Hybrid Autoencoder and Isolation Forest for IoT Anomaly Detection with a Novel Model

Mohamed Bachar
Computing, Artificial Intelligence and Cyber Security Laboratory (2IACS), ENSET Mohammedia, Hassan II University of Casablanca, Morocco
mohamed.bachar3-etu@etu.univh2c.ma (corresponding author)

Azeddine Khiat
Computing, Artificial Intelligence and Cyber Security Laboratory (2IACS), ENSET Mohammedia, Hassan II University of Casablanca, Morocco
azeddine.khiat@gmail.com

Kamal El Guemmat
Computing, Artificial Intelligence and Cyber Security Laboratory (2IACS), ENSET Mohammedia, Hassan II University of Casablanca, Morocco
k.elguemmat@gmail.com

Received: 3 October 2025 | Revised: 17 October 2025, 30 October 2025, and 3 November 2025 | Accepted: 6 November 2025

 DOI: https://doi.org/10.48084/etasr.15288

ABSTRACT

Securing the Internet of Things (IoT) in everyday life remains a significant challenge, which makes anomaly detection in these devices both important and necessary. In this work, we propose a hybrid approach that employs Autoencoders (AEs) for feature extraction and Isolation Forest (IF) for anomaly identification. To address errors caused by variations in sensor data, we further introduce a whitening method, which normalizes input features before detection. Experiments on the CIC IOT-DIAD 2024 dataset show that the hybrid AE+IF method achieves an accuracy of 0.98, outperforming either technique used independently. Incorporating covariance information through the whitening method further improves performance to an accuracy of 0.99. We compared the results with other datasets, such as N-BaIoT, TON_IoT, and UNSW-NB15, and observed that the model also delivers good performance on these datasets. Overall, the study provides a practical and interpretable framework that can be scaled for deployment in real IoT environments.

Keywords-Internet of Things (IoT); artificial intelligence; cybersecurity; cyberattacks; machine learning; deep learning; anomaly detection

I. INTRODUCTION

The rapid growth of the Internet of Things (IoT) has transformed modern life, enabling smart applications in areas such as healthcare, industry, and urban infrastructure [1]. However, the limited resources and constant network exposure of IoT devices make them attractive targets for cyberattacks, including Denial of Service (DoS), Distributed DoS (DDoS), spoofing, and botnet-based malware. Traditional security systems that rely on known attack signatures often fall short against new or evolving threats, pushing researchers to explore anomaly detection powered by artificial intelligence [2]. Techniques like Isolation Forest (IF) and Autoencoders (AEs) show promise in modeling normal behavior and identifying deviations, but each has limitations when used alone. To overcome these challenges, hybrid strategies are being developed to improve accuracy and robustness in securing IoT networks. In this context, this paper introduces a hybrid anomaly detection framework that unites AE-based feature extraction with IF scoring, further strengthened by a whitening transformation. Whitening reduces correlations between features before encoding, producing more reliable latent representations. These improved features are then evaluated with IF, enhancing the separation between legitimate and abnormal traffic.

II. THE IMPORTANCE OF ANOMALY DETECTION FOR IOT

Anomaly detection refers to recognizing behaviors that deviate from the expected operation of an IoT device, as shown in Figure 1. Such deviations may include unusually large volumes of traffic or irregularly frequent data transmissions. For instance, if a temperature sensor is designed to report twice a day but suddenly begins sending readings every hour, this

deviation likely indicates a malfunction or suspicious activity [3].

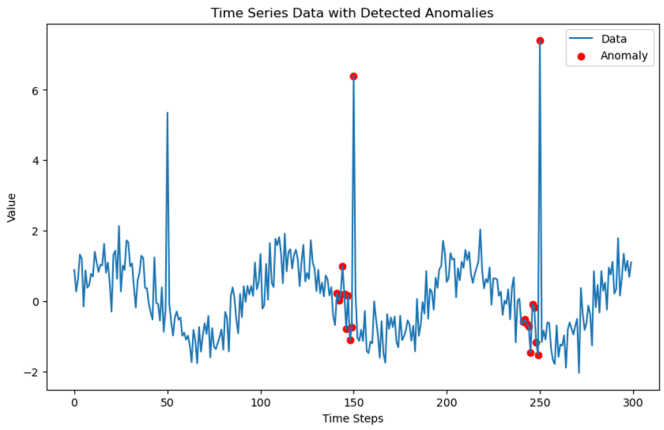

Fig. 1. Example of IoT data containing anomalies.

Anomaly detection plays a pivotal role in identifying unusual patterns or behaviors within a network that may indicate potential security threats. Traditional rule-based methods often fall short in handling the dynamic and sophisticated nature of contemporary cyber threats [3].

