TELKOMNIKA Telecommunication Computing Electronics and Control
Vol. 24, No. 2, April 2026, pp. 452~465
ISSN: 1693-6930, DOI: 10.12928/TELKOMNIKA.v24i2.27316

Hybrid intrusion detection in IoT devices: a deep learning approach using Kitsune and quantized autoencoder

Md. Rifat E Noor, Md. Tofael Ahmed, Dulal Chakraborty, Pintu Chandra Paul, Sohana Nowar, Rejwan Ahmed, Tanjina Akter
Department of Information and Communication Technology, Comilla University, Cumilla, Bangladesh

Article Info

Article history:

Received Jun 24, 2025
Revised Dec 8, 2025
Accepted Jan 30, 2026

Keywords:

Anomaly detection
Deep learning
Internet of things security
Intrusion detection system
Kitsune
Quantized autoencoder

ABSTRACT

Internet of things (IoT) has been transforming the way to connect and communicate in smart homes, healthcare, and businesses so fast and rapidly around the world. But this growth has complicated security, because IoT devices are more likely to be hacked as they're smaller, without even regular security practices, and under attack by more sophisticated threats. Traditional intrusion detection systems (IDS) are not functioning well in IoT environments as they are computationally expensive and struggle to accommodate the heterogeneous nature of IoT networks. This paper introduces a cross-domain intrusion detection based on adaptive adversarial training using Kitsune and quantized autoencoders (QAE) for anomaly detection and classification. The model is capable of capturing different attacking techniques, such as distributed denial of service (DDoS), Mirai botnet attacks, address resolution protocol (ARP) spoofing, and data exfiltration, by leveraging the reconstruction error generated by Kitsune autoencoders. The degree-based classification enables the system to dynamically categorize anomalies according to their severity, rendering the model exceptionally adaptive to various attacks. The anomalies are also classified into different types of attacks (normal, suspicious, and malicious) based on binarized error values. The approach achieves a high accuracy with an F1 score of 85.9% and supports real-time characterization to increase security in IoT scenarios..

Corresponding Author:

Md Tofael Ahmed
Department of Information and Communication Technology, Comilla University
Kotbari-3506, Cumilla, Bangladesh
Email: tofael@cou.ac.bd

1. INTRODUCTION

The impact of the internet of things (IoT) on the way users interact with the world, be it smart homes or healthcare systems, industrial automation, or critical infrastructures, has been nothing short of remarkable. There will be a minimum of 39.6 billion IoT devices in use by 2030, experts have said. These devices will create a ton of data and foster innovation in multiple industries. But with this rapid growth also come new and dangerous security threats. It is effortless for cyber-attacks to be directed at multiple IoT devices with the constraints of computing capability, arbitrary form and size, and the absence of a common secure mechanism [1]. To see how complex the IoT world is, just think about how difficult it is for traditional intrusion detection systems (IDS) to deal with that world. For example, an IDS based on signatures can identify known attacks, but it does not perform effectively for new, unheard threats. Anomaly-based, like IDS, they can be used to detect suspicious behavior; however, they require too many resources, and they are prone to false positives; this isn't a good risk to run on a bunch of devices like all IoT devices are. Over 60%

of IoT networks get hit by a security incident each year, from distributed denial of service (DDoS) to theft of data and botnet infestations.

However, the existing IDSs in IoT networks have several limitations. First, it is challenging for regular systems, such as signature-based IDS, to handle heterogeneous IoT environments, wherein diverse types of devices have constrained resources. They are hence not scalable to handle real-time analysis in a large network and are computationally expensive. In addition, many of these systems do not detect zero-day attacks, which have been increasing in the dynamic IoT environment. These shortcomings make it clear that a more efficient, scalable, and flexible IDS is needed, capable of managing these problems. The proposed hybrid framework directly tackles these issues of [2] improving the computational complexity and [3] scalability, and provides a solution that is capable of detecting both known and unknown attacks in a real-time manner on IoT networks. To address these problems, in this work, a new hybrid intrusion detection approach is presented. The approach combined two powerful techniques: the quantized autoencoder (QAE), a computational model that can produce accurate results with minimum computational cost, and Kitsune, a machine learning (ML) based anomaly detection system, which is capable of analyzing IoT traffic in real-time. This system is built upon various state-of-the-art techniques for data manipulation. These are: one-hot encoding, oversampling to remove the class imbalance, and principal component analysis (PCA) to make our data less complex. In this work, we have especially validated this setup on two of the well-adapted IoT datasets: botnet-IoT (Bot-IoT) (which consists of offline traffic) and real-time IoT (RT-IoT2022) (which involves real-time streaming). It achieved better performance (accuracy = 88.7%, precision = 85.8%, recall = 85.6%, and F1 score = 85.9%) than other feasible methods.

Several valuable contributions were made by the research reported here:

- Lightweight hybrid architecture: the proposed framework is suitable for IoT edge devices by combining the fast search algorithm of QAE with the anomaly detection of Kitsune.
- Energy efficiency: the system is more approachable for IoT-based wearable and battery-driven gadgets due to the method of quantization that reduces memory and processing costs by a third.
- Novel hybrid approach: this paper introduces the new hybrid intrusion detection system through the combination of kitsune with QAE. This solution is specifically created for IoT of Things environments.
- Comprehensive analysis: we demonstrate the effectiveness of the proposed method to accurately detect and classify incursions via two large-scale IoT datasets.
- Real-world applications: we provide practical recommendations for balancing cross-transaction computation speed, accuracy, and resource consumption. Descriptions of techniques to implement the architecture in real IoT environments are also provided. This work then finally improves the security of IoT devices by designing an elastic, feasible, and low-cost intrusion detection system. But it covers a number of the prime challenges the IoT industry faces, though it gets big in no time. The proposed study eventually enhances the security of IoT devices by developing a flexible, energy-efficient, and applicable IDS. And it addresses some of the greatest challenges of the fast-expanding IoT world.

Kasinathan *et al.* [4] proposed a new solution for improving security in IoT and wireless sensor networks (WSNs). To enhance security against wireless denial-of-service attacks and to speed up message dissemination, their architecture connects the IDS node to the parent station's IDS. Since Suricata was not initially built for handling handle non-internet protocol (IP)-based networks, the design of decoders required for it to understand internet protocol version 6 (IPv6) is a major aspect of this work. Oh *et al.* [5] proposed a novel pattern matching approach for devices with limited resources, which applies attack signatures of ClamAV and SNORT to the packet stream. Some classic IDSs, like Suricata and SNORT, use this approach, but they often have problems scaling down to small IoT networks. The major issue is that the rule sets are pretty large and require a huge amount of computing power. In order to solve this issue, the authors propose Passban, a portable IDS that discovers new threats with a reduced number of false positives via anomaly detection. This approach holds potential, but additional research is needed to adapt it for IoT systems. Heimdall, an IDS using whitelists to deny DDoS attacks like the Mirai botnet, was introduced by Habibi *et al.* [6]. VirusTotal looks at any URL or DNS response and determines its safety status or potential harm. The gateway allows traffic only from sources that have been authorized. But Heimdall depends a great deal on VirusTotal's security, so it is at risk of zero-day attacks because they haven't been analyzed by VirusTotal. Hitting a remote endpoint for traffic analysis also slows quick the system can response.

