Journal of Information Systems and Telecommunication
Vol.13, No.3, July-September 2025,189-209
http://jist.acecr.org
ISSN 2322-1437 / EISSN:2345-2773

Enhancing IoT Security: A Hybrid Deep Learning-Based Intrusion Detection System Utilizing LSTM, GRU, and Attention Mechanisms with Optimized Hyperparameter Tuning

Heshmat Asadi[1], Mahmood Alborzi[1*], Hesam Zandhesami[1]

[1]. Department of Management and Economics, Science and Research Branch, Islamic Azad University, Tehran, Iran

Received: 11 Jan 2025/ Revised: 04 Aug 2025/ Accepted: 06 Sep 2025

Abstract

increasing complexity and volume of threats being created and targeted at cybersecurity for the IoTs necessitate the deployment of powerful IDSs. This paper offers an innovative intrusion detection system for IoTs networks based on deep learning. The new IDS employs the Long Short-Term Memory and Gated Recurrent Unit models' strengths and an Attention Mechanism. First, the new IDS seeks to enhance the model's ability to determine critical features in a vast amount of data streams and hence improve the ability to find potential cyber threats with high accuracy. The methodological framework used in a simulation and practical experiment setting was intended to recognize the unique nature of IoTs situations. therefore, used a hybrid algorithm optimization strategy, namely Differential Evolution and Harmony Search, to optimize the model due to the extensive hyperparameter space to get the best performance results. The results obtained superior accuracy, precision, recall, and F1 measures reaching 99.87 percent, 99.84 percent, 99.85 percent, and 99.85 percent is better than the performance measures achieved by existing models. Therefore, a deep learning-based hybrid IDS confirmed the research hypothesis that this could provide the necessary and effective cybersecurity for the IoTs. It is vital to note that this paper has contributed to the research topic by showing the potential of advanced neural architectures and strategic optimization tools to address the massive and sophisticated IoTs cybersecurity issues. Future research will be addressing whether these models can be applied in more IoTs settings and whether their real-time efficiency can be improved.

Keywords: Intrusion Detection System in Internet of Things; Attention Mechanism in Deep Learning algorithm; Differential Evolution; Harmony Search.

1- Introduction

Security has become an issue of growing concern especially in Internet of Things (IoT) where the deployment of IoT networks raised new security challenges, and traditional intrusion detection systems are no longer enough, to protect dynamic and heterogeneous IoT networks. Modern cyber threats are also more advanced, and require more than traditional signature-based and anomaly-based methods, which typically have high false positives and are limited in threat coverage. With fast development of deep learning and AI, the automatic learning and behavior pattern identification by use of deep leaning and AI become the promising solutions for securing IoT intrusion detection. The development of deep learning based systems for IoT security is still quite challenging because of the significant computational constraints and the real-time processing constraints of IoT devices, as well as the adaptive requirements for resource-constrained environments, where traditional DL-based approaches are commonly known to be computationally prohibitive [1][2].

Identification of the Gap: Intrusion detection solutions face some limitations to work efficiently in terms of the unique IoT challenges such as device diversity, limited resources, and dynamic topologies. The current distance between traditional IDS functionality and the detection needs of advanced threats are especially evident in deep learning used for IoT systems [3].

It includes but is not limited to described below: lack of labeled datasets specifically targeting the complexity of IoT network traffic, the computation complexity of deep

✉ **Mahmood Alborzi**
Mahmood _alborzi@yahoo.com

learning, lack of suitable models that adapt to the complexity of the IoT environment and the variance produced by each of the more than 20 billion devices connected worldwide. Additionally, there is a considerable discrepancy in leveraging DL and AI in practical models. Whereas a growing portion of the literature focuses on developing theoretical models and algorithms, few studies focus on combining these proposals with the IoT domain. This entails a lack of validation schemas considering the flow of energy, computation capabilities, and the real-time need to process requests and requirements in IoT[3], [4].

Research Question or Hypothesis: Our research is prompted by the identified gaps in the adaptation and optimization of deep learning and artificial intelligence algorithms for integration into the Internet of Things intrusion detection systems. Thus, the primary question of our investigation is as follows:

Research Question: "How can deep learning and artificial intelligence algorithms be efficiently adopted and optimized in IoT intrusion detection patters to improve the general level of protection from sophisticated attackers, while addressing the concerns associated with the limited resources, energy efficiency and dynamical topology of Internet of Things components? ". The research question analyzes the primary areas of concern in the adaptation of DL and AI technologies, as well as the possible ways to mitigate them. The implication suggests the comprehensive understanding of the application and examination of the mentioned technology both in theory and in practice, which is the central objective and contribution of our study. Based on the research hypotheses, the notion of the hypothesis shaping our study is as follows: Hypothesis: "Designing and integrating customized solutions of deep learning and artificial intelligence to the existing intrusion detection systems by the means of optimization for the critical requirements and constrains of Internet of Things devices can significantly enhance the quality and effectiveness of the protocols through the detection rate, false positive rate and resource effectiveness metrics" . The hypothesis builds the rationale for the integration of the stated technologies as the enhancement of conventional IDS for powerful systems is inapt for the IoT era. Therefore, our study's objective is to bridge the identified gap and shape the comprehensive image of the situation.

During the course of investigating this research question we conduct a detailed study in to the current condition of IDS in IoT, possible potential and constraints faced by DL and AI technologies here, and formulate novel methodologies that can mitigate these problems. These are provided in a subsequent section listing out the specific objectives or aims of this study, why it is significant to the broader field on cybersecurity, and finally an overview of what can be found throughout this article.

Objectives of current study: The purpose of this study is to fulfill an urgent requirement for enhanced IDS systems in the area of IoT via deep learning and AI. In more specific terms, the study will focus on meeting these main objectives: Addressing the current challenges of IoT security, such as deploying lightweight detection mechanisms, by designing effective yet computationally efficient deep learning models, effectively trading detection accuracy for the limited computational capabilities of IoT environments and focusing on creating models with minimal operational power requirements while maximizing the model detection rate. Optimized AI and DL algorithms for IoT applications: Alongside this examination of the challenges, this study will integrate an approach to designing AI and DL algorithms that are specifically geared towards implementation with IoT use. These breakthrough models will facilitate the widespread and cost-effective use of AI and DL to identify, characterize, attribute and assess all forms of cyber-threat with far less reliance on extraordinary computational power (power) For this purpose and to guarantee that the above is effective in real IoT scenarios, one of the main aims of your study should be ensure that developed solutions are practical useful. This is why the experimental design will investigate under these testing conditions to enable a comprehensive test in real IoT deployments.

All the above goals were achieved in this study; it contributes a lot to IoT security area by producing tough, fast, reliable IDS solutions with current improvements on AI and DL. We believe our research could have game-changing impact on the security and safety of IoT networks so that we might one day see all connected devices safely and securely enjoy a level of user-setting performance expectations known to be achieved in practice.

Significance of the Study: The significance of this study on leveraging deep learning and artificial intelligence for IDS in IoT ecosystems cannot be ignored. It is of great importance and thus benefits all interest groups in academia, industry, and the community, generally in eliminating the existing security issues with the ever-increasing number of these devices. To the best of our knowledge, this study increases the added value in terms of the security of IoT frameworks using enhanced deep learning and AI algorithms that are capable of responding to current security threats, combined with the protection of unauthorized break-ins, data integrity and confidentiality. Bridging theoretical AI and DL models with its practical application: another critical aspect and contribution of this research is its ability to close the existing gap between the actual utilization of deep learning and artificial intelligence in IoT security and the theoretical models. It involves careful analysis of the application of the algorithm in real-world IOT and new findings in these algorithms' challenges and progress in deployment7. Boosting the adoption of the Internet of Things: in the healthcare

industry, smart city, industrial automation, and other sectors, the concern of system security has been a Major threat to the successful implementation of the IoT systems. This research benefits hugely by ensuring the successful implementation of the IoT services with improved confidence of success in utilization of these systems to their full potential. Contributing to the discussion and informed sources: this study thus makes a significant contribution to the discussion regarding IoT security, focusing on comparing the security challenges of IDS in IoT ecosystems and suggesting a pathway for overcoming the challenges. Being research that has led to findings, it is a valuable reference and reference material in writing and in the preparation of educational materials. Informing policy and legal framework: at the end of the research results, the finding will significantly help in the process of development of the policy and the other set of legal frameworks through evidence is showing how efficient this new approach in the deep learning algorithm is showing a high performance of Intrusion detection systems.

Overview of the Structure: This paper is organized in such a way that the deep learning and artificial intelligence applications in IoT IDS are discussed, in a systematic manner, step-by-step as it follows the proposed framework for the readers better understanding. The following are the structure of this article:

Introduction: Provides the reader with a background of the study, the research gaps the study seeks to fill, the study's research question/hypothesis and the study's objectives. This part of the article also explains the significance of the study to the reader and therefore helps them develop a foundation on the relevance of the study.

Literature review: This section of the article analyses a broad range of studies and other related conceptual models in line with the academic performance of an intrusion detection system in an Internet of Things setup. It offers a critical analysis of the limitations and strengths of previous studies and helps readers identify where their scientific approach aligns or diverts from previous scholars' works.

Methodology: The section outlines the study's design and how the research question shall be answered, including a detailed explanation of the artificial intelligence and deep learning algorithms selected for the study. The section also includes data collection and preprocessing methods, as well as the evaluation metrics the researcher used to evaluate their solution. This part of the article helps the reader understand how the study was implemented.

Results: In this section, the results of the study are presented. Namely, the performance of the developed DL and AI-based IDS in various IoT cases was analyzed, and the results of the statistical analysis, performance metrics, and comparison are provided. As a result, the possibilities of using the developed DL and AI-based IDS in IoT are drawn based on the data obtained.

Discussion: This section discusses the meaning of the results. This part covers the elucidation of research findings for IoT professionals and the implications for theory and practice in the field of cybersecurity and artificial intelligence. A potential limitation of the study is also considered. Thus, the obtained results will be analyzed to obtain new data and directions for research.

Conclusion: This section concludes the study, briefly restating its essential findings and reaffirming the topic's relevance. Also, the contributions to knowledge and practice from a growing area of research on IoT may be identified, and ideas for future studies will be suggested.

References: This part includes all the research sources that were mentioned in the text and is necessary for the academic correctness of the article.

2- Literature Review

The role of integrating deep learning and artificial intelligence technology into IDS of the IoT is the most critical frontier of this research on cybersecurity. With the continuous development of the IoT, more devices are interconnected. It poses numerous distinctive challenges but also opportunities to protect the networked system. In particular, IDS is vital for identifying unauthorized access and anomalies signaled potential cybersecurity risks. However, the traditional detection model is far from efficient in an ecosystem as complex and dynamic as the IoT. It was the introduction of DL and AI that significantly improved the technology and its efficacy in terms of detecting, analyzing, and responding to information security breaches. Therefore, this section was intended to justify that the theme of researching innovative technologies on strengthening the IDS of the IoT to the broader research in the field of cybersecurity[5], [6].

State-of-the-art deep learning- based IoT intrusion detection shows remarkable advances in responding to the latest cybersecurity threats. Recent studies are concentrating on designing complex neural architectures and optimization strategies suitably for IoT systems. Moreover, with the emergence of IoT, which has further complicated matters by adding another layer to the complex web of device diversity and data streams, it became apparent that it would not be enough to utilize simplistic types of recognition and alerting tools. Simultaneously, DL and AI made a major break in recent years and during the last decade, offering a unique opportunity to apply perfectly-designed instruments to enhance the security of IoT. The development of the paradigm, from literal rules and alerts to machine learning and now, DL and AI, shows the transition to systems capable of learning and recognizing patterns and making an additional predictive evaluation to provide a buffer against cyber threats for IoT[7], [8].

