Contents lists available at IdNNS
Indonesian Journal of Machine Learning and Intelligent Systems
Volume 1 ; Issue 1
journal homepage: https://journal.idnns.org/index.php/ijmlis

# TrustEdgeAI: A Lightweight, Calibrated, and Secure Deep Learning Framework for Intrusion Detection in Edge-IoT Environments

S Karthika[a], K Juliana Gnanaselvi[a], N S Sivakumar[b,*], M Gokilavani[c] and Hasan Koten[d]

[a]Department of Computer Science, Rathinam College of Arts and Science, Coimbatore, India
[b]Center for Data Science and Sustainable Technologies, Inti International University, Nilai, Negeri Sembilan 71800, Malaysia
[c]Guru Nanak Institutions Technical Campus Ibrahimpatnam, Hyderabad, India
[d]Department of Mechanical Engineering, Istanbul Medeniyet University, Istanbul, Turkey

ARTICLE INFO

*Keywords*:
edge-iot
intrusion detection system (ids)
deep learning
cnn-lstm
adaptive feature pruning
model calibration
temperature scaling
cybersecurity
lightweight models
real-time detection

ABSTRACT

The fast development of Edge-IoT systems has posed serious security threats because of its distributed and resource-limited characteristics. The conventional intrusion detection systems are normally ineffective in offering real time and accurate security against the changing cyber threats. To overcome this challenge, this paper presents TrustEdgeAI, a lean and safe deep learning system, which is specialized in effective intrusion detection in Edge-IoT systems. The proposed model will combine a hybrid CNN-LSTM to learn spatial and temporal information of network traffic as well as an adaptive pruning mechanism of features to achieve a reduction in the computational complexity. Moreover, a calibration method based on temperature scaling is also included in order to enhance the credibility of prediction probability in order to make reliable decisions. The experiment outcomes prove that TrustEdgeAI has a high detection rate of 97.8 and much less inference latency of 12 ms and resource use. The framework also has low rates of false positive on various types of attacks and can be very applicable in real time and secure edges deployment.

## 1. Introduction

The speed of the Internet of Things (IoT) expansion has revolutionized contemporary computing through the ability to connect billions of devices across different fields of use, including healthcare, smart cities, industrial automation, and smart transportation. However, under Edge-IoT environments, computing and processing is not done over a centralized cloud infrastructure but at the network edge, which lowers latency, saves bandwidth and makes real-time decisions possible. Nonetheless, this decentralized and resource-constrained design poses a considerable risk of IoT systems to many types of cyber attacks, such as denial-of-service (DoS) attacks, data manipulation, spoofing, and advanced persistent threats. These difficulties demand the work on creating effective, smart, and safe intrusion detection systems that could work within stringent computational and energy limitations.

Conventional Intrusion Detection Systems (IDS) such signature-based and rule-based systems are not suitable in the contemporary IoT context as they fail to identify unknown or dynamic attacks, as well as depending on patterns. On the other hand, deep learning methods have shown high-performance in detection of dynamic and complex attack patterns, through the automatic learning of feature representation in data. Specifically, CNNs are useful in learning spatial relationships of network traffic data whereas LSTM networks are good at learning temporal relationships. Architectures with CNN-LSTM hybrids have been shown to be promising with regards to intrusion detection, but are usually characterized by high complexity which renders them inappropriate in edge devices. In addition, the current models often have overconfident forecasts, and thus untrustworthy decision-making in vital security issues is observed.

The other major drawback of the existing intrusion detection frameworks is that there is no integration between the dependability of a model and the security mechanisms. Although a number of studies have been conducted to enhance the accuracy of the detection, most of them have ignored the fact that the accuracy of prediction is as well as the security of data. The use of overconfident models may give false probability estimates and this may result into high chances of unnoticed attacks. Moreover, most systems are prone to breaches of data and manipulation as it is not encrypted prior to processing. The challenges have created

*Corresponding author
siva.kumar@nusaputra.ac.id (S. Karthika); nssivakumarpsg@gmail.com (N.S. Sivakumar); mailme.gokilavani@gmail.com (M. Gokilavani); hasan.koten@medeniyet.edu.tr (H. Koten)
ORCID(s):

the necessity of having a single framework, which is not only supposed to guarantee high detection performance, but also greater trustworthiness, efficiency, and data security.

To solve these shortcomings, it is the proposal of TrustEdgeAI, a new lightweight, calibrated, and secure deep learning model that can be used to detect intrusion in Edge-IoT networks. The originality of the suggested study is its combined method, i.e. a hybrid CNN-LSTM network with adaptive feature pruning that should decrease the number of computations but preserve high accuracy. Moreover, the framework has a temperature scaling-based calibration mechanism, which enhances reliability of model predictions through reducing overconfidence and generating well-calibrated probability outputs. Moreover, a security-conscious preprocessor module is proposed, which incorporates the lightweight cryptographic principles that should support the integrity and confidentiality of the IoT data throughout the processing. TrustEdgeAI is an implementation of deep learning, model optimization, calibration, and security combined into one unified system that operates in resource constrained edge environments, unlike other existing systems that address these aspects separately.