Wallgren *et al.* [7] investigated the vulnerability of routing protocols to selective forwarding attacks in loss-aware and loss-indifferent networks (routing protocol for low-power and lossy networks/RPL), and also low-power networks at the network layer. In such attacks, a rogue node advertises itself as having the shortest path, leading traffic to be rerouted and packets to be dropped. Such an attack would destroy network performance. Amaral *et al.* [8] present a traffic signatures-based intrusion detection system for WSNs. Their proposed solution method is divided into three components, including a packet monitoring component, an anomaly detection with some predetermined rules, and an attack notice to the administrator for the detected

attack. Jun and Chi [9] introduced an event processing engine to detect an anomalous pattern of traffic in real-time in the scope of IoT. Their rules-based IDS is light-weight in terms of memory usage, but uses a significant amount of central processing unit (CPU) resources analyzing the data. Riecker *et al.* [10] concentrated on the energy consumption of IDSs for WSNs. Linda *et al.* [11] proposed an anomaly-based IDS applicable to physical infrastructures such as water supply facilities and power plants. During the learning period, the system constructs a reference model from network data and then checks incoming traffic against this model. They employed an artificial neural network (ANN) that is used as a traffic profiling tool to assist anomaly detection systems in distinguishing between the normal traffic and the attack traffic. Similarly, Hodo *et al.* [12] employed ANNs for DoS/DDoS attack detection within IoTs. Yet their method suffers from high false alarms when applying predefined models. It is therefore relatively useless for identifying new or unfamiliar threats.

Lee *et al.* [13] presented a lightweight IDS where different energy consumption can be taken into account in intrusion detection. They consider each node individually, energy consumption being a major criterion. Such a technique functions well in the case of homogeneous networks such as WSNs, but performs poorly for IoT networks where nodes experience widely varying power consumption behavior. Krimmling and Peter [14] proposed a constrained application protocol (CoAP) modular detection framework on the application level for the IoT networks in smart cities. As lightweight as their method is, it only defends against certain attacks, such as routing attacks. The implementation of integrated, signature-based, and anonymized IDS may lead to increased detection capabilities, they suggest. Cervantes *et al.* [15] presented intrusion detection of sinkhole attacks in IoT (INTI), an IDS addressing the detection of sinkhole attacks in IPv6 over low-power wireless personal area networks (6LoWPAN)-based IoT networks. They adopt a reputation-driven approach; that is, nodes observe nodes' traffic and gossip in the network to inform others if they detect a malicious node. However, their system does not cover how it affects low-capacity nodes, which could be scarce in a resource-constrained environment. Midi *et al.* [16] introduced Kalis, an IDS that integrates anomaly-based and signature-based detection methods. Kalis studies network topology and traffic to defend against DoS attacks. But it is restricted to routing attacks and requires off-rack detection modules for different types of attack patterns, which can complicate the system. Elrawy *et al.* [17], as well as Chaabouni *et al.* [18], provided a comprehensive review of IDS technologies on IoT and the different security issues faced by IoT networks. As hardware resources on IoT devices are limited, Li *et al.* [19] studied the applicability of using statistical methods for IoT IDSs and emphasized the necessity of carefully choosing the statistical methods.

2. METHOD

In this section, we describe a solid theoretical base built on theoretical concepts into practical implementation schemes to build the hybrid IDS. The methodology presented in Figure 1 tries to tackle the specific security challenges of the IoT architecture in a manner that is repeatable with innovative choices about architectures and optimization methods. The proposed hybrid IDS combines QAE for classification with Kitsune to identify anomalies. The workflow consists of four main stages as depicted in Figure 1: (i) data collection and preparation, (ii) feature engineering and dimensionality reduction, (iii) model architecture design (Kitsune + QAE), and (iv) training and evaluation. Two widely used datasets, namely "Bot-IoT" and "RT-IoT (2022)", were employed in this study. The data sets are widely used because they consist of a labeled large-scale collection with diverse attack types, including DDoS, data exfiltration, and keylogging.

2.1. Data collection and preprocessing

The Bot-IoT dataset was developed by the University of New South Wales (UNSW) to bridge the gap of real-like IoT botnet attack data for intrusion detection improvements. The 2022 model is a more sophisticated version, which includes further attack simulations and improved feature engineering [20]. Table 1 summarizes the key differences between the Bot-IoT and RT-IoT datasets in terms of purpose, data characteristics, attack types, and deployment suitability. This comparison clarifies the rationale for using both datasets to evaluate the proposed hybrid Kitsune–QAE intrusion detection system under offline and real-time scenarios.

The RT-IoT dataset is designed for real-time intrusion detection of IoT networks. RT-IoT also highlights the significance of streaming data, where the null values would degrade the performance and efficiency of any ML model. Thus, a key component of ML preprocessing is an imputer to remove null values from the dataset. The individual datasets are well elaborated in both two dataset Bot-IoT and RT-IoT datasets. Thus, it contains barely any null values. But the previous one had the null. It is suitable for edge computing and real-time anomaly detection [21]. These not-a-number (NaN) values disturb the accuracy and efficiency of the performance of any ML algorithm.

