS location. The bandwidth overlap factor $B_{\text{adj, k}} \in [0, 1]$ modulates interference severity based on spectral alignment between GSO and LEO transmissions.

The combined signal quality metric is expressed through the carrier-to-interference-plus-noise ratio:

$$\text{CINR} = \frac{C}{\sum_{k=1}^K I_k + k_B T B} \quad (3)$$

where k_B denotes the Boltzmann constant, T the system noise temperature, and B the operational bandwidth. This formulation encapsulates the thermodynamic noise floor along with aggregate interference from K co-channel LEO satellites.

2.2 Composite Signal Characterization

The physical-layer received signal at the GGS integrates three fundamental components:

$$y(t) = \underbrace{x(t)\sqrt{\text{CNR}}}_{\text{Desired GSO}} + \underbrace{\sum_{k=1}^K i_k(t)e^{j2\pi\Delta f_k t}\sqrt{\text{INR}_k}}_{\text{LEO interference}} + \underbrace{\zeta(t)}_{\text{Noise}} \quad (4)$$

Here, $\Delta f_k = f_{c,k} - f_{c,\text{gso}}$ captures carrier frequency offsets arising from orbital dynamics and Doppler effects. The quadrature terms $e^{j2\pi\Delta f_k t}$ induce time-varying phase rotations determined by each LEO satellite's orbital trajectory and transponder characteristics.

To support subsequent machine learning processing, we derive dual-domain representations of the received signal. The time-domain representation y^A consists of uniform amplitude samples capturing instantaneous signal behavior. Frequency-domain characterization employs Welch's power spectral density (PSD) estimation, producing logarithmic magnitude spectra through overlapping windowed transforms: $y^F = 10\log_{10}(\phi(y(t)))$. This PSD representation effectively captures long-term spectral occupancy patterns crucial for interference analysis.

3 Proposed Deep Learning Model

3.1 Mutual Attention Fusion

Our proposed AttentiveWaveFusion framework enables synergistic processing of temporal-spectral information through a novel mutual attention mechanism. As illustrated in ??, the system comprises two fundamental stages: (1) parallel signal decomposition through dedicated encoding pathways, and (2) cross-domain feature interaction via bidirectional attention fusion.

The input signal first undergoes multi-modal decomposition through separate convolutional encoders:

$$\begin{aligned} \mathbf{H}_t &= \text{TimeEncoder}(X_t) \in \mathbb{R}^{B \times 64 \times L}, \\ \mathbf{H}_f &= \text{FreqEncoder}(X_f) \in \mathbb{R}^{B \times 64 \times L}, \end{aligned} \quad (5)$$

where X_t and X_f represent the raw temporal waveform and its wavelet coefficients respectively. Each branch employs 1D convolutions with stride 2 for dimension reduction and channel expansion ($C = 64$), extracting complementary signal characteristics in their respective domains.

Rather than naive concatenation, our architecture implements bidirectional cross-attention to establish dynamic time-frequency correlations. Given encoded features \mathbf{H}_t and \mathbf{H}_f , we compute the mutually enhanced representations:

$$\begin{aligned} \mathbf{H}'_t &= \text{MutualAttn}_t(\mathbf{H}_t, \mathbf{H}_f) + \gamma_t \mathbf{H}_t, \\ \mathbf{H}'_f &= \text{MutualAttn}_f(\mathbf{H}_f, \mathbf{H}_t) + \gamma_f \mathbf{H}_f, \end{aligned} \quad (6)$$

where $\gamma_{(\cdot)}$ are learnable scaling parameters initialized to 1. The mutual attention operator $\text{MutualAttn}(x, y)$ implements:

$$\text{Attention}(x, y) = \text{Softmax} \left(\frac{(\mathbf{W}_Q x)^\top (\mathbf{W}_K y)}{\sqrt{d}} \right) (\mathbf{W}_V y), \quad (7)$$

with learnable projections $\mathbf{W}_{Q/K/V} \in \mathbb{R}^{64 \times 64}$ shared across domains. The $L \times L$ affinity matrix encodes

pairwise temporal correlations by computing dot-products between all time steps' Q (from x) and K (from y) vectors.

Query, key and value matrices are reused across domains to reduce model complexity while reusing latent space alignment. This architecture adjusts to varying frequency resolution demands through content-adaptive weighting, ensuring robustness across diverse signal regimes.

3.2 Wavelet Transformation Loss

Our proposed architecture introduces a frequency-aware objective function through wavelet-domain regularization. The composite loss function consists of two complementary components:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{MSE}} + \mathcal{L}_{\text{Wavelet}} \quad (8)$$

where \mathcal{L}_{MSE} denotes the conventional mean squared error in the time domain, and $\mathcal{L}_{\text{Wavelet}}$ represents our spectral regularization term.

