International Journal of Computer Sciences and Engineering
Vol.**13**, Issue.**5**, pp.**09-16**, May **2025**
ISSN: 2347-2693 (Online)
Available online at: www.ijcseonline.org

ISSN: 2347-2693 (E)

Research Article

An Integrated Approach to Optimize the Intrusion Detection System in SDN based on Feature Selection and Deep Learning Techniques

Pranjal Maurya[1*], Sangeeta Devi[2], Upendra Nath Tripathi[3]

[1,2,3]Dept. of Computer Science, Deen Dayal Upadhyaya Gorakhpur University, Gorakhpur, Uttar Pradesh, India

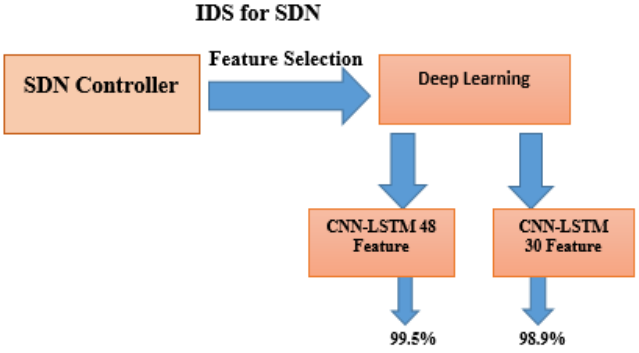
**Corresponding Author:* ✉

Received: 22/Mar/2025; **Accepted**: 23/Apr/2025; **Published**: 31/May/2025. **DOI**: https://doi.org/10.26438/ijcse/v13i5.916

Abstract: Software-defined networking (SDN) is a revolutionary innovation that has become known as an internet architecture. It allows for modular, adaptable, and effective network management alternatives by decoupling the management plane about the data plane. The centralized management and open interfaces that are present in software-defined networking (SDN) provide a number of new security issues, the most notable of which is the increased danger of network breaches. It is difficult for traditional security measures to adjust to the dynamic and programmable nature of SDN, which is why more intelligent solutions are required. The ability of deep learning (DL)-based intrusion detection systems (IDS) to acquire knowledge from data and distinguish complex attack trends in real time has led to their rising popularity. This research looks at software-defined networking (SDN) environments with intrusion detection systems (IDS) built on deep learning algorithms. Identify outliers, feature selection and architectures approaches are discussed. We have constructed and compared two models based on feature selection for comprehensive Intrusion Detection System (IDS) solution. One of these models uses CNN-LSTM architecture, using 48 features achieved the maximum accuracy, which was 99.5%. On the other hand, the CNN-LSTM model using 30 features achieved 98.9% accuracy.

Keywords: Software Defined Networking (SDN), Outlier, Feature Engineering, Intrusion Detection System, Deep Learning

Graphical Abstract: The research's methods and conclusions are graphically represented in the graphical abstract. It shows how data moves from the SDN Controller to a feature selection procedure and then into two CNN-LSTM models for deep learning-based categorization. The accuracy of the model with 30 selected features is 98.9%, whereas the model in 48 selected features is 99.5%. The impact of deep learning architecture and optimal feature selection on intrusion detection capability in SDN systems is highlighted in this picture.



Purpose- Through the use of deep learning, this research seeks to create an intelligent Intrusion Detection System (IDS) that will improve network security in Software-Defined Networking (SDN). The following are the main goals of the study:

- Addressing the security issues brought on by SDN's centralization.
- Utilizing deep learning models from CNN-LSTM to identify multifaceted attack behaviors.
- Reducing data dimensionality and enhancing the accuracy of models through the use of feature selection and outlier elimination.
- Two optimized models with an accuracy of up to 99.5% were evaluated on the InSDN dataset.

The result shows an adaptable and efficient IDS system for real-time SDN infrastructure security.

1. Introduction

The proliferation of network architectures gave rise to the idea of software-defined networking, or SDN [1]. SDN is a paradigm that centralises network management and encapsulates the infrastructure in order to enable capabilities for dynamic configuration. Although software-defined

networking (SDN) offers advantages such as scalability, flexibility, and simpler administration, it is vulnerable to new security vulnerabilities as a result of the architectural modifications it undergoes [2]. Intrusion Detection Systems, often known as IDS, are very important in the process of protecting networks since they are able to detect unauthorised access or behaviour that is irregular. Through the ability to enable systems to learn and identify threats that have not been observed before, deep learning (DL) has shown a great deal of promise in overcoming the limits of conventional intrusion detection systems (IDS) [3]. When incorporated with SDN, deep learning-based intrusion detection systems have the potential to greatly improve the identification and response to threats procedures.

