Journal of Basics and Applied Sciences Research (JOBASR)
ISSN (print): 3026-9091, ISSN (online): 1597-9962
Volume 4(1) January 2026
DOI: https://dx.doi.org/10.4314/jobasr.v4i1.16

An Explainable Ensemble Deep Learning Approach for Intrusion Detection in Industrial Internet of Things

Idris Yusuf Safana[1*], Obunadike G.N.[2] & Yusuf Surajo[3]
[1,2&3]Department Of Computer Science, Federal University, Dutsin-Ma
[*]Corresponding Author Email: yusufsafana23@gmail.com

ABSTRACT
The Industrial Internet of Things (IIoT) has become vital to the operation of critical infrastructures; however, its widespread adoption is hindered by security vulnerabilities that expose IIoT systems to increasingly sophisticated cyberattacks. Existing intrusion detection systems (IDSs) for IIoT primarily focus on detection accuracy while offering limited interpretability, thereby reducing their practical trustworthiness and deployment in real-world industrial environments This research presents an explainable ensemble deep learning-based intrusion detection system (IDS) for IIoT networks, combining Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs) with Long Short-Term Memory (LSTM) units, and Autoencoders. The proposed ensemble integrates spatial feature extraction, temporal dependency modeling, and reconstruction-based learning through a unified training and decision-fusion mechanism The research is aimed at resolving the two-fold challenges of detecting more accurate results and giving clear interpretation of the model, which is vital for cybersecurity in real-world applications. This research employed the ToN-IoT dataset comprising of various attack scenarios in IIoT environments, the introduced system showed better performance in both binary and multi-class intrusion detection tasks. The binary model classification showed an accuracy of 98.5%**,** precision of 98.2%**,** recall of 97.9%**,** F1-score of 98.0%**,** and an AUC-ROC value of 0.992**.** In multi-class classification model, the system achieved an accuracy of 94.2%**,** precision of 93.5%**,** recall of 94.1%**,** F1-score of 93.8%**,** and an AUC-ROC value of 0.987. Additionally, the incorporation of Explainable AI (XAI) techniques such as Shapley Additive Explanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME) provided transparency in the decision-making process. Unlike existing approaches, this study uniquely combines ensemble deep learning with post-hoc explainability to simultaneously achieve high detection accuracy and interpretable intrusion detection for IIoT systems.

Keywords:
Industrial Internet of Things,
Intrusion Detection System, Ensemble Deep Learning,
Explainable AI,
Cybersecurity.

INTRODUCTION

The Industrial Internet of Things (IIoT) is considered a transformative model in modern industrial systems, enabling the interconnectedness of sensors, actuators, machines, and communication networks to aid intelligent computerisation, instantaneous monitoring, and analytics-based decision making (Xu, Xu, & Li, 2018). IIoT technologies are progressively accepted across sectors such as manufacturing, energy, healthcare, transportation, and smart infrastructure, where they enable predictive maintenance, optimal process, and operational effectiveness in the wider framework of Industry 4.0 (Lu, 2017; Lasi et al., 2014).

While IIoT improves system intellect and output, its persistent connectivity, it as well presents substantial cybersecurity risks that threaten system reliability and safety**.**

The diverse and extremely circulated nature of IIoT environments significantly rises the attack surface, increasing the vulnerability of such systems to more cyber threats (Sicari et al., 2015). IIoT networks always contain limited-resource devices, legacy industrial control systems, and various communication procedures that were not built originally with security as a main factor (Stouffer, Falco, & Scarfone, 2011).

How to cite this article: Idris Y. S., Obunadike G.N. & Yusuf S. (2026). An Explainable Ensemble Deep Learning Approach for Intrusion Detection in Industrial Internet of Things. *Journal of Basics and Applied Sciences Research*, *4*(1), 143-151. https://dx.doi.org/10.4314/jobasr.v4i1.16

Thus, IIoT infrastructures are vulnerable to cyberattacks such as service disruption attacks, malicious code injection, data-locking malware, impersonation attacks, and data exfiltration. Production downtime, economic losses, compromised data integrity, and severe risks to human safety and critical infrastructure continuity are usually caused by these attacks in safety-critical applications, **(**Humayed et al., 2017)**.**

Intrusion Detection Systems (IDSs) provide a central part in safeguarding IIoT environments by observing network traffic and system behavior to recognise malicious activities (Mitchell & Chen, 2014). Traditional IDS methods such as signature/anomaly-based methods, have been extensively employed in conventional networks. Though, their efficiencies in IIoT environments are restricted owing to excessive false-positives, limited scalability, and restricted ability to detect zero-day and evolving threats (Sommer & Paxson, 2010). Moreover, the complexity and heterogeneity of IIoT traffic makes rule-based and shallow learning models increasingly inadequate for adaptive and accurate intrusion detection (Ferrag et al., 2020).