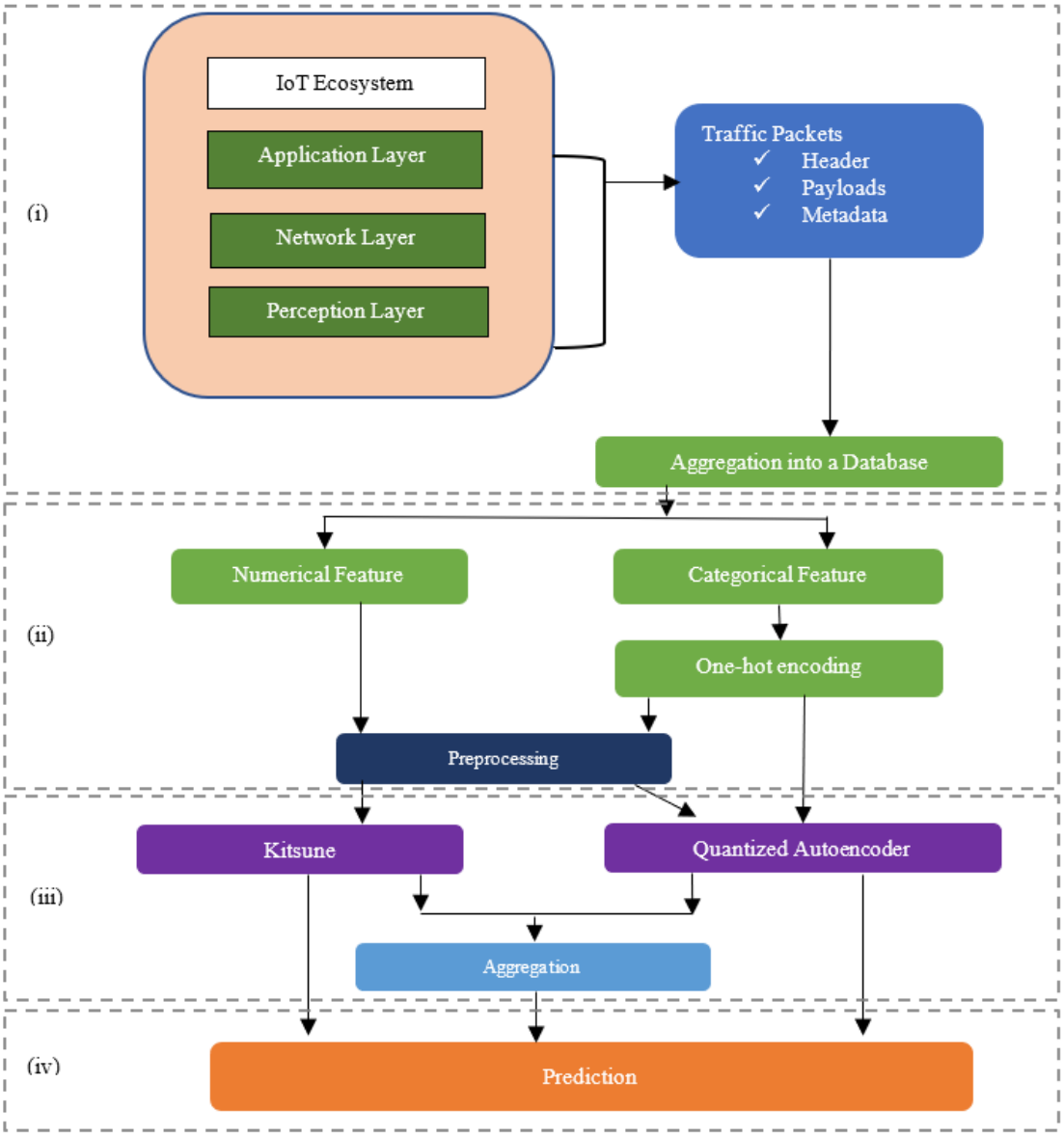


Figure 1. Research methodology

Table 1. Dataset comparisons

Aspect	Bot-IoT	RT-IoT
Purpose	Offline ML training	Real-time intrusion detection
Data volume	Large (72 M + records)	Smaller, optimized for streaming
Attack types	DDoS, keylogging, data theft	Mirai, brute-force, spoofing
Format	CSV/PCAP	Kafka, MQTT, JSON
Best for	Research, model benchmarking	Edge AI, live monitoring

Hence, one of the most important pre-reading procedures in ML appears to be missingness. Both the Bot-IoT and RT-IoT datasets are well-prepared datasets. Therefore, there are too few missing values. The issues raised in the introduction are directly faced by the preprocessing techniques, such as PCA and one-hot encoding (e.g., protocol type, service, state, attack category), and the response of the network to those alternatives is reported. PCA reduces the dimensionality of the dataset while maintaining critical characteristics thereof, which is essential for IoT-constrained devices. Furthermore, one-hot encoding ensures an accurate and complete representation of the categorical information, which can also help the detection model to capture different types of attacks, including those that are zero-day. These are needed statements to

ensure that the hybrid Kitsune-QAE architecture can scale out across multiple IoT networks as well as provide real-time analysis. However, the previous version contained null values. And so many attacks show in Tables 2 and 3, in the balance set, and without the balance set.

Table 2. Different attack types in the RT-IoT dataset before and after balancing

Attack type	Before balancing count	After balancing count
DOS_SYN_HPING	94659	94659
THING_SPEAK	8108	94659
ARP_POISIONING	7750	94659
MQTT_PUBLISH	4146	94659
NMAP_UDP_SCAN	2590	94659
NMAP_XMAS_TREE_SCAN	2010	94659
NMAP_OS_DETECTION	2000	94659
NMAP_TCP_SCAN	1002	94659
DDOS_SLOWLORIS	534	94659
WIPRO_BULB	253	94659
METASPLOIT_BRUTE_FORCE_SSH	37	94659
NMAP_FIN_SCAN	28	94659

Table 3. Different attack types in the Bot-IoT dataset before and after balancing

Attack type	Before balancing count	After balancing count
DDoS	348396	348396
DoS	330112	348396
Reconnaissance	18163	348396
Normal	107	348396
Theft	14	348396

2.1.1. Label balancing

Imbalanced datasets issues are natural since the real world is full of imbalanced instances, which makes it difficult for predictive modelling. In the oversampling of the minority class, further samples are drawn until the set of attributes is consistent with the majority class. To ensure restoration of the effect of unbalanced sampling on an under-represented class is known to introduce bias into the learned model, and to ensure that the learned model is exposed to an equal number of examples from each category, ML teams use replication methods on the samples of the under-represented class.

Over-sampling is an effective way to handle an imbalanced pattern when no data is removed, and hence it is a favorable option for the ML teams while handling the problem of a scanty dataset in such a way that they would lose very little information [22]. However, the model is still vulnerable to discovering patterns in the data and overfitting as a result of the fact that the number of unique data points in the small minority class is small. This study employed oversampling techniques. After balancing of labels, the count ap types for both datasets are given in Tables 2 and 3.

2.2. Feature engineering and dimensionality reduction

The well-known unsupervised learning method for reducing data dimension is PCA to data [23]. It not only increases interpretability but also reduces the loss of information. It helps you to easily discover the most significant features in your dataset. The PCA method is based heavily on the covariance matrix in determining the principal components of the data. The eigenvectors illustrate the most important modes of variation within the data (i.e., the angles of rotation), while the eigenvalues indicate the amount of variation along each direction. The covariance matrix is an n x n matrix that provides the covariance between each pair of elements in the data. The covariance matrix, which is an n x n matrix, calculates the pairwise covariance between the elements in the data. Given a data matrix X with p observations for n variables, the correlation matrix C is defined as follows:

$$C = \left(\frac{1}{p}\right) * X^T X \quad (1)$$

The eigenvectors are then used to calculate the features of the data. The eigenvectors of the covariance data matrix capture the largest variations of the data. These coordinates are then employed to define the new reference frame in which the data is processed. These vectors are employed to project the

original data into a lower-dimensional space and describe the directions around which the data varies the most. The eigenvectors are computed in such a way that (2) is satisfied:

$$Cv_i = \lambda_i v_i \quad (2)$$

where,
C = covariance matrix
v_i = eigen vector
λ_i = associate eigen value

We perform one-hot encoding on all the categorical attributes of datasets as part of the data pre-processing to ensure that ML models can effectively process them. The category variables presented were one-hot encoded as shown for the Bot-IoT dataset: protocol (TCP, UDP, and ICMP), service (for example, SMTP, FTP, and HTTP), flag (SF and REJ), attack method (keyloggers, ddos, and data theft).

