International Journal of Medical Toxicology & Forensic Medicine
2026, Volume 16, E52740

Research Paper

Drift-Aware Resilient Ensemble Deep Learning-based Anomaly Detection Model in Internet of Medical Things

Vishnu Prasad Devanathan*, **Saraswathi Selvarajan**

Department of Information Technology, Puducherry Technological University, Pillaichavady, Puducherry, India.

Citation Devanathan VP, Selvarajan S. Drift-Aware Resilient Ensemble Deep Learning-based Anomaly Detection Model in Internet of Medical Things. International Journal of Medical Toxicology and Forensic Medicine. 2026; 16:E52740.

https://doi.org/10.22037/ijmtfm.v16.52740

Article info:
Received: 09 July, 2026
Accepted: 20 July, 2026
Published: 23 August, 2026

Keywords:
Internet of medical things, Anomaly detection, Hybrid feature selection, Ensemble classifier, Hyperparameter tuning

ABSTRACT

Background: The rapid advancement of the Internet of Medical Things (IoMT) has facilitated continuous health monitoring and real-time physiological data collection. Artificial intelligence, particularly machine learning and deep learning techniques, has become widely employed for anomaly detection.

Methods: This research proposes a novel Drift-Aware Resilient Ensemble Deep Learning-based Anomaly Detection Model in Internet of Medical Things Environments (DREAM-IoMT). The proposed method employs a hybrid feature selection strategy integrating Gini-index, recursive feature elimination, and SelectFromModel. For anomaly detection and classification, the proposed DREAM-IoMT model employs a weighted-voting ensemble of a graph convolutional network, a bidirectional temporal convolutional network, and an improved conditional variational autoencoder. This ensemble approach enhances accuracy by accounting for each classifier's prediction confidence. The ensemble classification model is further optimized using AdamW.

Results: The DREAM-IoMT model achieved an accuracy of 98.74%, demonstrating superior anomaly detection performance compared with existing methodologies.

Conclusion: The proposed DREAM-IoMT effectively combines feature selection and weighted ensemble deep learning for reliable anomaly detection in IoMT environments.

*** Corresponding Author:**
Vishnu Prasad Devanathan, PhD
Department of Information Technology, Puducherry Technological University, Pillaichavady, Puducherry, India.
E-mail: *vishnuprasad@ptuniv.edu.in*

Introduction

The widespread integration of sensing and communication devices in healthcare environments has contributed to smart healthcare systems. The Internet of Things (IoT) serves as the foundation for interconnected healthcare applications by enabling communication among medical devices, wearable sensors, and healthcare platforms. A significant extension of IoT, known as the Internet of Medical Things (IoMT), has emerged to support continuous health monitoring, remote patient management, and healthcare service delivery. IoMT enables the integration of medical equipment, physiological sensors, and connected healthcare networks to collect and exchange health-related information in real time [1]. The combination of IoMT with conventional healthcare practices has improved accessibility, connectivity, and reliability in medical environments. Despite these advantages, the increasing connectivity and data exchange in IoMT also introduce security and privacy challenges. Anomaly detection in IoMT is a critical concern because of the exposure of IoMT networks to cyber-attacks. Nevertheless, traditional machine learning (ML)-driven anomaly detection methods frequently depend on shallow models, limited scalability, and handmade features that limit their utilization in large-scale IoT situations. Current improvements in DL present a promising option by allowing techniques to automatically learn expressive representations from higher-dimensional data [2].

A brief reviews prior related works on anomaly detection. Rodriguez et al. [3] addressed the vital requirement for powerful anomaly detection and categorization in IIoT models. Chen et al. [4] developed an optimum DL method for anomaly detection in IIoTs. Primarily, the XGBoost for FS is offered by establishing diverse thresholds. In [5], a new unsupervised model is proposed for time series data evaluation. This method primarily identifies anomalous models in IIoT sensor data through the extraction of local features. Riaz et al. [6] presented a two-phase generative oversampling system called Enhanced Optimisation of WGAN. Luo et al. [7] examined a self-adaptive BC sharding approach for combined IIoT devices, which mutually improves the security of model training and the count of shards. Wang et al. [8] presented an adaptable data streaming model that allows virtual anomaly detection for IIoT applications. The authors [9] implemented a time-series anomaly detection system based on a distributed knowledge distillation architecture. De Vita et al. [10] intended to design an innovative architecture that integrates ESN-AE to detect the anomalies in an industrial environment. The authors [11] developed an anomaly detection model by systematic comparison of various ML models. The LCA is applied for developing an efficient model for the applied dataset.