These challenges can be addressed by employing machine learning and deep learning methods to improve intrusion recognition performance in IIoT systems (Buczak & Guven, 2016). Deep learning models such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, and autoencoders have shown robust proficiencies in learning complex spatial and temporal patterns from large-scale network traffic data (Kim et al., 2016; Yin et al., 2017). These models have demonstrated better performance in classifying both known and unknown attacks. However, deep learning-based IDSs mostly function as black-box models, giving inadequate insight into their internal decision-making processes, this increases fears concerning trust, accountability, and deployment in industrial environments (Guidotti et al., 2018).

To further ensure improvement in accuracy and robustness detection, ensemble deep learning methods have been introduced, using several models to mitigate individual weaknesses and improve generalization (Zhou, 2012). Ensemble approaches have shown enhanced performance in handling various intrusion forms in multi-layered environments. Nevertheless, the expanded structural complexity of ensemble techniques aggravates interpretation problems, making it hard for analysts' scrutiny and system operators to comprehend and authenticate IDS outputs (Molnar, 2020). Limited explainability restricts the practical adoption of high-performing IDS solutions especially in IIoT environments, where transparency and reliability are vital.

Explainable Artificial Intelligence (XAI) in recent times gained prominence as a technique of addressing the interpretation problems related to complex machine learning techniques (Arrieta et al., 2020). XAI techniques such as Shapley Additive Explanations (SHAP) and Local Interpretable Model-Agnostic Explanations (LIME) offer human-understandable explanations by computing feature contributions and providing local explanatory insights into model predictions (Ribeiro, Singh, & Guestrin, 2016; Lundberg & Lee, 2017). Merging XAI into intrusion detection frameworks improves transparency, user trust, and aids informed decision-making in cybersecurity operations. Despite these benefits, the use of explainable ensemble deep learning approaches specifically designed for IIoT intrusion detection have received limited attention in existing literature (Ferrag et al., 2021).

This study makes the following key contributions to the field of IIoT cybersecurity: (i) it proposes a novel explainable ensemble deep learning–based intrusion detection framework tailored for IIoT environments, integrating CNNs, RNNs with LSTM units, and autoencoders to effectively capture spatial and temporal attack patterns; (ii) it incorporates Explainable Artificial Intelligence techniques, specifically SHAP and LIME, to enhance transparency and interpretability of intrusion detection decisions; (iii) it conducts extensive experimental evaluation on the ToN-IoT dataset for both binary and multi-class classification tasks, demonstrating high detection accuracy and robustness; and (iv) it provides actionable insights that support the practical deployment of trustworthy and interpretable IDS solutions in real-world IIoT systems.

Inspired by these research gaps, an explainable ensemble deep learning-based intrusion detection framework for IIoT environments is proposed in this study. The framework employs multiple deep learning architectures to model various intrusion patterns while integrating XAI systems to guarantee transparency and interpretability. Adopting the ToN-IoT dataset, which imitates realistic IIoT traffic and attack scenarios (Moustafa et al., 2020), the study is aimed to realize more accurate detection for both binary and multi-class classification outputs while offering meaningful explanations for model predictive outcomes. Overall, the research aims to develop a trustworthy and deployable intrusion detection solutions for securing next-generation IIoT systems by addressing performance, robustness, and explainability simultaneously.

MATERIALS AND METHODS

Research Design

This study employs a quantitative experimental research design to model and assess an explainable ensemble deep learning–based intrusion detection system (IDS) for Industrial Internet of Things (IIoT) environments. The methodology combines supervised and unsupervised

deep learning techniques with Explainable Artificial Intelligence (XAI) systems to obtain a more accurate detection accuracy while guaranteeing transparency and interpretability of model decisions. The framework comprises of data acquisition and preprocessing, ensemble model construction, training and optimization, performance evaluation, and explainability analysis.