The dynamic development of artificial intelligence (AI), machine learning, and data-driven technologies have been of great impact to different fields, such as healthcare, smart infrastructure, cybersecurity, and sustainable systems. Over the past few years, intelligent models have been more commonly utilized to manipulate large volumes of complex data and extract valuable information to create more adaptive and efficient systems. Machine learning-based traffic classification and digital twin-based cyber-physical systems are some of the applications that have shown that AI can model dynamic environments and optimize real-time decision-making processes. Simultaneously, the incorporation of secure computation methods, such as homomorphic encryption and cryptographic network modelling, has solidified the base of safe data processing, especially in one of the most sensitive areas such as healthcare. The increasing interest in smart healthcare systems has seen the emergence of real-time monitoring, wearable, and predictive models to aid in unceasing patient monitoring and the automatic detection of diseases. Developed deep learning methods have made successful predictions of chronic illnesses like diabetes, and AI diagnostic systems and models of personal medicine have enhanced the precision and fine-tuning of therapy. Additionally, medical imaging innovations such as sustainable AI systems and noise mitigation methods of MRI, and vision transformer-based disease detection have greatly improved the diagnostic and clinical decision-making systems.

In addition to healthcare, AI has been applied to various other fields such as emotion recognition based on multimodal neuroimaging, malware based on optimization algorithms, intrusion detection in smart grids, and voltage control in systems with electric vehicles. The applications demonstrate the scalability and flexibility of intelligent systems in a wide range of environments. Also, new developments in the field of machine learning, predictive maintenance, federated learning, and AI-based solutions to agriculture, tourism, and renewable energy systems demonstrate the growing range of intelligent technologies to real-life issues. The use of geospatial AIoT systems and microgrid-based solutions also help to highlight the importance of AI in ensuring sustainability and optimization of resources.

In spite of these developments, various issues still persist especially when it comes to providing data privacy, system security, interpretability and ethical responsibility. The greater the autonomy and the data intensity of the AI systems, the more acute will be the demand of the transparent, reliable, and fair models. In this respect, superior AI paradigms, such as neuro-symbolic learning and intelligent federated systems, are also being considered to improve performance and reliability. Moreover, the integration of multimodal sources of data such as clinical records and imaging data to sensor-based inputs has presented the opportunity of more holistic and context-sensitive decision-making systems.

The works of this work have many contributions. To begin with, a new TrustEdgeAI architecture is expected to provide effective and trustworthy intrusion detection within Edge-IoT. Second, a hybrid CNN-LSTM model is established that is useful to describe space and time features of network traffic data. Third, a dynamic feature pruning system is proposed which prunes features that are less critical so that the complexity of the model can be decreased and the computational efficiency of the model is enhanced. Fourth, there is an integration of a calibration method, which is temperature-scaling method, to increase the confidence and trustworthiness of the predictions. Fifth, security conscious preprocessing plan is applied under which sensitive information is secured by lightweight cryptographic techniques. Lastly, the suggested system also balances accuracy, latency, and security, which is why it is applicable in real-time resources-constrained environments. TrustEdgeAI, on the whole, is a major step towards the creation of robust, efficient and trustworthy intrusion detection systems of next-generation IoT networks.

The rest of this paper is structured in the following way. Section 2 includes the thorough review of the existing literature associated with the intrusion detection, deep learning models, and security mechanisms. Section 3 outlines the proposed TrustEdgeAI approach, the architecture of the system, the extraction of features, and the techniques of the calibration. Section 4 is devoted to the experimental arrangement, performance assessment, and comparative study of the new methods with the existing ones. Lastly, Section 5 is the concluding part of the paper and identifies possible future research directions.

## 2. Related Work

The creation of safe and effective intrusion detection systems has been highly researched in many areas of access

control, cryptography, data mining, and deep learning. The available literature indicates major progress in role-based access systems, privacy protection systems, and smart intrusion detection systems, which are the basis of the proposed TrustEdgeAI framework.

Initial studies on access control and role mining formed the background of the secure data management in the complex systems. Shuriya and Thenmozhi (2015) proposed a role-based access model with the method of constraint satisfaction to enhance the role mining efficiency. On the same note, Shuriya and Graduate (2015) improved on the RBAM frameworks further, and Shuriya (2013) provided a general study on the RBAC systems. Shuriya and Rajendran (2017b) suggested a tranquilizing role mining heuristic random approach and Shuriya (2019) created a fuzzy responsibility-based access organizer to ensure sensitive medical records. Moreover, a proxy re-signature-based approach to privacy protection was introduced by Shuriya (2015), and Shuriya and Sivaprakash (2023) introduced a cryptographic AES-based privacy preservation QoS model, which was further expanded by Shuriya et al. (2022) to the secure communication systems.

