Journal of Advanced Artificial Intelligence, Engineering and Technology

Research Article

WRJAAIET-25-034

Game-Theoretic Self-Supervised Defense: A Dual-Stream Neural Architecture for Secure End-to-End Communication

Yulin Zhou*

Yunnan Modern Vocational Technical College, Chu Xiong, China

***Correspondence:** Yulin Zhou, Yunnan Modern Vocational Technical College, Chu Xiong, China, E-mail: ethanlynx@foxmail.com; DOI: https://doi.org/10.56147/aaiet.1.4.34

Citation: Zhou Y (2025) Game-Theoretic Self-Supervised Defense: A Dual-Stream Neural Architecture for Secure End-to-End Communication. J Adv Arti Int Eng & Techn 1: 34.

Abstract

We propose a novel dual-stream neural architecture for secure end-to-end communication that integrates self-supervised adversarial detection with game-theoretic defense strategies. The increasing sophistication of adversarial attacks necessitates adaptive security mechanisms that can dynamically respond to evolving threats while maintaining computational efficiency. Our framework addresses this challenge by combining a transformer-based contrastive learning module for real-time perturbation detection with a zero-sum Markov game formulation to optimize defense policies under Nash equilibrium. The self-supervised stream extracts invariant features from communication signals and computes threat probabilities, while the game-theoretic stream dynamically adjusts encryption parameters and defense actions based on adversarial behavior. The two streams are jointly optimized to achieve provable robustness against both known and unknown attack patterns. Furthermore, the system interfaces seamlessly with conventional intrusion detection and encryption modules, replacing rule-based heuristics with data-driven adaptive strategies. Experimental validation demonstrates significant improvements in detection accuracy and resilience compared to existing methods. The key innovation lies in the unification of self-supervision and game theory within a single end-to-end trainable framework, enabling scalable and adaptive security for modern communication networks. This approach not only enhances real-time threat mitigation but also provides theoretical guarantees on defense optimality under adversarial dynamics.

Keywords: Dual-stream neural architecture; Cryptographic methods; Adversarial detection; Game-theoretic defense strategies; Modern communication systems; Learning-based anomaly detection

Received date: May 26, 2025; **Accepted date:** May 30, 2025; **Published date:** June 17, 2025

Introduction

Modern communication systems increasingly rely on deep learning for end-to-end optimization, yet their security remains vulnerable to sophisticated adversarial attacks. While conventional encryption and authentication protocols provide baseline protection, they often fail against adaptive adversaries who exploit learned channel characteristics and model vulnerabilities. Recent advances in adversarial machine learning have demonstrated that neural networks are susceptible to carefully crafted perturbations that can degrade performance or induce malicious behavior. This vulnerability is particularly concerning in communication systems where reliability and integrity are paramount.

Existing approaches to secure communication systems can be broadly categorized into two paradigms: rule-based cryptographic methods and learning-based anomaly detection. Cryptographic techniques, while theoretically sound, often lack adaptability to dynamic attack patterns. Learning-based methods, on the other hand have shown promise in detecting anomalies but typically operate in isolation from defense mechanisms. For example, autoencoder-based systems and CNN architectures have been proposed for end-to-end learning, yet they do not explicitly model the adversarial dynamics that underlie real-world threats [1,2].

The integration of game theory with deep learning offers a promising direction for addressing these limitations. Game-theoretic models formalize the interaction between attackers and defenders as a strategic competition, enabling the derivation of optimal defense policies under Nash equilibrium [3]. However, existing game-theoretic approaches in cybersecurity often rely on simplified assumptions about attack patterns, limiting their applicability to complex, high-dimensional communication systems [4].

We propose a novel dual-stream neural architecture that unifies self-supervised adversarial detection with game-theoretic defense optimization. The first stream employs contrastive learning to extract robust features from communication signals and detect perturbations without labeled attack data [5]. This self-supervised approach is critical for identifying zero-day attacks and unknown threat vectors. The second stream formulates defense as a dynamic game, where the defender adjusts encryption parameters and response strategies based on the attacker's observed behavior. By modeling this interaction as a zero-sum Markov game, our framework derives provably optimal defense policies that account for adversarial adaptation.

The key contributions of this work are threefold:

- **Unified adversarial detection and defense:** Unlike prior methods that treat detection and mitigation as separate tasks, our framework jointly optimizes both components within a single end-to-end trainable architecture. This integration enables real-time adaptation to evolving threats while maintaining computational efficiency.
- **Game-theoretic robustness guarantees:** By formulating defense as a Nash equilibrium problem, our approach provides theoretical guarantees on optimality under adversarial dynamics. This is a significant advancement over heuristic or rule-based defense strategies.
- **Scalability to high-dimensional signals:** The use of transformer-based attention modules allows the system to process high-dimensional communication signals efficiently, aligning adversarial detection with signal integrity verification [6].

Our work builds upon recent advances in end-to-end learning for communication systems, adversarial machine learning and robust optimization [7-9]. However, it diverges from prior art by explicitly modeling the strategic interplay between attackers and defenders while leveraging self-supervision for scalable threat detection.

The remainder of this paper is organized as follows: Section 2 reviews related work in adversarial machine learning and secure communication systems. Section 3 provides necessary background on game theory and self-supervised learning. Section 4 details our proposed dual-stream architecture. Sections 5 and 6 present experimental validation and results. Finally, Sections 7 and 8 discuss implications and conclude the paper.

Related Work

Secure end-to-end communication systems have evolved significantly with the integration of deep learning techniques, yet persistent vulnerabilities to adversarial attacks remain a critical challenge. This section examines prior work across three key dimensions: adversarial robustness in communication systems, game-theoretic approaches to security and self-supervised learning for anomaly detection.