1.1 Software-Defined Networking (SDN)

A game-changing method for building networks is software-defined networking, or SDN. Because of this partitioning, a programmed network management platform can be implemented, which enhances the efficiency, scalability, and adaptability of managing intricate networks.

Core Components of the Software-Defined Networking (SDN) [4]-

- **Control Plane (SDN Controller):** This component serves as the centralized brain of the network, administering flow control to the various networking devices (switches, routers) via the use of protocols such as OpenFlow. Policy enforcement and network configuration are both simplified as a result of this centralization.
- **The Data Plane**: The real-time packet relaying in response to controller instructions is performed by the Forwarding Equipment, which may be either physical or virtual switches or routers.
- **Northbound APIs:** These are interfaces that enable communication between the software-defined networking controller and the applications and business logic that are located above. They make it possible to program networks and integrate them with technologies that are used for orchestration and automation.
- **Southbound APIs:** Protocols that are used by the controller in order to connect with the data plane devices are referred to as southbound application programming interfaces (APIs). Southbound application programming interface (API) OpenFlow is frequently used because it enables direct control over the forwarding plane.

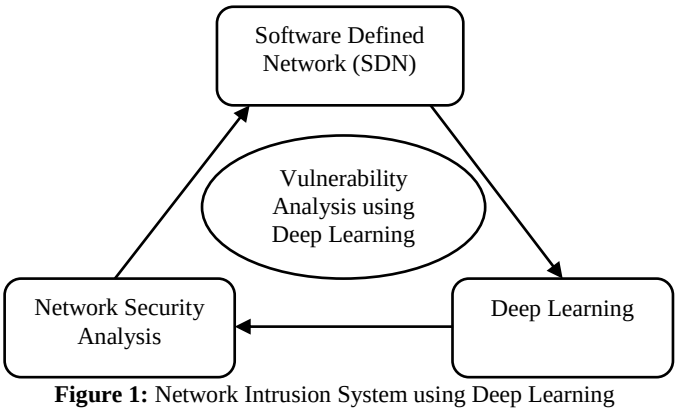


Figure 1: Network Intrusion System using Deep Learning

- Although it is convenient for administration, the centralized structure of software-defined networking (SDN) presents a unique set of security issues. The controller of the software-defined network (SDN) becomes a high-value target for attackers, and the inherently dynamic nature of SDN may be abused for nefarious reasons. It is vital to include Intrusion Detection Systems (IDS) into the design of the software-defined network (SDN) in order to solve these challenges as shown in figure 1.

The use of preexisting signatures or protocols is common in older intrusion detection systems (IDS), which makes them less effective against threats that are new or continue to evolve. Machine Learning is a subset of Deep Learning (DL), which provides improved capabilities for pattern identification and anomaly detection [5]. Because of these characteristics, Deep Learning is well suited for contemporary intrusion detection systems (IDS) implementations. DL-based intrusion detection systems are able to analyses enormous volumes of network data in SDN settings, which allows them to identify complicated attack patterns [6].

1.2 Problem in Network Security

Network security is considered one of the most critical challenges in contemporary times. For almost two decades, research on network defensive mechanisms has garnered increasing interest from the academic community. However, the issues pertaining to network security remain unresolved. Researchers have used game theoretic ways to tackle network security challenges; nevertheless, only a limited number of these methods effectively resolved the concerns. Game theory is mostly used when many participants with divergent agendas engage in competition [7].

It offers an analytical framework for tackling network security issues. Network security issues are often mitigated by the use of preventive or reactive technologies. The most prevalent defensive devices used are firewalls, although reactive systems such as Intrusion Detection Systems (IDS) are utilized.

1.2 Intrusion Detection System

Intrusion detection refers to the surveillance of activities inside a network or institution and the analysis of those activities for indications of assaults that may compromise dependability, accessibility, confidentiality, or circumvent security mechanisms. Intrusions may be caused by attackers entering the system, as well as by authorized users who misuse their privileges and those who unlawfully acquire new rights. Intrusion Detection Systems (IDS) are regarded as the intrusion alarms within the domain of technological security. Both internal and external attackers might compromise the network; consequently, the function of an Intrusion Detection System (IDS) is essential.

These systems are strategically placed throughout a network to monitor packets and detect possible threats. The IDS do this by aggregating data from multiple systems and network

sources of information, then analyzing it to identify possible threats [8]. The roles of an intrusion detection system (IDS) include delivering information on risks, implementing remedial measures upon threat detection, and documenting all key occurrences inside a network [9]. Figure 2 illustrates an example of an intrusion detection system.

