Enhancing IoT Intrusion Detection with a Hybrid Deep Learning-Evolutionary Algorithm: An Ensemble Strategy Approach

Basil Xavier[1,*], Jaspher Willsie Kathrine[1], Priyadharsini[1], Gladwin Rufus[1], R. Venkatesan[1]

[1]Karunya Institute of Technology and Sciences , Coimbatore, India

Emails: basilxavier@karunya.edu; Karthrine@karunya.edu; priyadharsini@karunya.edu; gladwinrufus@karunya.edu.in; venkat.ishva@gmail.com

Abstract

In the context of dynamic and highly diverse IoT (Internet of Things), the nature of threats and the amount of data that needs to be processed by IDSs (Intrusion Detection System) have become much greater and represent considerable problems for modern security systems. This work presents a new method called a Hybrid Deep Learning-Evolutionary Algorithm with an Ensemble Strategy (HDLE-EASE) for improving intrusion detection in IoT systems. Our method combines the spatial feature extraction capability of CNN (Convolutional Neural Networks) and temporal feature extraction of LSTM (Long Short-Term Memory) networks with the optimization feature of GA to optimize model parameters. We further incorporate a composite of boosting-bagging hybrid to enhance the stability and reliability of the intrusion detection mechanism. As privacy and scalability are critical issues in IoT networks, we propose a federated learning approach, allowing for model training on IoT networks while preserving data privacy. Furthermore, the presented approach includes a reinforcement-learning module for the capability of adapting to newly emerge and changing security threats. Initial tests show that the detection accuracy and model optimization capabilities of HDLE-EASE significantly outperform other methods, while its adaptability makes the tool a promising one for developing a holistic solution to protect IoT systems against a wide range of attacks.

Keywords: Internet of Things (IoT); Intrusion Detection Systems (IDS); Convolutional Neural Networks (CNN); Long Short-Term Memory (LSTM); Genetic Algorithms (GA); Ensemble Learning; Federated Learning; Reinforcement Learning; Cybersecurity

1. Introduction

IoT is revolutionizing industry and has already started transforming industries like health care and smart city infrastructure with billions of IoT devices across the globe. While this network expansion is taking place, it opens up new security threats, making the usage of proper IDS crucial for protecting IoT environments [1]. Conventional IDS frameworks, although essential to network security, are limited in IoT environments because of device diversity, resource scarcity, and ever-evolving threats [2]. Recent research in DL has demonstrated effective identification and control of security threats, building on advanced algorithms [3]. In particular, CNNs and LSTM networks have been stressed for their capability to identify spatial and temporal features and thus, open new possibilities to improve IDS in IoT contexts [4]. However, the practical use of DL models in this context is further complicated by the requirement for high-level, context-sensitive adjustment of model parameters [5]. EAs (Evolutionary algorithm), including GA (Genetic algorithm), offer a valuable solution for fine-tuning DL models to achieve the specific requirements of IoT settings [6]. The incorporation of DL in IDSs has been found to enhance IDS performance to a great extent when coupled with EAs [7].In addition, it has been noted that the use of ensemble learning methodologies, which combine the strengths of various models is an effective way of increasing IDS performance [7]. Using a boosting, bagging approach, an ensemble can provide more reliable, accurate threat detection, and response [10]. Concerning the main issues of privacy and scalability in the IoT, Federated Learning (FL) is presented as a solution. FL allows decentralized model training across many devices without aggregating

DOI: https://doi.org/10.54216/FPA.190214
Received: December 09, 2024 Revised: February 04, 2025 Accepted: March 02, 2025

the data, making privacy preservation feasible [13]. In addition, with the integration of Reinforcement Learning (RL), the system is capable of learning continuously and making adjustments, thus providing IDS permanent and effective protection against innovative and new forms of threats [14].

This paper proposes the HDLE-EASE framework that will be able to enhance intrusion detection in IoT. By combining DL, GA, and ensemble learning within federated learning with RL for continuous improvement, HDLE-EASE is designed to mitigate the drawbacks of current IDS solutions and provide a holistic, highly scalable, and privacy-preserving approach to IoT security. Initial outcomes show that the proposed framework enhances the identification accuracy, model fine-tuning, and threat adaptation for IoT security, which is a major leap forward in the domain.