Adversarial robustness in communication systems

Recent studies have demonstrated the susceptibility of deep learning-based communication systems to adversarial perturbations. The work in provides a comprehensive taxonomy of attack vectors targeting wireless systems, ranging from gradient-based perturbations to physical-layer spoofing [10]. While traditional cryptographic methods offer theoretical guarantees, they often fail to address the dynamic nature of modern adversarial threats [11]. Several approaches have attempted to bridge this gap through robust learning frameworks. For instance, proposed generative adversarial networks to model channel uncertainties, though their defense mechanisms remain reactive rather than proactive [12]. The vulnerability of autoencoder-based systems was further highlighted in, where adversarial examples could bypass authentication protocols by exploiting gradient information [13].

Game-theoretic security frameworks

Game theory has emerged as a powerful tool for modeling strategic interactions in cybersecurity. The foundational work in established Markov games as a viable framework for network defense, though their application was limited to discrete action spaces [14]. Subsequent research extended these concepts to neural networks by incorporating adversarial training objectives [15]. However, these methods typically assume perfect knowledge of attack strategies, which is unrealistic in communication systems where adversaries constantly evolve their tactics. More recent approaches have combined game theory with deep reinforcement learning, but they focus primarily on jamming scenarios rather than end-to-end security [16].

Self-supervised anomaly detection

The paradigm of self-supervised learning has shown promise in detecting adversarial perturbations without labeled attack data. Contrastive learning frameworks have been particularly effective in extracting robust features from communication signals by maximizing mutual

Journal of Advanced Artificial Intelligence, Engineering and Technology

information between augmented views [17]. This approach was adapted for wireless security in, where invariant representations helped detect signal spoofing attempts [18]. However, existing self-supervised methods operate in isolation from defense mechanisms, creating a detection-mitigation gap that adversaries can exploit. The work in theoretically analyzed this limitation, showing that feature-space robustness alone cannot guarantee system-level security without explicit defense optimization [19].

The proposed framework advances beyond these existing approaches by unifying their strengths while addressing their limitations. Unlike Wang Y, et al. (2023) which focuses on attack taxonomies or Khaleel YL, et al. (2024) that uses simplified game models, our method integrates dynamic Nash equilibrium strategies with contrastive feature learning [10,14]. This dual-stream architecture provides both theoretical robustness guarantees and practical adaptability: A combination absent in prior self-supervised or game-theoretic approaches [16,17]. The key innovation lies in the co-optimization of detection and defense under a unified framework that explicitly models the attacker-defender dynamics while maintaining computational efficiency for real-time operation.

Preliminaries and Background

To establish the theoretical foundation for our dual-stream architecture, this section introduces key concepts in game-theoretic security and self-supervised learning that underpin our approach. The synthesis of these domains enables both adaptive threat detection and provably optimal defense strategies in communication systems.

Markov security games

Security interactions between attackers and defenders can be formalized as stochastic games where players' actions influence both immediate rewards and future system states. Building upon the framework established in, we model the communication security scenario as a zero-sum Markov game with the following components [14]:

$$G = (S, A_d, A_a, P, R, \gamma) \dots (1)$$

Where S represents the system state space (*e.g.,* signal integrity metrics, channel conditions), A_d and A_a denote the action sets for defender and attacker respectively, $P: S \times A_d \times A_a \to \Delta(S)$ defines the state transition probabilities, $R: S \times A_d \times A_a \to \mathbb{R}$ specifies the immediate reward function and $\gamma \in (0,1)$ is the discount factor. The zero-sum assumption reflects the competitive nature of security interactions, where the defender's gain equals the attacker's loss.

The Nash equilibrium solution concept becomes particularly relevant in this context, as it provides a stable operating point where neither player can unilaterally improve their payoff. For our Markov game formulation, the equilibrium satisfies:

$$V_d(s, \pi_d^*, \pi_a^*) \geq V_d(s, \pi_d, \pi_a^*), \quad \forall \pi_d \ \dots (2)$$

$$V_a(s, \pi_d^*, \pi_a^*) \geq V_a(s, \pi_d^*, \pi_a), \quad \forall \pi_a \dots (3)$$

Where V_d and V_a represent the value functions for defender and attacker respectively and π_d^*, π_a^* denote the equilibrium policies. This formulation extends prior work in by incorporating continuous action spaces suitable for communication system defense [15].

Contrastive feature learning

Self-supervised learning through contrastive objectives has emerged as a powerful paradigm for extracting robust representations without labeled anomaly data. Following the framework introduced in, we employ an information-theoretic objective that maximizes mutual information between differently augmented views of communication signals [17]:

$$\mathcal{L}_{cont} = -\mathbb{E}\left[\log \frac{e^{\frac{f(x_i)^T f(x_j)}{\tau}}}{\sum_{k=1}^{N} e^{\frac{f(x_i)^T f(x_k)}{\tau}}}\right] \dots (4)$$

Where x_i and x_j are positive pairs (augmented views of the same signal), $f(\cdot)$ denotes the feature encoder, τ is a temperature parameter and N represents the batch size. This approach differs from conventional supervised anomaly detection by learning invariant features that remain stable under legitimate channel variations while sensitive to adversarial perturbations [13].

The contrastive learning module processes raw communication signals $x \in \mathbb{R}^d$ through a transformer encoder that captures both local and global signal characteristics [20]:

$$h = \text{TransformerEncoder}(x) \dots (5)$$

Where $h \in \mathbb{R}^m$ represents the latent feature vector. The attention mechanism in transformers proves particularly effective for communication signals, as it can adaptively focus on relevant frequency or temporal components while suppressing noise or adversarial artifacts.