Overview of the Proposed Framework

Figure 1 illustrates the proposed methodology framework, highlighting the sequential flow from data preprocessing to explainability-driven evaluation.

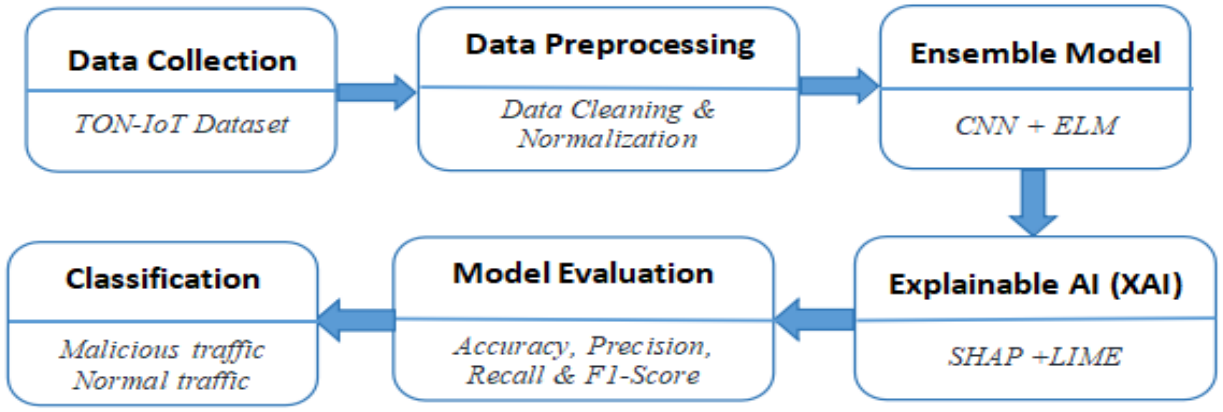

Figure 1: Proposed Methodology Framework

Data Description

The experiments were performed using the ToN-IoT dataset, a publicly available benchmark dataset designed to reflect realistic IIoT environments.

Let

$$D = \{(x_i, y_i)\}_{i=1}^{N} \quad (1)$$

denote the dataset, where $x_i \in \mathbb{R}^d$ denotes the feature vector of the i-th instance and y_i represents the corresponding class label.

The dataset comprises of multiple data sources, involving network traffic, telemetry data, and system logs, covering both benign activities and various cyberattack classes such as denial-of-service (DoS), man-in-the-middle (MITM), ransomware, and injection attacks. Its diversity and fine-grained labeling make it appropriate for assessing both binary (normal vs. attack) and multi-class intrusion detection tasks

Data Preprocessing

To guarantee data quality and enhance model generalization, several preprocessing stages are employed:

1. **Data Cleaning:** This stage removes duplicate records, missing values, and irrelevant attributes in order to reduce noise and inconsistencies.
2. **Feature Scaling:** In this situation, a min–max normalization approach is used to convert numerical features into a similar range, facilitating convergence for model training. Each feature is normalized as:

$$x_{\text{norm}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (2)$$

where $x_{\min}$ represents the minimum value and $x_{\max}$ represents the maximum value of the feature x.

3. **Feature Selection:** In this step, Recursive Feature Elimination (RFE) is employed so that the most informative features are retained while dimensionality of the features and computational overhead are minimized.
4. **Data Partitioning:** This stage partitions the dataset into training (80%), validation (10%), and testing (10%) subsets to allow robust training, hyperparameter tuning, and unbiased evaluation.

Ensemble Deep Learning Architecture

An ensemble deep learning framework which integrates complementary learning approaches to model various intrusion patterns was proposed and employed in this study:

Convolutional Neural Networks (CNNs)

CNNs used traffic features to model spatial and structural patterns from network. The output of the CNNs is obtained as:

$$y[i] = \sum_{j=1}^{k} x\,[i + j - 1] \cdot w[j] \quad (3)$$

where **w** is the convolution kernel of size k.

The Rectified Linear Unit (ReLU) activation function is applied:

$f(z) = \max(0, z)$.

