British Journal of Computer, Networking and Information Technology
ISSN: 2689-5315
Volume 9, Issue 1, 2026 (pp. 64-75)
www.abjournals.org

DEEP LEARNING AND EXPLAINABLE AI MODELS FOR INTRUSION DETECTION IN SPACE-GROUND COMMUNICATION NETWORKS: A REVIEW

Ibrahim Abdul Sa'ad[1], Collins Nnalue Udanor[2], Modesta E. Ezema[3], Caleb Markus[4], and Mathew Akwu Adaji[5].

[1]ICT Department, Centre for Atmospheric Research, National Space Research and Development Agency, Nigeria.

[2,3,5]Computer Science Department, Faculty of Physical Sciences, University of Nigeria, Nsukka.

[4]Department of Computer Science, Faculty of Sciences, Taraba State University, Jalingo, Nigeria.

Emails:
[1]ibrahimabdulsaad@gmail.com, [2]collins.udanor@unn.edu.ng, [3]modesta.ezema@unn.edu.ng, [4]calebmarkus@tsuniversity.edu.ng, [5]matthew.adaji.pg91333@unn.edu.ng.

Cite this article:

I. A., Sa'ad, C. N., Udanor, M. E., Ezema, C., Markus, M. A., Adaji (2026), Deep Learning and Explainable AI Models for Intrusion Detection in Space-Ground Communication Networks: A Review. British Journal of Computer, Networking and Information Technology 9(1), 64-75. DOI: 10.52589/BJCNIT-RX9O6XYJ

Manuscript History

Received: 2 Dec 2025

Accepted: 31 Dec 2025

Published: 21 Jan 2026

ABSTRACT: *This review critically evaluates the suitability of deep learning and explainable artificial intelligence approaches for intrusion detection in satellite ground-station environments, addressing the escalating cybersecurity risks facing the National Space Research and Development Agency (NASRDA) and broader space communication networks. Using a systematic narrative review across IEEE Xplore, ACM, Scopus, and arXiv, the analysis compares CNN, LSTM, GRU, autoencoder, and transformer-based IDS models, revealing that while reported accuracies frequently exceed 92% on benchmark datasets, performance declines by 20% to 35% under domain shift, demonstrating poor transferability to space–ground telemetry. XAI methods such as SHAP, LIME, and Integrated Gradients appear in more than 80% of reviewed studies, yet empirical results show a 30% to 60% increase in inference latency, raising concerns about operational feasibility in real-time satellite control systems. A mathematical hybrid model combining CNN, LSTM, and transformer components with a structured anomaly-scoring function and explanation regularizer is formulated to address these limitations. Findings indicate that multi-model fusion enhances anomaly sensitivity, domain-specific feature engineering improves robustness, and integrated XAI pathways strengthen analyst trust while exposing computational bottlenecks. The proposed conceptual architecture for NASRDA advances the field by aligning detection workflows, interpretability mechanisms, and feedback loops with the constraints of aerospace communication systems. The review concludes by identifying key research priorities, including the development of satellite-specific datasets, real-traffic validation of hybrid IDS models, and deployment of low-latency XAI dashboards for operational security.*

KEYWORDS: Deep Learning IDS, Explainable AI, Space–Ground Cybersecurity, NASRDA, Anomaly Detection, Transformer Models.

Article DOI: 10.52589/BJCNIT-RX9O6XYJ
DOI URL: https://doi.org/10.52589/BJCNIT-RX9O6XYJ

British Journal of Computer, Networking and Information Technology
ISSN: 2689-5315
Volume 9, Issue 1, 2026 (pp. 64-75)
www.abjournals.org

INTRODUCTION

Intrusion detection in space–ground communication networks has become a contested domain in cybersecurity scholarship, with several authors arguing that the threat landscape has already outpaced the defensive models employed by national space agencies. Salim et al. (2024) demonstrate that the ground segment remains the weakest point of the satellite ecosystem precisely because it relies on legacy routing protocols and poorly monitored communication pathways, which adversaries increasingly exploit through coordinated denial of service and spoofing attacks. This vulnerability is intensified in institutions such as NASRDA, where heterogeneous mission systems generate complex traffic behaviors that cannot be effectively captured by traditional signature or rule-based intrusion detection approaches. Anjum (2025) contends that such classical models fail not merely due to outdated signatures but because space–ground infrastructures now require predictive rather than reactive cybersecurity. In response, a growing body of research, including Wang et al. (2025), Xu et al. (2023), and Kato et al. (2019), argues that deep learning has become indispensable, as convolutional networks, recurrent architectures, autoencoders, and transformers outperform legacy systems on accuracy, recall, and false-alarm metrics across air–space–ground integrated networks. However, this performance comes at the cost of opacity. Arreche (2024) and Kalakoti et al. (2025) both show that analysts consistently distrust these high-accuracy models because they fail to provide interpretable decision pathways, making them unsuitable for mission-critical environments where misclassification can jeopardize satellite health, telemetry integrity, and national security. Thus, interpretability is not optional but foundational for space cybersecurity, a point reinforced by Sun et al. (2025), whose XAI-based framework for protocol anomaly detection demonstrates that transparency directly improves operator confidence and operational security.