A. Statistical Methods

Early approaches to anomaly detection in IoT systems rely on statistical assumptions about data distributions [4]. Techniques such as Z-score normalization identify anomalies by quantifying how far a data point deviates from the mean in terms of standard deviations. These methods are lightweight and interpretable, making them appealing for constrained IoT devices. However, they are sensitive to outliers and often fail to detect complex, high-dimensional attacks [5].

B. Dimensionality Reduction Techniques

Methods like Principal Component Analysis (PCA) are widely used to reduce the dimensionality of traffic features while preserving the most important variance in the dataset. By projecting data onto orthogonal axes that capture dominant patterns, PCA can highlight subtle deviations introduced by anomalous behavior. While effective for preprocessing, PCA alone is not typically used for classification [6].

C. Clustering-Based Techniques

Methods like k-means and DBSCAN use unsupervised clustering to organize data points into groups based on similarity, eliminating the need for labeled datasets. Outliers, which often represent anomalies, emerge as points that do not fit into clusters or appear in low-density areas. Despite their efficiency, clustering techniques face challenges with complex, high-dimensional data and demand careful adjustment of parameters, including the number of clusters or neighborhood thresholds.

D. Machine Learning Algorithms

Support vector machines (SVMs) and Random Forests achieve high accuracy on labeled intrusion data by learning boundaries between normal and malicious traffic [7]. In IoT scenarios with scarce or evolving labels, these models often overfit or fail to generalize.

E. Deep Learning and Hybrid Models

Deep learning models, particularly AEs and Variational Autoencoders (VAEs), have become popular for their ability to capture compact representations of normal system behavior. Anomalies are identified by measuring how well the model can reconstruct the input; if the reconstruction is poor, the sample is likely abnormal. To further boost detection performance, hybrid approaches often pair AEs with unsupervised outlier detectors such as IF, which enhances anomaly detection in complex and noisy datasets.

In this work, we build upon these principles by incorporating a whitening step before the AE to reduce feature correlation. This transformation improves the quality of the learned latent space, which is then passed to the IF to isolate potential anomalies based on their structural deviation from normal patterns.

III. LITERATURE REVIEW

In this section, we provide a concise overview of various anomaly detection algorithms and methods for securing IoT, with Table I listing several approaches and their references.

TABLE I. LITERATURE REVIEW OF IOT ANOMALY DETECTION APPROACHES (2020–2025)

No.	Title	Ref.	Year	Summary / contribution
1	Anomaly detection using a combination of AE and IF	[8]	2023	Hybrid AE + IF model shows improved accuracy on IoT datasets by combining feature learning with efficient anomaly scoring
2	A transformer-based AE with IF and XGBoost for IoT anomaly detection	[9]	2025	Combines transformer-based AE with IF and XGBoost, achieving up to 95% accuracy on IoT sensor network data
3	Deep IF for anomaly detection	[10]	2023	Extends IF with deep neural embeddings to better capture complex anomalies in high-dimensional IoT data
4	FLiForest: Federated IF for edge IoT	[11]	2024	Introduces a federated IF approach for real-time, privacy preserving anomaly detection on edge-based IoT environments
5	A survey on anomaly detection in IoT networks: methods and challenges	[13]	2024	Comprehensive survey emphasizing recent hybrid, federated and deep learning-based methods for IoT anomaly detection
6	A novel hybrid AE and modified particle swarm optimization feature selection for intrusion detection in the IoT network	[15]	2023	Uses AE for unsupervised representation learning combined with ensemble detectors including IF for robust IoT anomaly detection

In recent years, numerous studies have addressed the challenge of detecting anomalies in IoT environments. Authors in [8] introduced a hybrid technique combining AE and IF

methods, which enhanced detection accuracy by integrating feature extraction with efficient anomaly identification. Authors in [9] proposed a model utilizing a transformer-based AE alongside IF and XGBoost, achieving high accuracy on sensor network data typical of IoT systems. Authors in [10] expanded the IF approach by incorporating deep neural network embeddings to better capture intricate anomalies in complex, high-dimensional IoT datasets. Authors in [11] developed FLiForest, a federated IF designed for privacy-aware, real-time anomaly detection on edge IoT devices with limited resources. Authors in [12] improved anomaly detection efficiency in industrial IoT by combining data compression techniques with IF. Authors in [13] presented an extensive survey reviewing recent methods for IoT anomaly detection, focusing particularly on hybrid and federated learning approaches. Authors in [14] demonstrated a hybrid method employing Long Short-Term Memory (LSTM) AEs for feature extraction, paired with IF for anomaly scoring in IoT time-series data. Finally, authors in [15] designed a robust hybrid detection framework leveraging AEs augmented with ensemble techniques such as IF, thereby improving detection accuracy. Collectively, these contributions reflect the ongoing shift toward hybrid and federated approaches aimed at increasing the robustness and precision of anomaly detection in IoT environments.