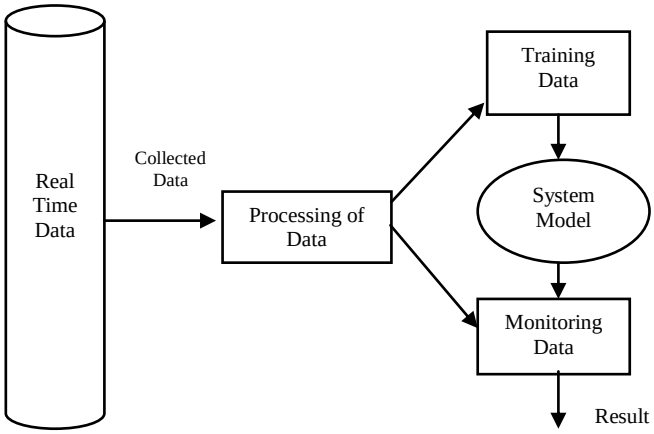


Figure 2: Intrusion Detection System Model

The protective scope of different IDS may range from a one computer to a vast network of computers. A Network Intrusion Detection System (NIDS) monitors entering network traffic, whereas a Host Intrusion Detection System (HIDS) oversees critical files inside the operating system as well [9]. An example of both kind of systems is one that evaluates incoming network data. The detection mechanism utilized by an IDS serves as an additional factor for its classification. This categorization may be further subdivided into various subcategories, the most prevalent of which include signature-based search and detection based on anomalies. An alternative is reputation-based detection. Currently, there are many intrusion detection systems (IDS) available that include the capability to react to identify intrusions.

1.3 Deep Learning Techniques

Deep learning approaches represent a significant improvement in Artificial Neural Networks (ANN), using several sophisticated calculations. Deep learning enables algorithms to comprehend information representations via several levels of generalization [10].

These deep learning approaches may be used in several applications such as network intrusion detection, network identification, visual object recognition, and numerous more domains [11]. The deep learning system may be trained in both unsupervised and supervised modes. CNN is a deep learning approach that is often learnt in a supervised manner. Numerous CNN models are used for intrusion detection, demonstrating enhanced detection accuracy and increased system efficiency.

CNN is currently considered the standard system for computer vision applications. A specific CNN framework is used for organizing two-dimensional pictures and is also applied in face recognition. An autoencoder exemplifies a deep learning approach whereby training occurs in an unsupervised manner. Autoencoders are mostly used to analyses the encoding or representation of a dataset, particularly for the goal of dimensionality reduction. These autoencoders are integrated with various additional techniques and used for network intrusion detection. A Deep Belief Network (DBN) is an deep learning algorithm that reconstructs its input using a collection of examples in an unsupervised manner, with these layers functioning as feature detectors for the inputs. Subsequent to the learning process, the DBN may be taught in a supervised manner for classification purposes [12].

Certain Deep Belief Networks (DBNs), including auto encoders and restricted Boltzmann machines (RBMs), may be used for regression, feature learning, collaborative filtering, topic modelling, and dimensionality reduction, among other applications. RNN exemplifies a deep learning methodology that is learnt using both supervised and unsupervised techniques. RNN can regulate internal memory to process arbitrary inputs. Numerous researchers use these deep learning approaches for effective intrusion detection [13].

1.4 Problem Statement

Intrusion detection for safeguarding networks is confronted with numerous obstacles, despite the numerous investigations conducted on IDS. The estimation of probabilities for high-dimensional information presents a challenge for the multidimensional statistical intrusion detection methods. The issues stem from the fact that it fails to identify objectionable behavior, which is why the FAR may be significant. The failure to identify intrusions which are still not public to the IDS and its dependence on an established security policy, and these may be absent, are among the issues. The following is a list of some of the most frequently confronted challenges by the network IDS:

- The detection techniques that are currently available are unable to reliably identify the numerous categories of intrusions.
- The detection of intrusions from encrypted traffic is a difficult task due to the fact that the majority of data is encrypted in the network.
- Certain intrusion detection systems (IDS) are capable of identifying network assaults, but they are incapable of detecting systemic attacks.
- The detection of intrusions from the high-speed network is challenging due to the adoption of multiple intrusion detection mechanisms.

It is exceedingly challenging to predict the timing and nature of actions that will transpire within a system.