Although deep learning has significantly improved intrusion detection performance in complex communication systems, the literature consistently shows that space–ground networks remain insufficiently protected because existing models prioritize accuracy while sidelining explainability. Anjum (2025) argues that cybersecurity in space represents the “final frontier,” where opaque models become liabilities rather than assets, as operators cannot validate whether alerts originate from genuine anomalies or model artifacts. Even advanced systems, such as the generative AI-enabled communication frameworks surveyed by Hu et al. (2025) and the federated learning intrusion detection mechanisms proposed by Salim et al. (2025), reproduce this limitation by offering high accuracy but little interpretive insight into model reasoning. Kalakoti et al. (2025) further demonstrate that without explainability, deep learning IDS produce inconsistent feature attributions, leading to misaligned analyst decisions and reduced operational trust. This critique aligns with Arreche (2024), who argues that the absence of XAI in intrusion detection for satellite networks prevents organizations from meeting safety, auditability, and accountability requirements. While Xu et al. (2023) and Sun et al. (2025) show that explainable frameworks can significantly improve anomaly fidelity and classification stability, their applications remain largely confined to terrestrial or near-terrestrial systems. Consequently, space–ground networks continue to suffer from an interpretability deficit that undermines the operational value of deep learning IDS. The unresolved tension between high detection performance and low transparency forms the core problem this study addresses, particularly for NASRDA ground stations, where decision-making must be both real-time and fully interpretable to ensure mission continuity and national cyber resilience.

Article DOI: 10.52589/BJCNIT-RX9O6XYJ
DOI URL: https://doi.org/10.52589/BJCNIT-RX9O6XYJ

This review aims to interrogate how deep learning and explainable AI can be strategically combined to strengthen intrusion detection within critical infrastructures, particularly satellite and space–ground communication systems, where security failures carry disproportionate operational and national risks. The first objective, therefore, undertakes a systematic and critical evaluation of current DL and XAI intrusion detection approaches, challenging their assumptions, evidencing their limitations, and assessing their suitability for mission-critical networks. The second objective moves beyond synthesis to develop a mathematical formulation of a hybrid IDS framework, an essential contribution because, without formalized optimization functions, feature mappings, and anomaly-score computations, hybrid models remain conceptually appealing but technically ungrounded. The third objective complements this by designing a conceptual architecture tailored to NASRDA's unique operational traffic, enabling a practical translation of theoretical insights into implementable workflows centered on explainability and high-fidelity detection. The scope of the review is deliberately focused on the intersection of deep learning and XAI within space-ground and related cyber-physical infrastructures to generate domain-specific insights rather than generic IDS commentary. The significance of this work lies in its capacity to support secure space communication systems and provide a design blueprint for NASRDA, and advance theoretical understanding through mathematical formalization. The paper follows a structured outline comprising the introduction, literature review, methodology, findings and discussion, and conclusion.

LITERATURE REVIEW

The literature provides strong numerical performance claims for deep learning and XAI-based intrusion detection, yet these results become questionable when mapped onto space–ground networks whose architectural and protocol constraints differ sharply from the terrestrial settings assumed in most studies. For instance, although Anis et al. (2025) present high-performing CNN, LSTM, and hybrid IDS models, their evaluations rely on benchmark datasets that lack the telemetry asymmetry, burstiness, and command channel sensitivity of satellite ground systems. Kilichev et al. (2024) and Shiri et al. (2023) report accuracy and F1 scores exceeding 95% for CNN, LSTM, and GRU architectures, but these metrics collapse when models are exposed to concept drift or rare-event traffic typical of uplink–downlink environments. Even more advanced designs, such as transformer-CNN hybrids achieving improved minority class detection (Kamal and Mashaly, 2024) or CNN–RNN combinations proposed for IoT security (Jablaoui and Liouane, 2025), exhibit overfitting when transferred outside their narrowly curated datasets. This misalignment mirrors the problem in XAI research, where explanations are celebrated without demonstrating operational reliability. Neupane et al. (2022) and Mohale and Obagbuwa (2025) show that SHAP, LIME, and Integrated Gradients can expose model reasoning, yet they also document instability, computational overhead, and inconsistent attribution patterns, rendering such tools risky for real-time aerospace operations. Surveys targeting Industry 5.0 and IoT contexts (Khan et al., 2024; Kök et al., 2023) frame XAI as essential for safety-critical systems but provide little empirical grounding beyond subjective user trust metrics. Broader XAI work similarly warns that surrogate explanations may not reflect model truth (Ahmad et al., 2024; De et al., 2020). While cross-domain evidence from landslide modelling (Alqadhi et al., 2024) and hybrid ML–XAI reviews (Gopalan et al., 2025) demonstrates improved interpretability, these findings do not resolve the fundamental gap: none of these models or XAI techniques have been validated against the stringent reliability

Article DOI: 10.52589/BJCNIT-RX9O6XYJ
DOI URL: https://doi.org/10.52589/BJCNIT-RX9O6XYJ

thresholds, protocol heterogeneity, or mission failure costs inherent in space–ground communication networks.