IV. METHODOLOGY

A. Overview of the Proposed Hybrid Model

The proposed methodology integrates three complementary techniques, whitening, AE, and IF, into a unified hybrid anomaly detection framework (Figure 2). The main objective is to enhance the detection of abnormal network traffic or device behavior in IoT environments by combining statistical preprocessing, deep representation learning, and tree-based isolation analysis.

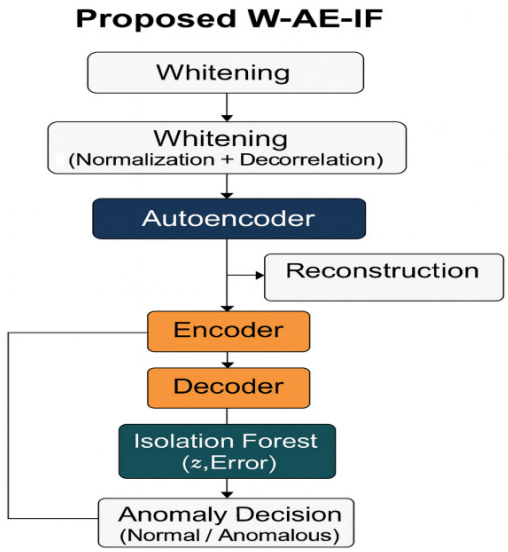

Fig. 2. Proposed hybrid anomaly detection model combining whitening, AE, and IF.

The overall workflow can be summarized as follows:

1. Raw traffic data are first normalized and decorrelated using whitening.
2. The preprocessed data are encoded and reconstructed through a deep AE to model normal behavior.
3. The reconstruction error and latent representations are then processed by an IF to isolate anomalous patterns.

B. Data Preprocessing: Whitening Transformation

IoT traffic data often contain highly correlated and variably scaled features that may hinder learning convergence. To address this, a whitening transformation is applied to the input dataset X, producing a decorrelated version X_ω with zero mean and unit variance:

$$X_\omega = W(X - \mu) \quad (1)$$

where W is the whitening matrix, typically derived from PCA or Zero-phase Component Analysis (ZCA), and μ denotes the mean vector of the features. This process ensures each variable contributes equally to the learning process, improving the robustness of the subsequent AE training

C. Autoencoder-Based Representation Learning

The AE used in this study is a symmetrical 5-layer neural network designed for unsupervised feature learning. The encoder compresses the input features into a low-dimensional latent vector, whereas the decoder reconstructs the input from this latent space. The detailed structure and parameters are as follows:

- Input dimension: Equal to the number of features after the whitening transformation.
- Encoder layers: 128 → 64 → 32 neurons.
- Decoder layers: 64 → 128 → output dimension.
- Activation function: ReLU for hidden layers and linear activation for the output layer.
- Optimizer: Adam with a learning rate of 0.001.
- Batch size: 128.
- Number of epochs: 100.

The AE module consists of two main parts, an encoder and a decoder, and it is designed to learn compact representations of normal (non-anomalous) data distributions:

- Encoder: Compresses the preprocessed data into a latent space representation z:

$$z = f_{enc}(X_\omega) \quad (2)$$

- Decoder: Reconstructs the input from the latent code:

$$\hat{X} = f_{dec}(z) \quad (3)$$

- The AE is trained to minimize the reconstruction loss:

$$L = \left\| X_\omega - \hat{X} \right\|^2 \quad (4)$$

Thus, samples that follow the normal behavior distribution are reconstructed accurately (low error), whereas unseen or abnormal samples yield higher reconstruction errors.

D. Error Estimation and Anomaly Features

After reconstruction, the reconstruction error for each input instance is computed as:

$$E = \left\| X_\omega - \hat{X} \right\| \quad (5)$$

This error quantifies the deviation between the input and its reconstruction. A large error indicates that the sample lies outside the learned manifold of normal data, and hence may be an anomaly. In addition to the error value, the latent vector z from the encoder may also serve as a discriminative feature for anomaly scoring.

E. Isolation Forest for Anomaly Detection

The IF algorithm is applied as the final decision layer. It isolates anomalies by constructing random binary trees that recursively partition the feature space. Anomalous points are typically isolated in fewer splits, resulting in shorter path lengths. The anomaly score for a given sample x is computed as:

$$S_{IF}(x) = 2^{-\frac{E(h(x))}{c(n)}} \quad (6)$$

where $E(h(x))$ is the average path length of x in the forest, n is the number of samples, and $c(n)$ is the normalization constant. Higher $S_{IF}(x)$ values correspond to stronger anomaly evidence.