To address the challenges, this study presents a novel Drift-Aware Resilient Ensemble Deep Learning-based Anomaly Detection Model in Internet of Medical Things Environments (DREAM-IoMT). The significant contribution of this study is as follows:

- Integrate a hybrid feature selection mechanism combining Gini-index, recursive feature elimination, and SelectFromModel (GI-RFE-SF) to retain beneficial features while removing redundant attributes.
- Design a weighted voting ensemble integrating graph convolutional network (GCN), a bidirectional temporal convolutional network (BiTCN), and an ICVAE to enhance anomaly detection precision.
- Present a drift-aware online adaptation system termed Extended ADWIN, integrating a variational Bayesian state estimator, ADWIN, and adaptive thresholding to detect and respond to concept drift in streaming IoMT data.
- Conduct evaluations on benchmark IoMT datasets and demonstrate its effectiveness under various measures.

Materials and Methods

This study presents a systematic approach for anomaly detection in IMoT environments. Figure 1 depicts the overall structural design of the DREAM-IoMT model. The proposed model encompasses multiple stages: feature selection, ensemble classification, optimization, drift adaptation, and performance evaluation. The detailed descriptions are presented in the following sections.

Feature Selection Strategy

A hybrid feature selection approach incorporating GI-RFE-SF models is deployed for selecting features for the classification process.

GI

The GI was verified to be effective in recognizing appropriate attributes. Nevertheless, it is subject to overfitting and sensitivity to feature interface [12].

