A Robust Hybrid Approach for Intrusion Detection in Dynamic and Heterogeneous IoT Environments

ARTICLE INFO

Received:
Accepted:

ABSTRACT

The rapid growth of the Internet of Things (IoT) has led to the deployment of billions of heterogeneous devices, significantly increasing the attack surface of IoT networks and exposing them to a wide range of cybersecurity threats. In this context, intrusion detection systems (IDS) play a key role in securing these environments. However, existing IDS solutions suffer from several limitations, including reduced generalization capabilities, poor adaptability to dynamic environments, sensitivity to class imbalance, and difficulties in processing massive volumes of heterogeneous data. In this article, we propose a hybrid approach to binary traffic classification (normal and malicious), specifically designed for dynamic IoT environments. This approach combines AutoEncoders (AE), convolutional neural networks (CNN), bidirectional long short-term memory (BiLSTM) neural net- works, and an attention mechanism. The model is based on a complete preprocessing pipeline that integrates feature selection based on five complementary techniques, namely analysis of variance (ANOVA), mutual information (MI), Kendall's rank correlation, minimum redundancy maximum relevance (mRMR), and the chi-square test. Outliers were handled using winsorization based on the interquartile range (IQR), and class imbalance was corrected by generating synthetic data using AE. The experiments were conducted on the massive NF-ToN-IoT-V2 dataset, comprising more than 13,135,881 observations divided into ten traffic classes. The results demonstrate the effectiveness of the AE–CNN–BiLSTM–Att model, achieving, with a first subset of 28 features (subset 28), an accuracy of 97.82%, a sensitivity of 97.87%, a specificity of 97.68%, a precision of 99.11%, an F1- score of 98.49%, and a Matthews correlation coefficient (MCC) of 94.60%. Additionally, the use of a second subset of 20 features (subset 20), obtained by a majority voting strategy from the variables selected by at least three of the five techniques, allows for further improvement in performance, with an accuracy of 98.18%, a sensitivity of 98.81%, a specificity of 96.49%, a precision of 98.68%, an F1-score of 98.74%, and an MCC of 95.41%, while maintaining an optimal balance between sensitivity and specificity. These results confirm the robustness, generalizability, and relevance of the proposed solution for securing highly dynamic and heterogeneous IoT environments.

Keywords: Internet of Things (IoT), intrusion Detection System (IDS), AutoEncoders (AE), convolutional neural networks (CNN), bidirectional long short-term memory (BiLSTM) neural networks, attention mechanism.

Keywords: Internet of Things (IoT), intrusion Detection System (IDS), AutoEncoders (AE), convolutional neural networks (CNN), bidirectional long short-term memory (BiLSTM) neural networks, attention mechanism.

1 INTRODUCTION

The Internet of Things (IoT) sector is experiencing rapid and continuous growth worldwide. The number of connected devices was estimated at around 12.5 billion in 2010, a number that exceeded 21.7 billion in 2020 [1], and according to Cisco's projections, the number of devices connected to the Internet is expected to exceed 500 billion by 2030 [2]. This expansion is driven by the strong potential of the IoT and the diversity of its applications, particularly in the fields of healthcare, industry, and smart cities. This massive growth is accompanied by increased interconnection of heterogeneous devices via the Internet, generating considerable volumes of data with varied characteristics. As a result, the attack surface of IoT networks is expanding significantly, rendering these environments particularly vulnerable. Attackers now have a growing number of entry points to target connected devices, increasing security risks. To address these threats, various intrusion detection systems (IDS) have been developed. Signature- based IDS effectively detect known attacks listed in databases, but remain ineffective against new or unknown attacks [3]. Anomaly detection approaches rely on learning normal system behavior, with any significant deviation generating an alert. However, the complexity of modeling the normal profile can lead to a high rate of false positives [4]. Specification-based methods, on the other hand, manually define the expected behavior of the system using expert rules, but suffer from a lack of adaptability and remain sensitive to specification errors [4]. Hybrid approaches, combining these different strategies, have therefore been proposed to take advantage of their respective benefits and improve the detection of both known and unknown attacks [4]. In recent years, artificial intelligence techniques, and in particular machine learning, have demonstrated their effectiveness in detecting anomalies within IoT environments [5]. Nevertheless, as highlighted in [6], traditional machine learning approaches still face difficulties in dealing with the volume, speed, and heterogeneity of data generated by the IoT. In this context, deep learning techniques have emerged as promising solutions, capable of processing large data streams in real time while extracting relevant representations to anticipate security threats. Furthermore, it has been reported that 99% of cyberattacks result from minor modifications to already known attacks, giving rise to new variants [12]. This reality highlights the need to design models with strong generalization capabilities, capable of extracting the essential characteristics needed to effectively detect new attacks in dynamic IoT environments. In this context, this study proposes a hybrid attack detection approach aimed at overcoming several limitations of existing IDS systems. The approach relies on the use of a large-scale, highly heterogeneous dataset and the application of advanced preprocessing techniques to extract relevant information and improve the overall performance of the model. The main contributions of this study are summarized as follows:

– Proposal of a new hybrid model (AE-CNN-BiLSTM-Att): Development of an innovative intrusion detection architecture that combines an AE, a CNN, a BiLSTM, and an attention mechanism, specifically designed for intrusion detection in dynamic and heterogeneous IoT environments.