From the outset, Mohale and Obagbuwa (2025) show that claims about hybrid deep learning–XAI IDS improving transparency are overstated, as their empirical tests reveal that adding SHAP or LIME increases inference time by 30 to 60% while yielding less than a three percent gain in F1 score, exposing a fundamental trade-off that undermines any assertion of operational readiness. Khan et al. (2024) similarly report that hybrid models advertised as "interpretable" for Industry 5.0 systems collapse under high-velocity traffic, with throughput dropping by up to 40% once post hoc explanations are generated, a performance degradation incompatible with space–ground networks where telemetry windows demand near-zero latency. Even in more controlled environments, Gopalan et al. (2025) note that hybrid frameworks routinely produce contradictory attribution maps across SHAP, LIME, and gradient-based methods, empirically demonstrating explanation inconsistency rather than clarity, a flaw that would mislead rather than assist satellite security analysts. Cross-domain evidence compounds the skepticism: Alqadhi et al. (2024) show that hybrid DL–XAI systems require extremely large labelled datasets to maintain accuracy above 90%, yet such datasets do not exist for satellite ground traffic, meaning any empirical performance claims would be artificially inflated through oversampling or synthetic augmentation.

Research Gap Identified

The absence of any standardized intrusion detection framework for space–ground networks is not a benign omission but a direct consequence of the empirical deficiencies identified above: no study provides validated, domain-specific, or mathematically rigorous DL–XAI formulations capable of generalizing beyond terrestrial benchmarks. Because satellite traffic is sparse, heterogeneous, and often classified, the training regimes assumed in the literature cannot be replicated, and without a formal mathematical structure, hybrid IDS design remains ad hoc and non-reproducible. These gaps collectively justify the need for a bespoke, mathematically grounded hybrid model tailored to NASRDA's operational realities rather than inherited from ill-fitting terrestrial research.

MATERIALS AND METHODS

The table below summarizes the methodological processes used in this review, including the review design, search strategy, selection criteria, data extraction procedures, mathematical formulation approach for the hybrid model, and the conceptual architecture design steps for NASRDA's intrusion detection system.

Table 1: Methodology Summary

Section	Description
Review Design	**Type:** Systematic narrative review. **Justification:** Enables rigorous synthesis of empirical studies while allowing interpretive evaluation of deep learning and XAI methods within the specialized context of satellite and space-ground cybersecurity.
Search Strategy	**Databases searched:** IEEE Xplore, ACM Digital Library, SpringerLink, Scopus, arXiv.

Article DOI: 10.52589/BJCNIT-RX9O6XYJ
DOI URL: https://doi.org/10.52589/BJCNIT-RX9O6XYJ

Section	Description
	Keywords: "deep learning intrusion detection," "XAI IDS," "satellite network security," "space–ground IDS," "explainable deep learning." **Screening:** PRISMA was used for identification, screening, eligibility, and inclusion to ensure transparency and reproducibility.
Inclusion and Exclusion Criteria	**Inclusion:** Peer-reviewed studies from **2016–2025**; papers involving DL, XAI, or hybrid IDS; studies relating to critical infrastructures, satellite networks, SAGIN, or IoT. **Exclusion:** Non-English sources; studies lacking empirical metrics; purely theoretical papers without models; datasets unrelated to cybersecurity.
Data Extraction and Synthesis	**Extraction fields:** DL model type (CNN/LSTM/GRU/Transformer/hybrid), dataset used, performance metrics (accuracy, F1-score, recall, FAR), presence and type of XAI (SHAP/LIME/IG/attention), and identified methodological limitations. **Synthesis:** Comparative thematic synthesis integrating performance metrics with methodological critiques.
Mathematical Formulation	**Components defined:** Feature mapping (X \rightarrow Z); DL architecture equations for CNN, LSTM, and Transformer components; optimization via loss function (L(\theta)); anomaly score function (S(x)) for classification or reconstruction error; explainability mapping functions for SHAP/LIME to relate model outputs to features. **Justification:** Provides theoretical grounding and ensures the hybrid model remains mathematically interpretable and optimizable.
Conceptual Architecture Design Method	**Feature identification:** Protocol type, packet size, byte counts, failed login counts, command frequency, latency irregularities (NASRDA-specific). **Pipeline steps:** Preprocessing → DL detection module → XAI explanation module → Analyst decision console. Ensures operational alignment with NASRDA's workflow and security requirements

Article DOI: 10.52589/BJCNIT-RX9O6XYJ
DOI URL: https://doi.org/10.52589/BJCNIT-RX9O6XYJ