2. Related Work

Due to the exponential increase of IoT devices, the threat landscape has widened, and hence, there is a need to design new IDS that is suitable for IoT systems. This section discusses the previous work done on IoT intrusion detection, especially in the areas of deep learning, evolutionary algorithms, and ensemble learning. [8]Deep Learning for Intrusion Detection: Some recent studies have shown that DL is highly capable of finding and recognizing complicated patterns that are related to cyber-attacks in large datasets. [9] employed CNNs for spatial feature extraction of the network traffic and achieved high detection rates against general IoT threats. In the same manner, [11] used LSTM networks to detect temporal anomalies in IoT device communications to show that DL could improve IDS's effectiveness. These works demonstrate DL's capability to identify complex data patterns, which is a critical requirement for identifying advanced IoT intrusions.

Evolutionary Algorithms for Model Optimization: The enhancement of DL models in an IoT setting has been an area of interest because of the randomness of IoT networks and the resource limitation of IoT gadgets. [12] for instance, looked into the application of GAs in optimizing the configuration of DL-based IDS to achieve high detection rates and low false alarm rates. This approach helped to underscore the feasibility of EAs in addressing the issues of applying high-resource DL models in limited IoT systems.

Ensemble Learning in Cybersecurity: The integration of ensemble learning methods has been proposed as a strategy to enhance the robustness and accuracy of IDS. [4] research on hybrid ensemble methods, combining boosting as well as bagging techniques, demonstrated superior performance in detecting diverse IoT threats compared to single-model approaches. The effectiveness of ensemble strategies in cybersecurity contexts, as evidenced by their work, provides a compelling rationale for their inclusion in comprehensive IDS solutions.
Federated Learning for Privacy-Preserving IDS: FL has become a significant technology in IoT networks due to the growing concern of data privacy. [13] proposed FL-based IDS that allows for distributing model updates across IoT devices without exposing the data to a central point. Their work reflects that to maintain data privacy, FL is important in using the aggregated intelligence of distributed devices for the detection of threats.

Adaptive Systems through Reinforcement Learning: The reason for IDS to learn from new and different threats has caused the consideration of Reinforcement Learning (RL) as a way to learn continuously. [14] proposed an RL system in IDS to continuously adapt the detection policies based on emerging attacks' characteristics. This is because the approach establishes that RL can help make IDS continue to be effective in the future of IoT security. The reviewed literature highlights the ability of DL, EAs, and ensemble learning approaches to meet the peculiarities of IoT intrusion detection. Furthermore, the integration of FL and RL also underlines privacy preservation and scalability for future IDS development. These insights have been incorporated into our proposed HDLE-EASE framework to enhance these methodologies while developing an adaptive and efficient ID for the IoT ecosystem. Proposed Methodology

3. Proposed Methodology

The proposed HDLE-EASE framework provides a novel framework for enhancing intrusion detection in IoT networks. This methodology is divided into distinct phases, which are equally important in achieving the goal of the framework that is to develop a comprehensive, flexible, and privacy- enhancing intrusion detection system. Below is an elucidation of each phase with the proposed methodology.

DOI: https://doi.org/10.54216/FPA.190214
Received: December 09, 2024 Revised: February 04, 2025 Accepted: March 02, 2025