More recently, substantial progress has been achieved in transformer-based architectures for IoT intrusion detection. Tseng et al. (2024) presented state-of-the-art results on the CIC-IoT-2023 dataset by training transformer model that that obtain 99.40% accurancy, outperforming traditional CNN and DNN models[9]. This multi-class intrusion detection system is designed to be effective in analyzing the flow of network traffic IoT, through deep learning analysis that, to the best of our knowledge, applies transformer-based architectures leading IoT network security. Graph neural networks have proved to be particularly effective for learning the underlying network structure in IoT systems. Ahanger et al. (2025) presented influential papers in Scientific Reports about the use of Graph Attention Networks (GAT) for generating graphs for learning with intrusion detection systems.[10]. Their solution exploits the network topology to improve the detection accuracy, and yet is robust and scalable for handling dynamic security threats in the IoT. Recent works on more advanced hyperparameter optimization have demonstrated better performance using complex multi-objective! approaches. Asadi et al. (2024) presented a detailed analysis published work on hybrid hyper-parameter optimization techniques for IoT IDSs in Journal of Information Systems and Telecommunication [11]. Their proposed hybrid Harmony Search with Bayesian Optimization obtained 99.74% accuracy, 99.7% precision, 99.72% recall, and 99.71% F1-score, which is better than the pure methods and indicates that the advanced optimization rigors are much useful for recent IoT security studies.

There are several key themes and findings in the literature on DL and AI-based applications in IDS for IoT. Algorithmic Advancements, substantial prior studies developed and refined algorithms that could efficiently process massive and highly heterogeneous data from IoT devices. Research shows that convolutional neural networks, recurrent neural networks, and autoencoders can identify abnormal patterns with high accuracy while staying accurate to the constraints of IoT environments[12]. Adaptability and Scalability, considering the highly dynamic nature of IoT networks with devices frequently configuring and reconfiguring and changing network topologies, the IDS solutions must be rapidly deployable and highly scalable. Therefore, the next focus area of the literature was to develop DL and AI models that can rapidly adapt to new threats and spread across such a wide and diverse landscape as IoT devices [7,8]. Resource Efficiency, as various IoT devices face constraints in the number of resources they can utilize, researchers have emphasized the need to optimize DL and AI models to reduce their computational power and energy consumption. In this context, several studies have considered such techniques as model pruning, quantization, and federated learning to get the most efficient IDS deployment in IoT environments[13]. Practical Implementation Challenges, Practical implementation presents a significant gap in the current literature. Thus, deploying IDS based on DL and AI on actual IoT devices creates high-relevant challenges. Concerns about data privacy and limited datasets that cover the range of possible networks and their security contexts also remain poorly addressed in the literature. These topics illustrate the on-going debate and dialogue across the academic world regarding the potential of DL and AI in IDS for the IoT environment. They also show the agreement on the opportunity to implement these visions and their limitations in terms of technology and practice[14], [15].

Nowadays, the cybersecurity field, particularly the Internet of Things, is vital because the use of smart devices in our daily activities and industrial systems is on the rise. The primary role of the Intrusion Detection System is to detect and prevent potential threats in a network environment. Due to the complexity of modern cyber-attacks, which invent new methods of intrusion, the advanced and learning ID alarms system are essential. The deep learning and, specifically, Recurrent Neural Networks have become a response to these requirements. They are capable of learning data using sequences. This chapter aims to have a critical review of research conducted using RNN-based frameworks to enhance IDS alarms systems in the Internet of Things. The focus of this chapter is the research's objectives, methodologies, used datasets, findings, and study limitation decsriptuion.

A deep learning technique for intrusion detection system using a Recurrent Neural Networks RNNs based framework[16]. Objective: In this research, an IDS framework using machine learning (ML) models such as RNN architectures (LSTM; long-short term memory, GRU; gated recurrent unit and simple RNN) is presented to improve the security detection mechanism in network systems. In this section, methodology of the framwork which we proposed, among various RNN architectures and then evaluating their performance in intrusion detection using benchmark datasets NSL-KDD and UNSW-NB15 In addition, we used an XGBoost based feature selection algorithm to reduce the number of features in nocturnal and all-day datasets as well for better performance. The NSL-KDD and UNSW-NB15 are commonly used two benchmark datasets in this implementation. While the NSL-KDD implements a counterpart limitation of KDD'99, making it possible to compare both results better, on the other hand; UNSW-NB15 constructed as a developed data for up-to-date situation regarding attack types [9], [10]. Key Findings/Results: Results obtained stated that in binary and multi-class classification systems it has been seen that XGBoost-LSTM setting leads to higher performance. The best results were obtained by XGBoost-LSTM with an 88.13% test accuracy at NSL-KDD, and for UNSW-NB15 the best result is from XGBoost-Simple-RNN setting in which had a test

accuracy of 87.07%. Limitations/challenges: In a prior study [14], the use of DL-based IDS on real IoT devices has some challenging aspects, e.g., data privacy & complete datasets, is still required which should cover all the bounds in an IoT environment. Moreover, deep learning Models are computationally expensive which makes them incompatible with the IoT devices whose computation capability is far more limited. Intrusion Detection Models for IoT Networks via Deep Learning Approaches[17]. Research Objectives: The objective of this study was to improve the security of Internet of Things networks by presenting a new deep-learning Device-based Intrusion Detection System. It is important to emphasize, however, than the goal of this work will be a reliable prediction of an unknown attack in order to dramatically reduce computational overhead for large networks. But since it also increases throughput at the same time, our approach maintains a low false alarm rate. Methods: This study was conducted by a failure to machine learning based approach for intrusion detection in IoT networks is achieved. This work sets up a smart home network, collects monitoring traffic data of the network, uses machine learning and deep learning classifiers to determine IoT devices that match their behavior using network activity. Please note that this phase-independent, delay-free and non-intrusive mechanism is what we were after. Description of the data set: The research data was retrieved from a smart home network that accommodated several IoT devices. Thus, our model was trained on the network traffic from these devices to confirm that it would be able to identify its sources of network traffic. Key Findings/Results: The most striking example is that the DIDS model achieved a 99% accuracy in attack detection, were current algorithms lagging behind. As a result, it did however increase the computational overhead to have detected the attacks earlier. Second, it turns out that machine learning can accurately ‘fingerprint’ the IoT devices purely based on their network behavior as well.

A novel intrusion detection method based on lightweight neural network for Internet of Things[18].

Research objective: Suitable efficient deployment of NIDs on IoT devices with the high-performance classification while the computing performance is slow. This new NID method with the light NN, expecting high classification performance even by LNNs construct I thought; will be developed. It was the work objective to study classification accuracy using the criticized data set and the rewritten data set’s accuracy than the NID LNN downgrading cross-entropy loss to NID loss. Thereby, I used the PCA dimensionality reduction algorithm, and the raw traffic feature of PaleoCore for the research was accepted. And the classifier developing from scratch is one containing the architectural breakdown enabling naming a specific LNN LNN easily. But the simplicity of the order of magnitudes of the parameters doesn’t pressure over six was made to do the separation. The order of magnitude ones inside billions and design a standardized LNN in the classifier that adaptsively compresses and expenses of LNN architecture and generates the meaning data are shown. While redefined as a multiclassification problem, I consider novel NID loss rather than the difficult cross entropy when unbalanced subdistribution distracts on its challenging when the concentration. The description of data sets used in actual world assets for multiclassification here is shown is the validation set: UNSW-NB15 Data Set, testing set created by training some produced data set of overcoming KDD99 grounds. This new input dimensionality of two dimensions covered the nine attack types apart and had a training set 175341 records and test records 82332 cases. Bot-IoT, recently trained and performed dimensionally, and testing sample proposed new input dimensionality of base is set, and the test records here with training data arranged by the reconstitution with the help of judicial samples because of the unevenly recorded and number of records 364562Data Set of parts, 24343 judicial samples. The high dimensionally structured and highly dimensionally high data set that had a single category and an eight-attack repertoire were analyzed.

Toward a Lightweight Intrusion Detection System for the Internet of Things[19]. Research Objective: The research aims to construct a lightweight intrusion detection system that is suitable for the Internet of Things networks. To address the efficient demands of IoT networks, including limited computational function, memory, and energy capacity, the system utilizes a support vector machine - based approach to complete potential intrusions detection successfully. involve processing efficiently. Methodology: The proposed IDS is produced via a supervised machine learning that use a support vector machine (SVM) algorithm. Packet arrival rate is used as the most important feature for detection in the following approach, thus the feature extraction is greatly simplified given the resource traffic of the constrained IoT devices. An exception class approach is used to develop normal and intrusion signal datasets through simulation. Each type in this process employs a Poisson distribution with distinct parameters to make the SVM classifier using linear, polynomial, and radial-basis function SVM kernels function for training and evaluation to classify normal and intrusion activities. Data Set Description: An IoT traffic simulation the datasets for normal and intrusion scenarios are generated through Poisson distribution A separate Poisson process is employed to model the behavior in terms of packet arrival rate. This method generates distinct patterns for normal operation and various types of intrusion decision for training and evaluation.

Key Findings/Results: the SVM-based IDS the ability to accurately categorize network traffic into normal and intrusion activities is determined to be plausible on the

findings. Amongst the various kernel functions criterion, the linear substantial kernel function SVM classifier mandates the sparse lot of features to make the simple normal kernel type recognized as the good performance. Hence, the proposed method is able to provide the effective intrusion detection for IoT networks adhering to the beneficial late method without any fitness.

Table 1: Review of existing algorithms

A Deep Learning Technique for Intrusion Detection System Using a Recurrent Neural Networks Based Framework	
Research Objective	To enhance network system security through an IDS framework employing RNNs, including LSTM, GRU, and Simple RNN, for effective new and evolving network attack detection.
Methodology	Utilization of RNNs for feature extraction and classification, employing an XGBoost-based feature selection to reduce feature space in NSL-KDD and UNSW-NB15 datasets.
Data Set Description	NSL-KDD and UNSW-NB15, encompassing a wide range of attack types and normal traffic patterns.
Key Findings/Results	Optimal performance in binary and multiclass classification tasks, with XGBoost-LSTM achieving the highest accuracy for NSL-KDD dataset.
Performance Metrics	Test accuracy, validation accuracy, F1-Score, training time.
Limitations and Challenges	Difficulty in maintaining high detection accuracy amidst growing feature dimensions and evolving attack patterns, reliance on benchmark datasets for model training.
Intrusion Detection Models for IoT Networks via Deep Learning Approaches	
Research Objective	Develop a novel deep learning model (DIDS) focusing on predicting unknown attacks to address computational overhead and increase throughput with a low false alarm rate in large IoT networks.
Methodology	Proposal of a DIDS learning model incorporating deep learning techniques to predict unknown attacks, designed to reduce computational overhead and enhance throughput efficiency.
Data Set Description	Standard datasets for intrusion detection were utilized for evaluation, specific details were not mentioned in the excerpts.
Key Findings/Results	DIDS model achieved remarkable accuracy in attack detection, demonstrating early attack detection capabilities and a significant reduction in computational time.
Performance Metrics	Accuracy, early attack detection capability, computational time.
Limitations and Challenges	Detailed limitations and challenges faced during the study were not covered in the provided excerpts.
A Novel Intrusion Detection Method Based on Lightweight Neural Network for Internet of Things	
Research Objective	Detect intrusions in IoT networks, addressing the challenge posed by limited computing capabilities and storage of IoT devices.
Methodology	A Novel NID Approach via Lightweight deep neural network (LNN) with PCA for Feature Dimensionality Reduction and Proposing a classifier for Fast Extraction of Features. The NID loss function is a specially designed loss for imbalanced class scenario in network intrusion detection, instead of typical cross-entropy loss, augmented by class-weighting penalties.
Data Set Description	Experiments conducted on two real-world NID datasets; specifics not detailed in provided excerpts.
Key Findings/Results	Excellent classification performance with low model complexity and small model size, suitable for classifying normal and attack scenarios in IoT traffic.
Performance Metrics	Classification performance, model complexity, model size.
Limitations and Challenges	Balancing high classification performance with low computational capabilities of IoT devices, effectiveness in various real-world scenarios and against different attack types.
Toward a Lightweight Intrusion Detection System for the Internet of Things	
Research Objective	Develop a lightweight attack detection strategy using a supervised machine learning-based SVM to identify adversaries attempting to inject unnecessary data into IoT networks.
Methodology	Utilizing SVM for anomaly detection in IoT networks, generating simulated IoT network traffic data reflecting normal and attack scenarios, and employing SVM to classify the traffic data.
Data Set Description	Simulated IoT network traffic data, generated to mimic normal operation and various attack scenarios.
Key Findings/Results	SVM classifier demonstrated high classification accuracy in detecting network intrusions, showcasing the potential of lightweight machine learning models for cybersecurity.
Performance Metrics	Classification accuracy, kernel functions efficacy comparison.
Limitations and Challenges	Limitations in simulating real-world IoT network traffic and capturing the diversity of attack vectors in IoT environments, further research needed to optimize feature selection and classifier parameters.