Intelligent models have been used in multiple fields as AI and machine learning continue to evolve. Babu et al. (2022) presented a prognostics and health management framework of IoT on aircraft engines that has predictive maintenance features. An IoT-powered fluid transportation monitoring system was developed by Bhaskaran et al. (2021), whereas Priyanka et al. (2022) investigated the digital twin as the technology to estimate the risks of an oil pipeline. These papers emphasize the need to incorporate IoT, machine learning, and digital twins to monitor the real-time and provide predictive analytics. Shuriya et al. (2023a) and Abinaya et al. (2024) provided further improvements in industrial applications, as they showed transformative potential of digital twins and cyber-physical systems towards smart manufacturing environment.

Some of the studies in the area of cybersecurity have aimed at enhancing intrusion detection and secure communication. Karthikeyan et al. (2023a) recommend an intrusion detector system based on auto cryptographic denoising, whereas Shuriya et al. (2023b) suggest a one-dimensional dilated learning approach to the resource-constrained environment. CNN-LSTM fusion A lightweight CNN-LSTM fusion network on edge-IoT real-time intrusion detection is developed by Sivaprakash et al. (2025c). There was also a proposal of an ideal malware detection method by Mani et al. (2025) and a review of intrusion detection systems in smart grids by Shuriya et al. (2025c). These articles have highlighted the importance of smart, dynamic, and effective security solutions in the contemporary online systems.

The healthcare uses of AI have attracted much interest because of their potential to enhance treatment and diagnosis of patients. Shuriya and Rajendran (2017a) implemented the data mining and ANFIS methods to cardiovascular disease diagnosis. Mouna et al. (2024) examined the data modeling in the Internet of Medical Things (IoMT), which allows analyzing healthcare data effectively. Umamaheswari et al. (2025b) and Umamaheswari et al. (2025a) designed a real-time monitoring system and wearable health tracking and battlefield awareness. Shuriya et al. (2025a) suggested an adaptive deep learning model of diabetes prediction, and Ezhil Vendhan et al. (2025) emphasized that AI can be used in predictive diagnostics and precision medicine. Besides, Sivaprakash et al. (2026) proposed the vision transformer-based model to address the issue of diabetic retinopathy and enhance the accuracy and reliability of the diagnostic process. On the same note, a deep learning-based brain tissue segmentation model was designed by Sivaprakash et al. (2025a), which is used to improve the analysis of medical images.

The use of AI has also brought significant contributions to medical imaging and signal processing. A noise-resilient homomorphic encryption model with the purpose of secure processing of healthcare data was suggested by Shuriya et al. (2024a), whereas Shuriya et al. (2024b) achieved better network security with the help of optimization algorithms. The article by Ranjith et al. (2025), discussed emotion recognition with the help of multimodal neuroimaging methods, which prove the unity of AI and neuroscience. These strategies emphasize the need to integrate data security, signal processing and intelligent analytics in healthcare applications.

Smart systems and infrastructure are also widely used in AI. Karthikeyan et al. (2023b) suggested a quantum model based on machine learning to classify traffic in a smart city. Yuvaraju et al. (2025) discussed the concept of voltage regulation in EV-related integrated smart grids, and Shuriya (2022) discussed the development in agricultural robotics. Also, Sivaprakash et al. (2025b) modeled a patent classification deep neural network-based service robots, which demonstrates the application of AI in automation and robotics. The research papers confirm the universal nature of AI usage.

In intelligent systems, recent developments in machine learning and the interdisciplinary uses have further increased the potentials of intelligent systems. Balajishanmugam et al. (2025) touched upon the development of machine learning methods, and Shuriya et al. (2025b) used graph convolutional networks to calculate material properties. Francis et al. (2025) and Ezhil Vendhan et al. (2025) gave insights on AI in orthopedics treatment and cancer treatment respectively. Moreover, Shuriya (2026), Shuriya (2026),

## 3. Proposed Methodology

The presented framework, TrustEdgeAI, will be aimed at offering a lightweight, dependable, and safe intrusion detection system that is adapted to the conditions of Edge-IoT, where the amount of computational resources is restricted, and real-time decision-making is paramount in Fig.1. The design approach combines the deep learning algorithm with model optimization and calibration in such a way that the

performance is efficient without affecting the detection accuracy. The system has the overall workflow comprising of various steps, which include data acquisition, preprocessing, feature extraction, temporal learning, adaptive feature pruning, calibration and secure classification.

At the first level, the information on network traffic of IoT devices is received at the edge layer. IoT data can be very heterogeneous, and it may be noisy, redundant, and missing values; hence, a powerful preprocessing plan is necessary. Data cleaning is done by the preprocessing module, which eliminates duplicated records and the data that has blank records with imputation methods. Numerical characteristics are brought to a normalized scale with a Min-Max scaling method to obtain equal statistics of data distribution, whereas categorical variables are transformed into numerical data with the help of the encoding technique (one-hot encoding). Also, it includes a lightweight encryption-aware preprocessing system to guarantee the protection of sensitive data at intermediate processing levels, which would increase the overall system protection.