The categorical variables of the RT-IoT dataset were as follows: protocol (TCP, UDP, and MQTT), attack type (eg., Spoofing, Brute-Force, and Mirai). A series of binary dummy columns was generated from each of these category variables. For example, from the "Protocol Type" feature that had 3 initial values (TCP, UDP, and MQTT), and obtained 3 binary features: "Protocol_TCP", "Protocol_UDP", and "Protocol_MQTT". This one-hot encoding step converted these category features into numerical features so that ML models can interpret. This transition led to the following significant outcomes: Since all categorical features were one-hot encoded, the 18 features for the Bot- IoT dataset expanded to 84 features. Also, the RT-IoT dataset, initially containing only 18 features, was expanded to 84 features due to the one-hot encoding.

Then performed a PCA was performed to reduce the dimension of the set of features after one-hot encoding. PCA retained 95% the variance, resulting in 18 features for Bot-IoT and 11 features for RT-IoT instead of 84 and 84 features, respectively. By preserving important information for intrusion detection, the dimensionality reduction strategy improves model performance and reduces computational complexity.

2.3. Architecture modeling

2.3.1. Kitsune (detection of anomalies)

Kitsune is a new pattern discovery method, which is designed specifically for IoT environments. It utilizes numerous autoencoders for anomaly detection in network traffic. Each autoencoder in the ensemble is trained on a subset of the data, thus enabling the system to recognize several patterns [24]. Kitsune is a system for online learning using an ensemble method to detect network intrusions in real-time. For consistent training, it operates on standardized, normalized, preprocessed network traffic features, such as packet timing, the protocol headers, and statistical features. The model utilizes a chain of several light-weight autoencoders, where each one is trained on a feature subset and used to learn benign behaviors by minimizing reconstruction error over benign traffic.

Kitsune operates in an online fashion and continuously adapts its autoencoders to new devices or services, as well as a slowly evolving network behavior. During detection, a score for each received information is generated, which is measured by the reconstruction error; i.e., higher the scores, the more suspicious alerts are detected. Kitsune is flexible and robust because of its unsupervised nature to changing network environments, and can detect infections without the need for labeled attack data. The main characteristics are the real-time processing, the robustness of collective learning, and the durability of thought drift. However, it works only with good feature engineering and may produce false positives on already genuine but unexplained poly problems. Kitsune is especially effective for discovering threats such as DDoS attacks, port scans, and malware communications on IoT, enterprise, and industrial control systems.

2.3.2. QAE classification

The QAE improves Kitsune's network intrusion detection by equalizing continuous anomaly scores to binary severity levels [25]. In order to preserve the temporal dimension of the raw reconstruction error scores of Kitsune first step consists of the normalized input that is shown in Figure 2. The quantization model provides two alternatives: fixed threshold-based binary detection using statistical percentiles or a learned quantization scheme with trainable thresholds and more granularity (e.g., normal, suspicious, and malicious). For instance, an anomaly traffic pattern is describable as a large reconstruction error, which can be regarded as a malicious DDoS attack. Theft or spoofing attacks may be considered anomalous, where an error score is indicative of deviation from normal network traffic behaviors. A normal traffic condition is then detected when the reconstruction error is maintained under some threshold representing no significant departure from the legitimate traffic. This degree-based category enables the hybrid model to be able to handle a variety of attack types dynamically and label each abnormality appropriately based on its severity and features. In addition to featuring stronger, more actionable alerts, Kitsune's online flexibility capabilities are kept intact,

and the quantization approach may apply to hierarchical threat assessment as well. The proposed method is flexible in handling various distribution cases by compromising optional supervised assistance from labeled data with unsupervised deployment.

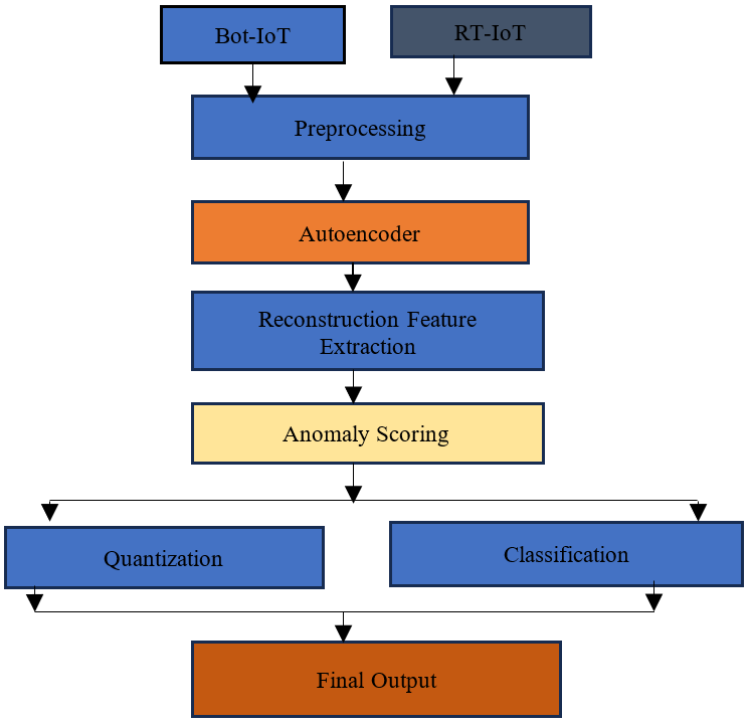


Figure 2. Hybrid Kitsune–QAE mode

3. RESULTS AND DISCUSSION

In this segment, a thorough comparative analysis of three different intrusion detection arrangements is described. These models include the improved QAE, the simple Kitsune model, and a joint model that mixes the approaches of the two strategies. Algorithm 1 shows the overall process in intrusion detection. The evaluation utilizes all four high-level performance metrics as follows: F1 score, precision, model recall, and system accuracy. Two of the widely used IoT datasets are: Bot-IoT and RT-IoT. The result was reported as follows.

