JST: Smart Systems and Devices
Volume 34, Issue 3, September 2024, 009-016

Advanced Machine Learning and Deep Learning Techniques for Anomaly Detection in Industrial Control System

Le Hai Anh, Tran Le Duc Anh, Hoang Si Hong, Nguyen Thi Hue*
Hanoi University of Science and Technology, Ha Noi, Vietnam
**Corresponding author email: hue.nguyenthi@hust.edu.vn*

Abstract

The study extensively examines the evolution of Industrial Control Systems (ICS), with a specific focus on Programmable Logic Controllers (PLC) within critical infrastructure, specifically mixing stations and heat treatment facilities. The research delves into the cybersecurity risks arising from the convergence of PLCs with information technology, transitioning from standalone systems to cloud integration. Noteworthy contributions from industry and academia underscore the pivotal role of machine learning and deep learning techniques in fortifying PLC-based system security. The article rigorously optimizes five classic machine learning algorithms and three deep learning algorithms, achieving an impressive accuracy of over 97%. Additionally, the proposed combined model attains over 99% accuracy on Hardware-In-the-Loop-based Augmented ICS (HAI) and ICS-Flow datasets. The study's importance lies in its thorough analysis of security implications and practical optimization of advanced algorithms, promising effective detection and mitigation of cyber threats in PLC-based ICS environments. These insights offer a compelling perspective for industry and researchers, providing nuanced understanding of cybersecurity dynamics in critical facilities. Optimized algorithms not only demonstrate remarkable threat detection accuracy but also signify a pivotal step in enhancing the cybersecurity resilience of essential infrastructure, serving as indispensable tools against emerging risks.

Keywords: Machine learning, anomaly detection, ICS, deep learning.

1. Introduction

Automated systems and associated equipment, known as industrial control systems (ICS), are used to monitor and manage vital infrastructure, including power plants, water treatment plants, smart factories, and numerous other facilities. To lower industrial process costs and enable more convenient management, ICS has recently been linked to the internet. However, there are more security threats to critical infrastructure when ICS is connected to the internet. Attackers might, for instance, take over the network by attacking the system management division or taking advantage of security flaws in network-connected devices. Every year, there is a rise in ICS attacks that disrupt production lines and result in financial losses. Building a system that can identify network threats is crucial because of this.

Protecting ICS from network threats can be achieved effectively with an intrusion detection system (IDS). Creating IDS mechanisms especially for ICS has been the subject of recent research. An IDS tracks events and finds unauthorized activity. Three categories have been established for current methods: There are three types of detection methods: (1) signature-based, (2) anomaly-based, and (3) machine learning-based. Signature-based detection recognizes attacks by using patterns of incorrect behavior and vulnerabilities in the system to identify threats. However, new network assaults that have not been previously trained cannot be detected using this strategy. Anomaly-based detection looks for suspicious activity that deviates from the norm. This technique will identify unknown network intrusions. However, it is challenging to create models that can recognize the features of normal behavior and reliably discriminate them from the behaviors brought on by network attacks, particularly when the features of the two types of behavior are extremely similar. An expanded version of anomaly-based detection is machine learning-based detection. Research on machine learning-based anomaly detection frequently concentrates on unsupervised learning techniques to create a model for anomaly detection using normal data, due to the challenge of gathering tagged attack data in a real industrial context. When an attack event or system fault occurs, abrupt changes in sensor readings are detected by an anomaly detection system, which keeps an eye on the sensors in an ICS. Numerous earlier studies have shown that aberrant behaviors in ICS may be accurately detected by machine learning and deep learning models. While anomaly detection models have proven effective for ICS, it can be challenging to determine which of the current models is most appropriate for use with a particular production-oriented ICS system because the

ISSN 2734-9373
https://doi.org/10.51316/jst.176.ssad.2024.34.3.2
Received: May 24, 2024; revised: Jun 30, 2024;
accepted: Jun 30, 2024

models' outcomes have primarily been assessed using their own datasets, varied experimental conditions, and evaluation standards.