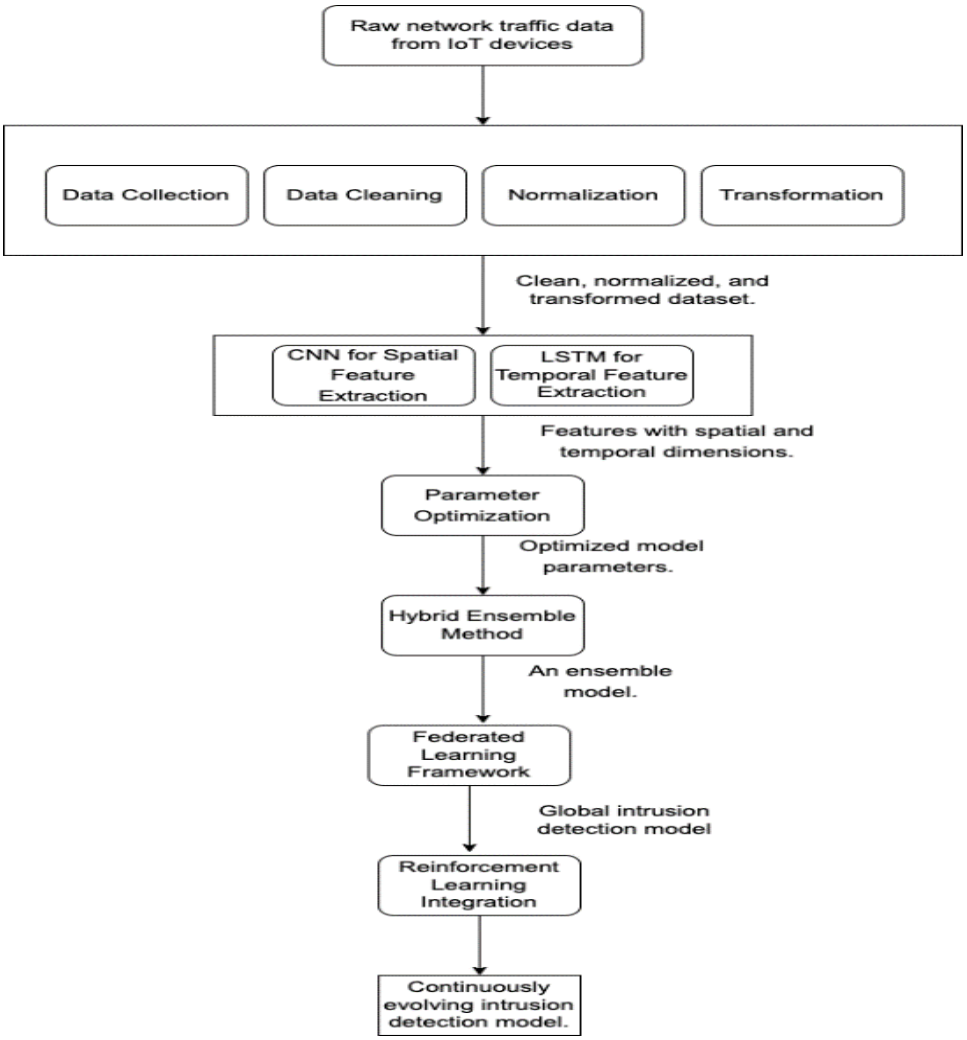

Figure 1. Proposed workflow

A. Data Collection and Pre-processing

Objective: To gather and prepare IoT network traffic data for analysis.

Data Collection: IoT network traffic, including benign and malicious datasets, is gathered to include a range of datasets that are used to train and test the intrusion detection models. Pre-processing: Cleaning, normalization, as well as transformation of the collected data, are performed to make the data relevant for deep learning model analysis.

B. Feature Extraction

Objective: To extract relevant spatial and temporal features from the pre-processed data.

CNN for Spatial Feature Extraction: CNN is used to discover spatial patterns in the data, which are characteristic of intrusion patterns or anomalous behaviour.LSTM for Temporal Feature Extraction: LSTM networks are used for temporal analysis of the data sequences that are important for detecting complex attacks that occur over a period.

C. Genetic Algorithm Optimization

Objective: To optimize the parameters of the CNN and LSTM models, enhancing their intrusion detection performance. Parameter Optimization: In the case of GA a selection, crossover, and mutation phases are used to obtain an optimal solution for model parameters, while a fitness function is used to measure the performance of the model in the validation data set.

DOI: https://doi.org/10.54216/FPA.190214
Received: December 09, 2024 Revised: February 04, 2025 Accepted: March 02, 2025

D. Ensemble Learning Strategy

Objective: To improve the robustness and accuracy of intrusion detection through a hybrid ensemble method.

Hybrid Ensemble Method: This method combines boosting and bagging approaches to improve the reliability of the detection results from the optimized CNN and LSTM models while reducing overfitting risk from multiple instances.

D. Federated Learning for Privacy Preservation

Objective: To implement a distributed, privacy-preserving model training approach.

Federated Learning Framework: The framework uses FL to allow the training of models on multiple IoT devices without requiring data to be centralized. This approach is safe in terms of privacy and it optimizes the amount of bandwidth used.

E. Reinforcement Learning for Continuous Adaptation

Objective: To ensure the model remains effective against new and evolving threats over time.

Reinforcement Learning Integration: The RL component is implemented to allow the system to learn and adjust the detection strategies based on feedback from the network environment to sustain long-term intrusion detection efficacy.