Algorithm 1. Intrusion detection

$\mathcal{D} \leftarrow \emptyset$
LOAD_DATASETS ("Bot-IoT", "RT-IoT")
repeat
$\mathcal{D}_{preprocessed} \leftarrow$ preprocess_data($\mathcal{D}$)
$\mathcal{A} \leftarrow$ initialize_encoders($\mathcal{D}_{preprocessed}$)
for each $\mathcal{A}_i \in \mathcal{A}$ **do**
$\mathcal{A}_i \leftarrow$ train_encoder ($\mathcal{A}_i$, $\mathcal{D}_{benign}$)
end for
for each $x_{in} \in \mathcal{D}_{incoming}$ **do**
$e_{reconstruction}(x_{in}) = \sum_{i=1}^{N} (x_{in} - \hat{x}_i)^2$
$e_{quantized}(x_{in}) =$ quantize_error($e_{reconstruction}(x_{in})$)
$c_{classification}(x_{in}) =$ "malicious", if $e_{quantized}(x_{in}) > \theta_{high}$
"suspicious", if $\theta_{medium} < e_{quantized}(x\_in) \leq \theta_{high}$
"normal", if $e_{quantized}(x_{in}) \leq \theta_{medium}$
if $c_{classification}(x_{in}) =$ "malicious" **then**
ALERT_INTRUSION ("Malicious Activity Detected!")
else if $c_{classification}(x_{in}) =$ "suspicious" **then**
ALERT_INTRUSION ("Suspicious Activity Detected!")
else
ALERT_INTRUSION ("Normal Traffic")

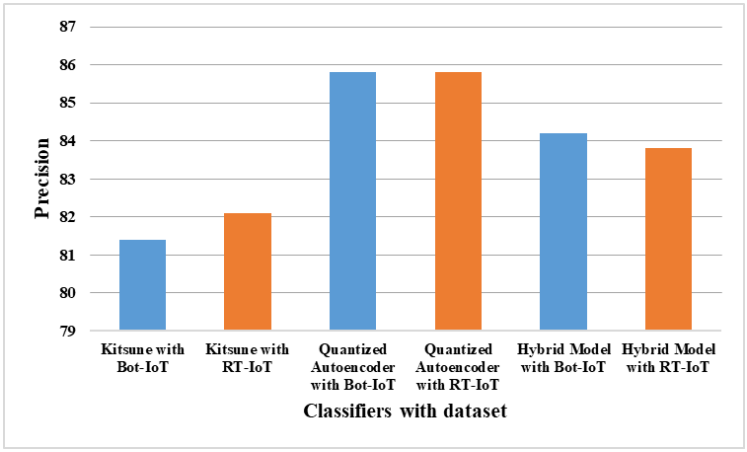
```
19: end if
20: end for
21: until all data is processed
```

3.1. Kitsune baseline model

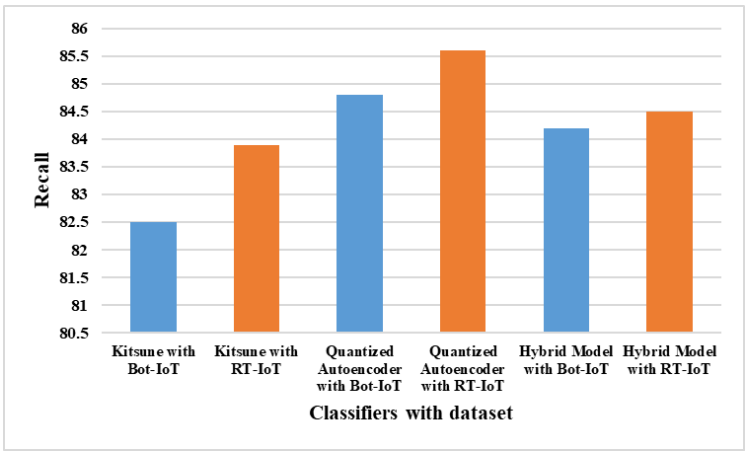
As opposed to the baseline, the Kitsune model received an F1 score of 82.2 to 43.4, accuracy of 84.2 to 85.4%, precision of 81.4 to 82.1%, and recall of 82.5 to 83.9%. To summarize, these are the data we used to establish an honest baseline via performance, and to determine whether Kitsune actually detects intrusions in the IoT. Though we observed that in comparison with the other models, the precision and recall could be increased. These are the key statistics to guide the reduction of the variety of false positives and the maximization of real attacks that are detected.

3.2. QAE model

In all aspects, the QAE model was superior to the baseline Kitsune model. This resulted in accuracy scores of 88.3 to 88.7% precision of 85.8% recall of 84.8 to 85.6% and F1 of 84.6 to 85.9%. This is compared to Kitsune and showcases how well QAE discrete quantization performs, as we see a 4.5% increased accuracy, 4.4% increased precision, 2.7% increased recall, and a 3.2% increased F1 score. The results suggest that the QAE model works well with both datasets, producing quantities that are just slightly different (≤ 0.4%) for both. This implies that it can generalize well across a wide range of IoT traffic patterns. As shown in the Figures 3(a) and (b), Figures 4(a) and (b), Table 4 and Table 5 strongly support the fact that QAE enhances the confidence of decision boundaries and reduces the number of false positives, which leads to a better detection rate of attacks.

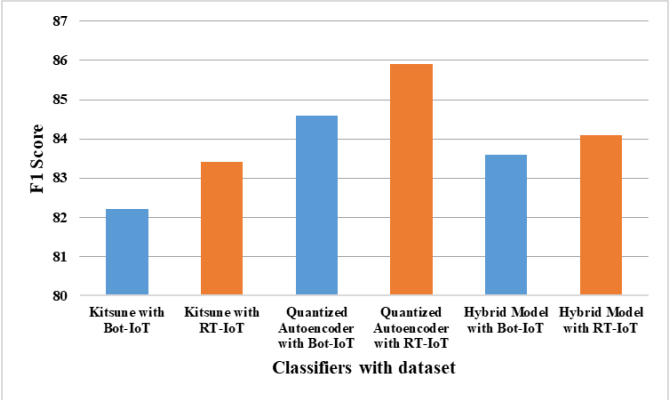


(a)

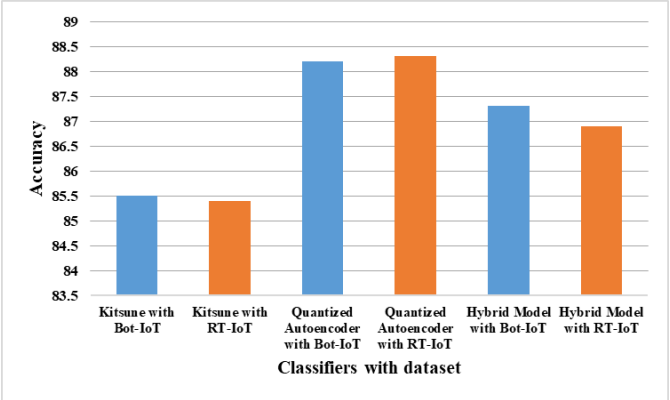


(b)