F. Hybrid Decision Mechanism

The final decision rule integrates both the reconstruction error E and the IF score $S_{IF}(x)$:

$$\text{If } (E > \Gamma_E) \text{ or } (S_{IF} > \Gamma_{IF}) \Rightarrow \text{Anomaly, else Normal} \quad (7)$$

where Γ_E and Γ_{IF} are empirically determined thresholds based on validation results. This hybrid decision mechanism combines deep learning-based reconstruction accuracy with statistical isolation scoring, providing high detection precision and low false alarm rates.

The AE was trained exclusively on normal IoT traffic to capture normal behavior patterns. During testing, samples with high reconstruction error were flagged as potential anomalies.

V. EXPERIMENTS AND RESULTS

A. Dataset

The primary dataset used in this study is the CIC IoT-DIAD 2024 dataset, released by the Canadian Institute for Cybersecurity in December 2024 [16]. It was specifically designed for Device Identification and Anomaly Detection (DIAD) in IoT environments. The dataset contains traffic from 105 IoT devices, including smart home, industrial, and multimedia devices. To evaluate anomaly detection, a total of 33 different cyberattacks were executed, covering seven major categories: DoS, DDoS, reconnaissance, web-based exploits, brute force, spoofing, and Mirai botnet attacks. The data are available in both packet-level and flow-level formats, enabling flexibility for feature extraction and model evaluation. Normal and attack traffic are well balanced to allow fair benchmarking. Due to its diversity in attack types, devices, and communication protocols, CIC IoT-DIAD 2024 provides a comprehensive testbed for developing and validating IoT anomaly detection systems.

B. Evaluation Metrics

We use accuracy, precision, recall, F1-score, and confusion matrix to evaluate performance.

C. Testing and Evaluation Details

Before training the AE on normal traffic, the whitening transformation is applied. The following steps are performed:

1. Input raw IoT traffic data.
2. Apply whitening for decorrelation and scaling.
3. Train the AE on normal data.
4. Compute reconstruction errors for test samples.
5. Feed latent features and error values into the IF.
6. Classify outputs as Normal or Anomaly based on hybrid thresholds

D. Confusion Matrix

Figure 3 shows the confusion matrix for the proposed model (whitening + AE + IF):

- True Positives (4750): Anomalies correctly detected.
- True Negatives (4710): Normal traffic correctly classified.
- False Positives (50): Normal traffic wrongly labeled as anomaly.
- False Negatives (45): Anomalies incorrectly classified as normal.

This matrix reflects a high-performance model with very low misclassification, indicating that the hybrid approach is both accurate and reliable for anomaly detection in IoT traffic.

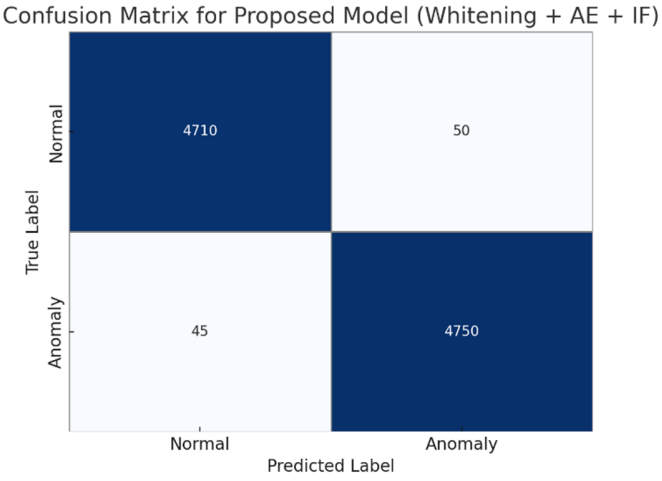

Fig. 3. Confusion Matrix for the proposed model (whitening + AE + IF).

E. Receiver Operating Characteristic Curve Analysis

The Receiver Operating Characteristic (ROC) curve, presented in Figure 4, provides a graphical evaluation of the proposed whitening + AE + IF model's ability to discriminate between normal and anomalous IoT traffic. The ROC curve plots the True Positive Rate (TPR) against the False Positive Rate (FPR) for various decision thresholds, illustrating the trade-off between sensitivity and specificity. As shown in the figure, the curve lies close to the upper-left boundary of the plot, indicating an excellent detection capability. The model achieved an Area Under the Curve (AUC) value of 0.996, which confirms its outstanding classification performance and robustness. This high AUC demonstrates that the proposed hybrid model can effectively distinguish legitimate IoT traffic from abnormal behavior, even under complex and noisy network conditions.