– Implementation of a comprehensive and rigorous preprocessing pipeline: A preprocessing pipeline integrating feature selection, outlier treatment, and class balancing to improve model robustness and generalization.

– Achieving high performance: The experimental results highlight remarkably high performance, with accuracy exceeding 98% and an F1 score exceeding 98.7% on subset 20. This performance underscores the importance of selecting discriminating features that can boost the model's generalization ability while improving detection reliability.

The rest of the document is organized as follows: Section 2 is devoted to related work. Section 3 describes the methods and materials used. Section 4 presents the experimental results, while Section 5 discusses them. Finally, Section 6 concludes the article.

2 RELATED WORK

Faced with the increasingly diverse threats that disrupt the proper functioning of IoT environments, numerous research efforts have been devoted to enhancing their security, leading to the development of a wide range of intrusion detection strategies. In this context, the authors of [7] presented a comprehensive systematic review of intrusion detection systems in IoT environments by analyzing the literature published between 1998 and 2018. Their study examined anomaly-based, signature-based, specification- based, and hybrid IDS approaches. Moreover, the challenges associated with each technique were thoroughly discussed in order to provide valuable insights for the design of more effective IDS solutions.

Building on these foundations, several studies have explored the application of machine learning techniques to improve intrusion detection performance. For instance, the effectiveness of three machine learning models, namely XGBoost, SVM, and DCNN, was evaluated on the IoT-23, NSL-KDD, and TON IoT datasets. The experimental results demonstrated that XGBoost achieved higher accuracy and overall effectiveness compared to the other evaluated methods [8].

Similarly, for intrusion detection in IoT environments, the authors of [9] adopted a hybrid feature selection strategy, followed by the application of four machine learning algorithms, namely Bagging, Multilayer Perceptron, J48, and IBk. The reported results were comparable to those obtained in other pioneering studies, thereby confirming the relevance of hybrid feature selection techniques.

In parallel, deep learning-based approaches have gained increasing attention due to their ability to automatically extract high-level features from complex network traffic. In this regard, an intelligent deep learning-based system named IoT-IDCS-CNN was proposed to detect cyberattacks within IoT networks. Experimental evaluations revealed that its detection performance surpassed that of several existing intrusion detection systems [10]. Furthermore, to enhance intrusion detection accuracy, features extracted from multiple layers of a deep CNN were subsequently used to train a linear support vector machine and a 1-nearest neighbor classifier. This hybrid approach was validated on the KDD 99 and NSL-KDD datasets and demonstrated superior performance compared to previously reported results for the same datasets [11].

Additionally, the authors of [12] leveraged the computational power of GPUs in conjunction with convolutional neural networks to detect cyberattacks in IoT environments. The simulation results indicated that the proposed method achieved higher classification accuracy than IDS solutions based on advanced machine learning techniques when evaluated on comparable datasets. Likewise, to address the challenge of malware intrusion detection, a deep learning-based approach relying on visual representations of IoT malware traffic was introduced. Using a dataset consisting of 1,000 PCAP files containing both normal and malicious traffic, the experimental results achieved an accuracy of 94.50% in detecting malicious traffic [13]. Moreover, deep learning models were also applied to the CICIDS2017 dataset for the detection of DDoS attacks in IoT networks, achieving a maximum accuracy of 97.16% [14].

More recently, attention-based and transfer learning techniques have been incorporated into IDS designs to further enhance detection performance. In this direction, an intrusion detection system based on transfer learning and the convolutional block attention module (CBAM) was proposed. By transforming tabular network traffic data into images, the approach enabled advanced feature extraction using convolutional neural networks, while the CBAM module improved feature representation. The experimental evaluation conducted on the BoT-IoT dataset achieved an accuracy of 99.93%, highlighting the effectiveness of the proposed IDS in securing IoT networks [15]. Similarly, for anomaly-based intrusion detection, CNN models were developed for both binary and multi-class classification tasks and validated on the BoT-IoT, IoT Network Intrusion, MQTT-IoT-IDS2020, and IoT-23 datasets. The obtained results outperformed existing deep learning-based solutions available at the time [16].

Furthermore, to identify malicious traffic more effectively, a self-supervised contrastive learning approach was introduced. This method combined a Transformer model with a bidirectional GLSTM to extract meaningful traffic features. When evaluated on the BoT-IoT dataset, the proposed approach surpassed recent state-of-the-art methods in terms of accuracy and F1-score [17]. In a similar vein, an innovative deep learning-based intrusion detection system targeting the most frequent IoT attacks was proposed, with experimental results demonstrating satisfactory performance and confirming its effectiveness in enhancing IoT network security [18]. Additionally, for DoS and DDoS attack detection, an approach based on transforming network traffic into images and applying a ResNet model was introduced. This method achieved remarkable performance, with an average accuracy for identifying eleven types of DoS and DDoS attack patterns exceeding the state of the art by 9% [19].

Finally, ensemble-based methods have also been investigated to improve intrusion detection in IoT environments. In [20], the authors proposed a multi-class IDS based on ensemble learning techniques, including XGBoost, Bagging, Random Forest, Extra Trees, and AdaBoost, using the ToN IoT dataset.