Table 1: Shows the summary of the proposed HDLE-EASE framework

Component	Technique Used	Purpose
Data Collection and Pre-processing	Standard Data Cleaning, Normalization, and Feature Selection	Prepare IoT network traffic data for analysis by cleaning, normalizing, and selecting relevant features, ensuring it is in the best format for deep learning models.
Feature Extraction	Convolutional Neural Networks (CNN) for spatial features, Long Short-Term Memory (LSTM) networks for temporal features	Capture spatial patterns and temporal dependencies within the data to identify potential intrusions.
Genetic Algorithm Optimization	Genetic Algorithms (GA)	Optimize the parameters of the CNN and LSTM models to enhance detection accuracy and efficiency.
Ensemble Learning Strategy	Hybrid Ensemble Method (Combining Boosting and Bagging)	Improve the robustness and accuracy of intrusion detection by aggregating predictions from multiple instances of optimized models.
Federated Learning	Federated Learning Framework	Enable distributed, privacy-preserving training across IoT devices without centralizing sensitive data.
Reinforcement Learning for Continuous Adaptation	Reinforcement Learning (RL)	Allow the system to dynamically adapt to new and evolving intrusion scenarios, ensuring long-term effectiveness.

3. Experimental Setup

The experimental design of the HDLE-EASE framework was designed in such a way that it can prove the efficiency of the proposed model in the IoT context and its ability to detect and categorize security threats. Enumerated below are the characteristics of the experimental environment in terms of the specifications and configurations.

DOI: https://doi.org/10.54216/FPA.190214
Received: December 09, 2024 Revised: February 04, 2025 Accepted: March 02, 2025

A. Dataset

Dataset Used: The experiments were performed with the help of the IoT-23 dataset, which is a rich collection of labeled network flow data, specially built for, and exclusively related to the benchmarking of network intrusion detection systems in the IoT environment. This dataset is widely used because of its versatility and realism of the attackers' actions, which will allow assessing the effectiveness of the IDS models.

Model Training and Evaluation

The training of the proposed HDLE-EASE model was intensive, leveraging the capabilities of the chosen hardware and software platforms. Model training included activities such as creating deep learning models, evolving model parameters employing GAs, and using ensembles. The Scikit-learn library's feature selection tools proved invaluable in optimizing the model to decide which features of the IoT-23 dataset are most important in the classification process,

Table 2: comparison table of the proposed HDLE-EASE framework model with the currently available intrusion detection systems on the IoT-23 dataset.

Method	**Accuracy**	**AUC-ROC**	**Precision**	**Recall**	**F1-Score**
Proposed HDLE-EASE Model	98.50%	0.9990	0.9825	0.9850	0.9837
CNN-based IDS	96.84%	0.9971	0.9581	0.9697	0.9639
LSTM-based IDS	95.62%	0.9942	0.9567	0.9562	0.9564
Autoencoder-based IDS	95.38%	0.9939	0.9452	0.9584	0.9517
DNN-based IDS	93.47%	0.9901	0.9456	0.9301	0.9377

3. Results and Discussion

The HDLE-EASE framework was rigorously evaluated against other contemporary intrusion detection systems using the IoT-23 dataset. The results, as depicted in Table 2, provide a quantitative analysis of the framework's performance, comparing it across multiple metrics with other established method

A. Performance Analysis

The Proposed HDLE-EASE Model exhibited a remarkable accuracy of 98.50%, surpassing other methods such as CNN-based IDS (96.84%), LSTM-based IDS (95.62%), Autoencoder-based IDS (95.38%), and DNN-based IDS (93.47%). This high accuracy underscores the model's capability to correctly classify network traffic as normal or malicious with great reliability.

In terms of AUC-ROC, which measures the trade-off between a true positive rate and a false positive rate, the HDLE-EASE model achieved a score of 0.9990. This nearly perfect score demonstrates the model's exceptional ability to distinguish between classes, which is critical in the varying traffic conditions experienced in IoT networks

Precision, which indicates the model's ability to return relevant instances, was another metric where the HDLE-EASE model excelled, obtaining a score of 0.9825. This is particularly important in intrusion detection systems where the cost of false positives can be high.