$$\text{Gini}(P) = 1 - \sum (P_i)^2 \qquad (1)$$

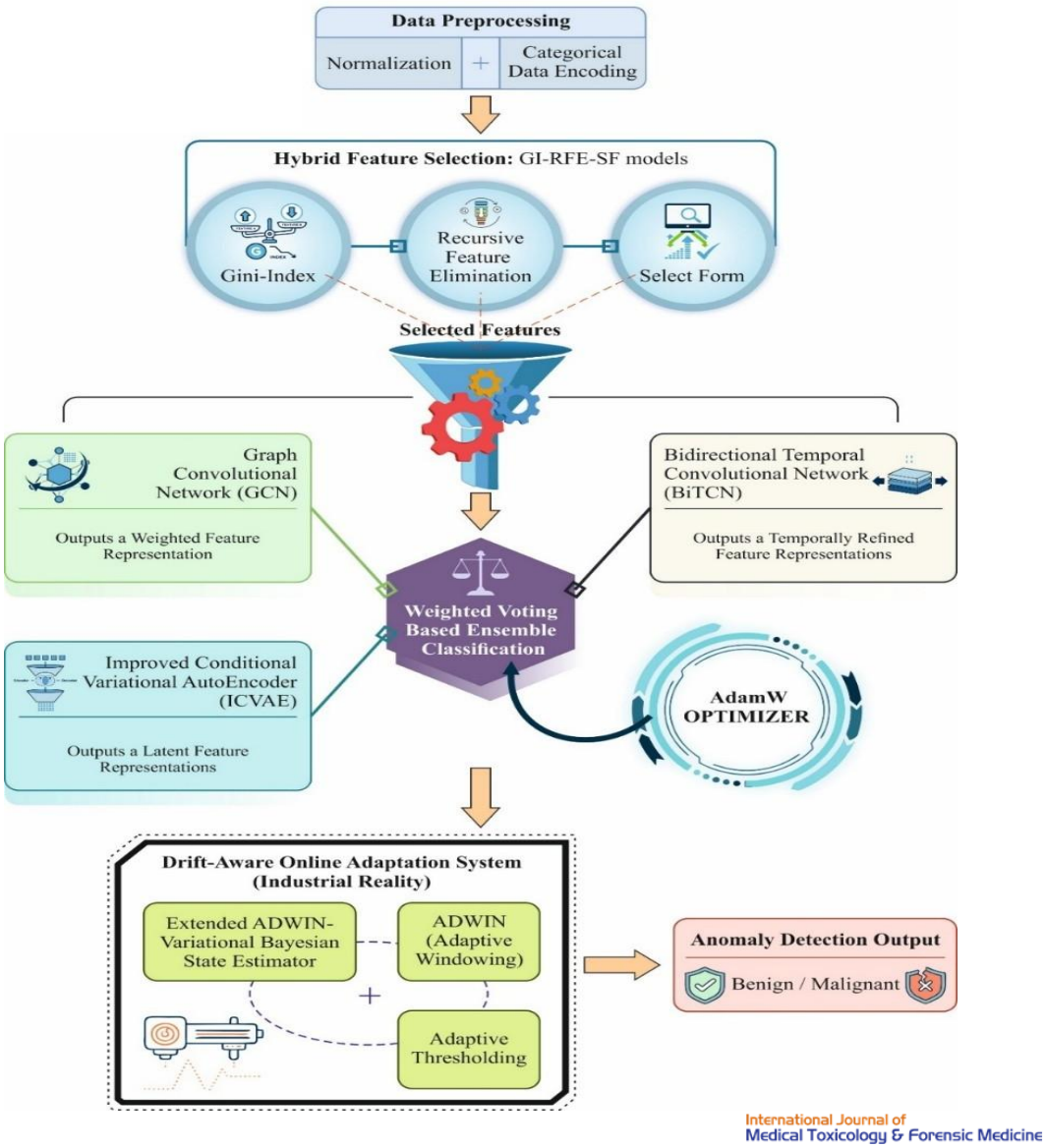


Figure 1. Overall structural design of the DREAM-IoMT model.

Here, $P_i$signifies the percentage of cases in the i^{th}class.

RFE

The FS model named RFE eliminates attributes one by one and builds methods using the remaining features.

$$F_{selected} = RFE(F_{all}) \qquad (2)$$

Here, F_{all}signifies the complete features, and $F_{selected}$denotes the chosen features.

SF

FS utilizing SF begins with defining the threshold value. Attributes with an importance score below the threshold will be removed.

Hybrid feature selection

The hybrid FS scheme combines GI, RFE, and SF methods.

Employ GI, RFE ranks feature importance and iteratively removes less significant attributes, and SF retains attributes above the threshold.

Ensemble Classification Module

For anomaly detection and classification, the proposed model employs a weighted-voting ensemble of three DL techniques: a GCN, a BiTCN, and an ICVAE. Numerous aggregation approaches for outcomes are discussed to enhance classifier performance [13].

$$H(Z) = \arg\max_{q \in C}\left(\sum_{p=1}^{3} V_p h_p^q(x)\right) \qquad (3)$$

where V_p represents the voting, weight allocated to the p-th base classifier, $h_p^{(q)}(\mathbf{x})$ defines the confidence score, $\mathcal{C}$ indicates the collection of target labels, and x is the input vector.

GCN

GCN is a neural network structure that can directly execute convolution-like processes on non-Euclidean space data. The input graph is signified as $G = (V, E, A, X)$, where Vindicates a group of nodes, Edenotes the group of edges, $A \in R^{n \times n}$denotes an adjacency matrix representing the connections among graph nodes, and $X \in R^{m \times n}$denotes the matrix of features.

$$(f * g)_G = T_{IGFT}[T_{GFT}(f) \otimes T_{GFT}(g)] \quad (4)$$

Where T_{GFT}signifies the graph Fourier transform, and T_{IGFT}denotes an inverse graph Fourier transform. Figure 2 illustrates the architecture of GCN.

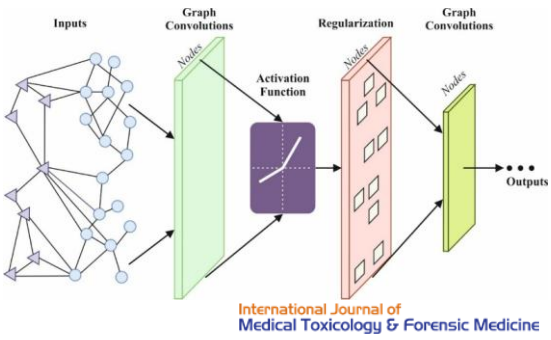


Figure 2. Architecture of GCN.