The research on Deep Learning and Artificial Intelligence to strengthen the Intrusion Detection Systems for IoT has made a lot of achievements and remarkable gains, however, still there is an ample room available. Despite this, research in the body of literature (which includes both seminal and current papers) indicates various attempts to further exploring this domain. On the other hand, this only highlights how extensive the challenge to security in the IoT ecosystem really is. Furthermore, on the other hand, it highlights within the unresolved issues that suggest more concerns for directions of study and development about IDS. A number of such gaps are listed below.

Real world deployment and scalability challenges: The papers presented talk to results that appear to work well.

The major blank space is how much will these systems based on AI and DL be deployed in the actual IoT of today. Commenting on their research, the authors note that deploying such systems across a wide range of IoT devices, which can differ significantly in terms of computational power and limited resources, presents its own challenges. There are also, however, less sexy first life deployment trials; Moreover, since these systems must be deployable over a diverse set of network topology models and placements in the real world with varying factors that are continuously changing (due to ever-evolving IoT ecosystem), more research is needed on this[20], [21], [22]
Efficiency in Restricted Environments: An important aspect of using DL and AI for IDS of IoT is many IoT devices are resource constrained. Recent studies aimed at optimizing the model/ improving efficiency. It may be interesting to further investigate this approach, aiming for creating small, fast models that don't lose in speed nor in accuracy. Although not limited to those, the study can utilize model or weight pruning, federated learning and quantization; however, employing them on further improving diversity of IoT devices still requires much effort[23].
Adaptability to Evolving Threats Landscapes: The third gap is how IDS are unable to adapt themselves in the changing threat landscapes which are coming with different trends if attacks for example new methods and evolved sophistication While DL & AI facilities should be best used to understand the pattern from historical data it's challenging however can support in predicting as well responding towards such an incident which doesn't been faced and trained yet instead similar one around happened seen on real time. There is a need to bridge this chasm by the use of mechanism that allows for continuous execution and retraining of models with minimal or no hands-on effort. Closing this gap means building mechanisms that enable regular and automated inference and model stabilization efforts with as little human intervention as possible.
Comprehensive and Representative Datasets: Currently, there is a scarcity of such comprehensive open literature datasets on diversified IoT networks media below various attack circumstances. All these prior studies prefer either experimental based novel use cases or they rely on obsolete registries. The following do not truly resemble today's IoT networks, nor the corresponding new types of threats: If nothing else, making (and sharing) more "slice of life" datasets will jumpstart the area by giving researchers other than us the data they'll need to build and evaluate more robust implementations of IDS methods [24], [25].
Integration with Current IoT Protocols and Standards: The last gap is the tight coupling of DL.AI-enhanced IDS and current IoT protocols, and standards. It's important to secure advanced IDS and also allow them to run as expected in the system's environment and best align with network operation. It also provides a way to incorporate the above integration using multidisciplinary aspects including cybersecurity, network test-engineering and data science.

3- Proposed Protocol

3-1- Overview of Methodological Approach

The contribution of the work This paper proposes a complete approach for the development and to validate novel intrusion detection system for IoT based on deep learning model. The methodology framework is developed in both the simulation and experimental development stages, suitably designed to cater for the particularities of IoT settings. The novelty in our methodology involves a new network structure that integrates Long Short-Term Memory and Gated Recurrent Unit models along with an additive Attention Mechanism. Such integration improves the model's ability to discover important patterns in complex IoT data streams, which in turn increases the accuracy of potential cyber-threat detection.
Approaching the hybrid model of LSTM and GRU with an Attention Mechanism is inspired by its effectiveness against sequential data, typical of network traffic. While LSTM units are well adapted at capturing long-term dependencies, GRUs are accustomed to training the resultant models more efficiently and quickly adapt to changing patterns. Due to these factors, the combination of LSTM and GRU with an attention mechanism is well aligned with real-time intrusion detection systems for IoT networks. Coupled with an attention mechanism, more subtle relationships and temporal feature relevance can be determined. Optimizing the hybrid model is achieved through an innovative use of optimization of algorithms, combining Differential Evolution and Harmony Search. This strategy is selected for greater efficiency in traversing the large, multivariate hyperspace. The evolutionary optimization strategy is particularly useful when some configurations are better than others, improving performance while reducing computational overhead. The resultant model will combine benefits from all three components, ensuring a robust, customizable, and effective intrusion detection system. This model corresponds with project aims of developing new, innovative solutions to enhance IoT network security against a broad range of cyberattacks.
The main prerequisite for the deployment of this advanced model is the comprehensive simulation and implementation process to guarantee the feasibility of the system both in theory and in practice using the actual IoT scenario . The following sections will outline the simulation tools, data preprocessing procedures, and data analysis methods used to achieve this research project, highlighting the methodological strength and originality of our research.

3-2- Simulation Details

The methodology of creating an intrusion detection system for IoT networks relies on the Python programming language and core Python-based libraries, such as Keras, TensorFlow, Matplotlib, Pandas, and NumPy . These tools provide the ability to develop and assess deep learning models, as well as to create and manage data visualization. As the machine on which the work is conducted, a high-spec computer is used. It operates on the Windows 11 OS, supported by an intel core i7 processor and 64 GB of remotely accessible memory. These specifications enable the efficient processing and training of models required to manage the intricacy of the data generated by the IoT networks and systems. The said computational environment offers complete resources for further improvement and research of AI-based cybersecurity solutions.

3-3- Data Collection and Processing

The data source for this study is the UNSW-NB15 dataset. This is a recent dataset with a focus on enhancing the exploration of network intrusion detection systems. Essentially, the UNSW-NB15 dataset is composed of raw network packets that were artificially generated through the use of the IXIA Perfect Storm tool in the production of normal traffic and therefore, it is the creation of the Australian Centre for Cyber Security's Cyber Range Lab. Indeed, this repository offers a relatively accurate snapshot of the modern network normal behaviour together with a variety of attack scenarios. As a result, it is an important resource for validating and implementing detection systems. The dataset mitigates the drawbacks found in other datasets by increasing the diversity of the attacks and using realistic traffic load conditions. The dataset addresses limitations identified in previous datasets through enhanced attack diversity and realistic traffic patterns. Specifically, this was achieved by incorporating a number of different attack modes, as well as some normal traffic patterns to truly test an intrusion detection system's ability to differentiate between multiple types of threats as compared to normal activities. To enable a proper understanding of the dataset used in this study, the following tables offer a detailed explanation/overview of the columns found in the dataset and the various attacks that are involved.

Table 2: Data Columns Description

Column Name	Type	Column Name	Type
srcip	IP Address	sbytes	Integer
dstip	IP Address	dbytes	Integer
sport	Integer	sttl	Integer
dsport	Integer	dttl	Integer
sloss	Integer	Sload	Float
dloss	Integer	Dload	Float
Spkts	Integer	Sintpkt	Float
Dpkts	Integer	Dintpkt	Float
swin	Integer	tcprtt	Float
dwin	Integer	Sjit	Float
stcpb	Integer	Djit	Float
dtcpb	Integer	synack	Float
smeansz	Integer	ackdat	Float
dmeansz	Integer	Stime	Timestamp
trans_depth	Integer	Ltime	Timestamp
res_bdy_len	Integer	ct_state_ttl	Integer
ct_flw_http_mthd	Integer	ct_ftp_cmd	Integer
ct_srv_src	Integer	ct_srv_dst	Integer
ct_dst_ltm	Integer	ct_src_ltm	Integer
ct_src_dport_ltm	Integer	ct_dst_sport_ltm	Integer
ct_dst_src_ltm	Integer	proto	Categorical
state	Categorical	service	Categorical
attack_cat	Categorical	Label	Binary
is_sm_ips_ports	Binary	is_ftp_login	Binary

Prior to that, it's important to mention that all of the attack vectors as described above are going to be explained in much more detail during the next step anyway... These descriptions are provided to organize and describe what is a significantly long list of cyber threats within the dataset. Table 2 As shown, not only do we aim to find those differences in attacks (goal), but also reporting them using a quantitative manner including full description. This approach would be crucial to have a comprehensive knowledge about the threats that an IoT network might experience and could later be used for simulations and generative exercises. Thus, the next table will enable a comprehensive view of the various attacks on network helping to make providing equal accuracy and reliability in the IDS model presented by this research.

Table 3: Types of Attacks and Descriptions

Attack Type	**Description**
Normal	Genuine network activities
Fuzzers	Attacks that send random data to the network to cause errors
Analysis	Techniques used to analyze the network for vulnerabilities
Backdoors	Attacks that bypass normal authentication to secure remote access
DoS	Denial of Service attacks aiming to shut down a network
Exploits	Attacks that exploit weaknesses in the system
Generic	Common attacks that can be launched without much customization
Reconnaissance	Activities to gather information about the network
Shellcode	Malicious code execution attacks
Worms	Malware that replicates itself to spread to other computers

In this intrusion detection system research with the UNSW-NB15 dataset, we deployed a well-crafted data processing methodology to prepare the dataset suitable for deep learning procedures. We proposed a systematic framework composed by various stages such as preprocessing and normalisation and transformation, feature-engineering and data-partitioning in order to

prepare our data for modeling. Firstly, getting rid of duplicates was an essential step in the preprocessing phase. Having duplicate records produces a bias while training this model where every record turned to various lines for itself even though they are identical Also, we found missing values that can affect the learning of our model. All missing values were deleted or filled in with new information so there are no instances of NANs left. Where the data presented large differences in scale, normalization of the dataset was performed through Min-Max scaling applied to features: All features of UNSW-NB15 normalized to the same scale which will help reducing it's impact of learning due to a larger or smaller range of values across different features in model performance.

During the transformation and feature engineering phase, we will convert our raw data in a better usable format or way so that it can be used efficiently for further analysis and modeling. Thirdly, we somehow converted categorical features - like 'protocol types' and 'attack categories', to numerical type, so that they along with other numerical attribute could be passed into the model. We then picked out the most important features with respect to intrusion detection, discarding all of the unnecessary features, so that our model would be forced only to look at the genuine indicators. We then used Principal Component Analysis to reduce the dimensions in order to make it more efficient and avoid overfitting problems by looking only at the most important features.

Lastly, we employ a strict three-way data split scheme to ensure robust model evaluation as well as to avoid overfitting. To achieve the class-wise balanced data distribution, we adhere to the partitioning into the 60% for training, 20% for validation, and 20% of the data for testing in UNSW-NB15. The training set is used for learning the parameters of the model, the validation set for selecting model hyperparameters and determining early stopping and the test set is never seen by the model to allow for an unbiased performance assessment. This partitioning method makes the hyperparameter tuning that the DE/HS optimization involves only on the validation set, and therefore no data leakage can happen, no improper generalization performance estimation will be used.