1.6 Objective of Research

Using Deep Learning (DL), the project aims to construct an Intrusion Detection System (IDS) for use in Software-Defined Networking (SDN) environments. Some of these goals are-

- The implementation of an Intrusion Detection System (IDS) in Software-Defined Networking (SDN)

environments using Deep Learning (DL) has well-defined research objectives, such as gathering and analyzing information gathered from network traffic to detect and mitigate security risks.

- Optimize feature selection and execute a deep learning model that proficiently categorizes network traffic inside SDN infrastructures using collected datasets.
- Evaluate the proposed Intrusion Detection System (IDS) in a Software-Defined Networking (SDN) setting for its ability to detect various types of intrusions using metrics including recall, accuracy, precision, and F1-score.

2. Literature Survey

In the contemporary linked world, where information and communication technology is progressing swiftly, a multitude of reliable online systems and services has arisen. Nonetheless, this expansion in digital connection has concurrently resulted in a rise in cybersecurity threats. Organizations and people face persistent attacks from malevolent entities aiming to exploit weaknesses in their networks. By monitoring network traffic, detecting anomalous behaviors, and delivering prompt notifications to administrators, Intrusion Detection Systems (IDS) perform an essential role in recognizing and mitigating security breaches. It seeks to tackle these difficulties by examining studies on SDN-based IDS and investigating methods to augment its functionalities.

Arevalo-Herrera et al. (2022) offered an artificial intelligence learning based anomalous network activity detection and categorisation approach. They analysed network traffic information and found irregularities using k-means clustering, support vector machines, and random forests. The models developed using machine learning have high detection rates and low false positive rates, proving that machine learning can identify network anomalies in software-defined systems.
Mostafa, N. et al. (2024) This research presents a literature review of SDN-based Intrusion Detection Systems (IDSs), their categories, and multiple deployments using contemporary machine and deep learning methodologies. The use of machine and deep learning in Intrusion Detection Systems (IDS) has markedly enhanced the efficiency of detection and introduced a new domain in information safety, including the employed learning models, characteristics, and learning variables. In the setting up of Intrusion Detection Systems (IDSs) using machine and deep learning approaches, enhancing detection accuracy, as well as optimizing and expediting inference time, is a prevalent objective. This study seeks to elucidate the significance of instituting stringent security protocols to safeguard important information and maintain the accuracy and accessibility of network assets.
Chouikik, M. et al. (2024) This study analyses a DDoS attack on an SDN system using OpenDaylight and Mininet simulator. Evaluation also includes DDoS attack repercussions and IDS effectiveness in reducing them. Network effectiveness under stress is tested using many performance measures, including bandwidth compared to delay time. The difference in performance curves comparing intrusion detection and non-intrusion detection shows its importance. Oscillations increased significantly without the IDS, suggesting network vulnerability. Instead, an IDS managed the environment, decreasing variances and improving network stability.

Chetouane, A. et al. (2025) addressed the collection of data including normal, DDoS, DoS, and probing attack flows is analyzed using several methods for selecting features and Machine Learning algorithms to evaluate its complexity. Coefficient-based feature selection (CFS) among filter techniques is distinguished by the use of a Decision Tree (DT) classifier, achieving good accuracy with the dataset while minimizing implementation time in identifying attacks. tenfold layered cross-validation is conducted for the proposed IDS, accompanied with a trust interval evaluation for the examination dataset generated from a business network utilize case, to demonstrate the model's sustainability support. A comparative examination of the information set and its capacity utilization is performed to assess its viability.

3. Methodology

A methodical strategy including data collecting, model construction, and evaluation is essential for the creation of an effective Intrusion Detection System (IDS) for SDN environments employing deep learning. The procedure starts with the selection of a suitable dataset that correctly represents the distinct features and problems of SDN systems. A deep learning model is then developed to analyse the obtained data, with the objective of identifying and classifying different network breaches. The model's performance is meticulously assessed using conventional criteria to verify its effectiveness in practical applications.

3.1 Dataset Collection

The In SDN dataset is a complete repository for SDN intrusion detection gathering data. Web attacks, password-guessing, botnet activities, probing, DoS, and DDoS attacks are among the more than 43,939 records in the data set, including both harmless and unwanted traffic [18]. Original dataset records included 84 characteristics, including packet-based- and flow-level information. We have selected and reduced dimensionality to 48 and, in some experiments, 30 key characteristics to improve models made up of deep learning. The In SDN dataset comprehensiveness and conformity to SDN-specific flow and attack trends make it ideal for training and assessing deep learning-based IDS algorithms for SDN settings.