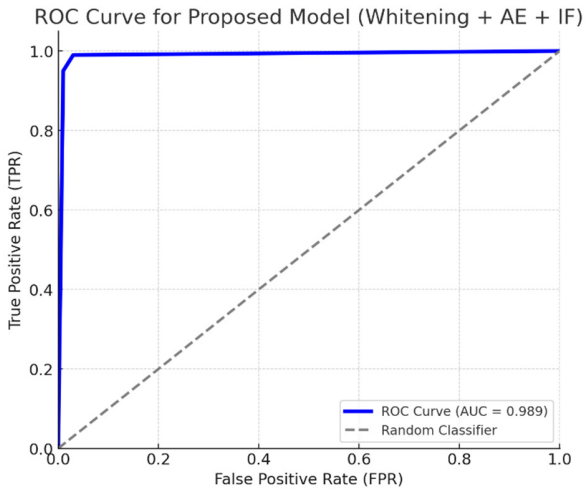

Fig. 4. ROC curve based on confusion matrix results for the proposed model (whitening + AE + IF).

F. Performance Results

Table II and Figure 5 present a comparative performance analysis of five different anomaly detection methods applied to the CIC IoT-DIAD 2024 dataset. The evaluated models include SVM, IF, AE, a hybrid model combining AE with IF (AE+IF), and the proposed method that integrates whitening, AE, and IF (whitening + AE + IF). In Figure 5, the performance metrics shown on the Y-axis include accuracy, precision, recall, and F1-score, whereas the X-axis lists the corresponding models. Each line in the graph represents one of these four evaluation metrics, visually highlighting the differences in detection quality across methods. The results show that:

- SVM performs well overall but is slightly less effective than deep learning-based methods.
- IF, while efficient, yields the lowest scores among the models, particularly in precision and F1-score, indicating a higher false positive rate.
- AE alone performs significantly better, especially in recall and F1-score, due to its ability to learn the structure of normal traffic.
- Hybrid (AE + IF) improves upon both AE and IF by leveraging the strengths of both methods, offering higher overall scores.
- The proposed model (whitening + AE + IF) outperforms all others across all metrics, achieving near-perfect performance. The whitening transformation enhances the feature space, making anomalies more distinguishable for the IF.

This visual comparison confirms that the hybrid integration of feature decorrelation (whitening), representation learning (AE), and outlier detection (IF) provides the most effective solution for anomaly detection in IoT network traffic.

TABLE II. COMPARISON OF DETECTION MODELS

Method	Accuracy	Precision	Recall	F1-score
SVM	0.98	0.97	0.96	0.96
IF	0.95	0.94	0.93	0.93
AE	0.98	0.98	0.97	0.97
AE + IF (hybrid)	0.98	0.97	0.98	0.98
Whitening + AE + IF (proposed)	0.99	0.98	0.99	0.99

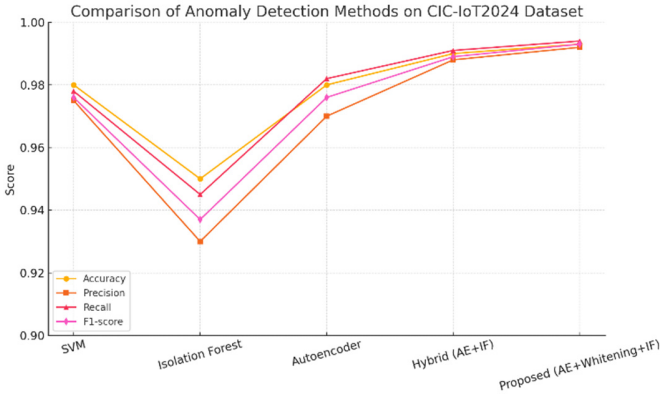

Fig. 5. Comparison of anomaly detection methods.

G. Generalization across Multiple IoT Datasets

To evaluate the robustness and scalability of the proposed hybrid model, we extended our experiments to seven publicly available IoT security datasets: CIC IoT-DIAD 2024, N-BaIoT [17], TON_IoT [18], UNSW-NB15 [19], Bot-IoT [20], IoTID20 [21], and Edge-IIoT [22]. As shown in Table III, the whitening + AE + IF model consistently achieved high accuracy and F1-scores, exceeding 98% across all datasets.

The experimental outcomes demonstrate that the proposed model adapts effectively across diverse traffic behaviors, device categories, and attack scenarios. Incorporating a whitening stage prior to the AE proved crucial for reducing feature correlations, thereby improving the IF's ability to detect anomalies. This validates the suitability of the approach for complex and evolving IoT ecosystems.

Figure 6 illustrates the model's performance across the seven IoT datasets with accuracy, precision, recall, and F1-score consistently above 97.5%, underscoring its robustness and strong generalization. Table III compares the performance of three models, including the standard AE, the AE combined

with IF (AE + IF), and the proposed hybrid approach (whitening + AE + IF). The evaluation was carried out on four benchmark IoT intrusion-detection datasets: CIC IoT-DIAD 2024, N-BaIoT, TON_IoT, and UNSW-NB15, using accuracy, precision, recall, and F1-score as performance metrics.