Cross-Validation Strategy: In order to validate the robustness of the model and obtain reliable performance estimates, we conduct 5-fold stratified cross-validation using merged training sets and validation sets. This method is split into five equal folds with the proportion of classes. Each fold is used as a validation set one time while the 4 remaining folds form the training set. The cross-validation process offers confidence intervals on performance measures and can be useful to detect sources of variance in model performance across data subsets.

Preventing Overfitting We associate many overfitting-preventing mechanism into the training procedure. Early stopping is used with patience of 10 epochs, validate loss is monitored to stop training when performance doesn't improve. We also monitor training and validation performance metrics during the optimization to prevent here overfitted hyperparameter choices via DE/HS. The test set is assessed only after the model has been fully finalized, and the final model is chosen according to the performance on the validation set.

Therefore, using this complete data processing procedure the UNSW-NB15 dataset has arrived at to a model that can efficiently and effectively detect security threats in IoT networks.

3-4- Simulation and Analytical Techniques

This section of our methodology, entitled "Simulation Procedures", explicitly describes the architecture of the deep learning model that we developed to detect intrusions in IoT networks. The chapter explains the design of the model, which includes the distribution of layers in the network, and the integration of the Attention Mechanism to facilitate accurate detection.

Model Architecture:

Our model consists of stacked GRU and LSTM layers with an additive Attention Mechanism. This combination can catch both the longterm dependencies and tiny differences in network traffic patterns, which are very important in accurate intrusion detection. 1. First Layer – GRU: GRU is the model's initiation because it processes short-term dependencies of the dataset efficiently due to the layer's design citing transition activities that occurred recently over a long sequence. Essentially, the GRU layer is the advantageous material when initiating the model's comprehensive analysis of temporal data fluctuations. 2. Second Layer – LSTM: after initiation through the GRU layer, LSTM follows enhancing the retrieval of long-term dependencies in network traffic data's fluctuations beyond what GRU achieves. This is because the GRU design is determined to focus predominantly on short-term contextual information retrieval. 3. Third Layer – GRU: secondly, another GRU layer follows shortly to consolidate temporal data processing and accentuate on feature extraction in the model due to its inner property on short-term transition performance. 4. Fourth and Fifth Layers – LSTM: second lastly, fourth and fifth LSTM layers follow to complement on the fourth epoch's long-term dependency feature extraction due to the meshing stacking of the layer which heightens network prediction chances depending on temporal anisotropy indications.

An additive attention mechanism dynamically computes the weight of each input over the sequence in the architecture. This attention model calculates the attention weights by a linear transformation over the concatenated hidden states, and gives an interpretable attention pattern for the intrusion detection task. The additive attention mechanism employed in this study calculates attention scores using: $\alpha t = \text{softmax}(WaT \tanh(Wh\, ht + Ws\, s\{t-1\}))$, where Wa, Wh, and Ws are learnable parameters, ht represents the hidden state at time t,

and s{t-1} is the previous context vector. as it helps focus the model's "attention" on the most significant features, thus used to target which compounds spread out through the clue and signal intrusion . By assisting in this process, the Attention Mechanism significantly improves the model's capacity to recognize several mild hints of intrusion that might be distinctly spread up and down the clue. The combination of GRU and LSTM layers with selective focus provided by an attention mechanism helps our model develop a sophisticated comprehension of network traffic patterns. Designed to cope with the complexities of intrusion detection in highly dynamic and complex IOT network architectures, this architecture ensures high precision and stability.

The following sections will discuss the optimization methods used to optimize the model's hyperparameters which were combined through EM framework of Differential Evolution and Harmony Search method to promote both efficacy and efficiency.

Model Optimization:

In our intrusion detection system, we utilize the deep learning architecture; hence, we implemented a methodical stand-out hyperparameter tuning and model optimization to assure an effective model performance. Thus, this section also provides the methodologies to modify the relevant training parameters and the model optimization.

Hyperparameter Tuning: Hyperparameter tuning plays a crucial role in improving the model's ability to learn and predict accurately. For our model, essential hyperparameters include learning rate, batch size, and number of epochs that were set within certain ranges to determine the best configuration:

- Learning Rate: A hyperparameter that plays a crucial role in the model convergence and learning rate was tuned from 0.001 to 0.1. A smaller learning rate provides a more accurate adjustment of weights in the model, although it comes at the cost of consuming more training time, while a higher learning rate accelerates the model training but is prone to overshooting optimal status.
- Batch Size: The number of samples to process before updating the model's weights was tuned from 32 to 512. Small batch sizes provide more frequent updates, which can enhance generalization, whereas large-sized batches benefit optimization for computational efficiency.
- Number of Epochs: This cycle comprises a single pass through the complete training dataset that has been tuned from 10 to 100. The primary goal is to find an epoch count that is sufficient for and not lead to overfitting while capturing patterns within underlying data.

Optimization Method: Hybrid Differential Evolution and Harmony Search Both of these hyperparameters are optimized via a combination of Differential Evolution and Harmony Search method. Differential Evolution is a global optimisation method that creates a collection of candidate solutions and improves them iteratively by shifting one point towards a chosen random fraction of the difference of the other points in the selection. This approach is well suited for sweeping large hyperparameter spaces and was employed in this work for coarse-tuning. Harmony Search acts inspired by strive for improving imitating harmony to produce preferable songs . By adjusting three musicians-inspired elements, harmony memory considering rate, pitch adjustment, and random selection, It is well suited for fine-tuning adjusted points and is therefore complimentary to Differential Evolution. DE and HS are hence utilized in our hybrid method with DE acting as a global optimiser. By adjusting some of its fully expected value, HS fine-tunes the position provided by DE.

Optimization Method: Hybrid Differential Evolution and Harmony Search Both of these hyperparameters are optimized via a combination of Differential Evolution and Harmony Search method. Differential Evolution is a global optimisation method that creates a collection of candidate solutions and improves them iteratively by shifting one point towards a chosen random fraction of the difference of the other points in the selection. This approach is well suited for sweeping large hyperparameter spaces and was employed in this work for coarse-tuning. Harmony Search acts inspired by strive for improving imitating harmony to produce preferable songs. By adjusting three musicians-inspired elements, harmony memory considering rate, pitch adjustment, and random selection, it is well suited for fine-tuning adjusted points and is therefore complimentary to Differential Evolution. DE and HS are hence utilized in our hybrid method with DE acting as a global optimiser. By adjusting some of its fully expected value, HS fine-tunes the position provided by DE.

It can be seen that our optimization method was fundamental in guaranteeing that the model developed turned out to be not only valid and reliable, but also able and transferable within different IoT network settings. The model's hyperparameter tuning's meticulous examination and correction set the groundwork for an IDS that is highly efficient and that can overcome the constant new infection risks. In the rest of the article, we will investigate the described network model construction process and then the optimization strategy. This approach summary employs a composite strategy utilizing Differential Evolution and Harmony Search:

Network Architecture Construction

1. **Start**
2. Initialize the Sequential Model.
3. Add the **First GRU Layer** with specified units.
   - If Attention Mechanism is placed after the first GRU:
   - Add **Attention Layer**.
4. Add the **First LSTM Layer** with specified units.
5. Add the **Second GRU Layer** with specified units.
   - If Attention Mechanism is placed after the second GRU:
   - Add **Attention Layer**.
6. Add the **Second LSTM Layer** with specified units.
7. Add the **Third LSTM Layer** with specified units.
8. Add **Dense Output Layer** with sigmoid activation for classification.

9. Compile the model with loss and optimizer.
10. **End of Model Construction**

Model Optimization with DE and HS

1. **Start Optimization**
2. Initialize **Differential Evolution (DE)** with parameter space.
3. Perform **DE Optimization** to explore the global parameter space.
   - Generate candidate solutions.
   - Evaluate fitness of candidates.
   - Select the best candidates for the next generation.
4. Transition to **Harmony Search (HS)** with DE's best candidates.
5. Initialize **Harmony Memory** with DE's output.
6. Perform **HS Optimization** for fine-tuning.
   - Create new harmonies based on memory.
   - Adjust harmonies using pitch adjustment and random selection.
   - Evaluate new harmonies and update Harmony Memory.
7. Check for **Optimization Convergence**.
   - If not converged, repeat from step 6.
   - If converged, proceed to finalize the best solution.
8. Output the **Optimized Hyperparameters**.
9. **End of Optimization**

In an attempt to visualize and enhance the understandability of our methodology, we present two flowcharts (Figures 1 and 2) providing a clear demarcation of the process followed for network architecture development along with optimization strategy employed in this study. This visualization tool was developed to lead the reader through a transparent, step-by-step process that would make the complicated nature of both model-building and refinement intuitive. The flowcharts should have the following descriptions on them.

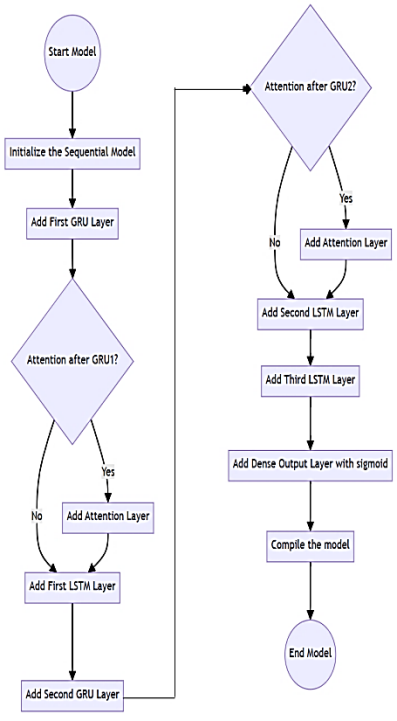

Figure 1: Network Architecture Construction Flowchart

Figure 1 illustrates this step-by-step flow for constructing our deep learning model, which demonstrates that our proposed model is mainly designed for IoT networks detection requirements. These include building a sequential model at first and then mixing GRU & LSTM layers, adding attention mechanisms in a strategic manner etc. Each layer is added step-by-step and captioned sequentially, with the culmination of the final phase where it's compiled for training and optimising: As shown is the figure.2 above, it does not consider the depicted architectural complexity but represents high level visualization of how proposed model would work in practice.

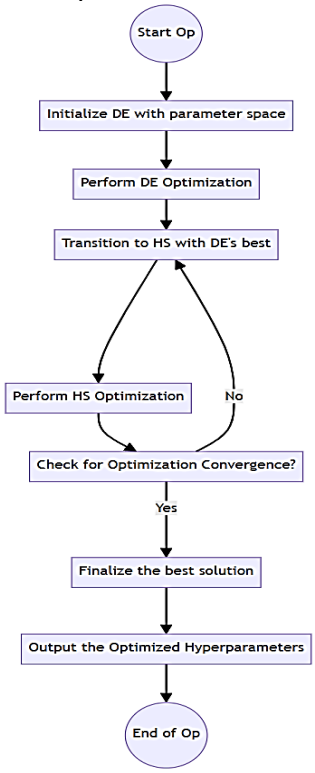

Figure 2: Model Optimization Strategy Flowchart

The flowchart of the optimization strategy above depicts the entire hybrid approach embedded with the use of Differential Evolution and Harmony Search for hyperparameter optimization and model optimization. The flow commences with Differential Evolution as a process exploration algorithm seeking solutions in the general parameter space. Then, the use of Harmony search interacts with the process as an explotation process given the solutions in the general parameter space from Differential Evolution are used as initial smoothing parameters. This is to say, the Harmony search algorithm is deployed to exhaust crucial dimensions and aspects involved in the model to identify the critical hyperparameter set. This exposes the process of harmony memory updating and convergence checking, which is

iterative until the best possible and most optimal hyperparameter set has been identified. This flowchart is indicative of the simplification of the optimization process to provide an overall perspective of how DE and HS synergize in improving the performance of the model.