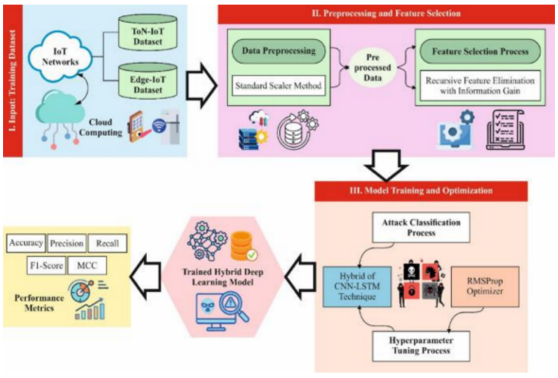


Figure 1: Architecture Diagram of IoT Cyber-Attack Detection Using a Hybrid CNN-LSTM Deep Learning Model

After preprocessing, the pre-cleaned and normalized data is inputted into the feature extraction module which is implemented using a Convolutional Neural Network (CNN). The CNN captures the spatial patterns and the correlations that exist in features of network traffic. Learning Multiple convolutional layers are applied to learn hierarchical codes of the input data with each layer learning more abstract features. The activation functions used are the Rectified Linear Unit (ReLU) to create non-linearity and max-pooling layers to decrease the dimensions and trim down features that are most significant. This step can help the model to discriminate successfully distinguishing normal and malicious traffic.

After extracting spatial features, the features are sent to the Long Short-Term Memory (LSTM) that do the temporal analysis. In multi-stage cyberattacks, IoT network traffic can tend to be time and sequence dependent particularly. LSTM network is quite appropriate to learn these temporal dependencies because it has the ability to retain long-term memory, via gated processes, such as input, forget, and output gates. The LSTM layer boosts the detection accuracy as it analyses the sequences of network traffic data to improve the ability of the model to detect complex attack patterns that change with time and therefore, the model is better placed to detect more attack patterns.

In order to make the model efficient and applicable to the deployment in the devices with limited resources (edges), an adaptive feature pruning mechanism is introduced. This element actively learns to recognize and drop minor features on the training process depending on their role in the output of the model. The model decreases computation complexity, memory and inference time by removing redundant or low impact features. Adaptive pruning unlike the other methods of feature selection that are static continuously calculates the importance of features and provides the feature with an opportunity to remain optimized even when distributions of data shift.

One of the most important novelties of the TrustEdgeAI framework is the calibration model of models, which helps to solve the problem of overconfidence that is typical of deep learning models. In intrusion detection systems the problem of overconfidence may result in the catastrophic security failure particularly when the model labels malicious traffic as innocuous. To curb this challenge, temperature scaling method is used on the model output logits. This is a post-processing step which modulates probability distribution of prediction without changing the classification accuracy. Subsequently, the actual likelihood of each of the classes is represented more accurately by the predicted probabilities, which enhance the credibility and comprehensibility of the system.

The framework also incorporates a secure classification layer in addition to the calibration step and hence the final predictions are protected and accurate. This layer in use integrates lightweight cryptographic concepts in the process of processing and transfer of data to block access and manipulation. Security and intelligence together render the system strong in the face of adversarial attacks and breach of data, which are an essential area of concern in the IoT.

The last categorization is determined with the help of fully connected dense layer and softmax activation function that assigns the probability of different classes that include normal traffic and types of attacks. These probabilities are made reliable by the calibrated output that is applicable in real-time applications in decision-making. It is built to work with low latency so it is extremely useful in the edge intrusion detection.

In general, the TrustEdgeAI methodology proposed results in an accuracy, efficiency, and reliability balance. The combination of CNN and LSTM is the key to the full learning of the features, whereas adaptive pruning is the key to the lightweight functioning. The calibration system increases the predictability, and the security system ensures data integrity. This is because TrustEdgeAI is a high-performing and scalable solution to next-generation edge-IoT intrusion detection systems through this holistic outlook.

Intelligent Framework: TrustEdgeAI Framework.

Input: Data of network traffic DDD.

Output: Prediction of intrusion detection (Normal / Attack)

Step 1: Data clean: Clean the data by eliminating all the missing and redundant values and normalize the features with Min-Max normalization:

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{1}$$

Output step 2: Convolution operation: This step will extract the spatial features on the input data.

$$F_i = \sigma\left(W * X' + b\right) \tag{2}$$

Step 3: Temporal dependencies of traffic data: LSTM capture:

$$h_t = \sigma\left(w_h h_{t-1} + w_x + b\right) \tag{3}$$

Step 4: Important features are picked out dynamically through the removal of low-weight features to reduce the complexity of the model.

Step 5: Use fully connected and Softmax layer to classify traffic.

Step 6: Verify prediction forecasts to enhance the accuracy.

Step 7: Train the model, and assess it with the help of such metrics as precision, recall, F1-score, and inference time.