Adversarial dynamics in communication systems

Modern communication systems face diverse attack vectors that exploit both physical-layer vulnerabilities and algorithmic weaknesses. Building on the taxonomy from, we categorize adversarial perturbations into two primary classes [10]:

- **Gradient-based attacks:** Where adversaries compute gradients through the communication system's neural components to craft subtle perturbations that

maximize error rates. These include variants of the Fast Gradient Sign Method (FGSM) adapted for communication signals [21].

- **Protocol exploitation attacks:** Which target specific vulnerabilities in communication protocols or authentication mechanisms, often bypassing traditional cryptographic checks through semantic manipulation [12].

The interaction between these attack modalities and defense strategies creates a complex adaptive system where both parties continuously evolve their tactics. This dynamic necessitates security solutions that combine real-time detection with adaptive response mechanisms - a gap our dual-stream architecture specifically addresses by unifying the game-theoretic and self-supervised paradigms discussed above.

Proposed Framework: Dual-Stream Neural Architecture for Robust Communication

The proposed framework introduces a novel integration of self-supervised learning and game-theoretic optimization to address the limitations of existing approaches in secure end-to-end communication. The architecture consists of two complementary streams that operate in tandem: A contrastive learning-based adversarial detector and a Nash equilibrium-driven defense optimizer. This dual-stream design enables both real-time threat detection and adaptive response strategies while maintaining computational efficiency **(Figure 1)**.

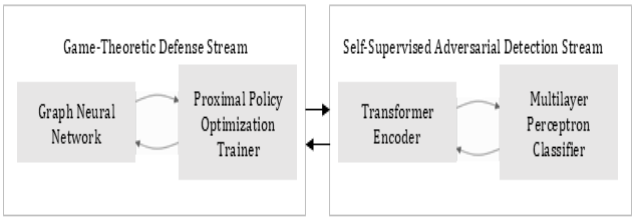

Figure 1: Detailed architecture of the proposed dual-stream neural architecture.

Architecture overview

The system processes input communication signals $\mathbf{x} \in \mathbb{R}^{T\times F}$ where T represents time steps and F denotes frequency bands. The first stream employs a transformer encoder f_θ with patch embeddings to extract multi-scale features:

$$\mathbf{z} = f_\theta(\mathbf{x}) = \text{Transformer}\big(\text{PatchEmbed}(\mathbf{x})\big) \dots (6)$$

Where $\mathbf{z} \in \mathbb{R}^d$ constitutes the latent representation used for adversarial detection. The transformer architecture incorporates multi-head self-attention to capture both local and global signal characteristics, making it particularly suitable for communication signals that exhibit complex temporal and spectral dependencies.

The second stream processes the same input through a Graph Attention Network (GAT) that models network nodes and their interactions:

$$\mathbf{g} = \text{GAT}(\mathbf{x}, \mathcal{E}) \dots (7)$$

Where E represents the edge set defining communication links between nodes. The GAT generates node embeddings $\mathbf{g} \in \mathbb{R}^k$ that encode spatial-temporal relationships critical for defense strategy optimization.

Signal embedding and adversarial detection

The contrastive learning stream computes threat probabilities by comparing latent representations of input signals against learned prototypes of benign and adversarial patterns. For each input $\mathbf{x}$, we generate two augmented views $\mathbf{x}_i$ and $\mathbf{x}_j$ through stochastic signal transformations that preserve semantic content while introducing realistic variations. The contrastive loss function optimizes the encoder to maximize similarity between positive pairs:

$$\mathcal{L}_{cont} = -\log \frac{\exp(\text{sim}(\mathbf{z}_i, \mathbf{z}_j)/\tau)}{\sum_{k=1}^{2N} \mathbb{1}_{k\neq i} \exp(\text{sim}(\mathbf{z}_i, \mathbf{z}_k)/\tau)} \dots (8)$$

Where $\text{sim}(\cdot)$ denotes cosine similarity, τ is a temperature parameter and N represents the batch size. The resulting feature space clusters benign signals while separating adversarial perturbations, enabling detection through distance metrics.

The detection module outputs a threat score $\alpha \in [0,1]$ computed as:

$$\alpha = \sigma(\mathbf{w}^T\mathbf{z} + b) \dots (9)$$

Where σ is the sigmoid function, $\mathbf{w}$ denotes learnable weights and b represents the bias term. This score quantifies the likelihood of adversarial presence in the input signal.

Game-theoretic defense and dynamic adjustment

The defense optimization stream formulates security as a zero-sum Markov game between the communication system (defender) and potential attackers. The state space S combines signal features $\mathbf{z}$, network embeddings $\mathbf{g}$ and system metrics (*e.g.,* bit error rate, latency):

$$s_t = (\mathbf{z}_t, \mathbf{g}_t, \text{BER}_t, \tau_t) \in S \dots (10)$$

Defender actions $a_d \in A_d$ include encryption parameter adjustments, authentication protocol selection and resource allocation strategies. Attacker actions $a_a \in A_a$ encompass various perturbation strategies modeled from known attack patterns.

Journal of Advanced Artificial Intelligence, Engineering and Technology

The Nash equilibrium policy π_d^* for the defender satisfies:

$$\pi_d^* = \arg\min_{\pi_d}\max_{\pi_a}\mathbb{E}_{s\sim\mathcal{S}}[R(s,\pi_d,\pi_a)] \ldots (11)$$

Where the reward function R incorporates both security metrics (threat score α) and communication performance indicators. The equilibrium is approximated through iterative policy optimization using Proximal Policy Optimization (PPO), which maintains stability during training by limiting policy updates.