BiTCN

TCN utilizes a dilated causal convolution, which can only convey data from the historical to the future.

$$Output = softmax\left(f\left(W\left(1DCov\left(\hat{Y} \oplus \overleftarrow{\hat{Y}_1}\right)\right) + b\right)\right) \quad (5)$$

Here, $(TCN)^{\rightarrow}$and $(TCN)^{\leftarrow}$denote forward and backward TCN; $\oplus$ represents addition; $\hat{Y}$and $\hat{Y}_1$mean the outcomes of forward and backward TCN; 1DCov denotes a 1D convolutional process; $b$and Wrepresent the bias term and weight matrix, and Output represents the Bi-TCN output. Figure 3 describes the architecture of the BiTCN model.

ICVAE

CVAE is an addition of VAE, which is demonstrated by conditioning the encoding and decoding to the Yclass. The encoding $Q(W, | UV)$is conditioned on dual variables such as $U$and V, and the decoding $P(U, | WV)$is conditioned on dual variables like $W$and V. ICVAE is organised into encoding $Q_\phi(W | U)$and decoding $P_\theta(U, | WV)$.

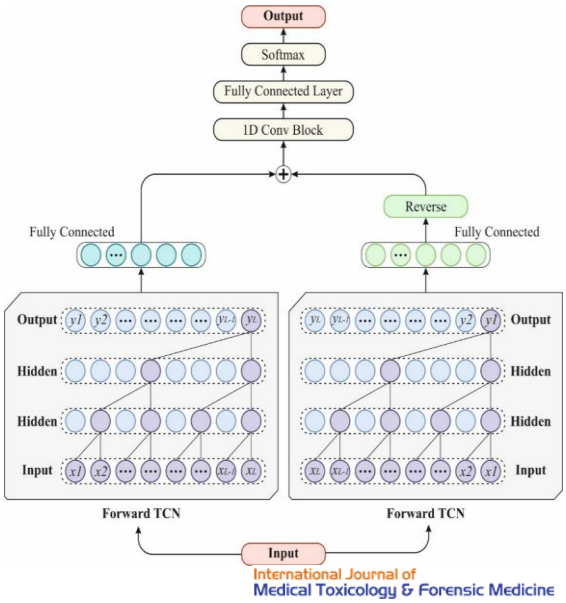


Figure 3. Model diagram of the BiTCN model.

$$L(\theta, \phi, U, V) = E[\log P(U, | WV)] - D_{KL}[Q(W | U) \| P(W | V)] \quad (6)$$

Here, $L(\theta, \phi, U, V)$consists of dual parts, such as log reconstruction probability of $P(U, | WV)$and a KL divergence $D_{KL}[Q(W | U) \| P(W | V)]$.

Model Optimization

To improve the ensemble model performance, this study employs the AdamW optimizer. AdamW is an enhanced variation of Adam optimizer, which improves the model's training by averting overfitting and enhancing the generalisation [14].

$$\theta_t = \theta_{t-1} - \eta\left(\frac{m_t}{\sqrt{v_t} + \epsilon} + \lambda\theta_{t-1}\right) \quad (7)$$

Here, θ_t represents an updated model parameter at the tth step; η is a rate of learning; m_t and v_t define an estimation of the first moment and second moment; ϵ is a smaller constant; λ represents a weight decay coefficient.

Drift-Aware Online Adaptation System

To ensure long-term reliability under changing data streams, the model integrates a drift-aware online adaptation system that combines an Extended ADWIN mechanism [15].

Variational Bayesian State Estimator

The objective of distributed Bayesian filtering is to capture uncertainty growth before drift. Let us take the linear state transition technique as;

$$u_t = A_t u_{t-1} + w_t \qquad (8)$$

whereas, $u_t \in R^p$means a target state at t^{th}time, $A_t \in R^{p\times p}$signifies a matrix of state-transition, and $w_t \sim N(0, Q_t)$denotes the unidentified covariance matrix $Q_t \in R^{p\times p}$.