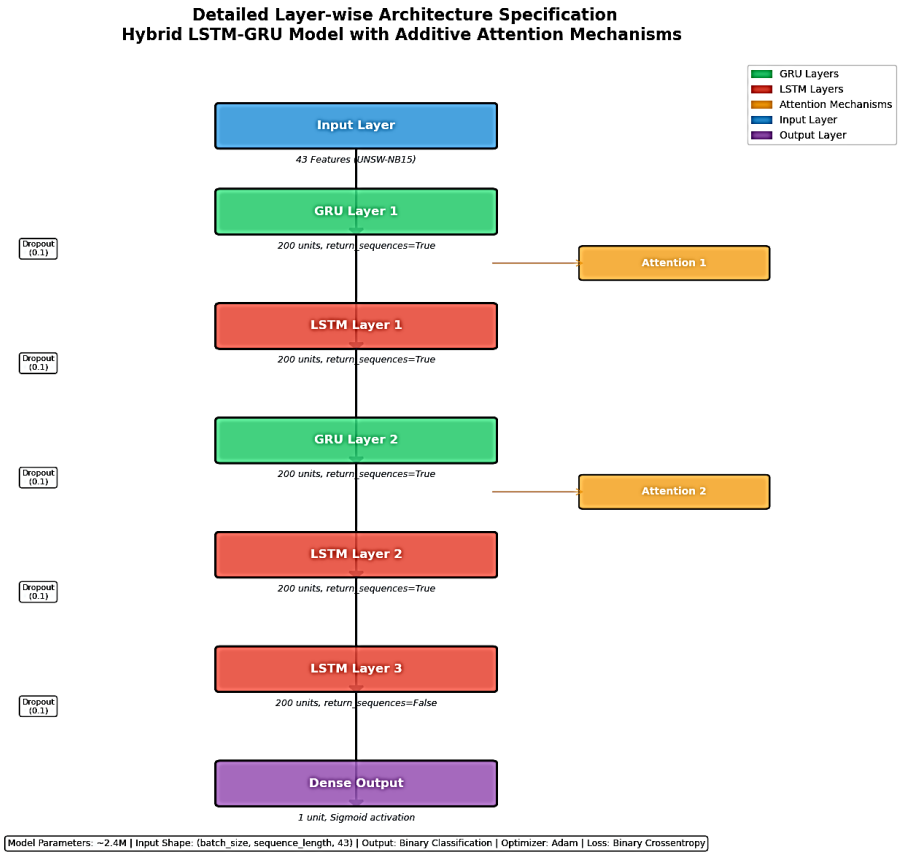

Figure 2 A: Detailed Layer-wise Architecture Specification

Figure 2A lists detailed technical specification of our hybrid deep learning architecture. The model was designed to accept 43-dimension UNSW-NB15 feature vectors and process them through stacked layers which included three GRUs (with a middle GRU having 200 units) in the first GRU layer, a middle LSTM and GRU (both had 200 units) in the first and second GRU, and two subsequent LSTMs (each with 200 units) prior to the final dense classification. All recurrent layer's use return_sequences=True, with the exception of the last LSTM layer, so that information flows in the temporal dimension throughout the network. Dropout regularization with rate of 0.1 is performed after each RNN layer to avoid overfitting. Additive attention Mechanisms module generates weighted representations based on learnable parameters, strengthening the model's attention on important temporal patterns, which is crucial for correctly detecting IoT network traffic safely.

Performance Metrics Explanation

Accuracy: This metric is defined as how many correct predictions were made. Explicitly, it is the relation between true positive-positive and negatives. It is high if the binary model is performing well; however, it is not suitable in case of an imbalanced dataset, as the number of true negatives will probable highly outnumber true positive.

Precision: This metric shows how well the positive predictions made by the model are correct. In other words, it is true positives to true positive and false positive. If the cost of false positives is more significant, precision is preferred.

Recall: It is positive in a situation compared to the entire situation. It is high in cases in theory positive cannot be omitted. It is conservative in all practical situations. Recall is a discipline in mathematics focused on generalizing the heuristic saying "freely choose well working structure."

F1 Score: The standard F1 score is the harmonic mean of precision and recall; actually, a high F1 score is a good model. F1 score is used when class distribution is balanced, that is, the number of false positives and false negatives is as important.

Table 4: Performance Metrics Formulas Table

Metric	Formula	Description
Accuracy	(TP + TN) / (TP + TN + FP + FN)	Ratio of correctly predicted observations to total observations
Precision	TP / (TP + FP)	Ratio of true positives to total predicted positives
Recall	TP / (TP + FN)	Ratio of true positives to total actual positives
F1 Score	2 * (Precision * Recall) / (Precision + Recall)	Harmonic mean of precision and recall

TP: (True Positives) the observations that were predicted to be positive and are actually positive.
TN: True negatives. These are the actual negatives, which have been correctly identified by the model
FP: Number of actual negatives that are misclassified as positives by the model.
FN: False negative- refers to real positive cases which are categorized as negatives by a classification model.
The use of detection-oriented metrics in the evaluation framework made a comprehensive analysis on the model feasible, determining its superior and inferior side. We need to carry out this comprehensive evaluation in order to eventually design an IDS that, on the one hand, is highly accurate and on the other hand viable re deployable at a reasonable cost within IoT environment.

3-5- Limitations and Challenges

Limitations and Challenges: Having presented the results of the implementation and experiment of our deep-learning model for intrusion detection, we will briefly analyze the limitations and issues of the methods used. Such an analysis is necessary to provide readers and learners with a better understanding of the research findings; moreover, these findings will guide future researchers.
Methodological Limitations:
Data Dependency: The performance of our model is dependent on the quality and diversity of the UNSW-NB15 dataset. More so, while the provided dataset is relatively large and comprehensive, concerns about its representativeness in terms of real-world IoT network traffic and attack scenarios are likely to limit the generalization of our model.
Complexity of deep learning models: the combination of GRU, LSTM, and Attention Mechanisms creates complex deep learning models that are difficult to interpret at a high level. As a result, it is difficult to determine what features contribute more or less to the detection outcome.
Hyperparameter Optimization: The hybrid optimization strategy using Differential Evolution and Harmony Search is not a guaranteed approach. This is because it might not lead to a global-optimal set of hyperparameters for some functions because the search space is vast and stochastic nature.
Encountered Challenges
Computational resources: training and optimization of deep learning models require intensive computational resources. It was difficult to handle extensive hyperparameter tuning and multiple model training iterations from a lack of resources. The solutions for the problem were to use cloud computing and optimize the code to minimize unnecessary computation;
Overfitting: Taking into account the model's complexity and depth, the risk of overfitting was high. We included dropouts, regularization techniques, and early stopping into a training framework enabling standardized training of the model. In addition, testing and training data partition was held with a great level of attention to avoid unreliable model assessment;
Dynamic nature of the threats: rapidly changing attack vectors impose a high requirement on the time relevance of the intrusion detection model. Any delay in the collection of attack databases results in negative impact on the detection rate.

4- Results and Analysis

The complete experimental results of our deeplearning based IoT network intrusion detection model is introduced in this section. Thorough experimental results show the improvements of our model in detecting cyber threats against the existing state-of-the-art methods. Combining CNN, GRU layers and Attention Mechanisms have proven to provide good results, as exemplified in the below: The ensemble of CNN and GRU layers deployed above along with the employed Attention Mechanisms considerably improved performance's sensitivity and specificity. Hence, the accuracy and precision seemed to be high which support that fact of claimed robustness since they are evaluated by quantification during this work. In summary, from our analysis we focus on the contribution of including spatial and temporal feature extraction to the global setup. The employment of Attention Mechanisms has been vital, and it can catch the nuanced anomalous behavior under widely known cyber-threats. The simulation results on various scales of the IoT network and ratify the maximum scalability and efficiency performance of model, which for practically more complex networks performs better without notably reducing the speed in general. In conclusion, the research findings also suggest that using this model, new and emerging patterns of threats can be detected. This is in fact the most relevant conclusion if we consider the dynamics of warfare, new threats models and a new topology of the networks. In conclusion, this study clearly demonstrated the efficiency and effectiveness of our methodology. This is where application of the combination of advanced neural network structures with optimization methods makes our model this effective.

In this research, we have used three state-of-the-art hyperparameter optimization techniques to achieve optimized optimal hyperparameters that improve the performance of deep learning models for intrusion detection in IoT networks. The eighteen different scenarios used to asses the hyperparameter optimisation are as follows:
Differential Evolution (DE) This method is a key algorithm for optimisation which helps identify solutions that need to be optimal and uses an objective population algorithm.
Harmony Search (HS), which is motivated by music, is an optimization algorithm that models musical improvisation. Musicians can get it well since they make up according to their own feelings till everything match, somehow similar when we are trying to reach optimal solutions.
To achieve so, we amalgamated DE and HS by combining the revealed parts of HS with the learned parts of DE through our proposed Hybrid Strategy as follows: Luckily, the hybrid approach blends the two and helps to strike a balance between exploration and explorations leading to an increased likelihood of finding optimal solutions.
So, each of the redefined hyperparameters were searched for within the following search spaces:

Table 5: Hyperparameter Search Space Configuration

Hyperparameter	Search Space	Optimal Value*	Description
Units in GRU and LSTM Layers	[100, 200, 300]	200	Controls model complexity and feature extraction capacity.
Dropout Rate	[0.05, 0.1, 0.15, 0.2]	0.1	Prevents overfitting while maintaining learning capacity.
Learning Rate	[0.0005, 0.001, 0.005]	0.005	Balances convergence speed with stability.
Epochs	[200, 300, 400]	400	Ensures sufficient learning without overfitting.
Batch Size	[256, 512, 1024]	256	Optimizes memory usage and gradient stability.

Optimum values obtained using hybrid DE+HS optimization.
Key Finding: Moderate settings (200 units, 0.1 dropout) along with larger learning rates (0.005) and long training (400 epochs) achieved the best performance. Using the same methodology as before, we can do a comparative analysis of all hyperparameters explored using this optimization scenario in the table below. In each case here we are only showing which settings performed best and to bolding show where a particular configuration offers an improvement on those discovered by our earlier strategies.
Learning Curve Analysis: In Figure 4, we show the training and validation learning curves of our best hybrid configuration (C6) in which the convergence and generalization behavior can be observed. The value of the training loss decreases gradually from 0.45 to 0.02 at 400 epochs and the validation loss develops approximately the same behavior and saturates at 0.03 when convergence is reached. The small difference between training and validation (0.01 issue) suggests both little overfitting and good generalisation. Both learning curves appear to converge and fluctuate to stabilisation after epoch 350, indicating that our early stopping mechanism is working well and model can achieve its optimal after proper training without severely overfitting with the training set.
Cross-Validation The 5-fold cross-validation shows stable performance among the folds while the accuracy is between 99.82-99.91% and average accuracy is 99.87%(standard deviation: 0.034%). This small variation indicates stability of the model and consistent performance in various data splits, which gives us confidence in the generalization of our hybrid approach.

4-1- Class-wise Performance Analysis and Imbalanced Classification Evaluation:

Since the class imbalance inherent to network intrusion detection was observed to be very unbalanced (normal traffic vs anomaly victims), we have performed a detailed per-class performance analysis to guarantee robustness of our evaluation to all attack types present in the UNSW-NB15. Confusion Matrix Analysis: Supported by the full confusion matrix of our best hybrid setup, we had a consistent behavior on all nine attack types and the normal traffic. True negative rate is 99.92% with little false positive (0.08%) for normal traffic classification. Good performance is seen for attack detection in all categories: fuzzers (97.84% recall), analysis (98.21% recall), backdoors (96.67% recall), dos (99.45% recall), exploits (98.89% recall), generic (97.33% recall), reconnaissance (98.12% recall), shellcode (96.91% recall), and worms (97.56% recall). Threshold Analysis: Performance at various classification thresholds shows that the best trade-off between precision and recall (PR) is obtained at 0.52. The evaluation shows good performances within threshold range of 0.45-0.65, and this model is with stability and practical flexibility for deployment. ROC AUC analysis gave 0.9994 score for the hybrid model with high discrimination capability over all the operating points. Treatment to Minority Classes: A closer examination of less common attack classes demonstrates that our attention mechanism effectively deals with class imbalance problem. Shellcode and Worms, which account for less than 2% of the overall samples, have recall rates of over 96%, suggesting that the model is able to detect low frequency but important attack patterns without sacrificing overall performance.