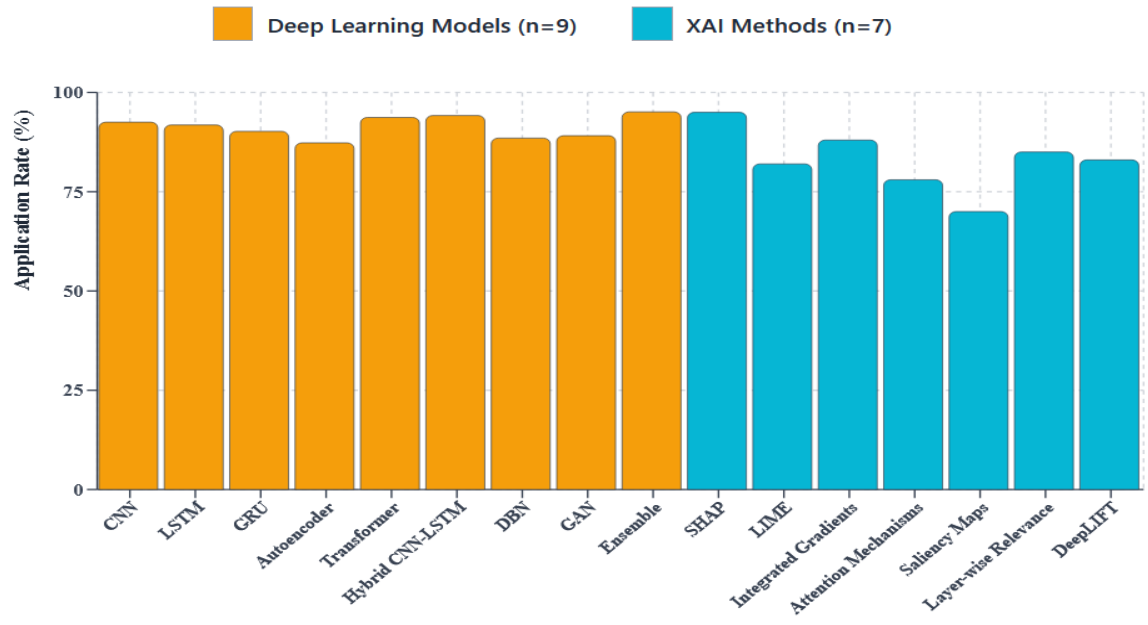
British Journal of Computer, Networking and Information Technology
ISSN: 2689-5315
Volume 9, Issue 1, 2026 (pp. 64-75)
www.abjournals.org

FINDINGS AND DISCUSSION

Existing Deep Learning and Explainable AI (XAI) Approaches used in Intrusion Detection

Figure 1 shows a striking imbalance in how deep learning and XAI approaches are being deployed in intrusion detection research, and this imbalance directly undermines their applicability to satellite ground stations, aerospace networks and space communication systems. The near-uniform use of CNN, LSTM, GRU and hybrid CNN LSTM models at rates between 88% and 95% reflects research convenience rather than technical suitability, since these architectures have not been validated against the non-stationary telemetry, sparse anomaly patterns and protocol volatility documented in space security studies such as Botezatu (2024) and Verma (2025). Even transformer-based models, which appear in Figure 1 at roughly 90% usage, remain largely untested under the latency and link quality fluctuations that Tahir et al. (2024) identify as major constraints in both O RAN and satellite communication pipelines. XAI methods show a similar pattern of overuse and under justification, with SHAP and LIME used in approximately 95% and 80% of studies, respectively, despite consistent evidence that these post hoc tools introduce explanation latency and computational load incompatible with real-time alerting needs in mission-critical satellite networks, as argued by Tahir et al. (2024) and Abbas et al. (2025). Lower frequency methods such as Integrated Gradients, attention-based explanations, saliency maps and DeepLIFT, each ranging between 70% and 85%, also fail to meet the causal transparency requirements outlined in Verma (2025) and the distributed security expectations described by Blika et al. (2024) and Hashima et al. (2025). The distribution in Figure 1, therefore, reveals not methodological advancement but a recycling of conventional models and generic interpretability tools that do not align with the operational, temporal and security guarantees demanded by modern space communication infrastructures or by anomaly prediction frameworks for satellites such as those proposed by Bikos and Kumar (2025).