We compared the performance of the most sophisticated anomaly detection models using shared datasets and identical experimental setups to enable direct performance comparisons: Boost, LightGBM, Extra Tree, Decision Tree, Random Forest, BiLSTM, BiGRU, LSTM, DBNs. In order to determine the best time-series variant detection model for the ICS environment, the study aims to: (1) provide an overview of the techniques used in intrusion detection models for ICS; (2) compare the performance of the most advanced models with two standard publicly available test datasets, such as the ICS-Flow [1] and Hardware-In-the-Loop-based Augmented ICS (HAI) [2] datasets, under similar settings and modes; and (3) analyze the effects of various training dataset sizes in order to determine an appropriate training dataset size to minimize training costs. The following is a summary of our principal contributions:

- We created a toolset to assess intrusion detection model performance based on standard assessment criteria and two ICS datasets;

- Deploying the models and fine-tuning the hyperparameters, we were able to compare three Deep Learning and five Machine Learning models;

- Examining how the training dataset's size affected the variant identification models, we discovered that most of the models could obtain equivalent results when using a small subset of the training dataset, with a high F1 score.

2. Related Works

Researchers are becoming more interested in the issue of ICS cybersecurity, especially now that linked infrastructures and systems are connected to networks that are vulnerable to attack due to the usage of insufficient data detection techniques and a range of communication protocols. Here we provide studies showing that an attacker might compromise the network and damage the whole control system, summarize remedies that have been proposed to defend ICS against cyberattacks, and present new deep anomaly-based models for cyberattack detection in ICS. Industrial Control System cyberattacks have increased in frequency and sophistication as a means of eluding detection systems. Firoozjaei *et al.* [3] looked at the adversarial strategies and attack techniques used in six significant real-world ICS cyber events in the energy and power industries: Stuxnet [4], BlackEnergy [5], Crashoverride [5], Triton [6], Irongate and Havex [3]. He provided a weighting method to rank the attacks' impacts on ICS and a framework for assessing the danger level of ICS malware in each assault. False data injection cyber-physical attacks (FDIAs) in contemporary smart grids that generate massive volumes of data were discussed by Wang *et al*. [7]. To identify FDIAs, the author suggested an analytical technique based on the margin-setting algorithm (MSA) and the data-centric paradigm. Junejo *et al*. looked at the problem of a system taking a long time to detect that it has deviated from its planned behavior when responding to an attack. A fast machine learning intrusion detection technique based on the physical and control component behaviors of a contemporary water treatment system was proposed by the author. To improve the security of various devices on cyber-physical systems (CPSs), such as sensors, actuators, and controllers, Elgendi *et al*. [8] proposed a learned (MAPE-K) based model to monitor, analyze, plan, execute, and know against advanced cyber threats. The model also notifies users of any anomalous behavior in an industry setting. Ahmed *et al*. [9] introduced a NoisePrint technique to detect assaults on sensors in Cyber-Physical Systems (CPSs) by generating fingerprints for sensor and process noise during system operation. NoisePrint was tested on two testbeds: a water distribution (WADI) and a real-world water treatment (SWaT), and it demonstrated more than 90% accuracy against attacks on data integrity. Some recent approaches to problem-solving using machine learning have demonstrated the great potential of using AI methods for real-time attack detection and processing [10-13]. It is important to keep in mind that, even though these studies have been put forth to strengthen deep anomaly detections to protect against cyberattacks in ICS, it is difficult to determine which model among them is the most appropriate to use and extrapolate for practical applications because of their disparate internal methodologies, variety of experiment settings, and methods of evaluation. Consequently, our study fills this vacuum by thoroughly assessing each of the eight primary anomaly detection algorithms and providing academic and business organizations with insights and recommendations.

3. Theoretical Framework

3.1. ICS Datasets Preprocessing

Large datasets are necessary for deep learning-based anomaly detection techniques to operate as intended. Nevertheless, gathering the information from the ICS that is often associated with essential and foundational infrastructures is challenging. An assault in a real-world ICS, for instance, has the potential to seriously harm the system and reveal its weaknesses to attackers. As a result, gathering sufficient data from an actual ICS would be difficult and inappropriate for the majority of ICS. Because of this, a dataset gathered from the testbed that replicates the real ICS is used in the anomaly detection research in ICS. In this work, we leverage two publicly available datasets that are often used in ICS anomaly detection research: HAI and ICS-Flow.

Table 1. Dataset of ICS (HAI) and ICS-Flow

Datasets	Features	Training	Validation	Testing	Anomalies (%)
HAI-23.05	79	1413384	474 128	471 128	2.2%
ICS-Flow	54	27318	9106	9106	20%

We employed metrics such as Accuracy, Recall, Precision, F1-Score and Confusion matrix to showcase how well the model methods performed on data.