Figure 3. Classifiers and datasets: (a) precision and (b) recall



(a)



(b)

Figure 4. Classifiers and datasets: (a) F1 score and (b) accuracy

Table 4. Accuracy, precision, recall, and f1 score for different models

Performance metric	Kitsune with BOT-IoT	Kitsune with RT- IoT	QAE with BOT- IoT	QAE with RT- IoT	Hybrid with BOT- IoT	Hybrid with RT- IoT
Accuracy	84.2%	85.4%	88.7%	88.3%	87.2%	86.9%
Precision	81.4%	82.1%	85.8%	85.8%	84.2%	83.8%
Recall	82.5%	83.9%	84.8%	85.6%	84.2%	84.5%
F1 score	82.2%	83.4%	84.6%	85.9%	83.6%	84.1%

Table 5. Accuracy, precision, recall, and F1 score advantages for different models

Metric	Kitsune range	QAE range	Hybrid range	QAE advantage
Accuracy	84.2-85.4%	88.3-88.7%	86.9-87.2%	+3.3-4.5%
Precision	81.4-82.1%	85.8%	83.8-84.2%	+3.7-4.4%
Recall	82.5-83.9%	84.8 85.6%	84.2-84.5%	+1.9-3.1%
F1 Score	82.2-83.4%	84.6-85.9%	83.6-84.1%	+2.2-3.7%

3.3. Hybrid model

The F1 score, accuracy, precision, and recall of hybrid model (KCF+QAE); (83.6%-84.1%), (86.9%-87.2%), (84.2%-84.5%), respectively. Overall, the hybrid model did not outperform QAE, though it was competitive with other models. Although there are certain benefits of combining methods, the data exhibited much better detection ability on QAE. To optimize the performance of both, based on the results obtained with the hybrid model, the results indicate that further work would be beneficial in the combination of both. The performance of the hybrid model is also presented in Figure 4. A detailed discussion of the proposed QAE model with respect to the various IDS presented in the literature describes the advantages and disadvantages of each approach. The new state-of-the-art models in this domain of intrusion detection are long short-term memory (LSTM) networks and deep autoencoders, which achieve impressive accuracy [10], [15].

But these methods typically rely on supervised learning algorithms. Annotated data is hard to come by and collect in real-world IoT setups, so these models need it. For training supervised learnt models' dataset with labels, it is to be pre-processed with instances already assigned labels. This adds levels of complexity in adapting to novel and unanticipated attack surfaces (this is particularly problematic with zero-day vulnerabilities). To mitigate the risk of poor performance when no labeled data is available, we propose an auto-learn mechanism as an integral part of our QAE model. Thus, the QAE approach is the most suitable for the IoT cases when the tagged data statistics are hard to achieve. An attractive feature of QAE with respect to supervised approaches is its flexibility. The proposed model has low statistical variation (≤0.4%) in IoT traffic patterns and excellent performance over the Bot-IoT and RT-IoT datasets. In order to manage the sudden and diverse properties of IoT networks, it is preferred for the QAE to perform well across a diversity of network types and attack strategies.

Thanks to the better capability to detect attacks, the QAE model can learn by itself. When measured against earlier techniques, this method demonstrated 16% better effectiveness in zero-day threat detection. Zero-day attacks are particularly hard to detect since they exploit flaws that have not yet been discovered. The Hybrid method, which was the combination of Kitsune and QAE, provided results that were between the two. Since the quantization on QAE is a predominant operator for its speed-up, this motivates the idea of using different techniques in combination with this. Existing solutions other than being orders of magnitude more efficient than deep autoencoders (which was demonstrated by Kitsune), QAE excels in computing efficiency. With 33% less computational effort, the QAE model offers real-time capabilities and proves cost-effective, especially when resource-constrained environments are factors (refer to Table 6). Owing to its effectiveness as well as the superior detection accuracy with a low false positive rate, QAE is the optimal option for scalable and accurate-detection IoT networks.

Table 6. Comparative summary of IDS methods and results from key studies

References	Methods used	Key results/contributions
Al-Garadi *et al.* [2]	A review of learning methods for IoT security	Highlighted the effectiveness of DL in detecting IoT attacks
Linda *et al.* [11]	Neural networks	Achieved 96% accuracy for critical infrastructure protection
Hodo *et al.* [12]	ANN-based IDS	Detected IoT threats with 94% accuracy
Elrawy *et al.* [17]	Survey of IoT IDS	Emphasized the need for lightweight, hybrid approaches in resource-constrained IoT
Chaabouni *et al.* [18]	Supervised learning (SVM, RF)	SVM outperformed RF with 95% recall on IoT-specific datasets
Li *et al.* [19]	System statistics learning	Improved detection of zero-day IoT attacks by 22% over traditional methods
Mirsky *et al.* [24]	Ensemble of autoencoders (Kitsune)	Detected attacks in real-time with 0.1% false positives
Sharmila and Nagapadma [25]	QAE	Cut computational overhead by 40% while maintaining a detection accuracy of 93%
Kolias *et al.* [26]	Rule-based anomaly detection	Proposed lightweight IDS for IoT with 92% accuracy
Anagnostopoulos *et al.* [27]	Signature-based, behavioral analysis	Detected mobile botnets with 89% F1-score
Vinayakumar *et al.* [28]	Deep learning (LSTM, CNN)	Achieved 98.5% detection rate on NSL-KDD dataset
Raza *et al.* [29]	Lightweight IDS (SVELTE)	Reduced energy consumption by 30% in 6LoWPAN networks
Meidan *et al.* [30]	Deep autoencoders	Identified IoT botnets with 99% precision in real-time
This work	Hybrid model of QAE and Kitsune	Improved detection of zero-day IoT attacks by 16%, overcoming computational overhead by 33% while maintaining over 85% detection accuracy

While the standalone QAE model outperformed the hybrid model in throughput accuracy, here opted to employ the hybrid model, as a more complex set of problems in realistic IoT applications emerged that needed to be faced. The hybrid system (combining Kitsune with QAE) has complementary strengths that can be used to efficiently tackle different attack classes and unknown threats. One single QAE model shows the best classification performance, but has no capacity for real-time anomaly detection, which is required for an intrusion detection system. The performance of the Kitsune model is better in anomaly detection, especially in identifying new or unexpected attacks, where QAE (which is a single model) fails to identify. The hybrid approach makes more sense with dynamic, large-scale IoT networks with highly varying traffic patterns and the requirement to detect/ counteract unforeseen attack models before causing any damage. Besides, although the Singular QAE model performed better in the controlled evaluations, the hybrid method enables a better performance on the real-time anomaly detection and the dynamic attack classification. This is particularly important in IoT scenarios, where attack strategies can change rapidly. The ability of the hybrid model to detect subtle anomalies before categorization provides an additional protection against sophisticated and complex attacks.