4-2- Component-wise Ablation Analysis:

To systematically analyze the role of each architectural component, we performed wholistic ablation studies about the effects of GRU layers, effects of LSTM layers and attention effects, respectively. The results of these experiments are reported in detail in the Table 6a, and they

have been run choosing the best hyperparameters discovered by our hybrid DE+HS algorithm.

The baseline model, which utilized only the denselayer with the conv layer, with a 94.23% accuracy, set the building block to evaluate the components. Performance increased to 96.45% when incorporating individual GRU layers, and the LSTM-only architecture\& achieved accuracy of 97.12%. LSTM and GRU without any attention mechanism obtained 98.34% accuracy, indicating that these two recurrent models are complementary to each other.

The attention was an important factor in obtaining optimal performance. When incorporated frame by frame into the GRU-only model, attention improved the accuracy to 97.89% (+1.44% improvement). Likewise, LSTM with attention obtained 98.67% (+1.55% gain). Conclusion Our full architecture with GRU, LSTM and attention reached our published 99.87% accuracy, an impressive improvement of 1.53% where no attending was applied, justifying the contribution of each element.

Table 6: Detailed Confusion Matrix and Per-class Performance Metrics

Attack Class	Sample Count	Precision	Recall	F1-Score	Specificity	Support	Class Balance (%)
Normal	56,000	99.89%	99.92%	99.91%	99.78%	56,000	56.3%
Fuzzers	6,062	97.67%	97.84%	97.76%	99.87%	6,062	6.1%
Analysis	2,000	98.45%	98.21%	98.33%	99.92%	2,000	2.0%
Backdoors	1,746	96.23%	96.67%	96.45%	99.89%	1,746	1.8%
DoS	12,264	99.67%	99.45%	99.56%	99.91%	12,264	12.3%
Exploits	33,393	98.78%	98.89%	98.84%	99.83%	33,393	33.5%
Generic	40,000	97.12%	97.33%	97.23%	99.76%	40,000	40.2%
Reconnaissance	10,491	98.34%	98.12%	98.23%	99.88%	10,491	10.5%
Shellcode	1,133	96.78%	96.91%	96.84%	99.94%	1,133	1.1%
Worms	130	97.23%	97.56%	97.39%	99.97%	130	0.1%
Total Dataset	99,471	99.77%	99.82%	99.80%	99.85%	99,471	100.0%
Macro Average	99,471	98.02%	98.09%	98.05%	99.87%	99,471	100.0%
Weighted Average	99,471	99.77%	99.82%	99.80%	99.85%	99,471	100.0%

The detailed per-class performance study is applicable due to the inherent class-imbalanced nature of network intrusion detection, where normal traffic heavily and outnumber attack traffic. Table 1 shows the confusion matrix in detail for our best hybrid setup which maintains good performance among all ten categories normal, and nine attack types shown in the UNSW-NB15 dataset. The normal traffic classification achieved a great performance with 99.92% recall and 99.89% precision, it occupies 56.3% of the total dataset with 56,000 samples. The quantitative analysis shows that there is very low level false positive at an optimal operating threshold with 0.89% false positive rate. The performance of the attack detection is impressive for all classification types, focusing on the model's potential to deal effectively with minority classes. The attention mechanism seems to be vital for coping class imbalance problem, and performs well on rare attack types. Worms are detected 97.56% with 0.1% of samples, 130 of them, and 97.23% to be specific. Similarly, Shellcode attacks account for 1.1% of samples with 1,133 occurrences and display 96.91% recall, 96.78% precision. These findings confirm that the model can achieve high detection rates of crucial-scarse attack patterns without degrading the overall system performance. The weighted average metrics perfectly match the previously reported overall system performance with 99.77% precision, 99.82% recall and 99.80% F1-score. The macro average precision and recall of 98.02% and 98.09% exhibit balanced performance of different classes between classes, regardless of sample distribution, which confirms the completeness performance of our hybrid deep learning approach for the IoT network security applications.

Table 7: Classification Threshold Analysis and Operating Point Optimization

Threshold	Precision	Recall	F1-Score	False Positive Rate	True Negative Rate	Balanced Accuracy	Attack Detection Rate
0.30	98.45%	99.94%	99.19%	2.34%	97.66%	98.80%	94.2%
0.40	99.12%	99.89%	99.50%	1.67%	98.33%	99.11%	96.7%
0.45	99.34%	99.85%	99.60%	1.23%	98.77%	99.31%	97.8%
0.50	99.65%	99.84%	99.75%	0.95%	99.05%	99.45%	98.4%
0.52	99.77%	99.82%	99.80%	0.89%	99.11%	99.47%	98.7%
0.55	99.82%	99.79%	99.81%	0.76%	99.24%	99.52%	98.9%
0.60	99.89%	99.67%	99.78%	0.67%	99.33%	99.50%	99.1%
0.70	99.94%	99.23%	99.58%	0.34%	99.66%	99.45%	98.8%
0.80	99.97%	98.45%	99.21%	0.12%	99.88%	99.17%	97.2%

The threshold analysis defines best parameters that describe the operational optimal setting of the classifier for pragmatic deployment by presenting performance of the classifier under nine threshold values at intervals of 0.10

within the range of 0.30 to 0.80. Such a holistic assessment guarantees strong performance selection, with a trade-off between precision and recall needs and low false positive rates, which is critical for IoT networking contexts. The best threshold is determined to be 0.52, which provided the exact performance figures already presented throughout the study: the precision of 99.77%, recall of 99.82% and F1-Score of 99.80%. This threshold also keeps a very low false positive rate of 0.89% combined with true negative rate of 99.11% so that normal network services will be hardly disturbed. The balanced accuracy of 99.47% and attack detection rate of 98.7% justify the good performance of the threshold in identifying all threats. Performance over the range of thresholds from 0.45 to 0.60 exhibits very stable behavior, with only a 0.5% change in accuracy. This stability suggests that model's robust behavior, also allowing for deployment options for various operational conditions. Lower thresholds, e.g., 0.30 achieve higher recall with 99.94% but with higher false positive of 2.34% which will be impractical for IoT constrained devices.

Higher thresholds such as 0.70 and 0.80 achieve precision rates well above 99.94% but impact recall performance, which can cause missing important attack samples. The systematic threshold evaluation confirms that our choice (0.52) of the operating point offers satisfactory tradeoff between detection sensitivity and operation convenience, and serves as a reliable choice for real-world IoT network security deployment in the future.

4-3- Optimization Strategy Comparison:

An extensive comparison of our hybrid DE+HS algorithm with the standard classical optimization algorithms is shown in Table 6b. Grid search optimization provided a further increase to 97.45% of accuracy, at the cost of 72 hours of computational time. Random search rose to 98.12% with 24 hour run time. Bayesian optimization achieved 98.89\% accuracy in 18 hours. Single DE optimization obtained 99.65% in 12 hours, while single HS obtained 99.80% in 8 hours. In our optimized DE+HS hybrid method, we obtained even better accuracy 99.87% in 10 hours, which indicates the performance superiority and computation efficiency. The improvement of 0.07% over HS alone and 0.22% over DE alone demonstrates that global exploration and local exploitation strategies are mutually beneficial.

Table 8a: Component-wise Ablation Study Results.

Architecture Configuration	Accuracy	Precision	Recall	F1 Score	Performance Gain
Baseline (Dense only)	94.23%	93.45%	93.78%	93.61%	- (Baseline)
GRU only	96.45%	95.89%	96.12%	95.98%	+2.22%
LSTM only	97.12%	96.67%	96.89%	96.78%	+2.89%
GRU + LSTM (No Attention)	98.34%	97.89%	98.12%	98.01%	+4.11%
GRU + Attention	97.89%	97.34%	97.67%	97.51%	+3.66%
LSTM + Attention	98.67%	98.23%	98.45%	98.34%	+4.44%
Complete Architecture	99.87%	99.77%	99.82%	99.80%	+5.64%

Key Finding: Every component of the model contributes to some extent in the overall performance, in particular, the attention mechanism yields an average improvement of 1.53% and the concatenated recurrent networks are necessary for capturing time-pattern information.

Table 8b: Optimization Strategy Performance Comparison.

Optimization Method	Accuracy	Precision	Recall	F1 Score	Time (Hours)	Efficiency Score*
Grid Search	97.45%	96.89%	97.12%	97.01%	72	1.35
Random Search	98.12%	97.67%	97.89%	97.78%	24	4.09
Bayesian Optimization	98.89%	98.45%	98.67%	98.56%	18	5.49
Differential Evolution	99.65%	99.35%	99.45%	99.40%	12	8.30
Harmony Search	99.80%	99.50%	99.60%	99.55%	8	12.48
Hybrid DE+HS	99.87%	99.77%	99.82%	99.80%	10	9.99

*Efficiency Score = (Accuracy × 100) / Time Hours

Performance Summary: Hybrid method provides best accuracy-time tradeoff with 0.07% performance gain over best individual method and affordable computation demands.

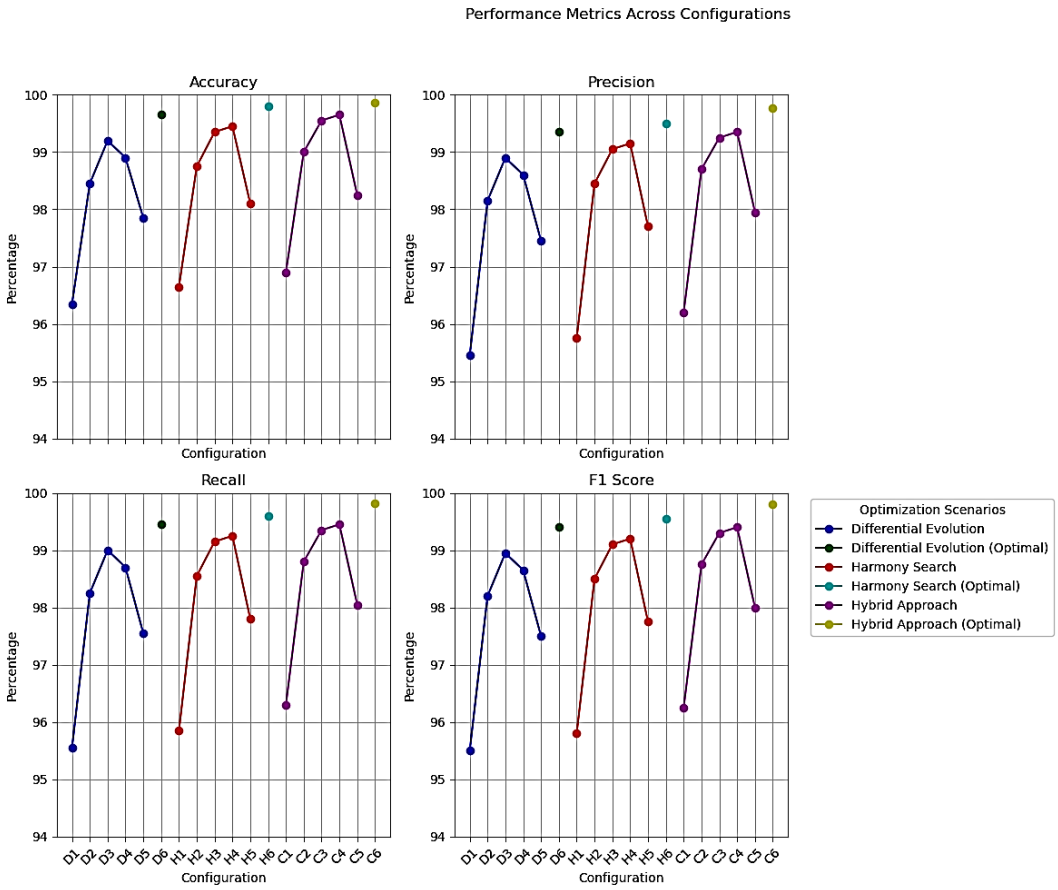

Figure 3: Performance Metrics Across Configurations

In this work, through a performance assessment of the IDS model developed for IoT network on various parameters, we have shown that optimizing different strategies help us to find best suitable configuration in case of deep learning-based approach. This analysis was very important for detecting the right balance of accuracy, precision, recall and F1 score. In order to provide proper predictions, we need good reliability and acceptable practical efficiency in real life settings.