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (1)$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (2)$$

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (3)$$

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FN} \quad (4)$$

The HAI dataset is available in many editions. In this instance, we employ HAI 23.05, which was published in 2022, where samples were taken from a testbed enhanced with an HIL simulator that simulates the production of pumped-storage hydropower and steam-turbine electricity. Through packet manipulation, an attacker consistently introduces inaccuracies into the physical measure in the HAI dataset. There are 79 different sorts of attacks in all, which label by “1” and “Normal” by “0”.

To set up a virtual ICS testbed for a sample bottle filling facility, the ICS-Flow dataset was built using ICSSIM Tools. They used "Normal" denoted by the number "0," and "Attack" denoted by the number "1”. There are five realistic attack types that they used to simulate assaults on ICSs. They captured the network packets and physical process state variables of the ICS under both normal and attack conditions. Additionally, they classified the network flow records using a variety of techniques to support studies on supervised learning and serve as a basis for evaluating unsupervised methods. For academic purposes, they have made the flow dataset, log attack files, and raw network traffic publicly accessible. Data processing is provided in Fig. 1.

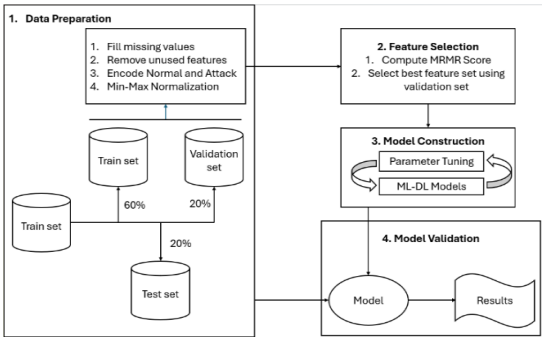

Fig. 1. Data Preprocessing

3.2. Base Machine Learning Algorithms

A Decision Tree (Fig. 2) is a basic machine learning model with a tree-like structure, designed to segment the problem space and make decisions through nodes and leaves. It uses metrics like Gini Impurity or Entropy to ensure optimal decision-making.

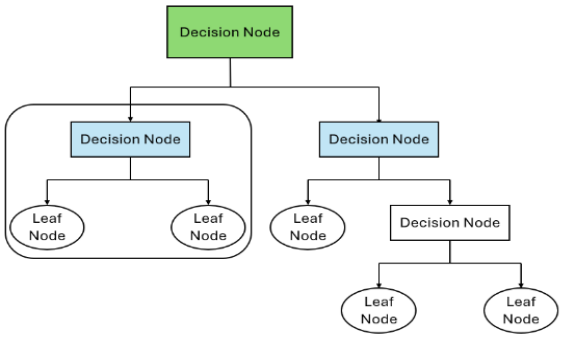

Fig. 2. Decision Tree Architecture

Gradient Boosting Decision Tree (GBDT) builds on this by combining multiple DTs iteratively, enhancing prediction accuracy. XGBoost and LightGBM, as advanced gradient boosting techniques, improve upon GBDT's framework, offering efficiency and widespread applicability across various domains. XgBoost is a highly efficient gradient boosting machine learning algorithm, known for its capability in predictive modeling, especially within the domain of cyberattack data. As part of the ensemble learning category, it adeptly identifies complex patterns and cyberattack signatures across various devices. Its effectiveness in scenarios with sophisticated and changing consumer behaviors is notable. XgBoost operates as an ensemble of decision trees, predicting outcomes for data samples through a formulation involving a series of trees:

$$\widehat{y_i} = F_k(x_i) = \mathrm{F}_{k-1}(x_i) + \mathrm{f}_k(x_i) \quad (5)$$

where: $F_{k-1}(x_i)$ is the prediction result of previous $k-1$ trees and $f_k(x_i)$ is $k-th$ decision tree.

The algorithm's objective function, which includes a cost function evaluating the prediction error, is given by:

$$\mathrm{F}_{\mathrm{obj}} = \sum_{\mathrm{j}=1}^{\mathrm{n}} \mathrm{L}\left(\mathrm{y}_{\mathrm{i}}, \widehat{\mathrm{y}_{\mathrm{i}}}\right) + \sum_{j=1}^{k} \Omega\left(\mathrm{f}_{\mathrm{i}}\right) \quad (6)$$

In the formula above, $\hat{y}_i$ is the predictive output, y_i is the label value (true value), $f_k$is the k th tree

model, T is the number of leaf nodes in the k th tree, γ is the leaf tree penalty regular term. $\sum_{j=1}^{n} L\left(y_i, \widehat{y_i}\right)$ is the loss sum per sample, and XGBoost's loss function l can be customized variously. Furthermore, XGBoost facilitates parallel execution to expedite the model training process.