ADWIN

ADWIN is the most general and precise supervised drift detector. Its central notion is upholding and identifying a variable-length window $W$with current data, like the error rate or its dual outputs with respect to misses and hits.

$$\varepsilon = \sqrt{\frac{2}{m}\sigma_W^2 \ln\left(\frac{2n}{\delta}\right)} + \frac{2}{3m}\ln\left(\frac{2n}{\delta}\right) \qquad (9)$$

Here, $n$represents the total number of samples, $m$represents the harmonic mean, and σ_W^2represents the variance of identified values in the window W.

Adaptive Thresholding Mechanism

To uphold robustness and responsiveness under the active condition, an Extended ADWIN employs a sliding window with a dynamic control size

$$w_{min}(t) = \begin{cases} \max(w_{min}, (t-1) - \lambda w_{lower}) & \text{if } \Delta_{\log A_2}(t) > \delta, \\ \min(w_{min}, (t-1) + \lambda w_{upper}) & \text{if } \Delta_{\log A_2}(t) < -\delta, \\ w_{min}(t-1), \end{cases}$$

Now, $w_{lower}$and w_{upper}imply the permitted bounds, λrefers to the fixed step size, and δregulates the sensitivity to comparative modifications.

$$k_n(t) = \begin{cases} k_n(t-1) + \delta_k, & \text{if } \Delta_{\log \bar{A}_2}(t) > \delta, \\ k_n(t-1) - \delta_k, & \text{if } \Delta_{\log \bar{A}_2}(t) < -\delta, \\ k_n(t-1), & \text{otherwise.} \end{cases} \qquad (11)$$

However, $\delta_k$implies a smaller size of gradual tuning.

Results

All experiments were conducted using the Kaggle cloud platform with a 4-core CPU, NVIDIA Tesla T4 GPU, and 29 GB RAM. The proposed model was implemented in Python.

Dataset Used

The DREAM-IoMT method can be inspected by using a dual dataset. Primarily, the WUSTL-IIOT-2021 Dataset [16] has 1194464 samples with two classes. Secondarily, the CyberTec IIoT Malware Dataset (CIMD-2024) dataset [17] comprises 45289 samples with six classes.

Binary class dataset

The WUSTL-IIoT-2021 dataset is an IIoT network traffic dataset. It encompasses numerous attack scenarios with labelled features for supervised learning. Out of the 49 features, only 42 features are chosen.

Multi-class dataset

The CyberTec IIoT Malware Detection Dataset (CIMD-2024) is a real-time dataset collected from an IIoT environment to identify malware categories. The dataset contains ransomware, botnets, spyware, trojans, worms, and normal traffic samples. Among the 37 available features, 32 relevant features are selected for analysis.

Figure 4 presents the confusion matrices of the proposed DREAM-IoMT model, showing accurate