3.2 Remove Outlier and Feature Selection

First, the technique of exploratory data analysis (EDA) was employed to analyse the dataset's layout and significant properties. To increase model precision and effectiveness, data preparation includes encoding categorical factors, scaling numbers, and deleting duplicate or constant-value features.

Risk identification utilising the Local Outlier Factor (LOF) method was crucial. LOF identifies dataset outliers by comparing the local density variations of a given input point

to its surroundings [19]. LOF compares a data point's number of points to its neighbours. A location is an outlier if its median density is considerably lower than its neighbours. This describes the LOF sequence as:

Recognizing Neighbours: LOF identifies the closest neighbours (k-nearest neighbours or k-NN) for each data point using distance metrics.

Reachability Distance: The Local Outlier Factor (LOF) calculates the reachability distance, which measures the distances between two locations while considering the density of neighbouring sites. This aids in ascertaining if a point is isolated from its neighbouring points.

Local Reachability Density (LRD): It is calculated by the LOF to estimate the density surrounding a site. Local Reachability Density (LRD) is going to be reduced if a point is encompassed by fewer points than those surrounding it.

Calculation of LOF Score: Points are rated by comparing their Local Reachability Density (LRD) to their neighbours' LRDs, as defined by LOF. The occurrence is unusual because its local density is much lower than its neighbours', as demonstrated by its high LOF score.

Equations (1), (2), and (3) explain outlier detection.

$$Reach\_Distance_{(p,o)} = \max(k - distance(o), distance(p,o)) \quad (1)$$

$$LRD(p) = \frac{1}{\frac{1}{k}\sum_{o \in n_{k(p)}} (Reach\_Distance_{(p,o)})} \quad (2)$$

$$LOF_{(p)} = \frac{\sum_{o \in n_{k(p)}} \frac{LRD_{(o)}}{LRD_{(p)}}}{\frac{1}{k}\sum_{o \in n_{k(p)}} (Reach_{Distance(p,o)})} \quad (3)$$

Distance (p, o) denotes the actual distance between points p and o, while k-distance(o) denotes the k-distance of positions. The local reachability density of neighbour o is denoted by LRD(p), while the reachability density of point p is represented by LOF(p).

The ideal number of neighbours (k) was determined using the Elbow strategy, with the optimum value for k being 20 for the LOF strategy. The objective is to ascertain the point where the mean LOF score stabilises and calculate the LOF scores for a variety of k values. This aids in the identification of the optimal number of neighbours that achieves a balance among the precise identification of genuine anomalies and the reduction of false positives.

The outlier values -1 (indicating risk) and 1 (indicating inliers) were transformed to binary form after the accumulation of predictions about LOF. This binary indicator denotes how much a data point is classified as risky or not.

Feature selection and reduction are crucial for improving the efficacy of Intrusion Detection Systems (IDS) that use the In SDN dataset in Software-Defined Networking (SDN) contexts [20]. Recursive Feature Elimination (RFE) is used to systematically eliminate the least important features according to model efficacy. The research used an advanced feature selection technique on the In SDN dataset, resulting in the identification of 48 features and 30 features. Figure 3 is showing working mechanism.