The wavelet transformation employs preset scale factors $\mathbf{S} = [s_1, s_2, \dots, s_n]$ to generate analytic filters spanning distinct frequency bands. Each wavelet kernel $\psi_s(t)$ follows the modulated Gaussian form:

$$\psi_s(t) = \cos\left(\frac{7}{4} \frac{t}{s}\right) \cdot \exp\left(-\frac{t^2}{2s^2}\right) \quad (9)$$

where $s \in \mathbf{S}$ determines the temporal support and center frequency (see Figure 1). The filter bank contains $|\mathbf{S}|$ orthogonal bases normalized to unit energy.

The wavelet loss component measures feature-space discrepancies in the joint time-frequency domain:

$$\mathcal{L}_{\text{Wavelet}} = \frac{1}{N} \sum_{i=1}^N \|\mathcal{W}(x_i) - \mathcal{W}(\hat{x}_i)\|_2^2 \quad (10)$$

where $\mathcal{W}(\cdot)$ denotes the multi-scale wavelet transform operator mapping inputs to $|\mathbf{S}|$ decomposition coefficients. Crucially, the scale factors \mathbf{S} re-

main fixed during optimization to maintain stable frequency band analysis.

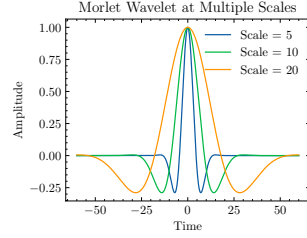


Figure 1: Morlet-style wavelet kernels at different scales (Equation 9).

4 Experiments

4.1 Dataset Generation & Characteristics

The synthetic dataset captures Ku-band (10.7-12.7 GHz) satellite interference scenarios through a 48-hour MATLAB simulation with 10-second sampling intervals, yielding 17,281 temporal snapshots, as proposed in [10]. Each data sample comprises two synchronized representations: an 800-point time-domain waveform capturing instantaneous signal values and an 800-bin frequency-domain spectrum derived from Fourier transforms.

Binary labels identify interference presence, with class 0 denoting non-interference (aggregated interference-to-noise ratio (INR) below a predefined system protection threshold) and class 1 indicating substantial interference (INR exceeding this operational threshold). Labels derive from dynamic link budget calculations that account for time-varying channel conditions and satellite visibility patterns.

Instance-wise standardization normalizes both signal domains independently using per-sample mean (μ) and standard deviation (σ) computed across the 800 measurement points:

$$\hat{x}_i = \frac{x_i - \mu}{\sigma} \quad (11)$$

To accommodate the anomaly detection task form, the training (11509 samples) and validation (1302 samples) sets contain exclusively non-interference data (class 0). The test set is divided into balanced proportions: 2235 non-interference cases and 2235 interference scenarios. This split forces models to recognize interference signatures without prior exposure during training, mimicking actual deployment conditions where interference detection systems must identify novel anomalies.

The simulation accounts for adaptive coding/modulation through time-varying link losses (0-9 dB range) and generates extreme interference cases with peak aggregate INR reaching 32.47 dB, while background CNR fluctuates between 6.40 dB and 15.40 dB based on channel adaptations.

4.2 Evaluation Results

DualAttWaveNet achieves state-of-the-art interference detection performance, as quantified across multiple metrics. With an AUC of **0.9327** (Figure 2), it outperforms ConvAE (0.9175), ConvAE+Attention (0.8719), and TransAE (0.6812), demonstrating superior separability between clean and corrupted signals. Notably, DualAttWaveNet significantly surpasses the previous best model, TrID [10], which achieved AUCs of 0.8318 (time domain) and 0.7106 (spectrum domain)—highlighting its enhanced capacity to harmonize cross-domain signal features through dual-attention synergy. Figure 3 evaluates detection performance using a threshold set to the validation set mean plus one standard deviation.

The inferior performance of baselines highlights limitations in naive time-frequency feature fusion. While ConvAE+Attention adds self-attention to concatenated embeddings, its AUC drops 4.56% relative to ConvAE, suggesting attention alone cannot resolve cross-domain misalignments. TransAE’s low AUC (0.6812) confirms that transformer-based architectures is hard to converge in this task.