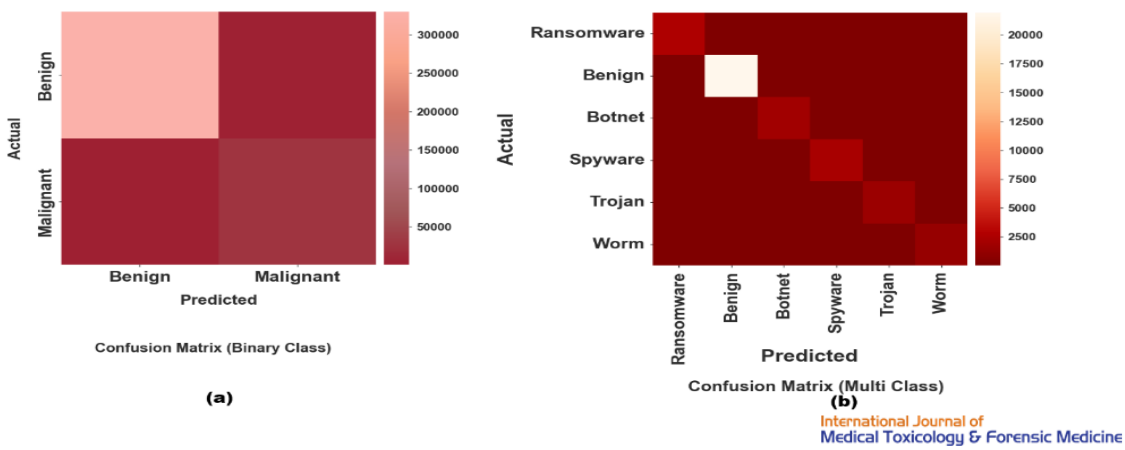


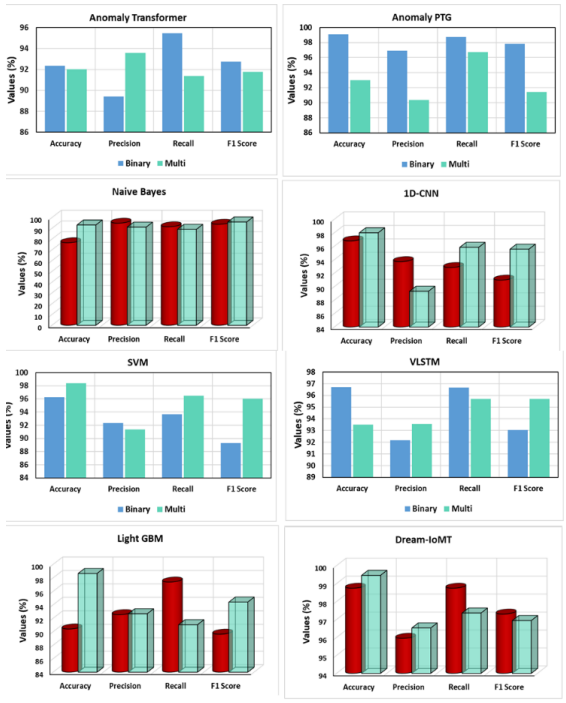
Figure 4. Confusion matrices of DREAM-IoMT method (a) binary class (b) multi class.

classification with strong diagonal values.

Table 1 implies the anomaly detection of the DREAM-IoMT method with binary and multi-class on 70:30 of TRAP/TESP. Under 70% TRAP, the DREAM-IoMT method provides average $accur_y$ of 98.73% for binary class, and 99.46% for multi-class. Under 30% TESP, the DREAM-IoMT method offers average $accur_y$ of 98.74% for binary class, and 99.42% for multi-class.

Discussion

Figure 5 present the experimental outcomes of the DREAM-IoMT model with existing methods [18-20]. The table shows the experimental outcomes of the DREAM-IoMT model with a binary class. The outcome determined that the proposed DREAM-IoMT method offers significant positive outcomes with $accur_y$ of 98.74%. Furthermore, the Anomaly PTG methods provided the lowest performance with $accur_y$ of 99.07%. Similarly, the SVM, VLSTM, and 1D-CNN methods attain moderately lesser $accur_y$ of 96.26%, 96.70%, and 96.80%, while the Anomaly Transformer



International Journal of Medical Toxicology & Forensic Medicine

Figure 5. Comparative outcomes of the DREAM-IoMT method under binary and multi-class.

Table 1. Anomaly detection of DREAM-IoMT method with binary and multi-class datasets.

Classes	$Accur_y$	$Preci_n$	$Recal_l$	$F1_{Score}$	MCC
Binary Class					
TRAP (70%)					
Benign	99.36	99.85	99.36	99.60	94.76
Malignant	98.10	92.29	98.10	95.11	94.76
Average	**98.73**	**96.07**	**98.73**	**97.36**	**94.76**
TESP (30%)					
Benign	99.33	99.85	99.33	99.59	94.65
Malignant	98.15	92.05	98.15	95.00	94.65
Average	**98.74**	**95.95**	**98.74**	**97.30**	**94.65**
Multi Class					
TRAP (70%)					
Ransomware	99.56	97.16	97.50	97.33	97.09
Benign	98.74	99.32	98.88	99.10	97.02
Botnet	99.59	95.80	97.33	96.56	96.35
Spyware	99.65	96.77	98.40	97.58	97.40
Trojan	99.47	95.17	94.02	94.59	94.31
Worm	99.72	94.94	98.04	96.47	96.34
Average	**99.46**	**96.53**	**97.36**	**96.94**	**96.42**
TESP (30%)					
Ransomware	99.60	97.37	97.72	97.55	97.33
Benign	98.67	99.27	98.81	99.04	96.87
Botnet	99.46	94.32	97.19	95.73	95.45
Spyware	99.62	97.08	97.67	97.38	97.17
Trojan	99.43	94.35	94.07	94.21	93.90
Worm	99.77	96.28	97.92	97.09	96.98
Average	**99.42**	**96.44**	**97.23**	**96.83**	**96.28**