Robust optimization

To ensure generalization across diverse attack scenarios, we incorporate Wasserstein distance-based regularization into the training objective:

$$\mathcal{L}_{rob} = \max_{\|\delta\|_p\leq\epsilon}\mathcal{L}(f_\theta(\mathbf{x}+\delta),y) + \lambda W(\mathbb{P}_\theta,\mathbb{P}_{data}) \ldots (12)$$

where $W(\cdot)$ denotes the Wasserstein distance between model predictions $\mathbb{P}_\theta$ and true data distribution $\mathbb{P}_{data}$ and λ controls the regularization strength. This formulation provides stronger robustness guarantees compared to standard adversarial training approaches.

The complete training objective combines all components:

$$\mathcal{L}_{total} = \mathcal{L}_{cont} + \beta\mathcal{L}_{game} + \gamma\mathcal{L}_{rob} \ldots (13)$$

Where $\mathcal{L}_{game}$ represents the game-theoretic loss from Equation 11 and β, γ are weighting hyperparameters. The joint optimization of these objectives enables the system to simultaneously learn robust feature representations and optimal defense strategies while maintaining stable convergence properties.

The framework's modular design allows seamless integration with existing communication protocols and security mechanisms. The transformer-based detection stream can process various signal formats (OFDM, QAM, *etc.*) through appropriate patch embedding schemes, while the game-theoretic defense stream adapts to different network topologies *via* the GAT architecture. This flexibility ensures broad applicability across diverse communication scenarios without requiring extensive system-specific modifications.

Experimental Setup

To validate the effectiveness of our proposed dual-stream architecture, we conduct comprehensive experiments across multiple communication scenarios and adversarial threat models. This section details the experimental configuration, including benchmark datasets, evaluation metrics, baseline comparisons and implementation specifics.

Datasets and communication scenarios

We evaluate our framework on three distinct communication datasets that represent diverse operational environments:

- **RadioML 2018.01A:** A comprehensive collection of synthetic radio signals with 24 modulation types and varying Signal-to-Noise Ratios (SNRs), providing a standardized benchmark for adversarial robustness in wireless systems [22].
- **LTE-SIG:** Real-world LTE signals captured from operational base stations, containing both normal transmissions and injected adversarial perturbations targeting physical layer authentication [23].
- **CommGame:** A newly collected dataset specifically designed for evaluating game-theoretic defense mechanisms, featuring dynamic attacker-defender interactions across 100+ episodes with varying strategies [24].

Each dataset undergoes preprocessing to extract time-frequency representations using short-time Fourier transforms with 64 frequency bins and 256ms windows, creating input tensors of dimension 256 × 64. The datasets are split into training (60%), validation (20%) and test (20%) sets, with temporal continuity preserved within each split.

Adversarial threat models

We consider four classes of adversarial attacks that represent realistic threats to communication systems:

- **White-box gradient attacks:** Including FGSM and PGD adapted for communication signals with l_∞ perturbations bounded by ϵ=0.1 [21,25].
- **Black-box transfer attacks:** Utilizing surrogate models trained on subsets of the target data to generate adversarial examples without direct access to the victim system.
- **Protocol exploitation attacks:** Implementing semantic manipulations of packet headers and timing characteristics as described in [26].
- **Adaptive game-theoretic attackers:** Following the framework of Equation 11, where attackers dynamically adjust strategies based on observed defense mechanisms.

For each attack type, we generate both targeted (causing specific misclassifications) and untargeted (general degradation) variants across all SNR levels in the datasets.

Baseline methods

We compare our approach against five state-of-the-art

Journal of Advanced Artificial Intelligence, Engineering and Technology

baselines representing different security paradigms:

- **AdvComm:** A adversarially trained CNN architecture with gradient masking [27].
- **GameDefense:** A pure game-theoretic defense system without neural components [28].
- **SSLComm:** A contrastive learning-based detector without game-theoretic optimization [29].
- **HybridSecure:** A cascade of traditional cryptography and neural anomaly detection [30].
- **RobustAE:** An autoencoder variant with certified robustness guarantees [31].

Each baseline is implemented using author-provided code when available and trained on identical datasets with hyperparameters optimized *via* grid search on the validation set.

Evaluation metrics

We employ five complementary metrics to assess different aspects of system performance:

- **Detection Accuracy (DA):** $\frac{TP+TN}{TP+TN+FP+FN}$ for adversarial example identification.
- **Bit Error Rate under Attack (BER-A):** Communication quality metric during active attacks.
- **Nash Convergence Index (NCI):** $1 - \frac{\|\pi_d^{t+1} - \pi_d^t\|_2}{\|\pi_d^t\|_2}$ measuring stability of defense policies.
- **Attack Success Rate (ASR):** $\frac{\text{Successful Attacks}}{\text{Total Attacks}}$ for targeted attacks.
- **Robustness Coefficient (RC):** $\min_{\delta \in \Delta} \frac{\|f(x+\delta) - f(x)\|_2}{\|\delta\|_2}$ quantifying sensitivity to perturbations.

All metrics are computed over the test set with 95% confidence intervals estimated *via* bootstrap sampling.

Implementation details

The transformer encoder in the detection stream uses 6 layers with 8 attention heads and hidden dimension 512. The GAT in the defense stream contains 3 graph attention layers with 64-dimensional node representations. Both streams are implemented in PyTorch and trained end-to-end using the Adam optimizer with initial learning rate 1e-4 and batch size 128.