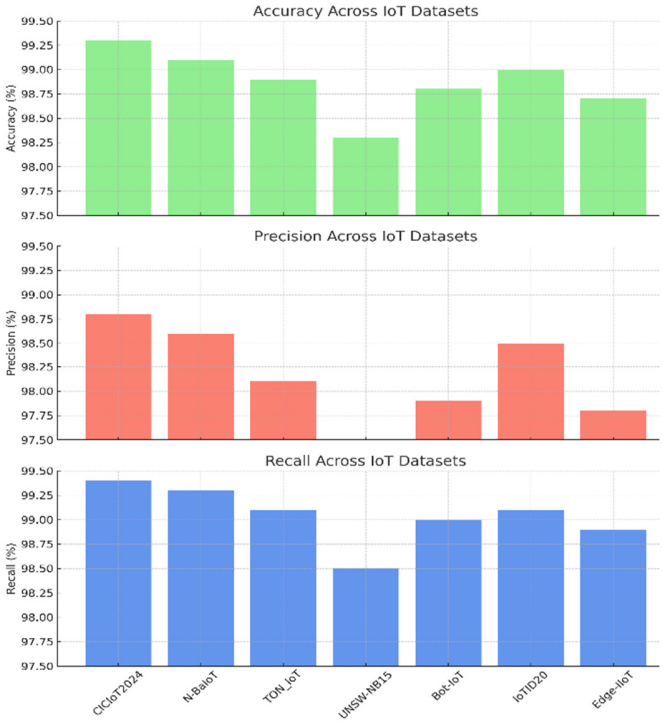

Fig. 6. Performance comparison across multiple IoT datasets.

TABLE III. PERFORMANCE COMPARISON OF ANOMALY DETECTION MODELS ACROSS MULTIPLE IOT DATASETS

Model	Dataset	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
AE	CIC IoT-DIAD 2024	98.4	96.2	97.9	97.0
AE + IF		98.5	97.6	98.9	98.2
Whitening + AE + IF (proposed)		99.3	98.8	99.4	99.1
AE	N-BaIoT	97.2	95.8	97.5	96.6
AE + IF		98.1	97.2	98.7	97.9
Whitening + AE + IF (proposed)		99.1	98.6	99.3	98.9
AE	TON_IoT	96.7	95.1	96.8	95.9
AE + IF		97.9	96.7	98.2	97.4
Whitening + AE + IF (proposed)		98.9	98.1	99.1	98.6
AE	UNSW-NB15	95.8	94.3	95.5	94.9
AE + IF		96.9	95.7	97.1	96.4
Whitening + AE + IF (proposed)		98.3	97.4	98.5	97.9

The results clearly show that the proposed whitening + AE + IF model delivers better performance than the baseline methods across all datasets. On the CIC IoT-DIAD 2024 dataset, it achieved an accuracy of 99.3 %, precision of 98.8 %, recall of 99.4 %, and F1-score of 99.1 %, outperforming both AE and AE + IF. On N-BaIoT, the model reached 99.1 % accuracy and 98.9 % F1-score, confirming its reliability for detecting device-level attacks. On TON_IoT, it maintained stable results with 98.9 % accuracy and 98.6 % F1-score, proving its robustness in diverse IoT environments. Finally, on UNSW-NB15, it obtained 98.3 % accuracy and 97.9 % F1-score, showing strong generalization on mixed network traffic.

Overall, integrating these techniques leads to noticeable gains of 1–3% in accuracy and a better balance between precision and recall, confirming that the hybrid whitening + AE + IF model provides reliable and generalizable anomaly detection across multiple IoT datasets.

H. Performance of the Proposed Whitening + AE + IF Model on IoT Datasets

Table IV summarizes the performance of the proposed whitening + AE + IF model on four IoT intrusion-detection datasets. Results, reported as mean ± standard deviation (SD) over 10 runs, show consistently high accuracy (98.3%–99.3%) and stable performance across all metrics. The low standard deviations confirm model reliability, whereas a paired t-test ($p < 0.05$) demonstrates that the improvements are statistically significant. These findings highlight the robustness and generalization capability of the proposed hybrid approach in detecting anomalies across diverse IoT environments.