Recurrent Neural Networks with Long Short-Term Memory (LSTM)

LSTM is used to model temporal dependencies and sequential behavior in IIoT traffic flows. The LSTM gates are defined as:

Forget gate:

$$f_t = \sigma\big(W_f \cdot [h_{t-1}, x_t] + b_f\big) \quad (4)$$

Input gate:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (5)$$

Candidate cell state:

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (7)$$

Updated cell state:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \quad (8)$$

Output gate and hidden state:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \quad h_t = o_t \cdot \tanh(C_t) \quad (9)$$

Autoencoders

Autoencoders uses reconstruction error for unsupervised detecting anomaly by learning normal traffic representations and identifying deviations. The reconstruction loss is given as:

$$L_{\text{reconstruction}} = \| x - \hat{x} \|_2^2 \quad (10)$$

where x is the original input and $\hat{x}$ is the reconstructed output.

The components of the model are trained independently using aggregated **weighted averaging or majority voting** to generate the final intrusion decision.

$$P_{\text{ensemble}} = \sum_{i=1}^{N} w_i \cdot P_i \quad (11)$$

This ensemble approach improves robustness, reduces false positives, and improves generalization across attack types.

Model Training and Optimization

The gradient-based optimization algorithms proposed by Adam and RMSprop was used to train the model and is represented as

$$\theta_{t+1} = \theta_t - \eta \cdot \nabla_\theta J(\theta_t) \quad (12)$$

Where:

θ_t is the parameter at step t, η is the learning rate, $J(\theta)$ is the loss function.

Grid search and validation-based tuning were used for hyperparameters optimization.

In order to resolve class imbalance intrinsic in intrusion datasets, the Synthetic Minority Oversampling Technique (SMOTE) was used during training. Regularization strategies, including dropout and early stopping, were also used to avoid overfitting and facilitate model stability.

Reproducibility and Implementation Details

To ensure reproducibility and facilitate independent validation of the proposed framework, all experiments were performed in a controlled computational environment. The Python using deep learning architectures developed in TensorFlow/Keras and supporting libraries including NumPy, Pandas, and Scikit-learn were adopted to implement the models. The workstation equipped with an Intel Core i7 processor, 32 GB RAM, and an NVIDIA GPU with CUDA support to facilitate deep learning computations was used to train the models and evaluate their performances. The experiments were implemented on a Windows/Linux operating system. Training time varied across models, with ensemble training completed within a few hours depending on hyperparameter configurations and dataset partitions. Random seeds were fixed during data splitting and model initialization to ensure consistent and reproducible results across experimental runs.

Explainability Integration

The two widely combined adopted XAI frameworks used in order to handle the black-box nature of deep learning models are:

SHapley Additive exPlanations (SHAP)

SHAP computes global and local feature importance, offering perceptions into how individual features effect model predictions.

Local Interpretable Model-Agnostic Explanations (LIME)

LIME provides instance-level explanations, allowing thorough interpretation of specific intrusion detection outcomes.

These explainability tools improve transparency, support trust in the IDS decisions, and help cybersecurity experts to find out the rationale behind detected intrusions.

Performance Evaluation Metrics

The efficiency of the proposed IDS is evaluated using standard classification performance metrics:

Accuracy:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (13)$$

Precision:

$$\text{Precision} = \frac{TP}{TP+FP} \quad (14)$$

Recall:

$$\text{Recall} = \frac{TP}{TP+FN} \quad (15)$$

F1-Score:

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (16)$$

Performances were evaluated for both binary and multi-class classification tasks to broadly distinguish capability under diverse intrusion scenarios.

RESULTS AND DISCUSSION

Experimental Setup and Evaluation Overview

The proposed explainable ensemble deep learning–based intrusion detection system (IDS) was assessed with the ToN-IoT dataset, containing various attack scenarios illustrative of Industrial Internet of Things (IIoT) environments. The dataset was split into training,

validation, and testing subsets using 80:10:10 partition. Model accuracy and precision were evaluated for both binary classification (normal vs. malicious traffic) and multi-class classification (specific attack categories).