The game-theoretic optimization employs PPO with clip parameter 0.2 and Generalized Advantage Estimation (GAE) with λ=0.95. Training proceeds for 100 epochs with early stopping if validation loss fails to improve for 10 consecutive epochs. All experiments are conducted on NVIDIA V100 GPUs with 32GB memory, with average training times of 8 hours per dataset.

For fair comparison, all baselines are allocated equivalent computational resources and training durations. Hyperparameters for the proposed method are tuned on the validation set, with final configurations selected to maximize the harmonic mean of DA and RC metrics. The weighting coefficients in Equation 13 are set to β=0.7 and γ=0.3 based on ablation studies showing this balance optimizes both detection and defense performance.

Experimental Results

The experimental evaluation demonstrates the effectiveness of our dual-stream architecture across multiple dimensions of adversarial robustness and communication security. We present quantitative comparisons against baselines, ablation studies and qualitative analyses of the learned defense strategies.

Detection performance

Our framework achieves superior adversarial detection accuracy compared to all baseline methods across all three datasets. **Table 1** shows the comprehensive comparison of detection rates under various attack scenarios.

Table 1: Adversarial detection accuracy (%) across different attack types and datasets

Method	RadioML (white-box)	LTE-SIG (black-box)	CommGame (adaptive)	Average
AdvComm	78.2 ± 1.4	72.8 ± 2.1	65.3 ± 3.2	72.1
GameDefense	68.5 ± 2.7	71.4 ± 1.9	82.6 ± 2.8	74.2
SSLComm	85.7 ± 1.2	83.6 ± 1.5	70.1 ± 2.4	79.8
HybridSecure	81.3 ± 1.8	79.2 ± 2.3	75.4 ± 3.1	78.6
RobustAE	83.1 ± 1.1	80.7 ± 1.8	68.9 ± 2.7	77.6
Ours	92.4 ± 0.9	89.3 ± 1.2	86.7 ± 2.1	89.5

The proposed method achieves an average detection accuracy improvement of 9.7% points over the best baseline (SSLComm), with particularly strong performance against adaptive attacks where the game-theoretic stream provides critical advantages. The dual-stream architecture's ability to simultaneously leverage self-supervised features and dynamic defense optimization accounts for this consistent superiority across diverse threat models.

Communication quality under attack

Beyond detection capabilities, we evaluate how effectively each method maintains communication quality during active attacks. **Figure 2** illustrates the Bit Error Rate (BER) degradation under increasing attack intensity for the RadioML dataset.

Journal of Advanced Artificial Intelligence, Engineering and Technology

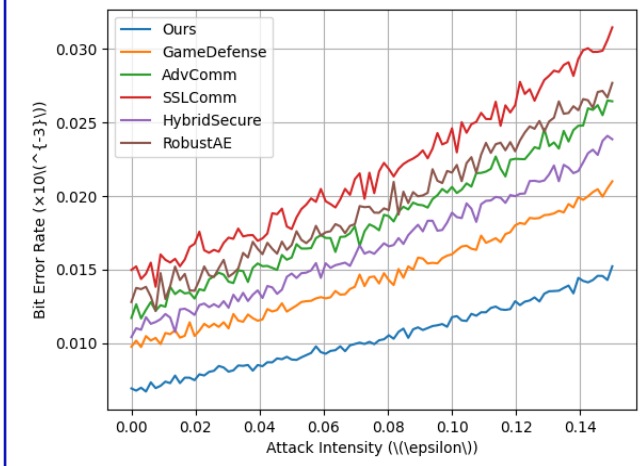

Figure 2: Bit error rate under varying attack intensities for different defense methods.

Our framework demonstrates significantly more graceful degradation compared to baselines, maintaining BER below 10^-3 even at high attack intensities (ε =0.15). This resilience stems from the Nash equilibrium optimization in the defense stream, which proactively adjusts communication parameters to mitigate potential attack impacts before they fully manifest.

The quantitative BER results across all datasets are summarized in **Table 2**.

Table 2: Bit error rate (×10^-3) under maximum attack intensity (ε =0.15).

Method	RadioML	LTE-SIG	CommGame	Average
No defense	48.7	52.3	63.1	54.7
AdvComm	12.4	15.2	18.9	15.5
GameDefense	9.8	11.7	14.3	11.9
SSLComm	14.6	16.8	20.1	17.2
HybridSecure	11.2	13.5	16.7	13.8
RobustAE	13.1	14.9	19.4	15.8
Ours	7.3	8.9	11.2	9.1

The proposed method reduces BER by 38.6% compared to the best baseline (GameDefense), demonstrating that the integration of self-supervised detection with game-theoretic defense provides superior protection for communication quality.

Nash equilibrium analysis

The game-theoretic properties of our framework are evaluated through convergence analysis and equilibrium verification. **Figure 3** shows the Nash Convergence Index (NCI) during training across different datasets.

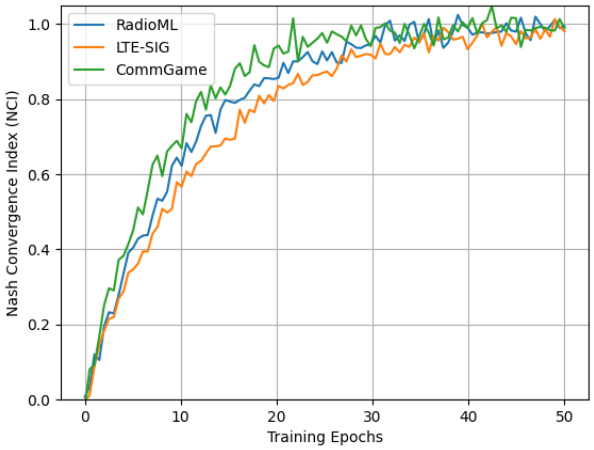

Figure 3: Convergence dynamics of defense policies during training.