The experimental results indicated that XGBoost outperformed the other evaluated models. However, a notable limitation of this approach lies in its inability to adequately address class imbalance. To overcome such limitations and enhance the detection of previously unseen attacks, the authors of [21] explored both machine learning and deep learning-based approaches. Their experimental findings showed that the medium-type neural network model achieved the highest accuracy, outperforming both the proposed methods and existing solutions. An hybrid CNN–BiLSTM model incorporating transfer learning (TL- BiLSTM) was introduced to detect various Mirai and BASHLITE attacks across nine categories of IoT devices. Using the N-BaIoT dataset with ten traffic classes, the proposed model achieved significantly higher accuracy than state-of-the-art techniques, with performance gains ranging from 3.2% to 16.07% [22].

3 METHOD AND MATERIAL

The dataset used in this study is NF-ToN-IoT-V2, which is available online. It is an expanded version of the NF-ToN-IoT dataset and is part of the NFV2 collection developed by the University of Queensland. This collection aims to standardize datasets dedicated to network security to improve their interoperability and facilitate large-scale analysis. In this enriched version, the number of characteristics has been increased from 8 to 43, and the data is divided into 10 distinct classes, for a total of 13,135,881 records, which represents a significant improvement over the initial version. Table 1 presents the distribution of records in the NF-ToN-IoT-V2 dataset according to the different classes.

Table 1: Distribution of different classes in the NF-ToN-IoT-V2 dataset

Class	Count	Description
Benign	3601284	Normalunmaliciousflows.
Scanning	3002169	Agroupthatconsistsofavarietyoftechniquesthataimtodiscover informationaboutnetworksandhosts,andisalsoknownasprobing.

Cross-siteScripting(XSS)	2449955	Cross-site Scripting is a type of injection in which an attacker uses web applications to send malicious scripts to end-users.
DDoS	1746590	AnattemptsimilartoDoSbuthasmultipledifferentdistributedsources.
Password	993718	covers a variety of attacks aimed at retrieving passwords by either bruteforce or sniffing.
Injection	660467	A variety of attacks that supply untrusted inputs that aim to alter the course of execution, with SQL and Code injections two of the main ones.
DenialofService(DoS)	654359	An attempt to overload a computer system's resources with the aim ofpreventingaccesstooravailabilityofitsdata.
Backdoor	16259	Atechniquethataimstoattackremote-accesscomputersbyreplyingto specific constructed client applications.
ManintheMiddle(MITM)	7723	ManInTheMiddleisamethodthatplacesanattackerbetweena victimandhostwithwhichthevictimistryingtocommunicate,with the aim of intercepting traffic and communications.
Ransomware	3357	An attack that encrypts the files stored on a host and asks for compensationinexchangeforthedecryptiontechnique/key.

3.1 Exploratory study of the dataset

The original datasets contain a large number of characteristics, some of which contribute effectively to the correct prediction of the class, while others can perturb this process. It is therefore essential to select characteristics that correlate significantly with the target variable to develop systems that can detect attacks with greater accuracy. The relationships between the different characteristics and the class to be predicted can be varied in nature. Consequently, it is important to use several selection techniques to identify a set of relevant characteristics that share a type of correlation with the target variable. We will therefore provide a brief description of the techniques used.

– ANOVA: is an inferential statistical method used to examine the effect of an explanatory variable on a target variable. It measures the statistical significance of the association between a numerical variable and a categorical target variable. In this study, this technique is used to identify explanatory variables that have a significant impact on the target variable [23, 24].

– MI: this is a feature selection technique widely used to improve the performance of intrusion detection systems. It allows the degree of dependence between each explanatory variable and the class label to be evaluated, while taking into account the non-linear relationships that may exist between variables and classes [25].

– Kendall's rank correlation: this is a nonparametric test used to measure the strength of the relationship between two variables. Compared to Spearman's correlation, this method is more robust and reliable [26].

– mRMR: is an approach that functions as a feature selection filter based on mutual information. It identifies the most essential subset of features, which contributes to improving the accuracy of learning models [27, 28].

– Chi test: This is a statistical test used to check for the existence of a significant relationship between two categorical variables [23, 29].

3.2 Analysis of the created subdata set

After applying the various feature selection techniques mentioned above, two subsets of data were created.

The first subset, noted as subset _28, is the result of combining all the features identified by each of the selection methods.

The second subset, noted as subset_20, was obtained by retaining only the 20 variables selected by at least three of the five feature selection techniques, according to a majority voting strategy.

The variables constituting each of the subsets are presented in Table 2.