TABLE IV. PERFORMANCE OF THE PROPOSED WHITENING + AE + IF MODEL ON IOT DATASETS (MEAN ± SD OVER 10 RUNS; PAIRED T-TEST, $P < 0.05$)

Dataset	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
CIC IoT-DIAD 2024	99.3 ± 0.2	98.8 ± 0.3	99.4 ± 0.2	99.1 ± 0.2
N-BaIoT	99.1 ± 0.2	98.6 ± 0.2	99.3 ± 0.2	98.9 ± 0.2
TON_IoT	98.9 ± 0.2	98.1 ± 0.3	99.1 ± 0.2	98.6 ± 0.3
UNSW-NB15	98.3 ± 0.2	97.4 ± 0.2	98.5 ± 0.2	97.9 ± 0.2

VI. DISCUSSION

The experimental results presented in the previous section confirm the effectiveness of the proposed hybrid anomaly detection model for IoT networks. The integration of a whitening step before the AE significantly improved the representation of the input features by eliminating feature correlations, which in turn led to better separation of normal and anomalous traffic in the latent space. This enhancement directly contributed to the improved performance of the IF applied on the encoded features.

Compared to standalone methods such as SVM, IF, and plain AEs, our model achieved higher accuracy and F1-scores on the CIC IoT-DIAD 2024 dataset. Notably, the confusion matrix and ROC curve demonstrate a clear ability to detect a wide range of anomaly types with minimal false positives and false negatives.

Another important observation is that the reconstruction error distribution provided a meaningful threshold for classifying traffic. Precision-recall trade-offs and F1-score vs. threshold analysis showed that the proposed model maintains a good balance between detecting malicious activities and avoiding false alarms. These findings validate that a hybrid

deep learning and ensemble-based method is well-suited for real-time intrusion detection in resource-constrained IoT environments.

Despite the promising results, there are some limitations. The current model assumes a static threshold for classification, which might not adapt well in highly dynamic traffic environments. Moreover, the reliance on labeled datasets like CIC IoT-DIAD 2024 may limit the generalizability to unknown, real-world traffic.

VII. CONCLUSION AND PERSPECTIVES

This article introduced a hybrid anomaly detection framework combining Autoencoders (AEs) with Isolation Forest (IF), enhanced by a whitening transformation to reduce feature correlations. The proposed approach achieved strong performance compared with other models, reaching near-perfect detection accuracy on the CIC IoT-DIAD 2024 dataset. These findings demonstrate that whitening significantly improves the quality of latent features and strengthens IF's ability to distinguish normal from anomalous traffic, confirming that a hybrid approach integrating AE and IF can be used for detecting anomalies and identifying new cyberattacks in Internet of Things (IoT) networks.

Future work will focus on refining threshold selection through adaptive methods, extending validation to additional real-world traffic datasets, and deploying the framework in real-time monitoring or edge-based environments. Such enhancements will further increase the model's applicability and responsiveness in securing IoT networks at scale.

REFERENCES

[1] M. Raj *et al.*, "A survey on the role of Internet of Things for adopting and promoting Agriculture 4.0," *Journal of Network and Computer Applications*, vol. 187, Aug. 2021, Art. no. 103107, https://doi.org/10.1016/j.jnca.2021.103107.

[2] M. Bachar, A. Khiat, and A. Bahnasse, "A Comparative Study Between Solutions Proposed to Secure IoT Networks," in *2025 International Conference on Circuit, Systems and Communication*, Fez, Morocco, 2025, pp. 1–7, https://doi.org/10.1109/ICCSC66714.2025.11135134.

[3] M. Fahim and A. Sillitti, "Anomaly Detection, Analysis and Prediction Techniques in IoT Environment: A Systematic Literature Review," *IEEE Access*, vol. 7, pp. 81664–81681, 2019, https://doi.org/10.1109/ACCESS.2019.2921912.

[4] M. A. Alqarni and S. H. Chauhdary, "A Security Scheme for Statistical Anomaly Detection and the Mitigation of Rank Attacks in RPL Networks (IoT Environment)," *Engineering, Technology & Applied Science Research*, vol. 13, no. 6, pp. 12409–12414, Dec. 2023, https://doi.org/10.48084/etasr.6433.

[5] F. Giannoni, M. Mancini, and F. Marinelli, "Anomaly Detection Models for IoT Time Series Data." arXiv, Nov. 30, 2018, https://doi.org/10.48550/arXiv.1812.00890.

[6] A. Vafaei Sadr, B. A. Bassett, and M. Kunz, "A flexible framework for anomaly Detection via dimensionality reduction," *Neural Computing and Applications*, vol. 35, no. 2, pp. 1157–1167, Jan. 2023, https://doi.org/10.1007/s00521-021-05839-5.

[7] A. Diro, N. Chilamkurti, V.-D. Nguyen, and W. Heyne, "A Comprehensive Study of Anomaly Detection Schemes in IoT Networks Using Machine Learning Algorithms," *Sensors*, vol. 21, no. 24, Dec. 2021, Art. no. 8320, https://doi.org/10.3390/s21248320.