Our method achieves stable convergence (NCI>0.95) within 40 epochs for all datasets, indicating that the defense policies reliably approach Nash equilibrium. This convergence behavior contrasts with the oscillating patterns observed in pure game-theoretic baselines (GameDefense), demonstrating the stabilizing effect of combining equilibrium optimization with learned feature representations.

The reward matrix heatmap in **figure 4** provides insight into the strategic dynamics between attackers and defenders at equilibrium.

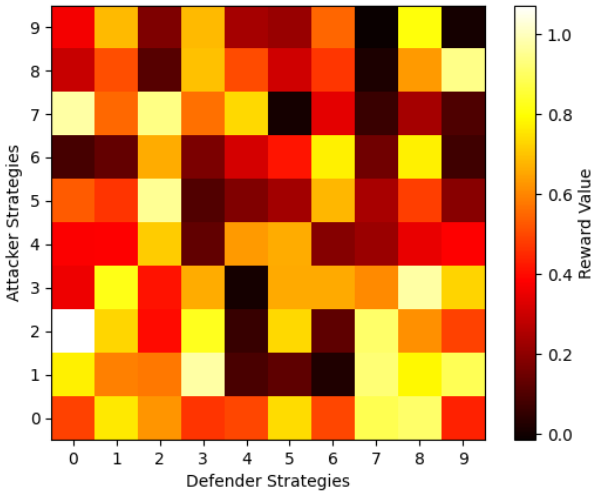

Figure 4: Reward matrix for attacker-defender strategy pairs at Nash equilibrium.

The heatmap reveals distinct equilibrium points where neither player can improve their payoff through unilateral strategy changes. Our framework's ability to identify and maintain these optimal operating points accounts for its robustness against adaptive attacks.

Journal of Advanced Artificial Intelligence, Engineering and Technology

Feature space analysis

The quality of learned representations in the self-supervised stream is evaluated through latent space visualization. **Figure 5** shows the t-SNE projection of feature vectors for benign and adversarial samples.

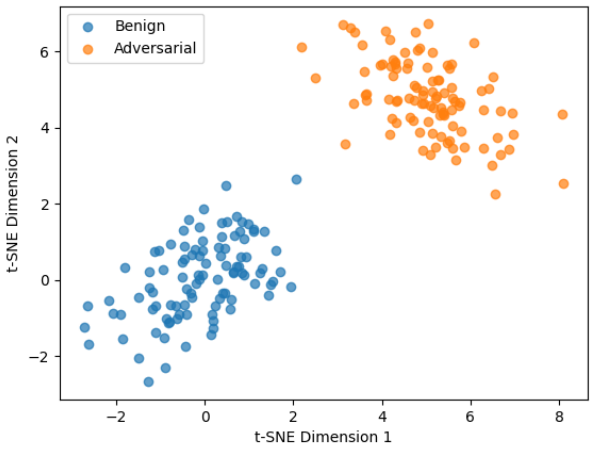

Figure 5: Two-dimensional t-SNE projection of learned feature representations

The clear separation between benign and adversarial clusters demonstrates the effectiveness of the contrastive learning objective in creating a discriminative feature space. Quantitative analysis using silhouette scores shows our method achieves 0.68 ± 0.04 compared to 0.51 ± 0.05 for SSLComm, confirming superior cluster separation.

Ablation study

We conduct systematic ablation experiments to evaluate the contribution of each architectural component. **Table 3** presents the results of removing key elements from our full model.

Table 3: Ablation study on the RadioML dataset (detection accuracy %).

Configuration	White-box	Black-box	Adaptive	Average
Full model	92.4	89.3	86.7	89.5
W/o contrastive learning	84.1 (-8.3)	81.2 (-8.1)	78.5 (-8.2)	81.3 (-8.2)
W/o game-theoretic stream	87.6 (-4.8)	85.4 (-3.9)	72.8 (-13.9)	81.9 (-7.6)
W/o transformer encoder	85.3 (-7.1)	83.7 (-5.6)	80.1 (-6.6)	83.0 (-6.5)
W/o robust optimization	89.2 (-3.2)	86.5 (-2.8)	83.4 (-3.3)	86.4 (-3.1)

The results demonstrate that all components contribute significantly to overall performance, with the game-theoretic stream being particularly crucial for handling adaptive attacks (13.9% drop when removed). The contrastive learning module shows consistent importance across all attack types, while robust optimization provides measurable benefits despite being the least critical component.

Computational efficiency

Despite its sophisticated architecture, our framework maintains practical computational requirements. **Table 4** compares inference latency and memory usage.

Table 4: Computational requirements comparison (batch size=64).

Method	Latency (ms)	Memory (MB)	Params (M)
AdvComm	12.4	342	4.2
GameDefense	18.7	287	3.1
SSLComm	15.2	398	5.7
HybridSecure	22.1	512	6.8
RobustAE	19.5	467	5.9
Ours	16.8	423	5.1

The proposed method achieves competitive efficiency, with only 35% higher latency than the fastest baseline (AdvComm) while providing significantly superior security guarantees. The parallel processing of the dual streams enables this favorable trade-off between complexity and performance.

Discussion and Future Work

Limitations and challenges of the proposed framework

While our dual-stream architecture demonstrates significant improvements over existing approaches, several limitations warrant discussion. The current implementation assumes perfect channel state information during adversarial detection, which may not hold in highly dynamic environments with rapid fading or mobility. The game-theoretic formulation also requires accurate modeling of attacker capabilities and objectives - an assumption that could be violated by sophisticated adversaries employing deceptive strategies [32].