Table 2: Selected variables for subset_20 and subset_28

subset_20	subset_28
MIN_IP PKT_LEN, LONGEST_FLOW_PKT, MAX_IP_PKT_LEN, L4 _DST_PORT, PROTOCOL, TCP _WIN_MAX IN, DNS _QUERY_ID, SRC_TO_DST_AVG_THROUGHPUT, TCP _WIN_MAX_OUT, FLOW _DURATION_MILLISECONDS, MIN _TTL, MAX _TTL, DST _TO_SRC_AVG_THROUGHPUT, CLIENT _TCP_FLAGS, SERVER _TCP_FLAGS, DNS _QUERY_TYPE, IN _BYTES, SHORTEST _FLOW_PKT, DURATION _IN, DURATION _OUT	MIN_IP PKT_LEN, LONGEST_FLOW_PKT, MAX_IP_PKT_LEN, L4 _DST_PORT, PROTOCOL, TCP _WIN_MAX IN, DNS _QUERY_ID, SRC_TO_DST_AVG_THROUGHPUT, TCP _WIN_MAX_OUT, FLOW _DURATION_MILLISECONDS, MIN _TTL, MAX _TTL, DST _TO_SRC_AVG_THROUGHPUT, CLIENT _TCP_FLAGS, SERVER _TCP_FLAGS, DNS _QUERY_TYPE, IN _BYTES, SHORTEST _FLOW_PKT, DURATION _IN, DURATION _OUT, NUM _PKTS_1024_TO_1514_BYTES, NUM _PKTS_128_TO_256 BYTES, NUM _PKTS_512_TO_1024_BYTES, OUT_BYTES, SRC _TO_DST_SECOND_BYTES, L4 _SRC_PORT, L7_PROTO,'NUM _PKTS 256_TOv512_BYTES

3.3 Interquartile range (IQR)

The IQR is one of the most commonly used methods for detecting outliers. It offers a simple, effective, and robust approach for identifying data points that deviate significantly from the majority of observations. By applying this technique, we succeeded in identifying and visualizing the evolution of outliers for each variable in our subset of data. The Figure 1 below illustrates this evolution by variable.

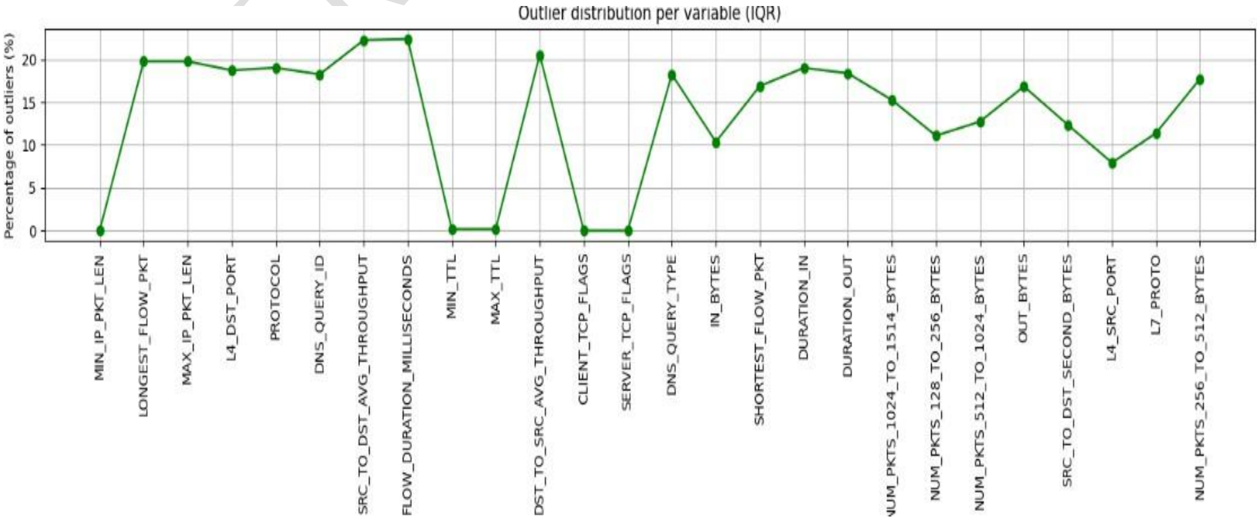

Fig. 1: Evolution of outliers before preprocessing

The Figure 1 clearly illustrates the evolution of outliers for each variable in our subset of data. Among the 28 selected variables, 26 have at least one outlier.

3.4 Class balance

To examine the distribution of classes (class 1 representing an attack and class 0 representing normal traffic), we used a Python function to visualize the proportion of each class. The results show that class 1 represents 72.58% of the 13,135,881 observations, while class 0 corresponds to 27.42%.

3.5 Data preprocessing

After identifying the variables correlated with the target variable and constructing the two data subsets, a preprocessing phase is implemented to correct any limitations that could affect the model's performance. Since subset_20 is included in subset_28, only the latter is considered in the preprocessing process. To this end, appropriate preprocessing techniques will be applied to address the various limitations identified within the dataset.

3.5.1 Winsorization method based on IQR

For the preprocessing of outliers, the winsorization method based on the interquartile range (IQR) was applied. This approach allows outliers to be processed while retaining all observations, thus preventing any loss of potentially useful information [30]. After applying this method, the analysis described in subsection 3.3 was performed once again to check for any outliers. The Figure 2 presents the graph indicating whether there is a variable with at least one outlier in our subset 28 of data after preprocessing. The Figure 2 below illustrates this evolution by variable.