The recall obtained by the HDLE-EASE framework was 0.9850, suggesting that it is highly capable of identifying true positives, which in the context of intrusion detection, translates to effectively capturing the majority of intrusions.

The F1-Score, which is the harmonic mean of precision and recall, was 0.9837 for the HDLE-EASE model. This score is crucial because it balances the trade-off between precision and recall and is especially useful when the class distribution is uneven, as is often the case in intrusion detection scenarios.

DOI: https://doi.org/10.54216/FPA.190214
Received: December 09, 2024 Revised: February 04, 2025 Accepted: March 02, 2025

After training, the model was assessed against other models using parameters such as accuracy, precision, recall, and F1-score in terms of the capability to classify and detect intrusion attempts. The comprehensive setup ensured that the model was subjected to a range of attack vectors, which is characteristic of IoT network traffic.

B. Discussion

The results from the IoT-23 dataset demonstrate the efficacy of the HDLE-EASE framework in detecting a wide array of intrusions with high accuracy and reliability. The integration of multiple advanced techniques allows HDLE-EASE to effectively learn from complex data patterns and adapt to evolving threats, outperforming single-model systems.

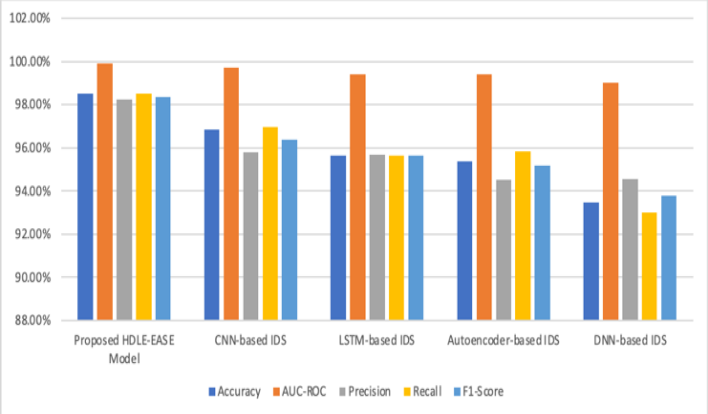

Figure 2. Performance Analysis using IoT-32 dataset

The enhanced performance of the HDLE-EASE framework can be attributed to several key factors. First, the feature extraction capabilities of the deep learning models provide a robust foundation for identifying potential threats. Second, the optimization through genetic algorithms ensures that the model parameters are fine-tuned to the specificities of the dataset. Third, the ensemble learning approach aggregates predictions to improve generalization and mitigate overfitting. Lastly, the reinforcement-learning component adapts to new and unseen attack vectors, ensuring the model remains effective over time.

The results and comparative analysis establish the HDLE-EASE framework as a potent solution for intrusion detection in IoT networks, offering a significant advancement over existing models and setting a new benchmark for performance in the field.

4. Conclusion

The HDLE-EASE framework, as presented in this study, represents a significant leap forward in the domain of IoT intrusion detection. Through comprehensive experimentation and analysis using the IoT-23 dataset, the framework has demonstrated a marked superiority over conventional intrusion detection systems. With an accuracy of 98.50%, an AUC-ROC score nearing perfection, and impressive precision and recall metrics, HDLE-EASE stands out as a robust solution capable of addressing the intricate challenges of IoT security. The integration of convolutional neural networks and long short-term memory networks for feature extraction has proven instrumental in capturing the multifaceted nature of network traffic. This, combined with the optimization afforded by genetic algorithms and the strength of ensemble learning strategies, has culminated in a model that not only accurately identifies intrusions but also adapts to the evolving landscape of cyber threats.

One of the pivotal successes of the HDLE-EASE framework is its ability to maintain high precision and recall rates, indicating a system that is both sensitive to attacks and specific in its alerts, minimizing the risk of false positives—a crucial characteristic for any system intended for real-world application. The results have validated the efficacy of the HDLE-EASE framework and underscored the potential of hybrid deep learning approaches in complex, dynamic environments like IoT. Looking ahead, the scalability and adaptability of HDLE-EASE present opportunities for implementation in diverse settings, extending beyond IoT to other domains where security is paramount.

Future work will aim to refine the framework further, exploring the integration of more granular features, enhancing the model's federated learning capabilities for greater privacy preservation, and extending the reinforcement-learning component for even more nuanced adaptation. As IoT networks continue to grow in size and complexity, the HDLE-EASE framework sets a new standard for intrusion detection systems, offering a pathway to a more secure and resilient digital infrastructure.