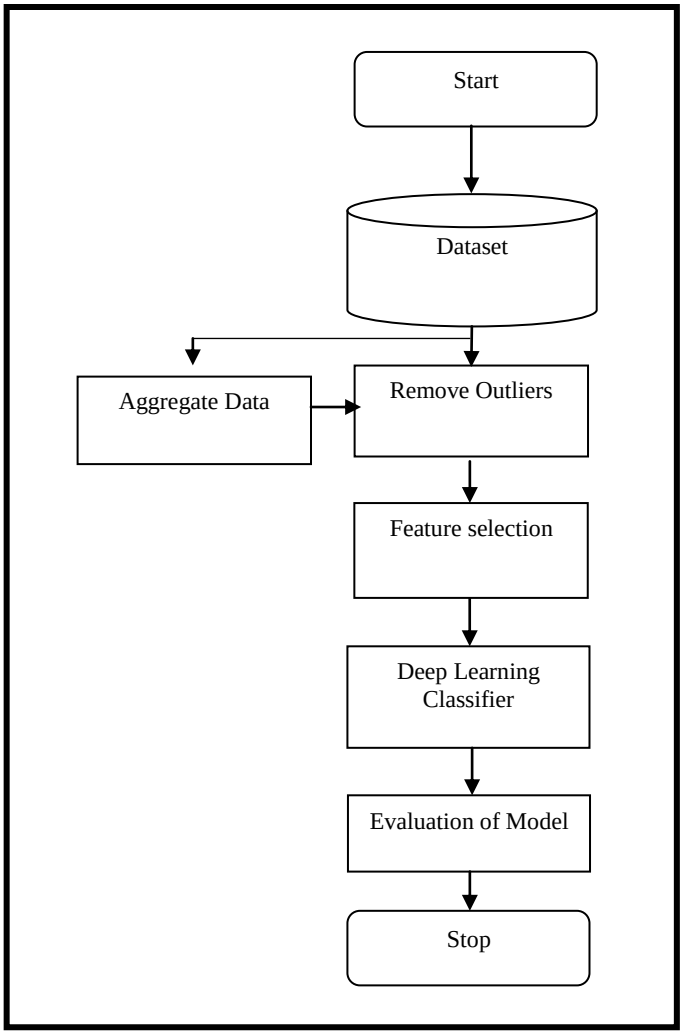


Figure 3: Intrusion Detection System using Deep Learning

3.3 Classification Algorithm

Following data preparation, which included pre-processing, outlier removal, selecting features, and splitting into training and test sets, we proceeded to model training using several deep learning techniques. This approach allows us to assess the effectiveness of numerous algorithms and select the most efficient option for intrusion detection. The methods used include Deep Belief Network, ResNet50, and CNN-LSTM.

4. Experimental Result and Discussion

Deep learning techniques were used to assess the proposed Intrusion Detection System's (IDS) performance on the InSDN dataset. To determine how feature selection affects

accuracy of detection, speed, and the effectiveness of resources, two separate feature sets were used for the evaluation: one with 48 features, and another with 30 features.

4.1 Comparison of Models with 48 Features

ResNet50, CNN-LSTM, and Deep Belief Network (DBN) are the three deep learning models whose results are shown in Table 1 employing 48 specifically selected features. Among the performance indicators are F1-Score, Accuracy, Precision, and Recall.

CNN-LSTM scored the highest, with a 99.5% accuracy, 99.6% precision, 98.3% recall, and a 99.0% F1-score. This demonstrates the model's exceptional ability to minimize false positives and false negatives while classifying both benign and malicious data.

DBN shown good performance as well, with a relatively high recall of 99.3% and an accuracy of 97.1%, indicating that it was successful in detecting real intrusions but had a little higher number of false positives than CNN-LSTM.

ResNet50, although efficient, it produced a strong accuracy of 97.0% while having the lowest metrics of the three.
These findings show that, particularly with powerful architectures like CNN-LSTM, adding additional features (in this case 48) improves detection rates and provides more contextual information for classification.

4.2 Comparison of Models with 30 Features

The analysis of the same models with a smaller set of 30 characteristics chosen by recursive feature removal is shown in Table 2. Testing the trade-off between detection performance and model complexity was the aim.

CNN-LSTM once more provided the top results, but somewhat less than the 48-feature model, with an F1-score of 98.9%, 99.1% accuracy, 99.2% precision, and 98.0% recall. This suggests that even with fewer input features, there is high adaptation capability.

DBN and ResNet50 achieved 96.3% and 96.7% accuracy, respectively. Even while the precision and recall were still good, there was a slight decrease, which could indicate that some important information was lost when characteristics were reduced.

The model results of the two feature sets are graphically compared in Figures 4 and 5. These graphic comparisons demonstrate CNN-LSTM's notable edge in all important criteria. CNN-LSTM is particularly suited for SDN situations where traffic behaviour is extremely dynamic because it uses deep sequential learning to efficiently capture spatial and temporal patterns in network traffic.

Table 1: Comparing Various Deep Learning Classifiers for CNN-LSTM, Deep Belief Network (DBN), and ResNet50 with 48 features

Model	Parameter	Values (%)
With 48 Features		
ResNet50	Accuracy	97.0
	Precision	97.0
	Recall	99.0
	F1-Score	98.0
Deep Belief Network (DBN)	Accuracy	97.1
	Precision	98.6
	Recall	99.3
	F1-Score	98.5
CNN-LSTM	Accuracy	99.5
	Precision	99.6
	Recall	98.3
	F1-Score	99.0

Table 2: Comparing Various Deep Learning Classifiers for CNN-LSTM, Deep Belief Network (DBN), and ResNet50 with 30 features

Model	Parameter	Values (%)
With 30 Features		
ResNet50	Accuracy	96.3
	Precision	96.5
	Recall	98.1
	F1-Score	98.0
Deep Belief Network (DBN)	Accuracy	96.7
	Precision	97.1
	Recall	98.3
	F1-Score	98.0
CNN-LSTM	Accuracy	99.1
	Precision	99.2
	Recall	98.0
	F1-Score	98.9

4.3 Optimizing Hyperparameters Using Optuna

Optuna, an autonomous hyperparameter optimization framework, was used to optimize the models' performance. By effectively exploring the hyperparameter space (learning rate, number of layers, dropout rates, etc.), this method reduced overfitting and training time while significantly improving the accuracy of the models. The hyperparameter-

tuned models regularly performed more than the default-setting models.