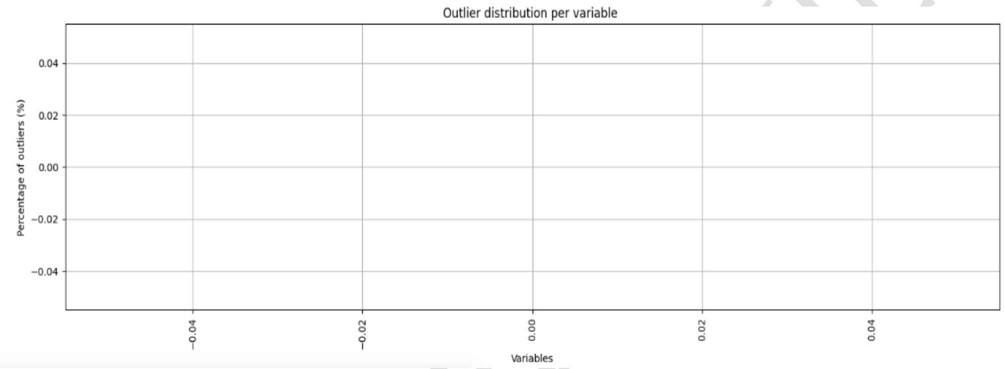

Fig. 2: Evolution of outliers after preprocessing

The Fig: 2 shows that no outliers are present in the subset_28 of data after applying this preprocessing technique.

3.5.2 Use of AutoEncoders

The distribution of observations shows a marked imbalance between class 0 (27.4156%) and class 1 (72.5844%), which can adversely affect the model's performance. To remedy this, data augmentation techniques are commonly used to balance the classes. Among these, AutoEncoders stand out by generating synthetic samples that are both diverse and complex, surpassing traditional oversampling, which can limit data diversity and promote overfitting. It has been reported that AutoEncoders can produce synthetic samples, significantly improving performance on imbalanced binary or multi-class datasets [31]. This approach was therefore adopted to correct the class imbalance in our subset of data.

3.5.3 The hybrid model: AE-CNN-BiLSTM-Att

For the design of our hybrid model, appropriate approaches were selected at each stage to build a model capable of addressing a wide range of limitations identified in the literature. The proposed model is thus based on four main components, namely:

– AE: is used to compress data by retaining only the most relevant information. Unlike principal component analysis (PCA), which relies on linear transformations to reduce the dimension of the data, autoencoders use nonlinear learning mechanisms to model complex relationships. This ability allows them to obtain a compressed representation that is more expressive and better suited to the structure of the data [32, 33].

– CNN: is used for feature extraction because of their ability to extract highly relevant information from data. Studies using this type of model for feature extraction have reported superior performance compared to state-of-the-art reference approaches [11, 34].

– BiLSTM: After feature extraction by the CNN, the representations obtained are sent to a BiLSTM layer, known for its effectiveness in modeling temporal dependencies within sequential data. The BiL-STM processes the sequence of features in order to learn the underlying dynamics of network traffic and detect attacks [35]. Several studies combining CNN and BiLSTM have reported performance superior to that of state-of-the-art reference models [35, 34].

– The attention mechanism is one of the major advances in deep learning over the past decade. Inspired by how human cognitive attention works, it allows the model to focus on the most relevant information. This approach is widely used to improve model performance byenhancing their representation and decision-making capabilities [36, 37].

3.5.4 Materials used

For this research, the Google Colab Pro environment was used for all stages, from data exploration to model training, including the preprocessing phases. This environment provides high-performance computing resources, enabling programs to be run efficiently and experimental results to be obtained.

4 Experimental results

This section evaluates the effectiveness of the proposed AE−CNN−BiLSTM−Att hybrid approach for intrusion detection in a dynamic IoT environment. The experiments were conducted using the NF-ToN- IoT-V2 dataset, which was first subjected to a comprehensive preprocessing pipeline including feature selection, outlier handling, and class balancing.

Feature selection is based on a combination of five complementary techniques. The variables thus identified present a significant correlation with the target variable, confirming their relevance for the attack detection task. The union of the feature sets resulting from these techniques was used to construct subset_28, while a majority voting strategy led to the definition of subset_20.

The model's performance was evaluated on subset_28 and subset_20, and the results are summarized in Table 3.

Table 3: Model performance on subset_28 and subset_20

Subset	Accuracy(%)	Sensitivity(%)	Specificity(%)	Precision(%)	F1-Score(%)	MCC(%)
Subset_ 28	97.82	97.87	97.68	99.11	98.49	94.60
Subset_ 20	98.18	98.81	96.49	98.68	98.74	95.41

Although all of the variables used in this study were selected based on their correlation with the target variable, the performance obtained from subset_20 slightly surpasses that observed with subset_28. This observation suggests that, among the variables correlated with the target class, there may be a more restricted and relevant subset offering marginally superior predictive capabilities. This observation is confirmed by the results obtained with subset_20, which performs slightly better overall than subset_28.

In addition, one of the main challenges associated with datasets dedicated to network intrusion detection systems (NIDS) is the absence of a standardized set of features. In fact, each public dataset is based on specific variables [38], and the degree of correlation between these variables and the target variable can vary considerably. Consequently, even when using the same number of features on the same dataset, it remains difficult to directly compare model performance if the selected features and preprocessing techniques are not exactly the same. This variability limits the reproducibility and comparability of results.

Thus, to illustrate more precisely the behavior of our model in classifying different types of traffic, Figure 3 and Figure 4 show the confusion matrices obtained with the two subsets.