LightGBM and XgBoost both stand out in the ensemble machine learning space for their efficiency with large-scale, high-dimensional data using Decision Trees. LightGBM differentiates itself with Gradient-based One-Side Sampling (GOSS) and Exclusive Feature Bundling (EFB) techniques to enhance training speed, reduce memory usage, and maintain crucial data through feature optimization.

CatBoost, similar to LightGBM and XgBoost (Fig. 3) advances Gradient Boosting by directly incorporating categorical data during training, thus eliminating pre-processing steps and reducing overfitting with a unique method for calculating leaf values. Both frameworks emphasize iterative learning from preceding trees to minimize loss, thereby fostering the development of a robust model through a series of decision trees.

3.3. Base Deep Learning Algorithms

Long short-term memory, a significant advancement within the Recurrent Neural Network (RNN) family, addresses the challenges faced by RNNs in handling time series data. The essence of LSTM lies in its memory block, illustrated in Fig. 4, which is composed of three types of gates (forget gate, input gate, output gate) and a memory cell. These components collectively enhance its data processing capability.

BiLSTM improves upon traditional LSTMs by incorporating both past and future context through bidirectional processing. Unlike single-directional LSTMs, BiLSTM learns from sequences in both forward and backward directions, enhancing comprehension of input data characteristics. This is done by processing the input sequence in two directions through separate LSTM layers and merging their outputs