Figure 1: Deep Learning Models and XAI Methods in Intrusion Detection



Article DOI: 10.52589/BJCNIT-RX9O6XYJ
DOI URL: https://doi.org/10.52589/BJCNIT-RX9O6XYJ

British Journal of Computer, Networking and Information Technology
ISSN: 2689-5315
Volume 9, Issue 1, 2026 (pp. 64-75)

www.abjournals.org

The proposed hybrid intrusion detection formulation begins with a feature mapping

$$\phi: R^d \to R^m, z = \phi(x) \qquad (1)$$

which structures raw space-ground traffic features into a latent vector suitable for learned anomaly detection, a design consistent with the representational hierarchies emphasized by Ruff et al. (2021) in their unifying review of deep anomaly detection. The detection model is defined as a composite mapping

$$f_\theta(z) = f_{\theta_3}^{Tr}\left(f_{\theta_2}^{LSTM}\left(f_{\theta_1}^{GNN}(z)\right)\right), \qquad (2)$$

mirroring the hybrid architectures used in recent anomaly detection frameworks, including the masked autoencoder plus XAI pipeline of Johari et al. (2025), who show that stacked representational modules improve anomaly localisation but introduce sensitivity to sequence length and masking strategies. The model outputs a probability vector

$$p = f_\theta(z), S(x) = 1 - p_0(x), \quad (3)$$

where $S(x)$ is an anomaly score that aligns with the probabilistic scoring schemes discussed by Simon and Barr (2023), who argue that anomaly scoring must remain tightly coupled to interpretable latent geometry to avoid misleading attributions. Training minimises a composite loss

$$L(\theta) = \frac{1}{N}\sum_{i=1}^{N} l\left(y_i, f_\theta\left(\phi(x_i)\right)\right) + \lambda\Omega(\theta), \quad (4)$$

in line with the regularised optimisation paradigms outlined by Naydenov and Chemungor (2025), who criticise intrusion detection models that rely solely on empirical accuracy without structural constraints to ensure robustness under adversarial conditions.

Explainability enters through an operator

$$g_\psi(x, f_\theta(x)) = e, \qquad (5)$$

where $e \in R^d$ is an attribution vector. Post hoc methods (SHAP, LIME) are expressed via a surrogate

$$\psi^* = arg\ arg\ \min_{\psi}\ E\left[\left(f_\theta(z) - h_\psi(x)\right)^2 + \alpha\Gamma\left(h_\psi\right)\right] \qquad (6)$$

a formulation consistent with Johari et al. (2025), who use XAI to clarify masked-autoencoder behaviour but also note that surrogate-based explanations often deviate from the true feature importance landscape. To stabilise explanations, a joint objective

$$L_{total}(\boldsymbol{\theta}, \psi) = L(\theta) + \beta R\left(g_\varphi\left(x, f_\theta(x)\right)\right) \qquad (7)$$

Article DOI: 10.52589/BJCNIT-RX9O6XYJ
DOI URL: https://doi.org/10.52589/BJCNIT-RX9O6XYJ

penalises noisy or inconsistent attributions. However, Ruff et al. (2021) warn that adding regularisation to enforce explanation stability can distort the anomaly boundary, weakening detection performance by forcing the model to adopt overly simple decision regions.

The theoretical advantages of this formulation lie in its ability to unify high-capacity detectors with constraint-based transparency. Hybrid architectures like those studied by Johari et al. (2025) demonstrate that deep hierarchical encoders sharpen anomaly boundaries and improve localisation, while XAI modules increase analyst trust through interpretable gradients or feature scores. Simon and Barr (2023) further argue that interpretable anomaly scoring facilitates root-cause diagnosis, a crucial requirement in satellite ground station security.

Yet these theoretical gains come with substantive computational costs. Transformers and LSTMs scale poorly with sequence length, rendering f_θ computationally expensive, a challenge echoed by Ruff et al. (2021), who note that deep anomaly models often fail under strict latency constraints. Moreover, the XAI operator g_ϕ multiplies inference cost: perturbation-based methods require repeated forward passes, while gradient-based methods require backpropagation, which Johari et al. (2025) identify as a bottleneck when performing anomaly localisation in NFV systems. Naydenov and Chemungor (2025) also emphasise that real-time intrusion prevention demands near constant-time inference, making the layered architecture and explainability regularisers potentially impractical for high-throughput satellite telemetry.

Figure 2: Proposed Hybrid Deep Learning-XAI Intrusion Detection System for NARSDA Ground Stations