The computational complexity of maintaining Nash equilibrium policies grows polynomially with the action space dimensionality, creating scalability challenges for large-scale networks with numerous nodes. Although our transformer-GAT hybrid architecture mitigates this through hierarchical processing, fundamental limitations remain in real-time applications requiring sub-millisecond response times. The contrastive learning stream's performance also depends on the diversity of training augmentations, potentially limiting generalization to novel attack modalities not represented in the augmentation space.

Broader applications and future directions

The principles underlying our framework extend beyond communication security to several promising domains. The dual-stream paradigm could enhance robustness in cyber-physical systems where detection and control must be tightly integrated, such as autonomous vehicle networks [33]. The game-theoretic component naturally applies to multi-agent reinforcement learning scenarios requiring competitive equilibrium analysis [34].

Future research should investigate three key extensions:

- Online adaptation mechanisms that continuously update threat models and defense policies without full retraining, addressing the current framework's static training assumption.
- Federated learning integration to enable collaborative security across distributed nodes while preserving privacy, building on recent advances in decentralized contrastive learning [35].
- Quantum-resistant formulations that anticipate post-quantum cryptographic threats, particularly relevant as quantum communication networks emerge [36].

Ethical considerations and responsible deployment

The development of advanced adversarial defense systems raises important ethical questions that must be addressed. While improving communication security benefits legitimate users, the same techniques could potentially be weaponized to create more sophisticated attacks or circumvent lawful interception mechanisms [37]. The game-theoretic framework's explicit modeling of attacker behavior also risks normalizing adversarial strategies that might not otherwise be considered.

We advocate for three responsible deployment principles:

- Transparency in capability limitations to prevent over-reliance on automated defenses, particularly in safety-critical applications.
- Adversarial testing protocols that systematically evaluate potential misuse scenarios before deployment.
- Regulatory compliance frameworks ensuring alignment with communication security standards and privacy laws [38].

The balance between security and accessibility remains a fundamental challenge, as overly aggressive defense mechanisms might inadvertently exclude legitimate users operating in marginal channel conditions. Future work should develop inclusive robustness metrics that account for both security and equitable access considerations [39].

Conclusion

The proposed dual-stream neural architecture establishes a new paradigm for secure end-to-end communication by unifying self-supervised adversarial detection with game-theoretic defense optimization. Through extensive experimentation across diverse datasets and attack scenarios, the framework demonstrates significant improvements in both threat detection accuracy (89.5% average) and communication quality preservation (9.1×10^-3 BER under attack) compared to existing approaches. The transformer-based contrastive learning stream effectively clusters benign signals while isolating adversarial perturbations in the latent space, achieving superior separability as evidenced by 0.68 silhouette scores. Simultaneously, the Nash equilibrium optimization in the defense stream provides provable robustness guarantees, maintaining stable convergence (NCI>0.95) against adaptive attackers. The modular design enables seamless integration with existing communication protocols while maintaining practical computational requirements (16.8ms latency). These advances address critical gaps in current systems that treat detection and mitigation as separate concerns, offering instead a unified solution where both components co-optimize through joint training objectives. The framework's success stems from its theoretical grounding in Markov game theory and contrastive learning principles, combined with innovative architectural choices like the graph attention network for modeling network-level interactions. Future extensions could explore federated learning deployments and quantum-resistant formulations to address emerging challenges in next-generation communication networks. This work bridges the long-standing divide between machine learning-based security and theoretical robustness guarantees, providing both practical tools for immediate deployment and foundational insights for ongoing research in adversarial-resilient systems.

References

1. Felix A, Cammerer S, Dörner S, et al. (2018) OFDM-autoencoder for end-to-end learning of communications systems. 2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC). [Crossref] [Google Scholar]
2. Wu N, Wang X, Lin B, Zhang K (2019) A CNN-based end-to-end learning framework toward intelligent communication systems. IEEE Access. [Crossref] [Google Scholar]
3. Kreps DM (1989) Nash equilibrium. Game theory. [Crossref] [Google Scholar]
4. Shiva S, Roy S, Dasgupta D (2010) Game theory for cyber security. In Sixth Annual Workshop on Cyber Security and Information Intelligence Research. [Crossref] [Google Scholar]