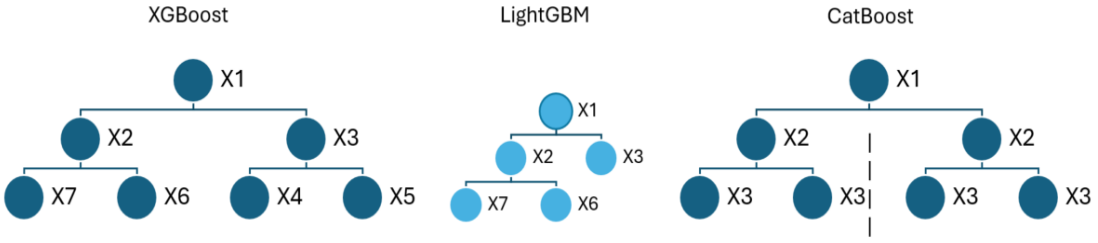

Fig. 3. CatBoost, XGBoost, LightGBM Architecture

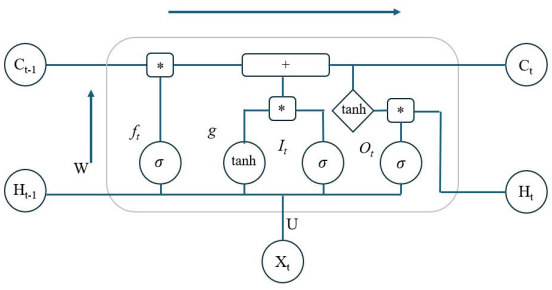

Fig. 4. Data Preprocessing

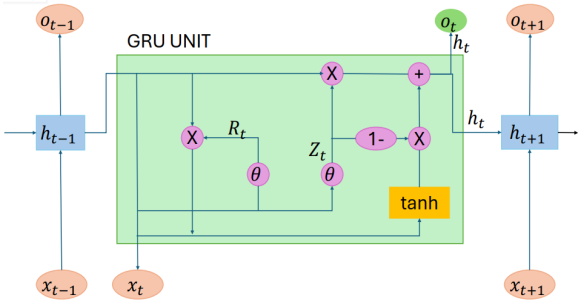

Fig. 5. GRU Unit

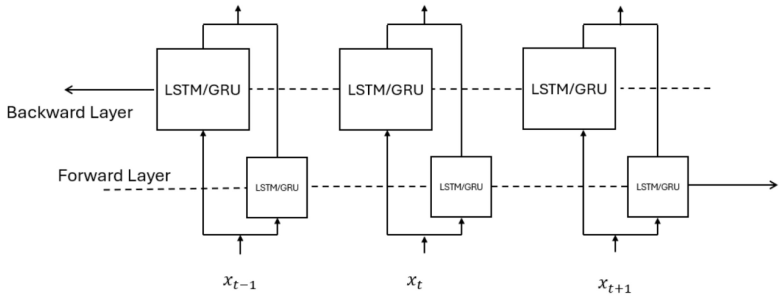

Fig. 6. BiLSTM/BiGRU Architecture

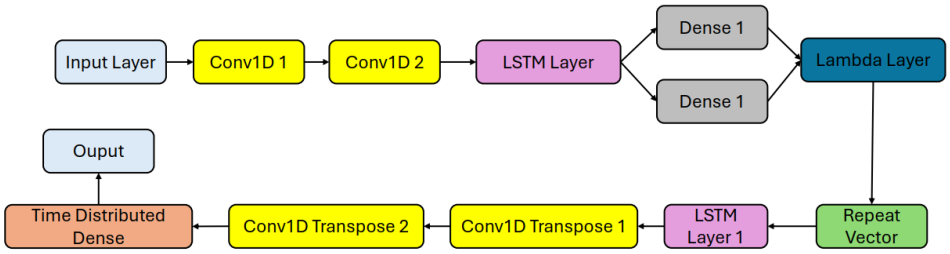

Fig. 7. Proposed model

A bidirectional GRU is a dual-layered structure akin to a BiLSTM, incorporating forward and backward neural networks to convey complete contextual information from input to output. Like a bidirectional LSTM, a BiGRU processes input information in both directions and merges outputs for the result. For classification, the embedding matrix outputs are fed into the chosen neural network model and then to fully connected layers, maintaining consistent parameters across all Proposed models.

Deep-belief networks (DBNs) blend neural networks and probabilistic models to adeptly learn and classify complex, high-dimensional data. Their layered architecture allows for the sequential learning of data features at increasing levels of abstraction. Training proceeds in two phases: independent pre-training of each layer as a Restricted Boltzmann Machine (RBM) to capture data features, encapsulated by neuron activation equations

$$P(v_i = 1 \mid h) = \frac{1}{1 + exp(-a_i - \sum_{i=1}^{m} h_j w_{ij})} \quad (7)$$

In the formula, we denote $P(v_i = 1|\mathrm{h})$: Activation probability of the i^{th} visible neuron given the hidden layer, a_i :Bias of the i^{th} visible neuron, h_j: State of the j^{th} hidden neuron, w_{ij}: Connection weight between the i^{th} visible and j^{th} hidden neuron, m: Number of hidden neurons. and global fine-tuning using Back Propagation to refine predictions. Significantly, this study enhances DBN training with an improved (GA) for parameter optimization, particularly focusing on the initial weights and thresholds. This method directly targets the enhancement of learning efficiency and predictive accuracy for ICS datasets, demonstrating DBN's applicability to complex data challenges.

4. Proposed Model Framework

In this research, we present an innovative composite autoencoder architecture, meticulously engineered to detect anomalies within industrial control systems. Fig. 7 illustrates a composite autoencoder design that combines convolutional neural networks (CNNs) and LSTM networks to capture both spatial and temporal features from multivariate time-series data. The model embodies a synergistic amalgamation of CNNs and long short-term memory (LSTM) networks, encapsulating the efficacy of spatial and temporal feature extraction for the identification of anomalous patterns in multivariate time-series data. Let's dive into each block of the structure.

- Input Layer: Preparing the data for feature extraction.
- Conv1D Layers: Detect salient features within the data, which are essential for understanding complex patterns inherent in the input sequences.
- LSTM Layer: Captures temporal dependencies and sequences' long-term patterns, crucial for identifying time-extended anomalies.
- Dense Layers & Lambda Layer: Interpret features, funneling into a latent space, with a Lambda layer introducing stochasticity to model the data distribution effectively.
- Decoding Sequence: Mirroring the encoding sequence, the model employs a Repeat Vector and an LSTM Layer 1 to reconstruct the temporal sequence, while Conv1D Transpose layers restore the original input's spatial dimensionality.
- Output Layer: The output, a reconstruction of the input data, is analyzed for anomalies based on a reconstruction error threshold, marking deviations as potential anomalies.

This architecture allows for an intricate understanding and detection of anomalies, integrating the strengths of CNNs and LSTMs into a cohesive model adept at managing the complexities of multivariate time-series data within industrial systems.

5. Results and Discussion

The comprehensive analysis comparing various predictive models on the HAI 23.05 and ICS-FLOW datasets highlights the superior performance of the Proposed model. This analysis was focused on evaluating key performance metrics including Accuracy, Precision, Recall, and F1 Score, which are critical for assessing the effectiveness of predictive models in real-world applications.