The suggested TrustEdgeAI algorithm will start with a network traffic data collection of Edge-IoT devices, followed by a preprocessing phase to eliminate noise and deal with missing values and normalizing features to allow efficient learning. This processed data is then fed through a CNN layer that identifies spatial features and then an LSTM layer to identify the temporal relationships in sequential traffic patterns. Adaptive feature pruning mechanism is used to reduce the number of important features and thus minimize the computational load. The refined features are then processed through a classification layer with softmax which determines whether the traffic is normal or malicious. In order to enhance the predictability accuracy, a calibration unit based on temperature scaling modulates output probabilities. Lastly, a decision based on the system is made and the performance is measured in terms of precision, recall, and inference time.

This step by step process is visually represented as a top to bottom flow chart beginning with the input of data and concluding with the performance evaluation as shown in Fig.2. Every block in the flow chart is associated with a significant step of the algorithm, and the data flow between the preprocessing, feature extraction, learning, optimization, classification, and the decision-making phases is evident. The vertical structure is also clear and is compatible with the IEEE standard, thus simple to know how the data is changed in every step. In general, the algorithm and flowchart as a whole give a good picture of how the TrustEdgeAI framework works and how efficient, scaled and suitable it is to real time intrusion detection in Edge-IoT applications.

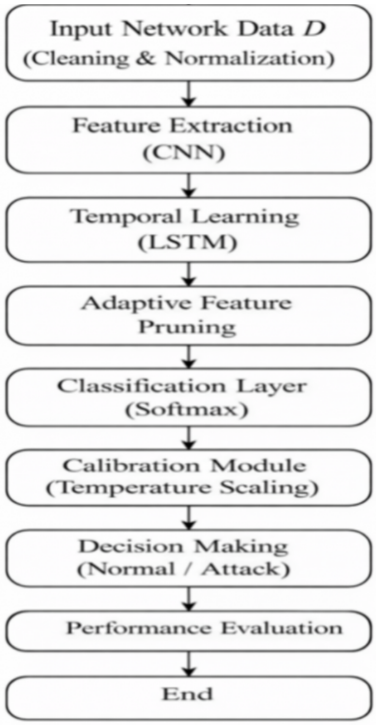


Figure 2: TrustEdgeAI flowchart for intrusion detection

## 4. Experimental Setup and Results

Python is used in the implementation of the proposed TrustEdgeAI framework with the deep learning libraries of Python (TensorFlow and Keras). The experiments are carried out on a system that has moderate computational resources to replicate the conditions of edge-level deployment. The hardware specifications of the system consist of an Intel i5 processor, 8 GB RAM and no graphics acceleration, so that the model can be applicable in resource-limited setting.To test it, common intrusion detection data sets are selected: NSL-KDD and CICIDS. These datasets are normalized and malicious network traffic instances, with different types of attacks, such as DoS, Probe, R2L, and U2R. Preprocessing is carried out to eliminate duplicate records, missing values and Min-Max normalization. Numerical features get the categorical features converted to numerical features, to enable them to suit deep learning models.

The data is separated into a training and testing set by 70:30. Training CNN-LSTM is done with Adam optimizer with the learning rate set to 0.001 and a batch size of 64. The epochs of training are going to be 30 to make sure that it converges. Adaptive pruning of features is used in the training stage to simplify the model and temperature scaling is used in the testing stage to normalize prediction probabilities.

Table.1 shows the performance of proposed TrustEdgeAI model under various types of attacks in an Edge-IoT environment. As can be seen, the model has the best model detecting rate of DoS attacks (99.2) and the lowest false positive rate of 1.2, which shows that it is good at detecting high volume

**Table 1**
Attack-wise Detection Performance

| Attack Type | Detection Rate (%) | False Positive Rate (%) |
|---|---|---|
| Normal | 98.5 | 1.8 |
| DoS | 99.2 | 1.2 |
| R2L | 95.4 | 3.5 |
| U2R | 94.8 | 3.9 |
| Average | 97.8 | 2.5 |

**Table 2**
computational Efficiency and Resource Usage

| Model | Memory Usage (MB) | Inference Time (ms) |
|---|---|---|
| CNN | 210 | 25 |
| LSTM | 240 | 28 |
| CNN-LSTM | 260 | 20 |
| TrustEdgeAI | 180 | 12 |

attacks. In the case of normal traffic, the model is quite accurate with the rate of detection of 98.5% and low false alarms. Nevertheless, R2L (95.4) and U2R (94.8) attacks are detected rather rarely and are more complex and less frequent. In general, the model has a false positive rate of 2.5 and 97.8 average detection rate, which indicates a stable and efficient intrusion detection.