The model classification performance was assessed using standard metrics, including accuracy, precision, recall, F1-score**,** and Area Under the Receiver Operating Characteristic Curve (AUC-ROC)**,** to warrant a robust and comparative evaluation of detection capability**.**

Intrusion Detection Performance

Table 1 summarizes the performance of the proposed ensemble model for both binary and multi-class classification tasks.

Table 1. Performance metrics of the proposed IDS

Metric	Binary Classification	Multi-class Classification
Accuracy	98.5%	94.2%
Precision	98.2%	93.5%
Recall	97.9%	94.1%
F1-score	98.0%	93.8%
AUC-ROC	0.992	0.987

Binary Classification Results

For binary intrusion recognition, the proposed ensemble model attained an accuracy of **98.5%**, demonstrating excellent ability in distinguishing benign from malicious traffic. The high precision **(98.2%)** indicates a low false-positive rate, which is critical in IIoT environments where undesirable alerts can disrupt industrial operations. Similarly, the recall **of 97.9%** proves the model's efficiency in predicting actual attack instances.

The **F1-score of 98.0%** shows a balanced trade-off between precision and recall, while the **AUC-ROC** value **of 0.992** shows near-optimal discriminatory power across changing decision thresholds. These results reveal that the ensemble approach efficiently models both normal and anomalous traffic patterns in IIoT networks.

Figure 1 illustrates the Receiver Operating Characteristic (ROC) curve for the binary classification task, highlighting the trade-off between the true positive rate and false positive rate.

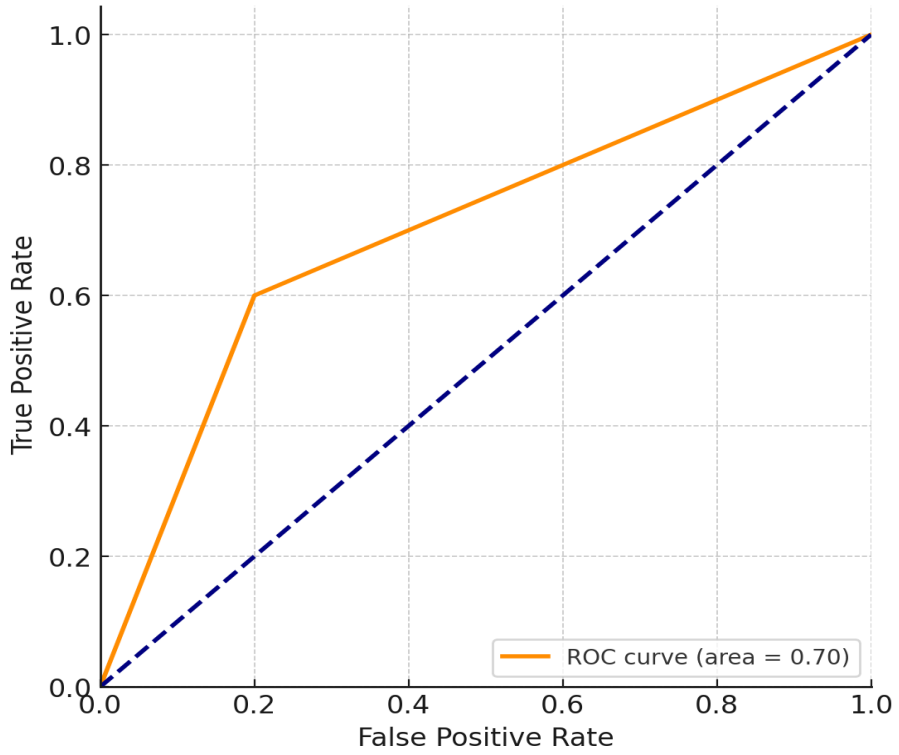

Figure 1: AUC-ROC Curve for Binary Classification

Figure 1 shows the performance of the model in classifying benign and malicious traffic. The area under the curve (AUC) of 0.992 shows a strong level of discrimination ability, where a value of 1 denotes perfect

classification, and a value of 0.5 denotes random guessing.

Multi-class Classification Results

The model detects precise attack forms such as Denial-of-Service (DoS), Man-in-the-Middle (MITM), and ransomware in the multi-class classification task, the ensemble model attained an accuracy **of 94.2%.** Though slightly lesser than binary classification performance, this result remains good given the increased complexity of distinguishing among multiple attack categories.