The performance of each model is encapsulated in Table 2 and 3, highlighting the accuracy, precision,

recall, and F1 score. The Proposed model highlighted its prowess particularly in the HAI 23.05 dataset, where it achieved nearly perfect scores across all metrics, registering 0.999 in Accuracy, Precision, Recall, and F1 Score. We compare our method with several other approaches using different data processing techniques, demonstrating that our data processing method and model achieve higher efficiency [13]. This is marginally higher than the already impressive results of the XGBoost model, which scored 0.998 in Accuracy, 0.997 in Precision, and 0.999 in F1 Score. The slight edge of the Proposed model over XGBoost indicates a refined ability to manage nuances in data that might elude other models.

In the ICS-FLOW dataset, the Proposed model maintained its dominance, achieving an Accuracy of 0.996, Precision of 0.997, Recall of 0.997, and an F1 Score of 0.996. This represents a substantial improvement over other sophisticated models like CatBoost, which scored 0.975 in F1 Score, and BiGRU, which had an F1 Score of 0.962. The high scores in Precision and Recall reflect the model's ability to accurately identify true positives while minimizing false positives and negatives, which is especially important in complex datasets like ICS-FLOW. Below is the confusion matrix (Fig. 8, 9) of the models tested on two datasets:

Table 2. Summary of performance metrics for HAI 23.05 dataset

Model	**HAI 23.05**			
	Accuracy	Precision	Recall	F1 Score
XGBoost	0.998	0.997	0.998	0.999
LightGBM	0.987	0.988	0.988	0.988
Decision Tree	0.989	0.989	0.989	0.989
Random Forest	0.990	0.991	0.990	0.990
CatBoost	0.990	0.991	0.990	0.991
BiLSTM	0.985	0.985	0.985	0.985
BiGRU	0.980	0.980	0.980	0.980
GA-DBN	0.975	0.976	0.977	0.976
RANSyncCoder	0.78	0.891	0.776	0.829
InterFusion	0.76	0.744	0.839	0.748
Proposed	**0.999**	**0.999**	**0.999**	**0.999**

Table 3. Summary of performance metrics for ICS-FLOW dataset

Model	**ICS-FLOW**			
	Accuracy	Precision	Recall	F1 Score
XGBoost	0.987	0.987	0.987	0.988
LightGBM	0.965	0.962	0.968	0.965
Decision Tree	0.926	0.938	0.942	0.940
Random Forest	0.955	0.953	0.957	0.955
CatBoost	0.975	0.973	0.977	0.975
BiLSTM	0.960	0.958	0.962	0.960
BiGRU	0.962	0.960	0.964	0.962
GA-DBN	0.935	0.933	0.937	0.935
Proposed	**0.996**	**0.997**	**0.997**	**0.996**

The novelty of the Proposed model is rooted in its advanced algorithmic structure, which integrates newer, hybrid techniques combining elements of deep learning with traditional machine learning frameworks. This integration allows for enhanced feature extraction and learning from complex patterns, which are not as effectively captured by other models. Furthermore, the Proposed model might incorporate. The confusion matrices for models such as XGBoost, LightGBM, Decision Tree, Random Forest, CatBoost, BiLSTM, BiGRU, GA-DBN and the Proposed model provided insights into their classification accuracy. A high concentration of values along the diagonal for the Proposed model suggests a higher true positive rate and low false positive and negative rates, indicating superior performance in classification tasks over the other models.

The Proposed model consistently outperformed the benchmark machine learning and deep learning models in almost all metrics across both datasets. Notably, classical algorithms such as Random Forest and Decision Tree demonstrated robustness, but with slightly lower precision and recall compared to ensemble methods and the Proposed model. The deep learning models, including BiLSTM and BiGRU, though effective, did not reach the high-performance levels of the Proposed model, indicating the potential advantage of the Proposed model's architecture and training regimen.