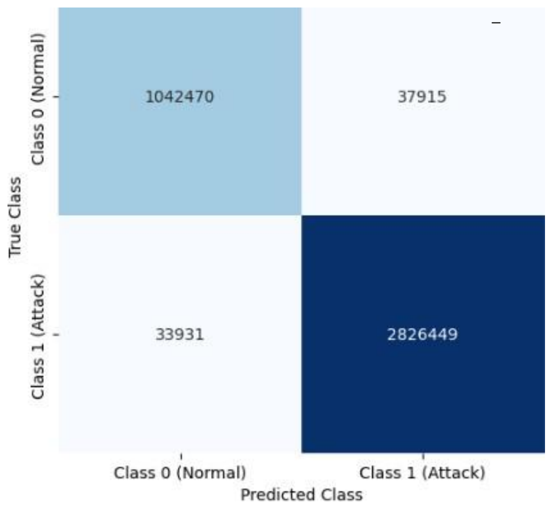

Fig. 3: Confusion matrix with subset_20.

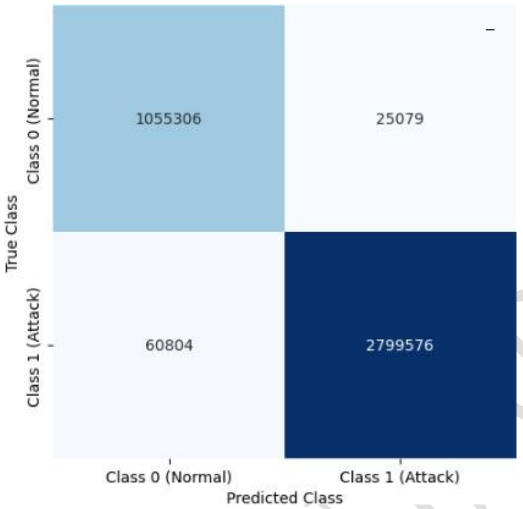

Fig. 4: Confusion matrix with subset_28.

The analysis of the confusion matrices (Fig 3 and Fig 4) demonstrates that the AE–CNN–BiLSTM–Att model achieves excellent performance in IoT traffic classification. With subset_20 (Fig 3), the majority of attacks are correctly detected and legitimate traffic is also almost always recognized, illustrating an optimal combination of sensitivity and specificity. When considering subset_28 (Fig 4), performance remains very high overall, but there is an increase in the number of undetected attacks. This difference highlights the presence of variables which are less essential to the correct prediction of the class, despite being correlated with the target variable. Thus, although the model performs very well with both subsets, the more restricted and targeted subset_20 maximizes the model's capacity to detect intrusions while maintaining a low false alarm rate, confirming the value of effective and discriminating feature selection.

5 DISCUSSION

The experimental results obtained in this study demonstrate clearly the effectiveness of the AE–CNN–BiLSTM–Att hybrid model for intrusion detection in IoT environments characterized by high dynamics and heterogeneity. The high performance observed, particularly in terms of accuracy, F1-score, and MCC, confirms the model's ability to learn rich and discriminative representations from complex and large network data. As highlighted in the literature, nearly 99% of cyberattacks derive from known attacks with minor modifications, giving rise to new attack variants [12]. This observation motivated the choice of a hybrid approach capable of effective generalization, going further than simply memorizing the patterns observed during the learning phase, to detect new or slightly modified attacks. It should also be noted that these results were achieved in a particularly challenging context, marked by a significant imbalance between classes, with malicious traffic accounting for 72.58% of the 13,135,881 observations,

compared to 27.42% for legitimate traffic. In order to correct this imbalance, the use of AE for the generation of synthetic samples made it possible to produce data that was both diverse and complex, promoting more balanced learning. The results show that this strategy effectively mitigated the bias associated with class distribution, as evidenced by high sensitivity and specificity values, reflecting a balanced and robust ability to discriminate between the two classes. The robustness of the model can also be attributed to the preprocessing pipeline that was implemented. In particular, the correction of outliers using IQR-based winsorization helped stabilize the distributions of the variables while retaining potentially informative observations. This step proved essential in improving the quality of the training data and enhancing the stability of the training process.

Furthermore, comparative analysis between the two subsets of features highlights the value of a reasoned reduction in dimensionality. The results obtained with subset_20 show that a smaller but more discriminating set of variables improves global performance, while promoting better model generalization and reducing computational complexity. In this context, direct comparisons between the results reported in different studies become less fiable, given that authors may rely on distinct sets of features, even when using the same dataset. This observation is supported by the findings of [38], which highlight that one of the main challenges of NIDS datasets is the lack of a standardized set of features, with each public dataset relying on specific variables. This makes it impossible to compare model performance. Examination of the confusion matrices confirms that the hybrid model maintains an effective balance between attack detection and legitimate traffic recognition. This property is essential in real-world IoT environments, where a high false positive rate can quickly render a detection system unusable, while a high false negative rate exposes the network to undetected attacks. Finally, although the results obtained are very encouraging, certain limitations must be taken into account. The experiments were carried out using a public dataset and in a computing environment with significant resources, which may differ from the constraints encountered during actual deployment on IoT devices with limited resources. Future work could therefore focus on optimizing and simplifying the model, as well as evaluating it in real-time scenarios or under real operating conditions. Overall, this study demonstrates that the AE–CNN–BiLSTM–Att approach, combined with an adapted preprocessing pipeline and effective feature selection, is a robust, high-performance, and promising solution for securing dynamic and highly heterogeneous IoT environments.