International Journal of Medical Toxicology & Forensic Medicine

approach obtains considerable performance with $accur_y$ of 92.33%. Also, the NB and light GBM reach a minimum $accur_y$ of 76.59% and 90.41%.

These experimental values under multi-class demonstrated that the Anomaly Transformer and Anomaly PTG methods offer poor performance with $accur_y$, $preci_n$, $recal_l$, and $F1_{score}$. Meanwhile, the VLSTM method attains moderately greater performance with $accur_y$ of 93.48%. Moreover, the NB method achieves a moderately maximum resolution with $accur_y$ of 93.13%. Likewise, the SVM, 1D-CNN, and light GBM methods accomplish a considerably closer solution with $accur_y$, $preci_n$, $recal_l$, and $F1_{score}$. Ultimately, the presented DREAM-IoMT approach achieved a superior performance with $accur_y$ of 99.46%, when compared to other approaches.

Table 2 compares the computational time of DREAM-IoM model with various deep learning models. Among the evaluated architectures, the proposed model exhibits the lowest training time of 9.08 ms, demonstrating superior computational efficiency.

Table 2. Computational efficiency of DREAM-IoMT method.

Methods	Computational Time (ms)	Inference Time (ms)	FLOPs (G)	GPU Memory (GB)
Anomaly Transformer	52.84	14.27	3.86	3.42
Anomaly PTG	39.76	11.58	2.74	2.96
Naïve Bayes	6.83	1.42	0.01	0.12
1D-CNN	18.64	5.37	0.84	1.18
SVM	28.91	8.63	1.37	0.74
VLSTM	34.28	9.74	2.16	2.24
LightGBM	11.42	2.36	0.09	0.28
Dream IoMT	**9.08**	**1.87**	**0.06**	**0.21**

International Journal of
Medical Toxicology & Forensic Medicine

In contrast, Anomaly Transformer and Anomaly PTG record the highest training times at 52.84 ms and 39.76 ms, respectively. Other models such as Naïve Bayes with 6.83 ms, 1D-CNN with 18.64 ms, SVM with 28.91 ms, VLSTM with 34.28 ms and LightGBM with 11.42 ms. Overall, the results highlight that HCNIVAE-ODD achieves faster convergence and reduced computational time while maintaining effective performance compared to other state-of-the-art methods.

Conclusion

In this manuscript, a novel DREAM-IoMT methodology has been designed for identifying evolving anomalies in dynamic IoMT environments. A hybrid feature selection approach incorporating GI-RFE-SF models is deployed. For anomaly detection and classification, the proposed DREAM-IoMT model leveraged a weighted voting ensemble of GCN, BiTCN, and ICVAE. The ensemble classification model is further optimized utilizing AdamW. The model integrated a drift-aware online adaptation mechanism termed Extended ADWIN. The experimental result analysis proved improved performance over other approaches in terms of diverse measures.

Ethical Approval

Not Applicable.

Acknowledgment

None.

Funding

No external funding was received for this research.

Conflicts of Interest

The authors report there are no competing interests to declare.

References

[1] Li X, Xie C, Zhao Z, Wang C, Yu H. Anomaly detection algorithm of industrial internet of things data platform based on deep learning. IEEE Transactions on Green Communications and Networking. 2024;8(3):1037-48. [DOI: 10.1109/TGCN.2024.3403102]

[2] Zeng GQ, Yang YW, Lu KD, Geng GG, Weng J. Evolutionary adversarial autoencoder for unsupervised anomaly detection of industrial internet of things. IEEE Transactions on Reliability. 2025;74(3):3454-68. [DOI: 10.1109/TR.2025.3528256]

[3] Rodríguez M, Tobón DP, Múnera D. A framework for anomaly classification in Industrial Internet of Things systems. Internet of Things. 2025;29:101446. [DOI: 10.1016/j.iot.2024.101446]