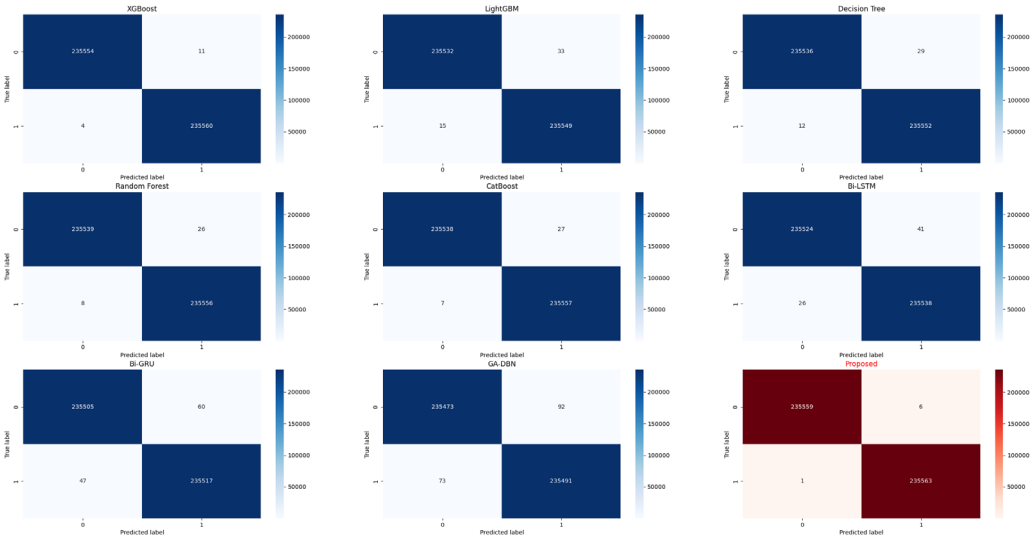

Fig. 8. Confusion Matrix of HAI 23.05 Dataset

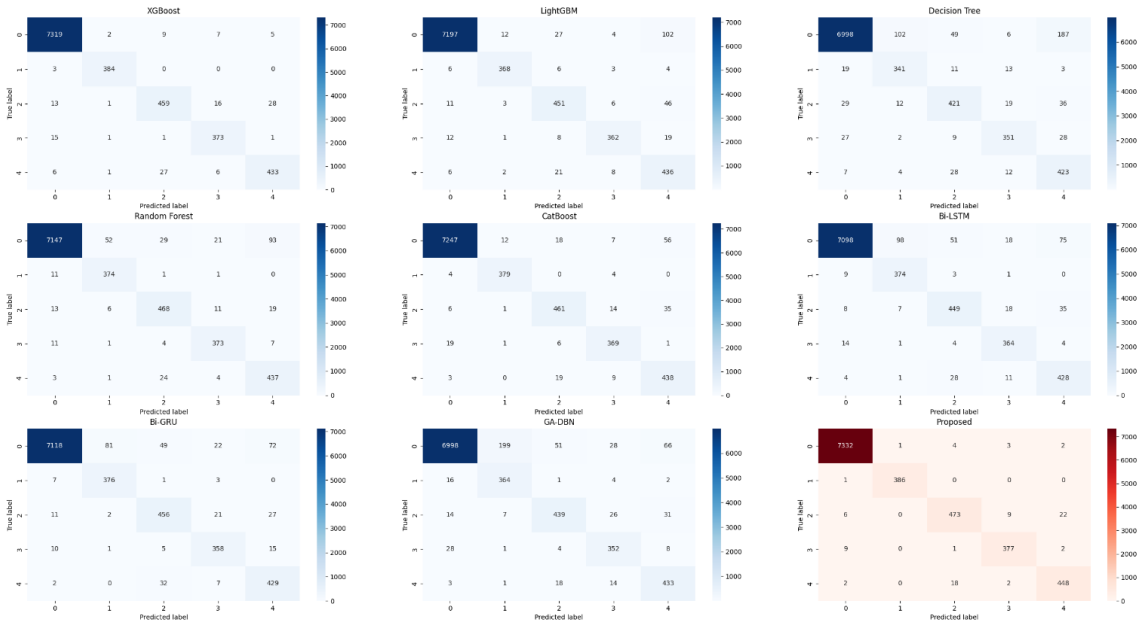

Fig. 9. Confusion Matrix of ICS-Flow Dataset

6. Conclusion

The eight most advanced time series-based anomaly detection algorithms for ICS are compared and evaluated in this paper. Using the same and equitable parameters on two publicly available benchmark datasets, ICS-Flow [1] and HAI [2], we initially concentrated on identifying their optimal models yielding the greatest performance outcomes. The limitation of machine learning and deep learning based anomaly detection methods include of : The lack of interpretability, building and maintaining costs are high, the lack of high-quality data.