6 CONCLUSION

This paper proposes a hybrid, high-performance attack detection system designed to address the security challenges of modern IoT environments, which are characterized by high heterogeneity, massive data volumes, and high traffic dynamics. The joint integration of AutoEncoders, CNNs, BiLSTMs, and an attention mechanism within the AE–CNN–BiLSTM–Att model effectively captures the spatial and temporal dependencies of network traffic, while focusing on the most discriminative features for at- tack detection. Rigorous preprocessing was implemented, combining several feature selection techniques, IQR-based winsorization for outlier handling, and synthetic data generated by AutoEncoders to correct class imbalance. Experiments conducted on the NF-ToN-IoT-V2 dataset show that the proposed model achieves excellent performance according to several performance metrics. The results also indicate that a limited but relevant subset of features can slightly improve overall performance while reducing the complexity of the model.

Despite these encouraging results, certain limitations must be taken into account.

The experiments were conducted using a public dataset and in a computing environment with significant resources, which may differ from the restrictions encountered during actual deployment on IoT devices with limited resources. Future work will therefore focus on optimizing and simplifying the model, evaluating it in real time, and validating it under real operating conditions. In addition, the adoption of standardized feature sets for NIDS datasets is an essential step toward improving the reproducibility and comparability of future work. Overall, this study shows that the proposed hybrid approach represents a robust, effective, and promising solution for improving the security of dynamic IoT networks.

REFERENCES

[1] ArtSurLeWeb. (n.d.). Historique, chiffres clés et portée, définition complète de l'IoT. [Online]. Available: https://requea.com/qu-est-ce-que-l-iot.html [Accessed Dec. 12, 2025]

[2] Khanh, Q. V., Hoai, N. V., Manh, L. D., Le, A. N., & Jeon, G. (2022). Wireless communication technologies for IoT in 5G: Vision, applications, and challenges. Wireless Communications and Mobile Computing, 2022(1), 3229294.

[3] Sharma, B., Sharma, L., Lal, C., & Roy, S. (2023). Anomaly based network intrusion detection for IoT attacks using deep learning technique. Computers and Electrical Engineering, 107, 108626.

[4] Asharf, J., Moustafa, N., Khurshid, H., Debie, E., Haider, W., & Wahab, A. (2020). A review of intrusion detection systems using machine and deep learning in internet of things: Challenges, solutions and future directions. Electronics, 9(7), 1177.

[5] Inuwa, M. M., & Das, R. (2024). A comparative analysis of various machine learning methods for anomaly detection in cyber attacks on IoT networks. Internet of Things, 26, 101162.

[6] Alsoufi, M. A., Razak, S., Siraj, M. M., Nafea, I., Ghaleb, F. A., Saeed, F., & Nasser, M. (2021). Anomaly- based intrusion detection systems in iot using deep learning: A systematic literature review. Applied sciences, 11(18), 8383.

[7] Hajiheidari, S., Wakil, K., Badri, M., &Navimipour, N. J. (2019). Intrusion detection systems in the Internet of things: A comprehensive investigation. Computer Networks, 160, 165-191.

[8] Balega, M., Farag, W., Wu, X. W., Ezekiel, S., & Good, Z. (2024). Enhancing IoT security: optimizing anomaly detection through machine learning. Electronics, 13(11), 2148.

[9] Albulayhi, K., Abu Al-Haija, Q., Alsuhibany, S. A., Jillepalli, A. A., Ashrafuzzaman, M., & Sheldon, F. T. (2022). IoT intrusion detection using machine learning with a novel high performing feature selection method. Applied Sciences, 12(10), 5015.

[10] Abu Al-Haija, Q., & Zein-Sabatto, S. (2020). An efficient deep-learning-based detection and classification system for cyber-attacks in IoT communication networks. Electronics, 9(12), 2152.

[11] Chowdhury, M. M. U., Hammond, F., Konowicz, G., Xin, C., Wu, H., & Li, J. (2017, October). A few-shot deep learning approach for improved intrusion detection. In 2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON) (pp. 456-462). IEEE.

[12] Al-Haija, Q. A., McCurry, C. D., & Zein-Sabatto, S. (2020, September). Intelligent self-reliant cyber-attacks detection and classification system for IoT communication using deep convolutional neural network. In International Networking Conference (pp. 100-116). Cham: Springer International Publishing.

[13] Bendiab, G., Shiaeles, S., Alruban, A., & Kolokotronis, N. (2020, June). IoT malware network traffic classification using visual representation and deep learning. In 2020 6th IEEE Conference on Network Softwarization (NetSoft) (pp. 444-449). IEEE.

[14] Roopak, M., Tian, G. Y., & Chambers, J. (2019, January). Deep learning models for cyber security in IoT networks. In 2019 IEEE 9th annual computing and communication workshop and conference (CCWC) (pp.0452-0457). IEEE.

[15] Abdelhamid, S., Hegazy, I., Aref, M., & Roushdy, M. (2024). Attention-driven transfer learning model for improved IoT intrusion detection. Big Data and Cognitive Computing, 8(9), 116.

[16] Ullah, I., & Mahmoud, Q. H. (2021). Design and development of a deep learning-based model for anomaly detection in IoT networks. IEEe Access, 9, 103906-103926.