4.5 Discussion

The outcomes of the experiment highlight the significance of:
• In IDS being, feature engineering is particularly important in high-dimensional information contexts such as SDN.
• CNN-LSTM has proven to be a highly successful model for both compact and big feature sets.
• A key factor in reaching nearly ideal settings for several classifiers was hyperparameter tuning.

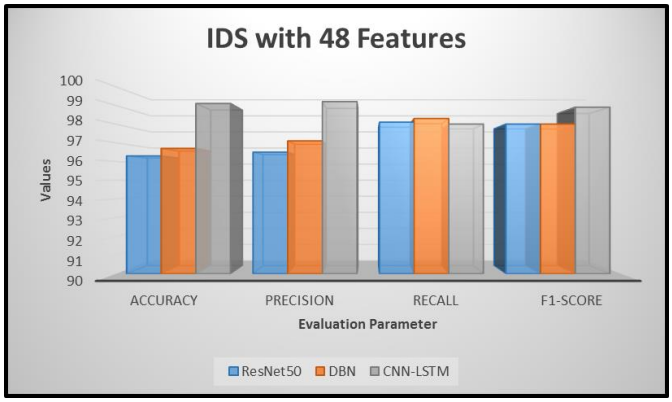


Figure 4: Comparing Outcome for Proposed Models using InSDN dataset with 48 Features

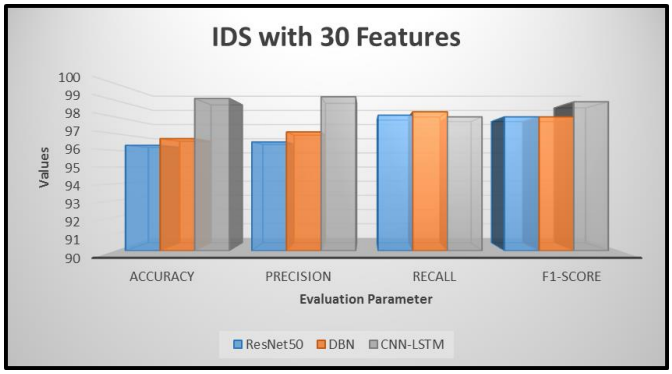


Figure 5: Comparing Outcome for Proposed Models using InSDN dataset with 30 Features

5. Conclusion and Future Scope

This study presents the creation of an efficient, adaptable, and dynamic intrusion detection system using a reduced feature selection approach and a deep learning algorithm. Feature engineering is a significant challenge in the development of an intrusion detection system. A straightforward feature selection approach is developed for dimensionality reduction to extract key features from the dataset. A Recursive Feature Elimination (RFE) combined with a Local Outlier Factor (LOF) simplified mutual information feature selection approach is suggested to identify the relevant features for classifiers. Each feature selection process exhibits little computational time complexity, averaging roughly 0.62 µs, while using less memory resources. The chosen subset of characteristics was then input into a deep learning method for multi-class classification models. The efficacy of the aforementioned models is assessed for multiclass applications. The performance of the extracted feature subset, when trained and evaluated with a deep learning algorithm for multiclass classification, was reported to be 99.5% using just 48 out of 84 features and 98.9% using only 30 out of 84 features from the InSDN dataset using CNN-LSTM. In the future, we will concentrate on the quantity of neurones in each layer and the number of layers inside each concealed layer, which is a primary topic of study.

Author's statements

Acknowledgements- I would like to express my sincere gratitude to my supervisor, Dr. Upendra Nath Tripathi, for his valuable guidance and continuous support throughout the course of this research. I am also thankful to my colleague, Ms. Sangeeta Devi, for her cooperation and helpful suggestions. I deeply appreciate the encouragement and support of my family, which played a vital role in the successful completion of this study. The authors are grateful for the reviewer's valuable comments that improved the manuscript.