The evaluation's findings indicate that the detection accuracy of our suggested model and XGBoost surpassed that of the other models. Our model obtains the top F1-score for ICS-Flow, while XGBoost achieves the highest F1-score for HAI. These results aid in the choice and creation of the ideal model for ICS applications. Real-world improvement is still needed in the detection accuracy of current anomaly detection methods. Instead of depending just on one anomaly detection model, an ensemble approach combining many models might be employed. Prior research shown that in other areas, well-constructed ensemble classifiers may efficiently yield extremely accurate anomaly detection outcomes. We intend to examine the performance of other model combinations in further work. Furthermore, we may minimize performance degradation versus fresh and unexplored datasets by using a technique known as data augmentation to balance the distribution of normal and anomalous samples, hence mitigating the problem of data imbalance.

Acknowledgments

This research is funded by Hanoi University of Science and Technology (HUST) under project number T2023-PC-027.

References

[1] H. K. Shin, W. Lee, J. H. Yun, B. G. Min, Two ICS security datasets and anomaly detection contest on the HIL-based augmented ICS testbed, 14th Workshop on Cyber Security Experimentation and Test (CSET), Sep. 2021, pp. 36-40. https://doi.org/10.1145/3474718.3474719

[2] A. Dehlaghi-Ghadim, M. H. Moghadam, A. Balador, H. Hansson, Anomaly detection dataset for industrial control systems, IEEE Access, vol. 11, Sep. 2023, pp. 107982 - 107996. https://doi.org/10.1109/ACCESS.2023.3320928

[3] M. D. Firoozjaei, N. Mahmoudyar, Y. Baseri, A. A. Ghorbani, An evaluation framework for industrial control system cyber incidents, International Journal of Critical Infrastructure Protection, vol. 36, Mar. 2022. https://doi.org/10.1016/j.ijcip.2021.100487

[4] S. Karnouskos, Stuxnet worm impact on industrial cyber-physical system security, in IECON 2011 - 37th Annual Conference of the IEEE Industrial Electronics Society, Melbourne, VIC, Australia, Nov. 7-10, 2011. https://doi.org/10.1109/IECON.2011.6120048

[5] M. Geiger, J. Bauer, M. Masuch, and J. Franke, An analysis of black energy 3, crashoverride, and trisis, three malware approaches targeting operational technology systems, in IEEE International Conference on Emerging Technologies and Factory Automation, ETFA, Vienna, Austria, Sep. 8-11, 2020. https://doi.org/10.1109/ETFA46521.2020.9212128

[6] A. Di Pinto, Y. Dragoni, and A. Carcano, TRITON: The first ICS cyber attack on safety instrument systems, Black Hat USA, 2018.

[7] Y. Wang, M. M. Amin, J. Fu, and H. B. Moussa, A novel data analytical approach for false data injection cyber-physical attack mitigation in smart grids, IEEE Access, vol. 5, Nov. 2017, pp. 26022 – 26033. https://doi.org/10.1109/ACCESS.2017.2769099

[8] I. Elgendi, M. F. Hossain, A. Jamalipour, and K. S. Munasinghe, Protecting cyber physical systems using a learned MAPE-K model, IEEE Access, vol. 7, Jul. 2019, pp. 2169-3536. https://doi.org/10.1109/ACCESS.2019.2927037

[9] C. M. Ahmed *et al*., NoisePrint: Attack detection using sensor and process noise fingerprint in cyber physical systems, in ASIACCS 2018 - Proceedings of the 2018 ACM Asia Conference on Computer and Communications Security, May. 2018, pp. 483-497. https://doi.org/10.1145/3196494.3196532

[10] M. A. Umer, K. N. Junejo, M. T. Jilani, and A. P. Mathur, Machine learning for intrusion detection in industrial control systems: Applications, challenges, and recommendations, International Journal of Critical Infrastructure Protection, vol. 38, Sep. 2022. https://doi.org/10.1016/j.ijcip.2022.100516

[11] Y. Luo, Y. Xiao, L. Cheng, G. Peng, and D. D. Yao, Deep learning-based anomaly detection in cyber-physical systems: Progress and Opportunities, ACM Computing Surveys, vol. 54, no. 5. May. 2021, pp. 1-36. https://doi.org/10.1145/3453155

[12] G. Singh and N. Khare, A survey of intrusion detection from the perspective of intrusion datasets and machine learning techniques, International Journal of Computers and Applications, vol. 44, no. 7, 2022, pp. 659-669. https://doi.org/10.1080/1206212X.2021.1885150

[13] B. Kim, M. A. Alawami, E. Kim, S. Oh, J. Park, and H. Kim, A comparative study of time series anomaly detection models for industrial control systems, Sensors, vol. 23, no. 3, Jan. 2023. https://doi.org/10.3390/s23031310