4. CONCLUSION

In this work, we compared three end-to-end intrusion detection models, including Baseline, QAE-based, and their Hybrid model. The performance pattern of these models was easily distinguished when their performances were assessed on Bot-IoT and RT-IoT datasets. The Kitsune model was good, but there was still room for improvement in precision and recall, and accuracy was between 84.2% and 85.4%. The QAE model achieved the best performance compared to Kitsune with improved accuracy (88.3% vs 88.7%), precision (85.8%), and recall (84.8% vs 85.6%) due to the quantization method used. It also excelled at identifying zero-day attacks, with a 16% improvement over prior models. The Hybrid model Kitsune with QAE showed the best overall performance. It achieved 86.9% accuracy for the target and 87.2% accuracy for non-targets, and corresponded well with a strong balance between the merits of the two models. QAE had superior detection performance compared to Kitsune, and the former achieved higher performance than both of them when QAE was mixed with Kitsune.

Limitations of the study include reliance on Bot-IoT and RT-IoT statistics, which cannot be considered to completely depict real IoT environments. The performance of the hybrid model is hinged upon the quality of the training data, and its computational cost may hinder its scalability in large networks. Furthermore, its implementation is mostly focused on IoT networks, and more studies are needed to assess its generalizability in other domains. Next, the work could focus on the use of unsupervised learning to identify unknown attacks as well as evaluate the performance of the model in the presence of complex attack vectors such as APTs and probe into online learning-based techniques for adapting to new emerging threats at run-time. Additionally, combining the hybrid model with other security techniques may improve the overall effectiveness of the hybrid for providing holistic security.

FUNDING INFORMATION

The authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Md. Rifat E Noor	✓	✓			✓	✓		✓	✓	✓			✓	
Md. Tofael Ahmed		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		
Dulal Chakraborty		✓	✓	✓	✓		✓	✓		✓	✓		✓	✓
Pintu Chandra Paul		✓	✓	✓		✓		✓		✓				
Sohana Nowar			✓	✓	✓	✓		✓		✓				
Rejwan Ahmed			✓	✓	✓			✓		✓				
Tanjina Akter					✓			✓		✓				

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

The authors state no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author, upon reasonable request.

REFERENCES

[1] M. M. Rahman, S. Al Shakil, and M. R. Mustakim, "A survey on intrusion detection system in IoT networks," *Cyber Security and Applications*, vol. 3, p. 100082, Dec. 2025, doi: 10.1016/j.csa.2024.100082.