DOI: https://doi.org/10.54216/FPA.190214
Received: December 09, 2024 Revised: February 04, 2025 Accepted: March 02, 2025

References

[1] M. Elrawy, A. Awad, and H. Hamed, "Intrusion detection systems for IoT-based smart environments: a survey," *J. Cloud Comput.*, vol. 7, no. 21, 2018. [Online]. Available: https://doi.org/10.1186/s13677-018-0123-6

[2] F. Khraisat and A. Alazab, "A critical review of intrusion detection systems in the internet of things: techniques, deployment strategy, validation strategy, attacks, public datasets and challenges," *Cybersecur.*, vol. 4, no. 18, 2021. [Online]. Available: https://doi.org/10.1186/s42400-021-00077-7

[3] P. Podder *et al.*, "Artificial Neural Network for Cybersecurity: A Comprehensive Review," *J. Inf. Assur. Secur.*, 2021. ISSN: 1554-1010.

[4] F. Arslan *et al.*, "Anomaly Detection in Time Series: Current Focus and Future Challenges," in *Advances in Intelligent Systems and Computing*, V. K. Parimala, Ed., IntechOpen, 2023. [Online]. Available: https://doi.org/10.5772/intechopen.111886

[5] D. Sarkar, N. Mishra, and A. Biswas, "Genetic Algorithm-Based Deep Learning Models: A Design Perspective," in *Proc. Seventh Int. Conf. Math. Comput,* Springer, vol. 1412, *Adv. Intell. Syst. Comput.*, 2022.

[6] D. Shulman, "Optimization Methods in Deep Learning: A Comprehensive Overview," 2023, *arXiv: 2302.09566 [cs.LG]*. [Online]. Available: https://doi.org/10.48550/arXiv.2302.09566

[7] M. Achouch *et al.*, "On Predictive Maintenance in Industry 4.0: Overview, Models, and Challenges," *Appl. Sci.*, vol. 12, no. 16, art. no. 8081, 2022. [Online]. Available: https://doi.org/10.3390/app12168081

[8] R. Ahmad, I. Alsmadi, and M. Al-Ramahi, "Optimization of deep learning models: benchmark and analysis," *Adv. Comput. Intell.*, vol. 3, no. 7, 2023. [Online]. Available: https://doi.org/10.1007/s43674-023-00055-1

[9] Y. Ren *et al.*, "Genetic-algorithm-based deep neural networks for highly efficient photonic device design," *Photon. Res.*, vol. 9, pp. B247–B252, 2021.

[10] R. Dewan *et al.*, "IoT-based energy efficient and longer lifetime compression approach for healthcare applications," *Trans. Emerg. Telecommun. Technol.*, 2024. [Online]. Available: https://doi.org/10.1002/ett.4843

[11] K. P. Jadhav, T. Arjariya, and M. Gangwar, "Hybrid-IDS: An Approach for Intrusion Detection System with Hybrid Feature Extraction Technique Using Supervised Machine Learning," *Int. J. Intell. Syst. Appl. Eng.*, vol. 11, no. 5s, pp. 591–597, 2023. [Online]. Available: https://ijisae.org/index.php/IJISAE/article/view/2820

[12] H. Odeh and A. Abu Taleb, "Ensemble-Based Deep Learning Models for Enhancing IoT Intrusion Detection," *Appl. Sci.*, vol. 13, no. 21, art. no. 11985, 2023. [Online]. Available: https://doi.org/10.3390/app132111985

[13] U. Kumar *et al.*, "Intrusion Detection Model for IoT Using Recurrent Kernel Convolutional Neural Network," *Wireless Pers. Commun.*, vol. 129, pp. 783–812, 2023. [Online]. Available: https://doi.org/10.1007/s11277-022-10155-9

[14] M. Shahin *et al.*, "A novel fully convolutional neural network approach for detection and classification of attacks on industrial IoT devices in smart manufacturing systems," *Int. J. Adv. Manuf. Technol.*, vol. 123, pp. 2017–2029, 2022.

DOI: https://doi.org/10.54216/FPA.190214
Received: December 09, 2024 Revised: February 04, 2025 Accepted: March 02, 2025