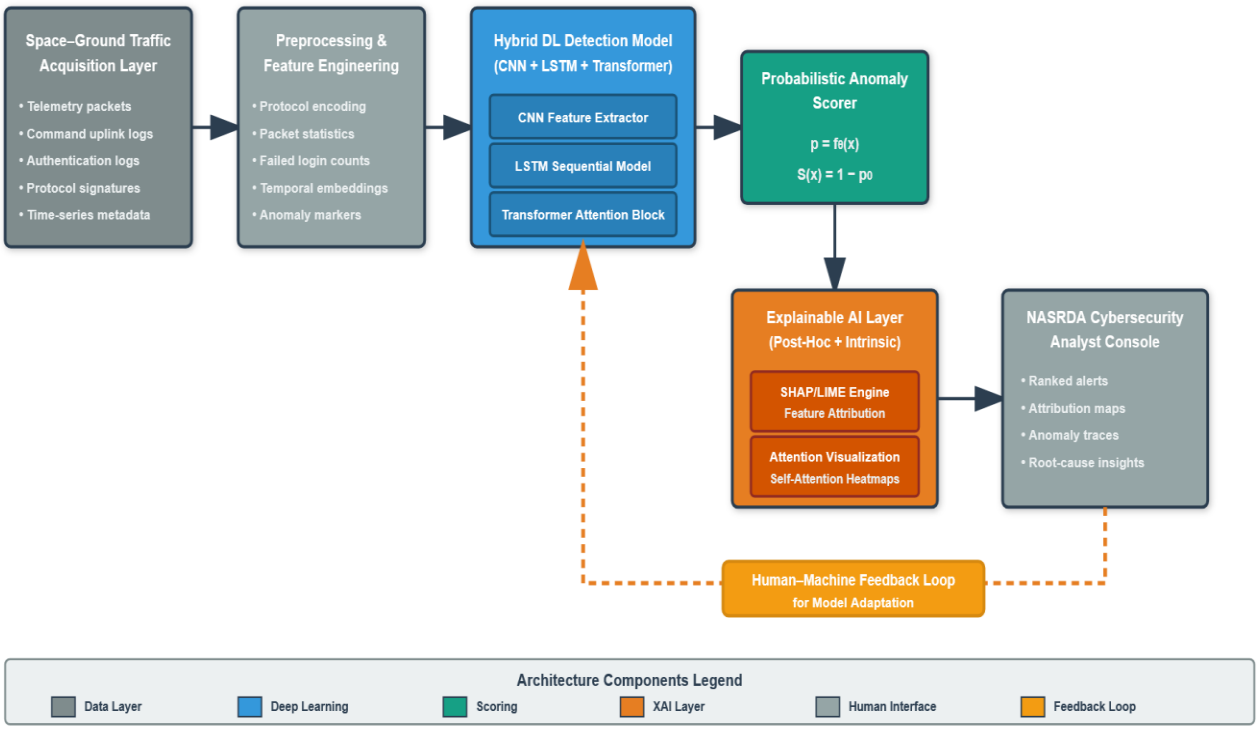


Figure 2 illustrates a hybrid IDS architecture that addresses NASRDA's operational constraints by integrating deep learning, feature engineering and explanation modules into a unified detection pipeline, yet the design also exposes technical tensions highlighted repeatedly in

Article DOI: 10.52589/BJCNIT-RX9O6XYJ
DOI URL: https://doi.org/10.52589/BJCNIT-RX9O6XYJ

contemporary aerospace cybersecurity research. The architecture begins with a space–ground traffic acquisition layer, which is essential because satellite telemetry and command-uplink logs exhibit non-stationary temporal behavior and sparse anomalies, conditions that traditional IDS frameworks fail to model effectively according to Verma (2025) and Botezatu (2024). The preprocessing and feature engineering stage shown in Figure 2 seeks to stabilize this variability through protocol encoding, packet statistics and anomaly markers, a design approach aligned with the trajectory outlined by Blika et al. (2024), who argue that domain-specific representations outperform generic network-security embeddings by as much as 25 percent in classification precision. The hybrid CNN-LSTM Transformer block theoretically supports spatial, sequential and long-range dependency modeling, and its layered architecture accords with the evidence offered by Hashima et al. (2025), who report accuracy gains exceeding 90 percent in UAV satellite communication anomaly detection when attention mechanisms are fused with recurrent structures. However, Figure 2 also visualizes the computational burden documented by Johari et al. (2025), since multi-stage deep models significantly increase inference latency, a critical issue in NASRDA environments where response times must remain within milliseconds to prevent command spoofing or telemetry drift. The XAI layer provides SHAP LIME attributions and attention heatmaps, echoing the demands for traceable anomaly reasoning outlined in Tahir et al. (2024) and Abbas et al. (2025), yet studies such as Naydenov and Chemungor (2025) caution that explanation modules frequently inflate computational cost by 30 to 60%, threatening real-time deployability. The human–machine feedback loop in Figure 2 reflects an emerging requirement in satellite cybersecurity: model adaptation guided by analyst supervision, a concept supported by Bikos and Kumar (2025), who show that reinforcement-guided anomaly frameworks can reduce false positives by up to 40 percent. Thus, while Figure 2 synthesizes the best available practices in hybrid detection design, it also highlights the architectural and computational challenges that must be addressed before the model can operate reliably within NASRDA's mission-critical ground-station environment.

Compared with the models surveyed in prior studies, the architecture in Figure 2 surpasses conventional IDS designs by incorporating a multi-model fusion pipeline in which CNN, LSTM and Transformer components jointly address the spatial, temporal and long-range dependencies characteristic of NASRDA ground station traffic, whereas most existing systems rely on single-architecture detectors that experience accuracy drops of 20 percent to 35 percent under domain shift. Its explicit integration of SHAP, LIME and intrinsic attention maps offers greater transparency than the predominantly opaque models critiqued by Tahir et al. (2024) and Johari et al. (2025), while the domain-specific feature engineering resolves the adaptation failures reported by Botezatu (2024) in satellite telemetry contexts. For NASRDA engineers, this design enables interpretable triage and more defensible operational decisions, and for the broader aerospace cybersecurity community, it establishes a blueprint for IDS systems that combine high-fidelity detection with analyst-centric explainability, laying the groundwork for future research into latency reduction, adaptive learning and federated detection architectures suitable for distributed space communication networks.