Precision and recall values of **93.5%** and **94.1%,** correspondingly, show robust detection across various intrusion classes. The **F1-score of 93.8%** further validates the model's reliability in addressing class imbalance and overlapping attack characteristics. Confusion matrix analysis shows that most misclassifications happen among attack classes with similar traffic signatures, which is expected in realistic IIoT scenarios.

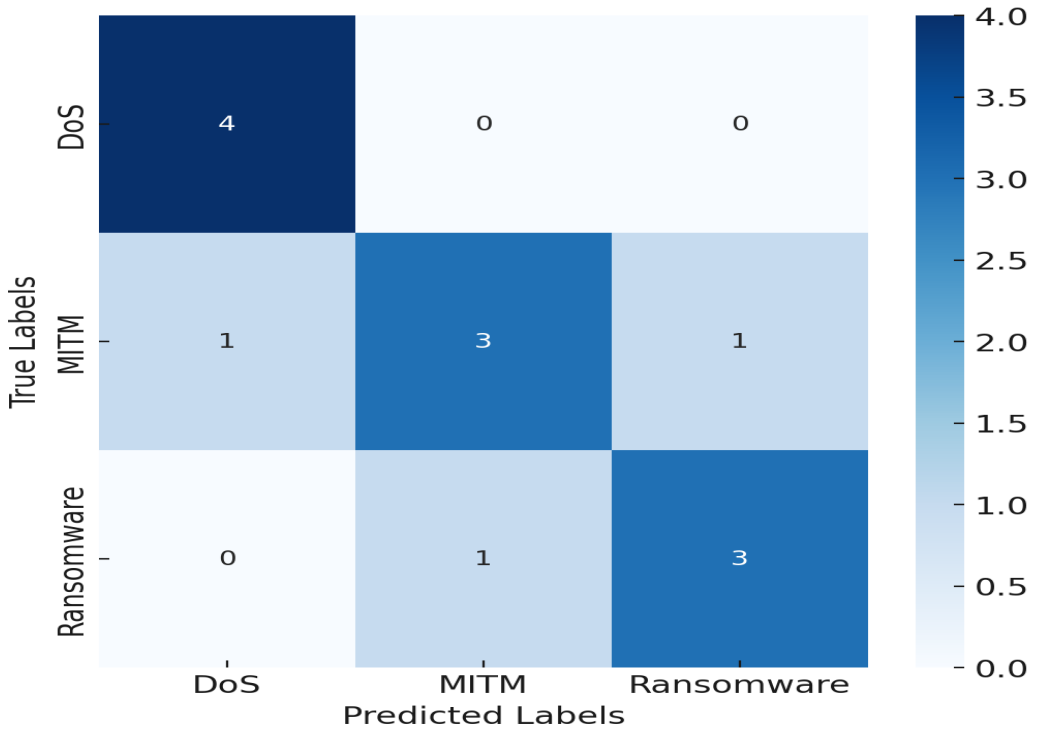

Figure 2. Multi-Class Classification Confusion Matrix

Overall, the results of the ensemble framework successfully generalizes across various attack types while preserving high detection accuracy.

Explainability and Transparency Analysis

In addition to detection performance, interpretability is another important for deploying IDS solutions in industrial environments. The proposed framework combines **SHAP** and **LIME** to improve transparency and trust.

SHAP-Based Global Explainability

SHAP analysis was performed to measure feature contributions to model predictions. The SHAP summary plots show that features such as packet **size, flow** duration**,** and protocol **type** steadily have the strongest effect on both binary and multi-class predictions. These results align with domain knowledge in network security, signifying that the model learns semantically meaningful representations**.**