Table.2 shows the computational efficiency of various models in the form of memory usage and time of inference. It is also possible to note that the proposed TrustEdgeAI model has the minimal memory (180 MB), which makes it highly applicable in the resource-constrained Edge-IoT environment. On the contrary, CNN-LSTM is the one that consumes the most memory (260 MB), which means that its computational overhead is larger. TrustEdgeAI is the fastest to respond with inference time of 12 ms as the company allows intrusion detection in real-time. In general, the findings indicate that in comparison to traditional models, TrustEdgeAI is more effective as it has a balanced approach in terms of low resource usage and high-performance rate.

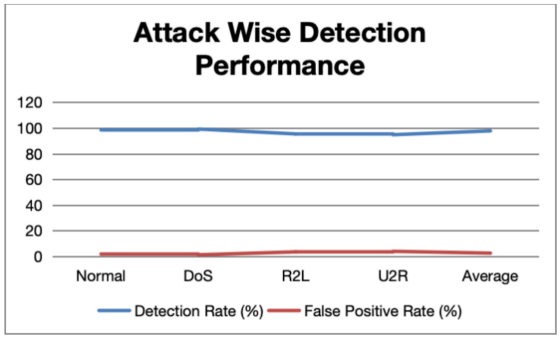


Figure 3: Attack-wise Detection and False Positive Analysis of TrustEdgeAI.

Table.3 shows the main performance indicators of the suggested TrustEdgeAI model, which mean high classification performance. The intrusion detection performance is accurate and consistent that is reflected in high precision

**Table 3**
Advanced Performance Metrics

| Metric | Value (%) |
|---|---|
| Precision | 96.9 |
| Recall | 96.5 |
| Specificity | 97.2 |
| F1-Score | 96.7 |

(96.9) and recall (96.5). Also, the balanced F1-score (96.7) and high specificity (97.2) indicate the utility of the model with the lowest number of false alarms and high accuracy to detect normal traffic.

The performance of the proposed TrustEdgeAI model on the various categories of attacks is shown in Fig.3 based on the detection rate and false positive rate. It demonstrates the highest rates of detection of all types of attacks, the highest rate being observed of DoS attacks and slightly lower rates of R2L and U2R with their complexity. Simultaneously, the false positive rate is very low in all categories thus showing quality classification of normal and malicious traffic. All in all, the graph shows that the model can be successful in achieving a high detection accuracy with a minimal number of false alarms in Edge-IoT settings.

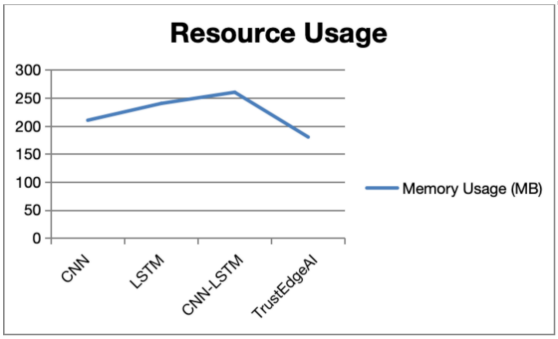


Figure 4: Memory Utilization Comparison of Deep Learning Models in Edge-IoT

Fig.4 graph shows the memory usage of each of the models used such as CNN, LSTM, CNN-LSTM and the proposed TrustEdgeAI framework. It is noted that CNN-LSTM uses the most memory and then LSTM and CNN, meaning that it needs more computational resources. TrustEdgeAI on the other hand is much more lightweight in its consumption of memory, showing an edge-oriented design. This minimization shows the efficiency of the adaptive feature pruning and the model can be used in the resource-constrained Edge-IoT systems.

Fig.5 graph shows the inference time of various models, which points to their efficiency in the computation of Edge-IoT environments. As it can be seen, the LSTM model incurs the longest inference time, then CNN, and the delay is somewhat decreased in the hybrid CNN-LSTM. The proposed TrustEdgeAI model has the shortest inference time of 12 ms, which proves that it can be used to detect intrusion in real-time and at a high rate. This has been greatly enhanced by the

fact that it is a lightweight architecture, which is optimized to maximize features, thus it is very appropriate in time-sensitive edge applications.