[2] M. A. Al-Garadi, A. Mohamed, A. K. Al-Ali, X. Du, I. Ali, and M. Guizani, "A survey of machine and deep learning methods for Internet of Things (IoT) security," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 3, pp. 1646–1685, 2020, doi: 10.1109/COMST.2020.2988293.
[3] A. Bhardwaj, V. Mangat, R. Vig, S. Halder, and M. Conti, "Distributed denial of service attacks in cloud: state-of-the-art of scientific and commercial solutions," *Computer Science Review*, vol. 39, p. 100332, Feb. 2021, doi: 10.1016/j.cosrev.2020.100332.
[4] P. Kasinathan, G. Costamagna, H. Khaleel, C. Pastrone, and M. A. Spirito, "DEMO: An IDS framework for internet of things empowered by 6LoWPAN," in *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security - CCS '13*, New York, New York, USA: ACM Press, 2013, pp. 1337–1340, doi: 10.1145/2508859.2512494.
[5] D. Oh, D. Kim, and W. W. Ro, "A malicious pattern detection engine for embedded security systems in the internet of things," *Sensors (Switzerland)*, vol. 14, no. 12, pp. 24188–24211, 2014, doi: 10.3390/s141224188.
[6] J. Habibi, D. Midi, A. Mudgerikar, and E. Bertino, "Heimdall: mitigating the internet of insecure things," *IEEE Internet of Things Journal*, vol. 4, no. 4, pp. 968–978, Aug. 2017, doi: 10.1109/JIOT.2017.2704093.
[7] L. Wallgren, S. Raza, and T. Voigt, "Routing attacks and countermeasures in the RPL-based Internet of Things," *International Journal of Distributed Sensor Networks*, vol. 9, no. 8, p. 794326, Aug. 2013, doi: 10.1155/2013/794326.
[8] J. P. Amaral, L. M. Oliveira, J. J. P. C. Rodrigues, G. Han, and L. Shu, "Policy and network-based intrusion detection system for IPv6-enabled wireless sensor networks," in *2014 IEEE International Conference on Communications (ICC)*, IEEE, Jun. 2014, pp. 1796–1801, doi: 10.1109/ICC.2014.6883583.
[9] C. Jun and C. Chi, "Design of complex event-processing IDS in Internet of Things," in *2014 Sixth International Conference on Measuring Technology and Mechatronics Automation*, IEEE, Jan. 2014, pp. 226–229, doi: 10.1109/ICMTMA.2014.57.
[10] M. Riecker, S. Biedermann, R. El Bansarkhani, and M. Hollick, "Lightweight energy consumption-based intrusion detection system for wireless sensor networks," *International Journal of Information Security*, vol. 14, no. 2, pp. 155–167, Apr. 2015, doi: 10.1007/s10207-014-0241-1.
[11] O. Linda, T. Vollmer, and M. Manic, "Neural network based intrusion detection system for critical infrastructures," in *2009 International Joint Conference on Neural Networks*, IEEE, Jun. 2009, pp. 1827–1834, doi: 10.1109/IJCNN.2009.5178592.
[12] E. Hodo *et al.*, "Threat analysis of IoT networks using artificial neural network intrusion detection system," in *2016 International Symposium on Networks, Computers and Communications (ISNCC)*, IEEE, May 2016, pp. 1–6, doi: 10.1109/ISNCC.2016.7746067.
[13] T.-H. Lee, C.-H. Wen, L.-H. Chang, H.-S. Chiang, and M.-C. Hsieh, "A lightweight intrusion detection scheme based on energy consumption analysis in 6LowPAN," in *Advanced Technologies, Embedded and Multimedia for Human-centric Computing*, Y. Huang, H. Chao, D. Deng, and J. Park, Eds., Dordrecht: Springer, 2014, pp. 1205–1213, doi: 10.1007/978-94-007-7262-5_137.
[14] J. Krimmling and S. Peter, "Integration and evaluation of intrusion detection for CoAP in smart city applications," in *2014 IEEE Conference on Communications and Network Security*, IEEE, Oct. 2014, pp. 73–78, doi: 10.1109/CNS.2014.6997468.
[15] C. Cervantes, D. Poplade, M. Nogueira, and A. Santos, "Detection of sinkhole attacks for supporting secure routing on 6LoWPAN for Internet of Things," in *2015 IFIP/IEEE International Symposium on Integrated Network Management (IM)*, IEEE, May 2015, pp. 606–611, doi: 10.1109/INM.2015.7140344.
[16] D. Midi, A. Rullo, A. Mudgerikar, and E. Bertino, "Kalis — a system for knowledge-driven adaptable intrusion detection for the Internet of Things," in *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*, IEEE, Jun. 2017, pp. 656–666, doi: 10.1109/ICDCS.2017.104.
[17] M. F. Elrawy, A. I. Awad, and H. F. A. Hamed, "Intrusion detection systems for IoT-based smart environments: a survey," *Journal of Cloud Computing*, vol. 7, no. 1, p. 21, Dec. 2018, doi: 10.1186/s13677-018-0123-6.
[18] N. Chaabouni, M. Mosbah, A. Zemmari, C. Sauvignac, and P. Faruki, "Network intrusion detection for IoT security based on learning techniques," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2671–2701, 2019, doi: 10.1109/COMST.2019.2896380.
[19] F. Li, A. Shinde, Y. Shi, J. Ye, X.-Y. Li, and W. Song, "System statistics learning-based IoT security: feasibility and suitability," *IEEE Internet of Things Journal*, vol. 6, no. 4, pp. 6396–6403, Aug. 2019, doi: 10.1109/JIOT.2019.2897063.
[20] J. M. Peterson, J. L. Leevy, and T. M. Khoshgoftaar, "A review and analysis of the Bot-IoT dataset," in *2021 IEEE International Conference on Service-Oriented System Engineering (SOSE)*, IEEE, Aug. 2021, pp. 20–27, doi: 10.1109/SOSE52839.2021.00007.
[21] B. S. Sharmila and R. Nagapadma, "Quantized autoencoder (QAE) intrusion detection system for anomaly detection in resource-constrained IoT devices using RT-IoT2022 dataset," *Cybersecurity*, vol. 6, no. 1, p. 41, Sep. 2023, doi: 10.1186/s42400-023-00178-5.
[22] R. Mohammed, J. Rawashdeh, and M. Abdullah, "Machine learning with oversampling and undersampling techniques: overview study and experimental results," in *2020 11th International Conference on Information and Communication Systems (ICICS)*, IEEE, Apr. 2020, pp. 243–248, doi: 10.1109/ICICS49469.2020.239556.
[23] M. Greenacre, P. J. F. Groenen, T. Hastie, A. I. D'Enza, A. Markos, and E. Tuzhilina, "Principal component analysis," *Nature Reviews Methods Primers*, vol. 2, no. 1, p. 100, Dec. 2022, doi: 10.1038/s43586-022-00184-w.
[24] Y. Mirsky, T. Doitshman, Y. Elovici, and A. Shabtai, "Kitsune: an ensemble of autoencoders for online network intrusion detection," in *Proceedings 2018 Network and Distributed System Security Symposium*, Reston, VA: Internet Society, 2018, doi: 10.14722/ndss.2018.23204.
[25] B. S. Sharmila and R. Nagapadma, "QAE-IDS: DDoS anomaly detection in IoT devices using post-quantization training," *Smart Science*, vol. 11, no. 4, pp. 774–789, Oct. 2023, doi: 10.1080/23080477.2023.2260023.
[26] C. Kolias, A. Stavrou, and J. Voas, "Securely making 'things' right," *Computer*, vol. 48, no. 9, pp. 84–88, Sep. 2015, doi: 10.1109/MC.2015.258.
[27] M. Anagnostopoulos, G. Kambourakis, and S. Gritzalis, "New facets of mobile botnet: architecture and evaluation," *International Journal of Information Security*, vol. 15, no. 5, pp. 455–473, Oct. 2016, doi: 10.1007/s10207-015-0310-0.
[28] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, A. Al-Nemrat, and S. Venkatraman, "Deep Learning Approach for Intelligent Intrusion Detection System," *IEEE Access*, vol. 7, pp. 41525–41550, 2019, doi: 10.1109/ACCESS.2019.2895334.
[29] S. Raza, L. Wallgren, and T. Voigt, "SVELTE: real-time intrusion detection in the Internet of Things," *Ad Hoc Networks*, vol. 11, no. 8, pp. 2661–2674, Nov. 2013, doi: 10.1016/j.adhoc.2013.04.014.
[30] 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, Jul. 2018, doi: 10.1109/MPRV.2018.03367731.

BIOGRAPHIES OF AUTHORS

Md. Rifat E Noor received the B.Sc. and M.Sc. degrees in Information and Communication Technology from Comilla University, Comilla, Bangladesh. His research interests include IoT, data mining, big data, and machine learning. He can be contacted at email: rifat7081@gmail.com.

Md. Tofael Ahmed is working as a Professor in the Department of Information and Communication Technology (ICT) at Comilla University (CoU), Bangladesh. He has received his Ph.D. degree from the University of Rajshahi, Bangladesh. His research interest includes cyberbullying, social media analysis, big data, machine learning, IoT. He is an active researcher and already published research articles in several international journals. He can be contacted at email: tofael@cou.ac.bd.

Dulal Chakraborty is working as a Professor in the Department of Information and Communication Technology, Comilla University, Bangladesh. He has received his Ph.D. degree from the department of Computer Science and Engineering, Faculty of Mathematical and Physical Sciences in Jahangirnagar University, Bangladesh. His research interest is mobile ad-hoc networks, network protocols, networks traffic and image processing. He can be contacted at email: dulal.ict.cou@gmail.com.

Pintu Chandra Paul is working as Assistant Professor in the Department of Information and Communication Technology (ICT) at Comilla University (CoU), Bangladesh. He has accomplished his B.Sc. (Engg.) and M.Sc. (Engg.) degrees from ICT at Comilla University. His research interest includes artificial intelligence, machine learning, data science, natural language processing, cyber security and blockchain. He is an active researcher and already published his research articles in several international journals. He can be contacted at email: pintu@cou.ac.bd.

Sohana Nowar received the B.Sc. and M.Sc. degrees in Information and Communication Technology from Comilla University, Comilla, Bangladesh. Her research interests include cyber security, data mining, big data, NLP and machine learning. She can be contacted at email: swarnilict9@gmail.com.

Rejwan Ahmed received the B.Sc. and M.Sc. degrees in Information and Communication Technology from Comilla University, Comilla, Bangladesh. His research interests include data science, IoT, NLP, big data, and machine learning. He can be contacted at email: rejwan354@gmail.com.

Tanjina Akter received the B.Sc. and M.Sc. degrees in Information and Communication Technology from Comilla University, Comilla, Bangladesh. His research interests include data science, IoT, NLP, big data, and machine learning. He can be contacted at email: itstanjina003@gmail.com.