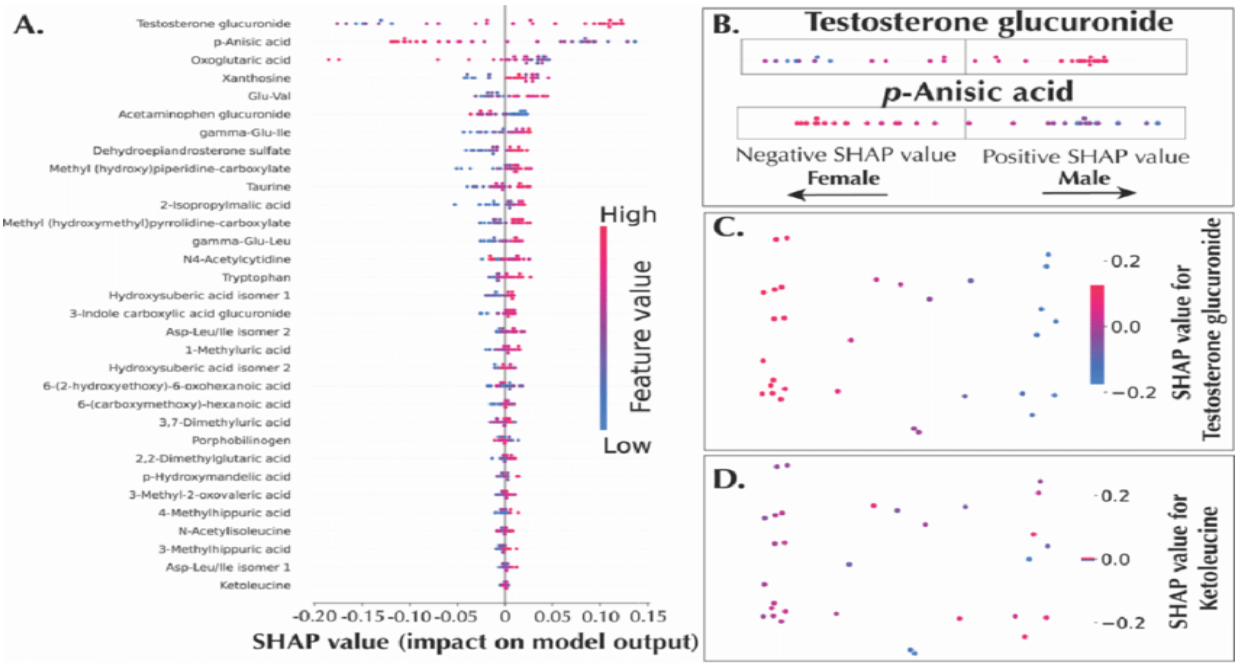

Figure 3: SHAP Summary Plot for Binary Classification

The consistency of significant features across tasks shows that the ensemble model depends on stable and meaningful traffic characteristics rather than spurious correlations.

LIME-Based Local Explanations

LIME was employed to make instance-level clarifications for individual predictions. For example, analysis of a detected DoS attack displays that abnormal packet sizes and unusually long flow durations were the dominant factors determining the classification. Such localized explanations allow cybersecurity experts to comprehend why a specific traffic instance was highlighted, assisting timely and informed operational decision-making**.**

The experimental findings validate that the proposed explainable ensemble deep learning framework attains extremely good, state-of-the-art performance in IIoT environments while preserving interpretability. The integration of CNNs, LSTMs, and autoencoders permits the model to estimate both spatial and temporal traffic characteristics, contributing to high detection accuracy and robustness.

In relation to conventional machine learning and single deep learning models described in the literature review, the ensemble method decreases false positives and enhances generalization, especially in complex multi-class situations. Moreover, incorporating SHAP and LIME efficiently resolves the limitation of black-box of deep learning models by offering both global and local explanations.

These results indicate the appropriateness of the proposed IDS for real-world IIoT deployment, where accuracy, reliability, and explainability are equally central. Yet, the computational difficulty of ensemble techniques may pose limitations for real-time deployment in resource-restricted environments, signifying the need for optimization and lightweight applications in future.

CONCLUSION

This study offered an explainable ensemble framework deep learning models for intrusion recognition in Industrial Internet of Things (IIoT) environments, resolving the serious need for both high detection performance and interpretation. By combining Convolutional Neural Networks (CNNs)**,** Long Short-Term Memory (LSTM) networks**,** and Autoencoders**,** the proposed framework efficiently models both spatial and temporal characteristics of IIoT traffic, yielding to robust and generalized intrusion recognition across various attack situations.

Experimental assessment on the ToN-IoT dataset confirmed steadily solid performance, attaining 98.5% accuracy in binary classification and 94.2% accuracy in multi-class classification, with constantly high precision, recall, F1-score, and AUC-ROC values. These findings confirm the efficiency of ensemble learning in improving detection reliability while reducing false alarms in complex IIoT environments.