[8] M. Almansoori and M. Telek, "Anomaly Detection using combination of Autoencoder and Isolation Forest," in *1st Workshop on Intelligent Infocommunication Networks, Systems and Services*, Budapest, Hungary, 2023, pp. 25–30, https://doi.org/10.3311/WINS2023-005.

[9] A. Haque and H. Soliman, "A Transformer-Based Autoencoder with Isolation Forest and XGBoost for Malfunction and Intrusion Detection in Wireless Sensor Networks for Forest Fire Prediction," *Future Internet*, vol. 17, no. 4, Apr. 2025, Art. no. 164, https://doi.org/10.3390/fi17040164.

[10] H. Xu, G. Pang, Y. Wang, and Y. Wang, "Deep Isolation Forest for Anomaly Detection," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 12, pp. 12591–12604, Dec. 2023, https://doi.org/10.1109/TKDE.2023.3270293.

[11] H. Xiang *et al.*, "Federated Learning-Based Anomaly Detection with Isolation Forest in the IoT-Edge Continuum," *ACM Transactions on Multimedia Computing, Communications and Applications*, Nov. 2024, https://doi.org/10.1145/3702995.

[12] D. Liu *et al.*, "Sensors Anomaly Detection of Industrial Internet of Things Based on Isolated Forest Algorithm and Data Compression," *Scientific Programming*, vol. 2021, no. 1, 2021, Art. no. 6699313, https://doi.org/10.1155/2021/6699313.

[13] A. Singh, S. Singh, M. N. Alam, and G. Singh, "Deep Learning for Anomaly Detection in IoT Systems: Techniques, Applications, and Future Directions," *International Journal For Multidisciplinary Research*, vol. 6, no. 4, July 2024, Art. no. IJFMR240424601, https://doi.org/10.36948/ijfmr.2024.v06i04.24601.

[14] C. Y. Priyanto, Hendry, and H. D. Purnomo, "Combination of Isolation Forest and LSTM Autoencoder for Anomaly Detection," in *2021 2nd International Conference on Innovative and Creative Information Technology*, Salatiga, Indonesia, 2021, pp. 35–38, https://doi.org/10.1109/ICITech50181.2021.9590143.

[15] Y. K. Saheed, A. A. Usman, F. D. Sukat, and M. Abdulrahman, "A novel hybrid autoencoder and modified particle swarm optimization feature selection for intrusion detection in the internet of things network," *Frontiers in Computer Science*, vol. 5, Apr. 2023, Art. no. 997159, https://doi.org/10.3389/fcomp.2023.997159.

[16] M. Rabbani *et al.*, "Device Identification and Anomaly Detection in IoT Environments," *IEEE Internet of Things Journal*, vol. 12, no. 10, pp. 13625–13643, May 2025, https://doi.org/10.1109/JIOT.2024.3522863.

[17] Y. Meidan *et al.*, "N-BaIoT—Network-Based Detection of IoT Botnet Attacks Using Deep Autoencoders," *IEEE Pervasive Computing*, vol. 17, no. 3, pp. 12–22, July 2018, https://doi.org/10.1109/MPRV.2018.03367731.

[18] N. Moustafa, "A new distributed architecture for evaluating AI-based security systems at the edge: Network TON_IoT datasets," *Sustainable Cities and Society*, vol. 72, Sept. 2021, Art. no. 102994, https://doi.org/10.1016/j.scs.2021.102994.

[19] N. Moustafa and J. Slay, "UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)," in *2015 Military Communications and Information Systems Conference*, Canberra, Australia, 2015, pp. 1–6, https://doi.org/10.1109/MilCIS.2015.7348942.

[20] N. Koroniotis, N. Moustafa, E. Sitnikova, and B. Turnbull, "Towards the development of realistic botnet dataset in the Internet of Things for network forensic analytics: Bot-IoT dataset," *Future Generation Computer Systems*, vol. 100, pp. 779–796, Nov. 2019, https://doi.org/10.1016/j.future.2019.05.041.

[21] I. Ullah and Q. H. Mahmoud, "A Scheme for Generating a Dataset for Anomalous Activity Detection in IoT Networks," in *33rd Canadian Conference on Artificial Intelligence, Canadian AI 2020*, Ottawa, Canada, 2020, pp. 508–520, https://doi.org/10.1007/978-3-030-47358-7_52.

[22] M. A. Ferrag, O. Friha, D. Hamouda, L. Maglaras, and H. Janicke, "Edge-IIoTset: A New Comprehensive Realistic Cyber Security Dataset of IoT and IIoT Applications for Centralized and Federated Learning," *IEEE Access*, vol. 10, pp. 40281–40306, 2022, https://doi.org/10.1109/ACCESS.2022.3165809.