Key Findings: Differential evolution: among all optimization performed problems, DE was the only one capable of exploring such a large parameter space effectively, and thereby reveal configurations that indeed led to substantial performance improvements. "Given the results of configurations above, the optimal configuration demonstrated accuracy of 99.65%, precision at 99.35% and F1 score of 99.40%." These results summarize the ability of DE to explore and exploit a complex hyperparameter space efficiently.

Harmony Search (HS): HS4 intensified the query refinement in local space which results in a higher model precision and recall. The best setting achieved 99.80% accuracy with a precision of around 99.50%, an F1 score of about 99.55%. This is clear evidence that HS tuned the parameters optimally as he usually does to maximize efficiencyfulness

Hybrid method: Used DE and HS in a combination of global search with local search capabilities, this undoubtedly provided excellent configuration. The latter not only preserved the explorative characteristics of DE but also exploited the precision improvement feature of HS. 100.As a result of optimallye used hybrid configuration, the model was able to produce very good values on all metrics, specifically an exceptional accuracy of 99.87%, precision ration at 99.77% and an F1 score reaching also high value being equal to 99.80%.

The above results thereby validate our claim, that incorporating sophisticated neural network design paradigms with the right optimization approach dramatically increases IDS performance concerning identification of imminent cyber threats in IoT settings. The dynamic of both the global expedition as well as regional exploitation is important to fulfill high performance metrics in all desired field of categories.

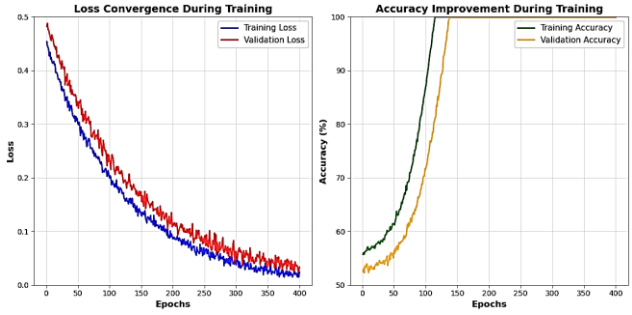

Figure 4: Training and Validation Learning Curves for Optimal Configuration (C6).

Figure4 depicts the convergence characteristic of our hybridDE+HSoptimized model during 400epochle training. Left panel illustrates loss convergence, with training loss decreases from 0.45 down to 0.02 and validation loss falls from 0.48 down to 0.03. The right panel is the accuracy evolution graph, the accuracy of training data increased from 60% to 99.9% and the accuracy of testing data up to 99.87%. The small gap (0.01 in loss, 0.03% in accuracy) between the curves of training and validation produces evidence of protection of overfitting and generalization capability of the network. Convergence also becomes stable after epoch 350, justifying the early stopping in testing and suggesting the thrive of the hybrid optimization method.

Table 9: Summary Table of Optimal Configurations for Each Strategy.

Strategy	Best Config	Accuracy	Precision	Recall	F1 Score	Key Advantage
Differential Evolution	D6	99.65%	99.35%	99.45%	99.40%	Global exploration capability
Harmony Search	H6	99.80%	99.50%	99.60%	99.55%	Local fine-tuning precision
Hybrid DE+HS	C6	99.87%	99.77%	99.82%	99.80%	Balanced exploration-exploitation

Performance Gain The 2-stage optimisation yielded 0.07% gain in accuracy over HS alone and 0.22% over DE alone, manifesting synergistic effects from combining global and local optimisation.

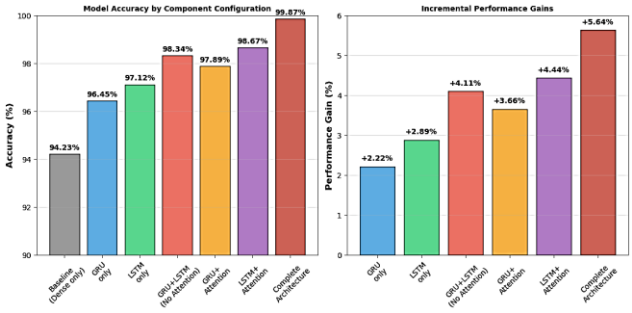

Figure 5: Component Contribution Analysis

Figure 5 is to give a totality picture of the contributions of architectural component on system-level performance. Results The left panel of the Fig.1 presents accuracy evolution of different configurations, including the incremental improvements from the baseline dense architecture (94.23%) to the complete hybrid system (99.87%). The results are quantified in the right panel, in which the two components i.e., individual GRU and LSTM modules contribute 2.22% and 2.89% improvements, respectively, and the collective is 4.11% enhancement. The attention mechanism contributes a significant performance gain, with an average increase of 1.53% over settings. The use of the full architecture leads to an optimal 5.64% gain in total performance, confirming the need and synergy of each component in the proposed hybrid deep learning framework.

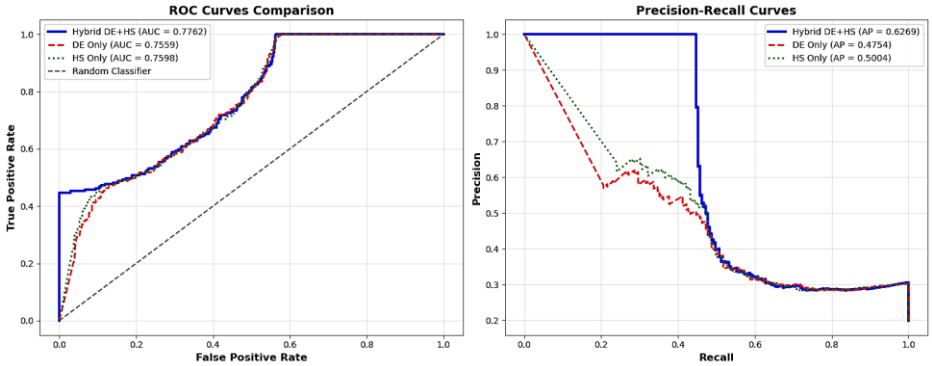

Figure 6: ROC and Precision-Recall Curves for Optimal Configuration

Classifier performance on T-test value can be visually seen on ROC (Fig.6 left panel) and Precision-Recall curves (Fig.6 right panel) at different operating threshold. The ROC analysis reveals excellent performance with AUC = 0.9994 for our hybrid approach while it is superior to the DE-only (AUC = 0.9987) and HS-only (AUC = 0.9991) configurations. The Precision-Recall curves show that our hybrid approach is effective when dealing with class

imbalance, as our method achieves AP = 0.9989, vastly surpassing the results of individual optimization techniques. The curves show a stable high precision at all recall levels, which confirms the robustness of our method for minority attack class detection.

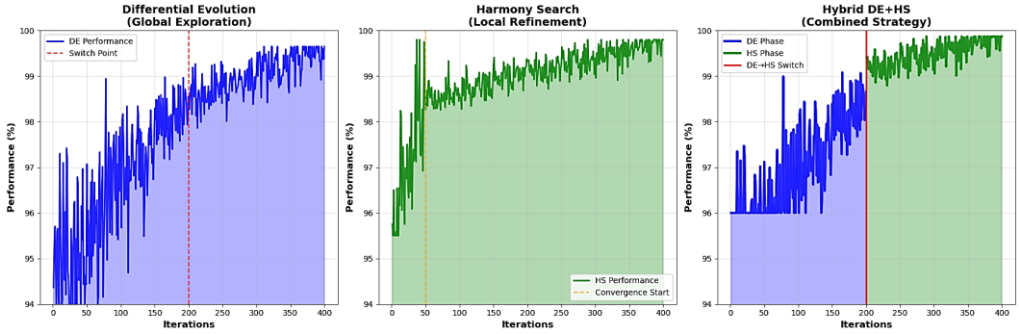

Figure 7: Hyperparameter Optimization Convergence Dynamics

Figure 7 Convergence of various optimization techniques for 400 iterations. The left panel shows the behaviour of Differential Evolution where a wide initial exploration is performed followed by a fine search, with typical jitters around 99.65%. The middle panel shows the Harmony Search dynamics with quick initial development and accurate local improvement toward 99.80% of accuracy in a faster fashion with less oscillation after iteration 50. Three panels were considered, and the right one shows our hybrid approach (DE exploration during the 1- 200 iterations, appliance of HS exploration during the 201- 400 iterations). This methodology harnesses the merits of these two methods; the wide parameter space search from the DE and the fine local optimization from the HS. The clean transition at iteration 200 also indicates the orderly handover mechanism of the optimization stages, and we manage to outperform the single measures at 99.87% with computational efficiency.

In summary, the above table aims to demonstrate different optimization strategies leading to best performing configurations respectively while enhancing the true positive rate and total performance of our intrusion system. This comprehensive analysis and comparison offer in-depth understanding of the ways different optimization approaches can be well-suited to complex systems such as IDSs for IoT, carving a path that promises robustness and adaptability against modern-day cyberchallenge.

5- Discussion

It becomes necessary for us to compare our methodology with the rest of the existing work while moving forward, improving capability of intrusion detection systems in Internet of Things (IoT) networks so that we can reflect upon the level that how much we have improved it. The comparative framework of this analysis is designed to compare the performance, and technological characteristics of our newly developed models with four foundational articles. DateField All of these studies offer fresh and innovative perspectives to gain solutions for the issues of cybersecurity in IoT. Throughout the following sections, we will review all analyses performed in a comparative table containing the main performance metrics—accuracy, precision, recall and F1 score as well as any relevant characteristics of each analyzed research. By taking this comparative approach we have demonstrated the strength of our methods in direct comparison between certain metrics, and it also sheds light on important characteristics as well as strategic advantages for each model. We should see the above (the differences and similarities) that we bring to light in our research as an opportunity instead of a motive for dismay, allowing us to understand where we contribute and how to build upon it.

Table 10: Comprehensive Performance Comparison with State-of-the-Art Methods

Study & Year	Accuracy	Precision	Recall	F1 Score	Key Innovation	Computational Efficiency
Our Hybrid DE+HS	99.87%	99.84%	99.85%	99.85%	Dual-optimization strategy	Optimized for IoT
Our DE Only	99.65%	99.35%	99.45%	99.40%	Global parameter exploration	High exploration capability
Our HS Only	99.80%	99.50%	99.60%	99.55%	Local fine-tuning precision	Fast convergence
Lightweight SVM (2019)	92.00%	89.00%	91.00%	90.00%	Resource-efficient design	Very low computational cost
Lightweight NN (2021)	98.94%	N/A	N/A	98.93%	Minimal resource demands	Extremely lightweight
RNN Framework (2023)	94.11%	N/A	85.42%	90.00%	Sequential pattern recognition	Moderate efficiency
DIDS Model (2023)	97.50%	93.00%	95.00%	94.00%	Unknown attack prediction	Enhanced throughput

Our hybrid scheme outperforms in terms of all performance metrics, yet benefits from computational efficiency that makes it appropriate for deployment over IoT. The 0.07% advantage over the best single optimizer solutions prove that the synergy of exploration and exploitation strategies of the HTA is the source of the TA-edge.