[4] Chen Z, Li Z, Huang J, Liu S, Long H. An effective method for anomaly detection in industrial Internet of Things using XGBoost and LSTM. Sci Rep. 2024;14(1):23969. [DOI: 10.1038/s41598-024-74822-6]

[5] Abudurexiti Y, Han G, Zhang F, Liu L. An

explainable unsupervised anomaly detection framework for Industrial Internet of Things. Comput Secur. 2025;148:104130. [DOI: 10.1016/j.cose.2024.104130]

[6] Riaz R, Han G, Shaukat K, Khan NU, Zhu H, Wang L. A novel ensemble Wasserstein GAN framework for effective anomaly detection in industrial internet of things environments. Sci Rep. 2025;15(1):26786. [DOI: 10.1038/s41598-025-07533-1]

[7] Luo S, Zeng P, Ma C, Wei Y. Anomaly detection for industrial Internet of Things devices based on self-adaptive blockchain sharding and federated learning. Journal of Communications and Networks. 2025;27(2):92-102. [DOI: 10.23919/JCN.2025.000019]

[8] Wang R, Qiu H, Cheng X, Liu X. Anomaly detection with a container-based stream processing framework for industrial internet of things. Journal of Industrial Information Integration. 2023;35:100507. [DOI: 10.1016/j.jii.2023.100507]

[9] Tang L, Xue C, Zhao Y, Chen Q. Anomaly detection of service function chain based on distributed knowledge distillation framework in cloud-edge industrial internet of things scenarios. IEEE Internet Things J. 2023;11(6):10843-55. [DOI: 10.1109/JIOT.2023.3327795]

[10] De Vita F, Nocera G, Bruneo D, Das SK. A novel echo state network autoencoder for anomaly detection in industrial IoT systems. IEEE Transactions on Industrial Informatics. 2022;19(8):8985-94. [DOI: 10.1109/TII.2022.3224981]

[11] Prasad DV, Saraswathi S. The role of anomaly detection in industry 4.0: a survey of techniques and applications. Journal of Trends in Computer Science and Smart Technology. 2024;6(2):125-38. [DOI: 10.36548/jtcsst.2024.2.003]

[12] Bouke MA, Abdullah A, Frnda J, Cengiz K, Salah B. BukaGini: a stability-aware Gini index feature selection algorithm for robust model performance. IEEE Access. 2023;11:59386-96. [DOI: 10.1109/ACCESS.2023.3284975]

[13] Toor MS, Shahbaz H, Yasin M, Ali A, Fitriyani NL, Kim C, Syafrudin M. An optimized weighted-voting-based ensemble learning approach for fake news classification. Mathematics. 2025;13(3):449. [DOI: 10.3390/math13030449]

[14] Kingma DP, Ba J. Adam: a method for stochastic optimization. In: Proceedings of the 3rd International Conference on Learning Representations (ICLR); 2015 May 7-9; San Diego, CA. [DOI: 10.48550/arXiv.1412.6980]

[15] Hua J, Li C. Distributed robust Bayesian filtering for state estimation. IEEE Transactions on Signal and Information Processing over Networks. 2018;5(3):428-41. [DOI: 10.1109/TSIPN.2018.2889579]

[16] Annamalai U. WUSTL-IIOT-2021 dataset [dataset]. Kaggle. [Link]

[17] Dataset Engineer. Cybertec IIoT malware dataset (CIMD 2024) [dataset]. Kaggle. [Link]

[18] Zia S, Bibi N, Alhazmi S, Muhammad N, Alhazmi A. Enhanced anomaly detection in iot through transformer-based adversarial perturbations model. Electronics. 2025;14(6):1094. [DOI: 10.3390/electronics14061094]

[19] Zakariah M, Almazyad AS. Anomaly detection for IoT systems using active learning. Appl Sci. 2023;13(21):12029. [DOI: 10.3390/app132112029]

[20] Rhachi H, Balboul Y, Bouayad A. Enhanced anomaly detection in IoT networks using deep autoencoders with feature selection techniques. Sensors. 2025;25(10):3150. [DOI: 10.3390/s25103150]