[17] Yang, J., Jiang, X., Liang, G., Li, S., & Ma, Z. (2023). Malicious traffic identification with self-supervised contrastive learning. Sensors, 23(16), 7215.

[18] Awajan, A. (2023). A novel deep learning-based intrusion detection system for IOT networks. Computers, 12(2), 34.

[19] Hussain, F., Abbas, S. G., Husnain, M., Fayyaz, U. U., Shahzad, F., & Shah, G. A. (2020, November). IoT DoS and DDoS attack detection using ResNet. In 2020 IEEE 23rd International Multitopic Conference (INMIC) (pp. 1-6). IEEE.

[20] Awotunde, J. B., Folorunso, S. O., Imoize, A. L., Odunuga, J. O., Lee, C. C., Li, C. T., & Do, D. T. (2023). An ensemble tree-based model for intrusion detection in industrial internet of things networks. Applied Sciences, 13(4), 2479.

[21] Mohamed, R. H., Mosa, F. A., & Sadek, R. A. (2022). Efficient intrusion detection system for IoT environment. International Journal of Advanced Computer Science and Applications, 13(4).

[22] Nandanwar, H., &Katarya, R. (2024). TL-BILSTM IoT: transfer learning model for prediction of intrusion detection system in IoT environment. International Journal of Information Security, 23(2), 1251-1277.

[23] Rihan, S. D. A., Anbar, M., &Alabsi, B. A. (2023). Approach for detecting attacks on IoT networks based on ensemble feature selection and deep learning models. Sensors, 23(17), 7342.

[24] Kassel, R. (2024, December 30).ANOVA (Analysis of Variance) : un outil fondamental pour l'analyse des donn´ees. DataScientest. [Online]. Available: https://datascientest.com/analyse-de-la-variance-anova [AccessedDec. 24, 2025]

[25] Alalhareth, M., & Hong, S. C. (2023). An improved mutual information feature selection technique for intrusion detection systems in the internet of medical things. Sensors, 23(10), 4971.

[26] Dasari, K. B., & Devarakonda, N. (2022). TCP/UDP-based exploitation DDoS attacks detection using AI classification algorithms with common uncorrelated feature subset selected by Pearson, Spearman and Kendall correlation methods. Revue d'IntelligenceArtificielle, 36(1), 61-71.

[27] Shirley, J. J., & Priya, M. (2024). Hybrid MRMR-PCA BagDT-An Effective Feature Selection based Ensemble Model for Real-Time Intrusion Detection in IoT Environment. IEEE Access.

[28] Houkan, A., Sahoo, A. K., Gochhayat, S. P., Sahoo, P. K., Liu, H., Khalid, S. G., & Jain, P. (2024). Enhancing security in industrial IoT networks: Machine learning solutions for feature selection and reduction. IEEe Access.

[29] IBM SPSS Statistics. (n.d.). [Online]. Available: https://www.ibm.com/docs/fr/spss-statistics/31.0.0?topic=features-chi-square-test [Accessed Dec. 24, 2025]

[30] Da Silva, C. M. C. (2025, October 7). Winsorization: Handling outliers in machine learning. Train in Data's Blog. [Online]. Available: https://www.blog.trainindata.com/winsorization-handling-outliers-in-machine-learning/[Accessed Dec. 26, 2025]

[31] Zhou, A., Liu, B., Wang, J., Sun, K., & Liu, K. (2024, August). AEMLO: AutoEncoder-Guided Multi-label Oversampling. In Joint European Conference on Machine Learning and Knowledge Discovery in Databases (pp. 107-124). Cham: Springer Nature Switzerland.

[32] Belaidi, N. (n.d.). Auto-encodeurs en Deep Learning : tout savoir. Formation Tech et Data en ligne — Blent.ai. [Online]. Available: https://blent.ai/blog/a/auto-encodeurs-deep-learning [AccessedDec. 26, 2025]

[33] Thor, W. (n.d.). Application: Data Compression using Autoencoders. [Online]. Available: https://apxml.com/courses/applied-autoencoders-feature-extraction/chapter-7-applying-autoencoder- features-practical-guidance/application-data-compression-autoencoders [Accessed Dec. 26, 2025]

[34] Sadhwani, S., Khan, M. A. H., Muthalagu, R., Pawar, P. M., & Suresh, K. (2025). A hybrid BiLSTM-CNN approach for intrusion detection for IoT applications. Scientific Reports.

[35] Jouhari, M., &Guizani, M. (2024, May). Lightweight cnn-bilstm based intrusion detection systems for resource-constrained iot devices. In 2024 International Wireless Communications and Mobile Computing (IWCMC) (pp. 1558-1563). IEEE.

[36] Liu, C., Liu, Y., Yan, Y., & Wang, J. (2020). An intrusion detection model with hierarchical attention mechanism. IEEE Access, 8, 67542-67554.

[37] Laghrissi, F., Douzi, S., Douzi, K., &Hssina, B. (2021). IDS-attention: an efficient algorithm for intrusion detection systems using attention mechanism. Journal of Big Data, 8(1), 149.

[38] Sarhan, M., Layeghy, S., & Portmann, M. (2022). Towards a standard feature set for network intrusion detection system datasets. Mobile networks and applications, 27(1), 357-370.