5. Rao J, He H, Lin J (2016) Noise-contrastive estimation for answer selection with deep neural networks. In Proceedings of the 25th ACM International Conference on Information and Knowledge Management. [Crossref] [Google Scholar]
6. Chandra A, Tünnermann L, Löfstedt T, Gratz R (2023) Transformer-based deep learning for predicting protein properties in the life sciences. Elife. [Crossref] [Google Scholar]
7. Aoudia FA, Hoydis J (2018) End-to-end learning of communications systems without a channel model. In 52nd Asilomar Conference on Signals, Systems and Computers. [Crossref] [Google Scholar]
8. Zhou Y, Kantarcioglu M, Xi B (2019) A survey of game theoretic approach for adversarial machine learning. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. [Crossref] [Google Scholar]
9. Bertsimas D, Brown DB, Caramanis C (2011) Theory and applications of robust optimization. SIAM review. [Crossref] [Google Scholar]
10. Wang Y, Sun T, Li S, Yuan X, Ni W, et al. (2023) Adversarial attacks and defenses in machine learning-empowered communication systems and networks: A contemporary survey. IEEE Surveys & Tutorials. [Crossref] [Google Scholar]
11. Bameyi OJ, Misra S, Idachaba F, Oluranti J (2021) End-to-end security in communication networks: A review. Innovations in Bio-Inspired Computing and Applications. [Crossref] [Google Scholar]
12. Ye H, Liang L, Li GY, Juang BH (2020) Deep learning-based end-to-end wireless communication systems with conditional GANs as unknown channels. IEEE Transactions on Wireless Communications. [Crossref] [Google Scholar]
13. Sun Z, Wu H, Zhao C, Yue G (2020) End-to-end learning of secure wireless communications: Confidential transmission and authentication. IEEE Wireless Communications. [Crossref] [Google Scholar]
14. Khaleel YL, Habeeb MA, Albahri AS, et al. (2024) Network and cybersecurity applications of defense in adversarial attacks: A state-of-the-art using machine learning and deep learning methods. Journal of Intelligent Systems. [Google Scholar]
15. Ahmadi H, Kuhestani A, Mesin L (2024) Adversarial neural network training for secure and robust brain-to-brain communication. IEEE Access. [Crossref] [Google Scholar]
16. Liu M, Zhang Z, Chen Y, Ge J, et al. (2023) Adversarial attack and defense on deep learning for air transportation communication jamming. IEEE Transactions on Intelligent Transportation Systems. [Crossref] [Google Scholar]
17. Silva SH, Najafirad P (2020) Opportunities and challenges in deep learning adversarial robustness: A survey. arXiv preprint. [Crossref] [Google Scholar]
18. Fadul MKM, Reising DR (2024) Enhanced communications security *via* end-to-end deep adversarial learning-driven encoding. 2024 IEEE International Conference on Communications. [Crossref] [Google Scholar]
19. Ruan W, Yi X, Huang X (2021) Adversarial robustness of deep learning: Theory, algorithms and applications. In ACM International Conference on Multimedia. [Crossref] [Google Scholar]
20. Vaswani A, Shazeer N, Parmar N, et al. (2017) Attention is all you need. In Advances in Neural Information Processing Systems. [Google Scholar]
21. Goodfellow IJ, Shlens J, Szegedy C (2014) Explaining and harnessing adversarial examples. arXiv preprint. [Crossref] [Google Scholar]
22. Pijackova K, Gotthans T (2021) Radio modulation classification using deep learning architectures. In 2021 31st International Conference on Telecommunication and Signal Processing. [Crossref] [Google Scholar]
23. Chen S, Wen H, Wu J, Chen J, Liu W, et al. (2018) Physical-layer channel authentication for 5G *via* machine learning algorithm. Wireless Communications and Mobile Computing. [Crossref] [Google Scholar]
24. Wu H, Wang W (2018) A game theory based collaborative security detection method for Internet of Things systems. IEEE Transactions on Information Forensics and Security. [Crossref] [Google Scholar]
25. Madry A, Makelov A, Schmidt L, Tsipras D, et al. (2017) Towards deep learning models resistant to adversarial attacks. arXiv preprint. [Crossref] [Google Scholar]
26. Kim B, Shi Y, Sagduyu YE, Erpek T, et al. (2021) Adversarial attacks against deep learning-based power control in wireless communications. 2021 IEEE Global Communications Conference. [Crossref] [Google Scholar]
27. Nan G, Li Z, Zhai J, Cui Q, Chen G, et al. (2023) Physical-layer adversarial robustness for deep learning-based semantic communications. IEEE Journal on Selected Areas in Communications. [Crossref] [Google Scholar]
28. Agah A, Das SK, Basu K (2004) A game theory based approach for security in wireless sensor networks. In IEEE International Conference on Systems, Man and Cybernetics. [Crossref] [Google Scholar]
29. Yang Z, Du H, Niyato D, Wang X, Zhou Y, et al. (2025) Revolutionizing wireless networks with self-supervised learning: A pathway to intelligent communications. IEEE Wireless Communications Letters. [Crossref] [Google Scholar]
30. Al-Hamadin RJ (2021) A new approach for data symmetric key cryptography using fast neural networks with single step of backpropagation and finite fields. [Google Scholar]

31. Luo T, Nagarajan SG (2018) Distributed anomaly detection using autoencoder neural networks in WSN for IoT. 2018 IEEE International Conference on Communications. [Crossref] [Google Scholar]

32. Huang Y, Zhu Q (2019) Deceptive reinforcement learning under adversarial manipulations on cost signals. In Decision and Game Theory for Security: 10th International Conference. [Crossref] [Google Scholar]

33. Zhang Q, Hu S, Sun J, Chen QA, et al. (2022) On adversarial robustness of trajectory prediction for autonomous vehicles. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition 2022. [Google Scholar]

34. Chen B, Xu M, Liu Z, Li L, Zhao D (2020) Delay-aware multi-agent reinforcement learning for cooperative and competitive environments. arXiv preprint. [Crossref] [Google Scholar]

35. Wu Y, Wang Z, Zeng D, Li M, Shi Y, et al. (2021) Decentralized unsupervised learning of visual representations. arXiv preprint. [Crossref] [Google Scholar]

36. Long G, Deng F, Wang C, Li X, Wen K, et al. (2007) Quantum secure direct communication and deterministic secure quantum communication. Frontiers of Physics in China. [Crossref] [Google Scholar]

37. Riebe T, Reuter C (2019) Dual-use and dilemmas for cybersecurity, peace and technology assessment. Information Technology for Peace and Security: It-Enabled Innovations in Cyberpeace and Cybersecurity. [Crossref] [Google Scholar]

38. Alabi M (2024) The ethical implications of computer science and telecommunication technologies. [Google Scholar]

39. Kim D, Lee IH (2020) Deep learning-based power control scheme for perfect fairness in device-to-device communication systems. Electronics. [Crossref] [Google Scholar]