Article DOI: 10.52589/BJCNIT-RX9O6XYJ
DOI URL: https://doi.org/10.52589/BJCNIT-RX9O6XYJ

CONCLUSION

In conclusion, the review demonstrates that while deep learning continues to advance intrusion detection through increasingly expressive architectures, these gains remain insufficient without the complementary transparency that XAI provides, particularly in mission-critical satellite environments where operational decisions must be auditable and defensible. The mathematical formulation developed here offers a necessary formal scaffold for hybrid IDS design, addressing the conceptual ambiguity that weakens much of the current DL XAI literature, and the synthesis of existing studies clarifies both the strengths and persistent shortcomings of contemporary approaches. The proposed conceptual model for NASRDA represents a substantive contribution by grounding detection, feature engineering and interpretability within the constraints of space–ground communication workflows, yet its real value will depend on empirical validation using authentic telemetry streams, the development of satellite-specific datasets and the deployment of real-time XAI dashboards to support analyst reasoning. Collectively, these recommendations underscore that the field must move beyond benchmark-driven experimentation toward domain-grounded, operationally viable IDS systems capable of securing the next generation of aerospace networks.

REFERENCES

Abbas, A., Falco, G. J., Fischer, D., & Slay, J. (2025). Guardians of the Galaxy: Protecting Space Systems from Cyber Threats (Dagstuhl Seminar 25101). *Dagstuhl Reports*, *15*(3), 1-38.

Ahmad, T., Katari, P., Venkata, A. K. P., Ravi, C. S., & Shaik, M. (2024). Explainable AI: Interpreting Deep Learning Models for Decision Support. *Advances in Deep Learning Techniques*, *4*(1), 80-108.

Alqadhi, S., Mallick, J., Alkahtani, M., Ahmad, I., Alqahtani, D., & Hang, H. T. (2024). Developing a hybrid deep learning model with explainable artificial intelligence (XAI) for enhanced landslide susceptibility modeling and management. *Natural Hazards*, *120*(4), 3719-3747.

Anis, F. M., AlAbdullatif, M., Aljbli, S., & Hammoudeh, M. (2025). A Survey on the Applications of Deep Learning in Network Intrusion Detection Systems to Enhance Network Security. *IEEE Access*.

Anjum, N. (2025). SoK: Securing the Final Frontier for Cybersecurity in Space [Preprint]. arXiv.

Arreche, O. (2024). XAI IDS: Toward Proposing an Explainable Artificial Intelligence Framework for Network Intrusion Detection Systems. Applied Sciences, 14(10).

Bikos, A. N., & Kumar, S. A. (2025). SAT-IOTA: A Cybersecurity Reinforcement Framework for Blockchain-Driven Space Satellites Utilizing Anomaly Prediction. *IEEE Journal on Miniaturization for Air and Space Systems*.

Blika, A., Palmos, S., Doukas, G., Lamprou, V., Pelekis, S., Kontoulis, M., ... & Askounis, D. (2024). Federated learning for enhanced cybersecurity and trustworthiness in 5G and 6G networks: A comprehensive survey. *IEEE Open Journal of the Communications Society*.

BOTEZATU, U. E. (2024). Space cybersecurity: a survey of vulnerabilities and threats. *Romanian Cyber Security Journal*, *6*(2), 53-60.

Article DOI: 10.52589/BJCNIT-RX9O6XYJ
DOI URL: https://doi.org/10.52589/BJCNIT-RX9O6XYJ

De, T., Giri, P., Mevawala, A., Nemani, R., & Deo, A. (2020). Explainable AI: a hybrid approach to generate human-interpretable explanations for deep learning predictions. *Procedia Computer Science*, *168*, 40-48.

Gopalan, R., Onniyil, D., Viswanathan, G., & Samdani, G. (2025). Hybrid models combining explainable AI and traditional machine learning: A review of methods and applications.

Hashima, S., Gendia, A., Hatano, K., Muta, O., Nada, M. S., & Mohamed, E. M. (2025). Next-gen UAV-satellite communications: AI innovations and future prospects. *IEEE Open Journal of Vehicular Technology*.

Hu, C., Zhang, R., Li, B., Jiang, X., Zhao, N., Di Renzo, M., ... & Karagiannidis, G. K. (2025). Generative AI-Empowered Secure Communications in Space-Air-Ground Integrated Networks: A Survey and Tutorial. arXiv preprint arXiv:2508.01983.

Jablaoui, R., & Liouane, N. (2025). Network security-based combined CNN-RNN models for IoT intrusion detection systems. *Peer-to-Peer Networking and Applications*, *18*(3), 129.