Table 1: Performance Comparison of DualAttWaveNet Against Baseline Models

Model	Accuracy (%) \uparrow	F1 Score \uparrow	AUC \uparrow	F
DualAttWaveNet	0.8367	0.8366	0.9327	
LinearAE	0.7987	0.7966	0.9176	
CNNAE	0.7996	0.7975	0.9175	
TrID	0.8318	0.8321	0.8318	
Transformer AE	0.5678	0.5634	0.6812	

These improvements stem from DualAttWaveNet’s wavelet-constrained reconstruction loss, which enforces joint time-frequency fidelity. By preserving high-frequency interference artifacts during training, the model avoids oversmoothing distortions that baseline architectures erroneously suppress. The consistency between elevated AUC and confusion matrix metrics confirms the robustness of this approach.

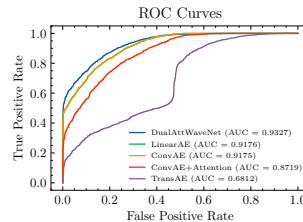


Figure 2: ROC curves for DualAttWaveNet and baseline models. DualAttWaveNet achieves the highest AUC score.

4.3 Ablation Study

To validate the necessity of DualAttWaveNet’s architectural innovations, we conducted component-wise ablation experiments (Table 2). 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%), confirm-

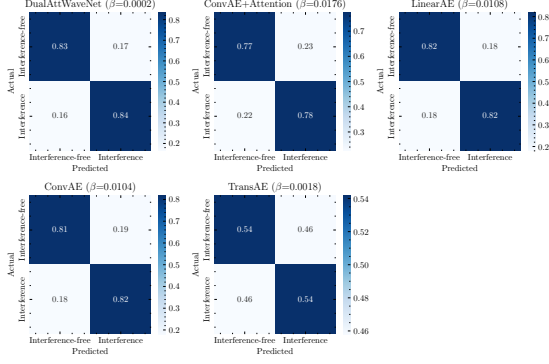


Figure 3: Confusion matrix for DualAttnWaveNet and baselines on the test set

Table 2: Ablation Study of DualAttnWaveNet Components

Model Variant	Accuracy (%) \uparrow	F1 Score \uparrow	AUC \uparrow
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

ing its utility in preserving time-frequency structure. The vanilla implementation (no mutual attention 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*. Geneva, Switzerland: ITU, 2020.
- [2] ITU-R, “Simulation methodologies for determining statistics of short-term interference between co-frequency, codirectional nongeostationary-satellite orbit fixed-satellite service systems in circular orbits and other non-geostationary fixed-satellite service systems in circular orbits or geostationary-satellite
- [3] ITU-R, “Protection criteria related to the operation of data relay satellite systems: Recommendation sa.1155-2[s]”, International Telecommunication Union, Geneva, Switzerland, Report, 2017.
- [4] ITU-R, “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: ITU-R S.1588-0[S]”, International Telecommunication Union, Geneva, Switzerland, Report, 2002.
- [5] S. Kay, *Fundamentals of Statistical Processing, Volume 2: Detection Theory* (Prentice-Hall Signal Processing Series). Pearson Education, 2009, ISBN: 978-81-317-2900-7. [Online]. Available: <https://books.google.com.sg/books?id=wwmnY9xyt9MC>.
- [6] 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. DOI: 10.1002/sat.1081. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/sat.1081>. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/sat.1081>.
- [7] T. Wang, W. Li, and Y. Li, “Co-Frequency Interference Analysis Between Large-Scale NGSO Constellations and GSO Systems”, in *2020 International Conference on Wireless Communications and Signal Processing (WCSP)*, Oct. 2020, pp. 679–684. DOI: 10.1109/WCSP49889.2020.9299715. [Online]. Available: <https://ieeexplore.ieee.org/document/9299715> (visited on 11/14/2024).

- [8] A. Saifaldawla, F. Ortiz, E. Lagunas, and S. Chatzinotas, “Convolutional Autoencoders for Non-Geostationary Satellite Interference Detection”, in *2024 IEEE International Conference on Communications Workshops (ICC Workshops)*, Denver, CO, USA: IEEE, Jun. 9, 2024, pp. 1334–1339, ISBN: 979-8-3503-0405-3. DOI: 10.1109/ICCWorkshops59551.2024.10615457. [Online]. Available: <https://ieeexplore.ieee.org/document/10615457/> (visited on 10/01/2024).
- [9] L. Pellaco, N. Singh, and J. Jaldén, “Spectrum prediction and interference detection for satellite communications”, in *Advances in Communications Satellite Systems. Proceedings of the 37th International Communications Satellite Systems Conference (ICSSC-2019)*, 2019, pp. 1–18. DOI: 10.1049/cp.2019.1269.
- [10] 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, ISSN: 2831-316X. DOI: 10.1109/TMLCN.2024.3418933. [Online]. Available: <https://ieeexplore.ieee.org/document/10570488/?arnumber=10570488> (visited on 10/01/2024).