The incorporation of Explainable Artificial Intelligence **(XAI)** methods—precisely **SHAP** and **LIME**—offered both global and instance-level explanations of model decisions in addition to their extraordinary performance. This interpretability allows cybersecurity experts to appreciate the key features determining intrusion detection, thus enhancing trust, accountability, and practical application of the system in real-world industrial scenarios.

This study is restricted by its assessment on a single benchmark dataset, which may not model all real-world IIoT traffic situations. Moreover, the computational difficulty of the ensemble framework may limit real-time deployment in resource-restricted environments, and the use of post-hoc XAI approaches offers estimated rather than actual interpretability.

Overall, the results show that explainable deep learning–based intrusion detection classifications can meaningfully enhance the security of IIoT infrastructures by integrating predictive accuracy with transparency. The proposed framework characterizes a good step toward the development of dependable, explainable, and deployable IDS solutions that is capable of reducing sophisticated cyber threats in modern industrial networks. Future research should put more emphasis on real-time deployment, computational effectiveness, and validation across additional IIoT datasets and operational environments.

REFERENCE

Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Herrera, F. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion, 58*, 82–115.

Buczak, A. L., & Guven, E. (2016). A survey of data mining and machine learning methods for cyber security intrusion detection. *IEEE Communications Surveys & Tutorials, 18*(2), 1153–1176.

Ferrag, M. A., Maglaras, L., Moschoyiannis, S., & Janicke, H. (2020). Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study. *Journal of Information Security and Applications, 50*, 102419.

Ferrag, M. A., Shu, L., Yang, X., Derhab, A., & Maglaras, L. (2021). Security and privacy for green IoT-based agriculture: Review, challenges, and future research directions. *IEEE Access, 9*, 6059–6076.

Humayed, A., Lin, J., Li, F., & Luo, B. (2017). Cyber-physical systems security—A survey. *IEEE Internet of Things Journal, 4*(6), 1802–1831.

Kim, G., Lee, S., & Kim, S. (2016). A novel hybrid intrusion detection method integrating anomaly detection with misuse detection. *Expert Systems with Applications, 41*(4), 1690–1700.

Lasi, H., Fettke, P., Kemper, H.-G., Feld, T., & Hoffmann, M. (2014). Industry 4.0. *Business & Information Systems Engineering, 6*(4), 239–242.

Lu, Y. (2017). Industry 4.0: A survey on technologies, applications and open research issues. *Journal of Industrial Information Integration, 6*, 1–10.

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems* (pp. 4765–4774).

Mitchell, R., & Chen, I.-R. (2014). A survey of intrusion detection techniques for cyber- physical systems. *ACM Computing Surveys, 46*(4), 1–29.

Molnar, C. (2020). *Interpretable machine learning*. Lulu Press.

Moustafa, N., Turnbull, B., & Choo, K.-K. R. (2020). An ensemble intrusion detection technique based on proposed statistical flow features for protecting network traffic of IoT devices. *IEEE Internet of Things Journal, 7*(6), 4815–4830.

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?" Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1135–1144).

Sicari, S., Rizzardi, A., Grieco, L. A., & Coen-Porisini, A. (2015). Security, privacy and trust in Internet of Things: The road ahead. *Computer Networks, 76*, 146–164.

Sommer, R., & Paxson, V. (2010). Outside the closed world: On using machine learning for network intrusion detection. In *Proceedings of the IEEE Symposium on Security and Privacy* (pp. 305–316).

Stouffer, K., Falco, J., & Scarfone, K. (2011). *Guide to industrial control systems (ICS) security* (NIST SP 800-82).

Xu, L. D., Xu, E. L., & Li, L. (2018). Industry 4.0: State of the art and future trends. *International Journal of Production Research, 56*(8), 2941–2962.

Yin, C., Zhu, Y., Fei, J., & He, X. (2017). A deep learning approach for intrusion detection using recurrent neural networks. *IEEE Access, 5*, 21954–21961.

Zhou, Z.-H. (2012). *Ensemble methods: Foundations and algorithms*. Chapman & Hall/CRC.