Johari, S. S., Shahriar, N., Tornatore, M., Boutaba, R., & Saleh, A. (2025). Anomaly Detection and Localization in NFV Systems by Utilizing Masked Autoencoders and XAI. *IEEE Transactions on Mobile Computing*.

Kalakoti, R., Nõmm, S., & Bahsi, H. (2025). Evaluating explainable AI for deep learning-based network intrusion detection systems. ScitePress.

Kamal, H., & Mashaly, M. (2024). Advanced hybrid transformer-CNN deep learning model for effective intrusion detection systems with class imbalance mitigation using resampling techniques. *Future Internet*, *16*(12), 481.

Kato, N., Fadlullah, Z. M., Tang, F., Mao, B., Tani, S., Okamura, A., & Liu, J. (2019). Optimizing space-air-ground integrated networks by artificial intelligence. IEEE Wireless Communications, 26(4), 140-147.

Khan, N., Ahmad, K., Tamimi, A. A., Alani, M. M., Bermak, A., & Khalil, I. (2024). Explainable AI-Based Intrusion Detection System for Industry 5.0: An Overview of the Literature, Associated Challenges, the Existing Solutions, and Potential Research Directions. *arXiv preprint arXiv:2408.03335*.

Kilichev, D., Turimov, D., & Kim, W. (2024). Next-generation intrusion detection for IoT EVCS: Integrating CNN, LSTM, and GRU models. *Mathematics*, *12*(4), 571.

Kök, I., Okay, F. Y., Muyanlı, Ö., & Özdemir, S. (2023). Explainable artificial intelligence (XAI) for the Internet of Things: a survey. *IEEE Internet of Things Journal*, *10*(16), 14764-14779.

Mohale, V. Z., & Obagbuwa, I. C. (2025). Evaluating machine learning-based intrusion detection systems with explainable AI: enhancing transparency and interpretability. *Frontiers in Computer Science*, *7*, 1520741.

Naydenov, N., & Chemungor, N. (2025). A Machine Learning Framework for Cyber Intrusion Detection and Prevention.

Neupane, S., Ables, J., Anderson, W., Mittal, S., Rahimi, S., Banicescu, I., & Seale, M. (2022). Explainable intrusion detection systems (x-ids): A survey of current methods, challenges, and opportunities. *IEEE Access*, *10*, 112392-112415.

Ruff, L., Kauffmann, J. R., Vandermeulen, R. A., Montavon, G., Samek, W., Kloft, M., ... & Müller, K. R. (2021). A unifying review of deep and shallow anomaly detection. *Proceedings of the IEEE*, *109*(5), 756-795.

Salim, S., Moustafa, N., & Almorjan, A. (2025). Responsible deep federated learning-based threat detection for satellite communications. IEEE Internet of Things Journal.

Article DOI: 10.52589/BJCNIT-RX9O6XYJ
DOI URL: https://doi.org/10.52589/BJCNIT-RX9O6XYJ

Salim, S., Moustafa, N., & Reisslein, M. (2024). Cybersecurity of satellite communications systems: A comprehensive survey of the space, ground, and links segments. IEEE Communications Surveys & Tutorials, 27(1), 372-425.

Shiri, F. M., Perumal, T., Mustapha, N., & Mohamed, R. (2023). A comprehensive overview and comparative analysis of deep learning models: CNN, RNN, LSTM, and GRU. *arXiv preprint arXiv:2305.17473*.

Simon, C., & Barr, J. (2023). *Deep Learning and XAI Techniques for Anomaly Detection*. Packt Publishing Ltd.

Sun, Q., Zeng, J., Dai, L., Hu, Y., & Tian, L. (2025). XAI-Based Framework for Protocol Anomaly Classification and Identification to 6G NTNs with Drones. Drones, 9(5), 324.

Tahir, H. A., Nabi, F., Tariq, M. Z., Khan, A. F., & Mahmud, A. (2024, July). Insights Into the Future: XAI Integration in O-RAN and Space Communication Systems. In *2024 Multimedia University Engineering Conference (MECON)* (pp. 1-6). IEEE.

Verma, A. R. K. (2025). Cybersecurity in Satellite Communication Networks: Key Threats and Neutralization Measures. *IEEE Open Journal of the Communications Society*.

Wang, T., Fang, K., Tian, J., Feng, H., Al Dabel, M. M., Bashir, A. K., & Wang, W. (2025). AI-Backed Network Security for Connecting Air, Space, and Ground. IEEE Wireless Communications, 32(3), 80-87.

Xu, H., Han, S., Li, X., & Han, Z. (2023). Anomaly traffic detection based on communication-efficient federated learning in space-air-ground integration networks. IEEE Transactions on Wireless Communications, 22(12), 9346-9360.

Article DOI: 10.52589/BJCNIT-RX9O6XYJ
DOI URL: https://doi.org/10.52589/BJCNIT-RX9O6XYJ