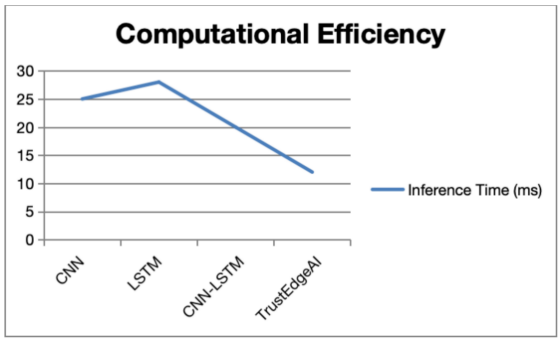


Figure 5: Inference Time Analysis for Computational Efficiency of Models

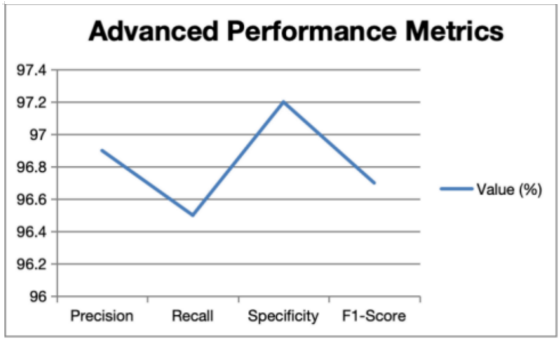


Figure 6: Evaluation of Advanced Performance Metrics for TrustEdgeAI

Fig.6 shows the results of the advanced performance measures of the proposed TrustEdgeAI model, which is precision, recall, specificity, and F1-score. The values are constantly high in all metrics, with specificity being maximum, which demonstrates a high ability to identify normal traffic correctly. There is a minor change in recall because of the complexity of the detection of some attack patterns, whereas precision and F1-score are balanced. The graph indicates a strong and stable performance of the model in the intrusion detection of Edge-IoT environments.

## 5. Conclusion

In the given paper, a new framework, called TrustEdgeAI, has been suggested to overcome the issue of intrusion detection in resource constrained Edge-IoT settings. The system incorporates a hybrid CNN-LSTM network to ensure the successful capture of the spatial and temporal patterns in the network traffic, and an adaptive feature pruning scheme decreases the complexity of the computation and improves efficiency. Moreover, the use of temperature scaling enhances calibration of prediction probabilities so that a more reliable and trustworthy decision-making will be achieved. Data protection during processing is also enhanced by the fact that there is a security-aware preprocessing module. As shown by experimental results, the proposed model attains high detection performance, low latency, and consumes lesser resources hence may be deployed in edge-based systems in real time. TrustEdge AI, in general, offers an equal measure of accuracy, efficiency, reliability, and security with one platform.

To expand the suggested system to future work, it is possible to incorporate federated learning to support decentralized model training on several edge devices, which is privacy-preserving. It is also possible to improve the framework to identify zero-day attacks by adding anomaly-based and unsupervised learning methods. Also, further enhancements in the interpretability of model decisions may be achieved by adding Explainable AI (XAI) techniques, which will result in a more transparent system to be used in real life. Future studies can be done to integrate the model with actual IoT hardware platforms and make it more energy efficient. A further way to enhance the strength and flexibility of the framework in the dynamic IoT ecosystems is expanding it to multi-modal data and cross-domain generalization.

## Credit authorship contribution statement

S Karthika – Methodology, Analysis. , K Juliana Gnanaselvi – Resources, Conceptualization. N S Sivakumar – Draft Writing and reviewing, Resources. Anggy Pradiftha Junfithrana– Software, Data Curation. Hasan Koten – Draft Writing and reviewing

**Funding:** Not Applicable

**Data Availability Statement:** The proposed study did not report any data.

**Acknowledgments:** Not Applicable

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

Abinaya, P., Prem Kumar, S., Sivaprakash, P., Arun Kumar, K., Shuriya, B., and Bhatia, S. (2024). Smart manufacturing with a digital twin–driven cyber-physical system. In *Digital Twins in Industrial Production*, pages 247–276. Wiley.

Babu, R. M. H., Sivaprakash, P., Kumar, K. A., Shebana, M., and Sowndharya, S. G. (2022). Deep phm: Iot-based deep learning approach on prediction of prognostics and health management of an aircraft engine. In *Big Data Analytics in Smart Manufacturing*, pages 121–134. Chapman and Hall/CRC.

Balajishanmugam, V., Shuriya, B., et al. (2025). Advancements in machine learning. In *Evolutionary Artificial Intelligence*. Springer.

Bhaskaran, P. E., Maheswari, C., Thangavel, S., Ponnibala, M., Kalavathidevi, T., and Sivakumar, N. S. (2021). Iot based monitoring and control of fluid transportation using machine learning. *Computers & Electrical Engineering*, 89:106899.

Ezhil Vendhan, K., others, and Shuriya, B. (2025). Ai in predictive diagnostics and precision medicine. *Journal of Applied Bioanalysis*, 11(8):563–570.

Francis, C., Naidu, V. S., and Shuriya, B. (2025). Comment on ai in orthopedics. *Journal of Robotic Surgery*, 19(1):586.

Karthikeyan, N., Sivaprakash, P., and Karthik, S. (2023a). Classification model for ids using auto cryptographic denoising technique. *Computer Systems Science & Engineering*, 45(1).

Karthikeyan, N., Sivaprakash, P., Periyakaruppan, K., and Shobana, M. (2023b). Machine learning-based quantum modeling to classify the

traffic flow in smart cities. In *Quantum-Safe Cryptography Algorithms and Approaches: Impacts of Quantum Computing on Cybersecurity*, page 49.

Mouna, P. S., Sivaprakash, P., Kumar, K. A., Samuel, P., Shuriya, B., and Sharma, V. (2024). Data modeling and analysis for the internet of medical things. In *Wireless Communication Technologies*, pages 197–223. CRC Press.