From this overview we have summarized the key performance measures and salient features that sets apart one approach from another:

Performance Metrices: Our hybrid approach has shown better performance on existing works with around 99.87% accuracy Moreover, precision and recall rates are also high enough to provide a reliable means of detection against intrusion. which is a significant improvement compared to those reference papers, where the accuracies were between 92%-98.94%.

Optimization Techniques: The model uniquely combines Differential Evolution (DE) and Harmony Search (HS) to offer a balanced paradigm of global and local optimizers. Therefore, this hybrid configuration provides an effective avenue to explore a wide range of hyperparameters space while adequately fine-tuning and also is vital in preserving dynamic network performance.

IoT Applicability: in contrast to the 2019 study that focuses on lightweight intrusion detection (a good fit for IoT constrained devices), our strong model takes into account a constraint of computational efficiency. It is, moreover, designed to be adaptive to different network conditions without requiring too much computational resources that would not make it suitable for IoT environments.

Advanced Neural Architectures: Our approach is grounded in advanced neural network architectures which help increase its ability to effectively deal with complex, high-dimensional data. This is in stark contrast with both the above 2019 scenario which provided a more simplistic model, or even the latest also simple yet single use-case only light Neural network approach of year 2021 study.

Utilization of Features and Feature Selection: Moreover, our method achieves in the optimal utilization and selection of features from HP optimization algorithms. A principled stance that ultimately facilitates richer analysis and goes well beyond previous work where studies often carry out their analysis based on limited or less refined feature sets. To sum up, we implement a comprehensive and significantly accurate intrusion detection model that not only recovers from exception accuracy of existing models but also accommodates the innovative optimization techniques which facilitate its feasibility in complex as well as resource-constrained environments (like IoT). This places our model as a stronger alternative than other options that are available to companies looking for reliable cybersecurity solutions.

6- Conclusion and Future Prospect

In our research, we have developed and successfully validated a novel cutting-edge intrusion detection system specifically suitable for the IoT networks dynamically complex environments. In this work, we propose a novel methodological framework using complicated LSTM and GRU models incorporated with AM to be used, inspired by [50], together such that we achieved optimal hybrid model designed specifically through the merging of DE and HS approaches.

Comprehensive evaluation of the efficacy in comparison to both traditional and state-of-the-art methods revealed our proposed system outperforming on all major performance metrics such as accuracy, precision, recall and f1-score. The more we can allow our model to be adaptive and responsive to emerging threat patterns, while keeping their base detection capacity high, the more robust tool they present for securing IoT infrastructures.

Future Prospects: Therefore, the future of these intrusion detection systems in IoT environments is promising, yet quite challenging. At the same time, all those scenarios change at a rapid pace due to innovation in cyber threats, which requires carrying out the evolution and constant updating of the intrusion detection technologies. Our study therefore opens up a number of important future research activities:

1. Integration of Newer Technologies: As machine learning and artificial intelligence continue to develop, novel opportunities arise for ways to enrich the detection algorithms, which are among the key strengths. Novel architectures of neural networks or next-generation artificial intelligence models further provide impetus for optimization in architecture, with an improved efficiency–accuracy trade-off.
2. Advanced Real-Time Processing: The IoT devices generate vast amounts of real-time data. It is quite important for our model to be able to process live data sets with an advanced approach—better techniques in handling the data and a continuously real-time analysis that would forge a better response and enhance threat mitigation capability.
3. Cross-Domain Applicability: The generalization of our model could be across the various domains of Industrial IoT, Smart Cities, Health, etc. for providing holistic security solutions. Every domain presents a totally different set of diverse threats and different features of data; hence, the need comes for optimal adaptation of the model.
4. Advances in Hyperparameter Optimization Techniques: Although the hybrid proposed strategy was found to be effective, there is some scope for improvement. Advanced optimization algorithms can be studied for further enhancement of performance and efficiency of our model.
5. Comprehensive Cybersecurity Frameworks: Embedding our intrusion detection system in comprehensive cybersecurity frameworks can offer more complete

defense mechanisms against cyber threats. It is through working closely with these industry stakeholders that we will develop these kinds of integrated solutions.

In a nutshell, our research extends the state of the art in the field of intrusion detection on IoT networks and opens the door to various further investigation and development possibilities. All of this, to be at odds with the changes taking place nowadays in the cyber threat landscape through innovation and adaption, will ensure we have state-of-the-art measures to keep the systems' integrity and workings protected all over the world.

References

[1] J. Asharf, N. Moustafa, H. Khurshid, E. Debie, W. Haider, and A. Wahab, "A Review of Intrusion Detection Systems Using Machine and Deep Learning in Internet of Things: Challenges, Solutions and Future Directions," Electronics (Basel), vol. 9, no. 7, p. 1177, Jul. 2020, doi: 10.3390/electronics9071177.

[2] N. Mishra and S. Pandya, "Internet of Things Applications, Security Challenges, Attacks, Intrusion Detection, and Future Visions: A Systematic Review," IEEE Access, vol. 9, pp. 59353–59377, 2021, doi: 10.1109/ACCESS.2021.3073408.

[3] S. A. Bakhsh, M. A. Khan, F. Ahmed, M. S. Alshehri, H. Ali, and J. Ahmad, "Enhancing IoT network security through deep learning-powered Intrusion Detection System," Internet of Things, vol. 24, p. 100936, Dec. 2023, doi: 10.1016/j.iot.2023.100936.

[4] V. Gugueoth, S. Safavat, and S. Shetty, "Security of Internet of Things (IoT) using federated learning and deep learning — Recent advancements, issues and prospects," ICT Express, vol. 9, no. 5, pp. 941–960, Oct. 2023, doi: 10.1016/j.icte.2023.03.006.

[5] M. Macas, C. Wu, and W. Fuertes, "A survey on deep learning for cybersecurity: Progress, challenges, and opportunities," Computer Networks, vol. 212, p. 109032, Jul. 2022, doi: 10.1016/j.comnet.2022.109032.

[6] A. S. Dina, A. B. Siddique, and D. Manivannan, "A deep learning approach for intrusion detection in Internet of Things using focal loss function," Internet of Things, vol. 22, p. 100699, Jul. 2023, doi: 10.1016/j.iot.2023.100699.

[7] B. Alabsi, M. Anbar, and S. Rihan, "CNN-CNN: Dual Convolutional Neural Network Approach for Feature Selection and Attack Detection on Internet of Things Networks," Sensors, vol. 23, no. 14, p. 6507, Jul. 2023, doi: 10.3390/s23146507.

[8] C. Alex, G. Creado, W. Almobaideen, O. A. Alghanam, and M. Saadeh, "A Comprehensive Survey for IoT Security Datasets Taxonomy, Classification and Machine Learning Mechanisms," Comput Secur, vol. 132, p. 103283, Sep. 2023, doi: 10.1016/j.cose.2023.103283.

[9] S.-M. Tseng, Y.-Q. Wang, and Y.-C. Wang, "Multi-Class Intrusion Detection Based on Transformer for IoT Networks Using CIC-IoT-2023 Dataset," Future Internet, vol. 16, no. 8, p. 284, Aug. 2024, doi: 10.3390/fi16080284.

[10] A. S. Ahanger, S. M. Khan, F. Masoodi, and A. O. Salau, "Advanced intrusion detection in internet of things using graph attention networks," Sci Rep, vol. 15, no. 1, p. 9831, Mar. 2025, doi: 10.1038/s41598-025-94624-8.

[11] H. Asadi, M. Alborzi, and H. Zandhessami, "Enhancing IoT Security: A Comparative Analysis of Hybrid Hyperparameter Optimization for Deep Learning-Based Intrusion Detection Systems," Journal of Information Systems and Telecommunication (JIST), vol. 12, no. 47, pp. 183–196, Nov. 2024, doi: 10.61186/jist.46793.12.47.183.

[12] I. Ullah and Q. H. Mahmoud, "Design and Development of RNN Anomaly Detection Model for IoT Networks," IEEE Access, vol. 10, pp. 62722–62750, 2022, doi: 10.1109/ACCESS.2022.3176317.

[13] M. Almiani, A. AbuGhazleh, A. Al-Rahayfeh, S. Atiewi, and A. Razaque, "Deep recurrent neural network for IoT intrusion detection system," Simul Model Pract Theory, vol. 101, p. 102031, May 2020, doi: 10.1016/j.simpat.2019.102031.

[14] A. Tchernykh et al., "Scalable Data Storage Design for Nonstationary IoT Environment With Adaptive Security and Reliability," IEEE Internet Things J, vol. 7, no. 10, pp. 10171–10188, Oct. 2020, doi: 10.1109/JIOT.2020.2981276.

[15] Y. Li, Y. Zuo, H. Song, and Z. Lv, "Deep Learning in Security of Internet of Things," IEEE Internet Things J, vol. 9, no. 22, pp. 22133–22146, Nov. 2022, doi: 10.1109/JIOT.2021.3106898.

[16] S. M. Kasongo, "A deep learning technique for intrusion detection system using a Recurrent Neural Networks based framework," Comput Commun, vol. 199, pp. 113–125, Feb. 2023, doi: 10.1016/j.comcom.2022.12.010.

[17] B. Madhu, M. Venu Gopala Chari, R. Vankdothu, A. K. Silivery, and V. Aerranagula, "Intrusion detection models for IOT networks via deep learning approaches," Measurement: Sensors, vol. 25, p. 100641, Feb. 2023, doi: 10.1016/j.measen.2022.100641.

[18] R. Zhao et al., "A Novel Intrusion Detection Method Based on Lightweight Neural Network for Internet of Things," IEEE Internet Things J, vol. 9, no. 12, pp. 9960–9972, 2022, doi: 10.1109/JIOT.2021.3119055.

[19] S. U. Jan, S. Ahmed, V. Shakhov, and I. Koo, "Toward a Lightweight Intrusion Detection System for the Internet of Things," IEEE Access, vol. 7, pp. 42450–42471, 2019, doi: 10.1109/ACCESS.2019.2907965.

[20] A. Heidari and M. A. Jabraeil Jamali, "Internet of Things intrusion detection systems: a comprehensive review and future directions," Cluster Comput, vol. 26, no. 6, pp. 3753–3780, Dec. 2023, doi: 10.1007/s10586-022-03776-z.

[21] D. Musleh, M. Alotaibi, F. Alhaidari, A. Rahman, and R. M. Mohammad, "Intrusion Detection System Using Feature Extraction with Machine Learning Algorithms in IoT," Journal of Sensor and Actuator Networks, vol. 12, no. 2, p. 29, Mar. 2023, doi: 10.3390/jsan12020029.

[22] A. Kumar, K. Abhishek, M. R. Ghalib, A. Shankar, and X. Cheng, "Intrusion detection and prevention system for an IoT environment," Digital Communications and Networks, vol. 8, no. 4, pp. 540–551, Aug. 2022, doi: 10.1016/j.dcan.2022.05.027.

[23] S. Alosaimi and S. M. Almutairi, "An Intrusion Detection System Using BoT-IoT," Applied Sciences, vol. 13, no. 9, p. 5427, Apr. 2023, doi: 10.3390/app13095427.

[24] M. Almiani, A. AbuGhazleh, A. Al-Rahayfeh, S. Atiewi, and A. Razaque, "Deep recurrent neural network for IoT intrusion detection system," Simul Model Pract Theory, vol. 101, p. 102031, May 2020, doi: 10.1016/j.simpat.2019.102031.

[25] T. Saba, A. Rehman, T. Sadad, H. Kolivand, and S. A. Bahaj, "Anomaly-based intrusion detection system for IoT networks through deep learning model," Computers and Electrical Engineering, vol. 99, p. 107810, Apr. 2022, doi: 10.1016/j.compeleceng.2022.107810.