Priyanka, E. B., Thangavel, S., Gao, X. Z., and Sivakumar, N. S. (2022). Digital twin for oil pipeline risk estimation using prognostic and machine learning techniques. *Journal of Industrial Information Integration*, 26:100272.

Shuriya, B. (2013). Study of an rbac system. *International Journal of Applied Research*, 3(3):104–106.

Shuriya, B. (2015). Privacy preserving and securing using proxy re-signatures. *International Journal on Engineering Technology and Sciences*, 2(2).

Shuriya, B. (2019). A fuzzy responsibility-based access organizer for leukemia record protection using kwatts algorithm. *International Journal of Applied Mathematics & Information Sciences*, 13(6):1047.

Shuriya, B. (2022). A review on weeding robots and its key technologies. In *International Conference on Computational Linguistic and Natural Language Processing*.

Shuriya, B., Balajishanmugam, V., and Sivaprakash, P. (2025a). Towards accurate diabetes prediction using adaptive deep learning. In *ICAECA*. IEEE.

Shuriya, B., Balajishanmugam, V., Sivaraju, S. S., and Mythili, S. (2025b). Predicting mechanical properties of ceramic-modified concrete using gcn. *Journal of the Australian Ceramic Society*.

Shuriya, B. and Graduate, M. E. (2015). RBAM with constraint satisfaction problem in role mining. *International Journal of Innovative Research and Development*, 4(2):100–102.

Shuriya, B., Kumar, S. V., and Bagyalakshmi, K. (2024a). Noise-resilient homomorphic encryption: A framework for secure data processing in healthcare. *arXiv preprint*, arXiv:2412.11474.

Shuriya, B., Prakash, P., and Kiruthikka, D. C. (2022). Qos based aes cryptography network model. In *Proceedings of the International Conference on Innovative Computing & Communication (ICICC)*.

Shuriya, B. and Rajendran, A. (2017a). Cardiovascular disease diagnosis using data mining techniques and anfis approach. *International Journal of Applied Engineering Research*, 13(21):15356–15361.

Shuriya, B. and Rajendran, A. (2017b). Tranquilize role mining using hr (heuristic random) approach. *Asian Research Consortium*, 7.

Shuriya, B., Santhamani, V., Balaji Shanmugam, V., and Subashini, S. (2024b). Enhancing network security through viper optimization algorithm. *Frontiers in Health Informatics*, 13(8):1747–1761.

Shuriya, B. and Sivaprakash, P. (2023). Qos based aes cryptography network model. SSRN.

Shuriya, B., Sivaprakash, P., Kumar, K. A., Saravanakumar, M., and Rajendran, A. (2023a). Digital twins and artificial intelligence: transforming industrial operations. In *Digital twin for smart manufacturing*, pages 111–133. Academic Press.

Shuriya, B., Sivaprakash, P., Mythili, S., Karthiga, S., and Juliana, G. K. (2025c). Intrusion detection systems in smart grids: A review-based study. *i-Manager's Journal on Communication Engineering and Systems*, 14(2).

Shuriya, B. and Thenmozhi, S. (2015). Rbam with constraint satisfaction problem in role mining. *International Journal of Innovative Research and Development*, 4(2).

Shuriya, B., Umamaheswari, S., Rajendran, A., and Sivaprakash, P. (2023b). One-dimensional dilated hypothesized learning method for intrusion detection system under constraint resource environment. In *Proceedings of ICAECA*. IEEE.

Sivaprakash, P., Banumathi, J., and Jayapriya, P. (2025a). Deep-bts: Deep learning based brain tissue segmentation using resu-net model. *Measurement Science Review*, 25(6):371–379.

Sivaprakash, P., Priya, S. S., Maheswari, K., Rubini, B., Karthikeyan, N., and Shuriya, B. (2025b). Patent search classification model for service robots field using deep learning approach. *International Journal of Robotics and Automation*, 40(1):15–22.

Sivaprakash, P., Thenmozhi, S., and Shuriya, B. (2026). Diabetic retinopathy classification using vision transformers. *Tropical Journal of Pharmaceutical Research*, 24(12):1479–1491.

Sivaprakash, P., Thenmozhi, S., Shuriya, B., and Karthika, S. (2025c). Eclnet-a lightweight cnn-lstm fusion network with adaptive feature pruning for real-time intrusion detection in edge-iot environments. In *2025 International Conference on Next Generation Computing Systems (ICNGCS)*, pages 1–7. IEEE.

Umamaheswari, S., Jagannath, V., and Shuriya, B. (2025a). Integrated wearable system for battlefield awareness. In *ICAECA*. IEEE.

Umamaheswari, S., Lingeswaran, G., and Shuriya, B. (2025b). Integrated real-time monitoring for soldier health. In *ICAECA*. IEEE.

Yuvaraju, M., Lekha, K., and Shuriya, B. (2025). Voltage regulation in ev-integrated smart grids. In *CMSS*, pages 285–292. IEEE.