Funding Source: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Authors' Contributions: Pranjal Maurya: Conceptualization, Data Collection, Formal Analysis, Writing – Original Draft. **Sangeeta Devi:** Methodology, Literature Review, Visualization, Writing – Review & Editing. **Dr. Upendra Nath Tripathi (Supervisor):** Supervision, Project Administration, Validation, Critical Revision of the Manuscript. All authors reviewed and edited the manuscript and approved the final version of the manuscript.

Conflict of Interest: The author declares that there is no conflict of interest regarding the publication of this research paper.

Data Availability: The data used in this study comes from the InSDN dataset, which is publicly available and specifically designed for intrusion detection research in Software-Defined Networking (SDN) environments. The dataset includes a wide range of traffic types, including benign traffic and various attack scenarios such as DoS, DDoS, probing, botnet, and password-guessing attacks. It contains over 43,000 records with 84 features, including both packet-level and flow-level attributes.
The InSDN dataset can be accessed freely at: http://iotseclab.ucd.ie/datasets/SDN/

References

[1] A. Kaur, C. R. Krishna, and N. V. Patil, "A comprehensive review on Software-Defined Networking (SDN) and DDoS attacks: Ecosystem, taxonomy, traffic engineering, challenges and research directions," *Computer Science Review*, Vol.**55**, pp.**100692**, **2025.**

[2] D. S. N. Wijesekara and P. Arachchige, "Intrusion detection using blockchain in Software-Defined Networking: A literature review," *Journal of Engineering Science & Technology Review*, Vol.**18**, No.**1**, **2025**.

[3] A. Vijayan and A. Anitha, "A review of intrusion detection systems in Software-Defined Networks," in *Proc. 2025 Int. Conf. Electronics and Renewable Systems* (ICEARS), Feb., pp.**859–864, 2025.**
[4] A. H. Janabi, T. Kanakis, and M. Johnson, "Survey: Intrusion Detection System in Software-Defined Networking," *IEEE Access*, Vol.**12**, pp.**164097–164120, 2024.**
[5] A. O. Salau and M. M. Beyene, "Software-defined-networking-based network traffic classification using machine-learning techniques," *Scientific Reports*, Vol.**14**, No.**1**, **2024.**
[6] D. Nuñez-Agurto, W. Fuertes, L. Marrone, E. Benavides-Astudillo, C. Coronel-Guerrero, and F. Perez, "A novel traffic classification approach by employing deep learning on Software-Defined Networking," *Future Internet*, Vol.**16**, No.**5**, **2024**.
[7] F. Ahmed, I. A. Sumra, and U. Jamil, "A comprehensive review on DDoS attack in Software-Defined Network (SDN): Problems and possible solutions," *Journal of Computing & Biomedical Informatics*, Vol.**7**, No.**1**, pp.**353–363, 2024.**
[8] M. Zhong, M. Lin, C. Zhang, and Z. Xu, "A survey on graph neural networks for intrusion detection systems: Methods, trends and challenges," *Computers & Security*, Vol.**141**, **2024.**
[9] H. Satılmiş, S. Akleylek, and Z. Y. Tok, "A systematic literature review on host-based intrusion detection systems," *IEEE Access*, Vol.**12**, pp.**27237–27266, 2024.**
[10] D. Srivastava, V. Sharma, P. Kumar, and Y. K. Sharma, "A comparative analysis of liver-cancer detection using neural-network algorithm and correlation algorithm," in *Proc. 2nd Int. Conf. Computational and Characterization Techniques in Engineering & Sciences* (IC3TES), Nov., pp.**1–5, 2024.**
[11] R. A. Abed, E. K. Hamza, and A. J. Humaidi, "A modified CNN-IDS model for enhancing the efficacy of intrusion detection system," *Measurement: Sensors*, Vol.**35**, **2024.**
[12] D. Srivastava *et al.*, "Deep ensemble model for sequence-based prediction of PPI: Self-improved optimization assisted intelligent model," *Multimedia Tools and Applications*, Vol.**83**, No.**26**, pp.**68135–68154, 2024.**
13] F. Zhao, H. Li, K. Niu, J. Shi, and R. Song, "Application of deep-learning-based intrusion detection system (IDS) in network anomaly traffic detection," *Applied and Computational Engineering*, Vol.**86**, pp.**231–237, 2024.**
[14] J. Arevalo-Herrera, J. Camargo Mendoza, J. I. Martínez Torre, T. Zona-Ortiz, and J. M. Ramirez, "Assessing SDN controller vulnerabilities: A survey on attack typologies, detection mechanisms, controller selection, and dataset application in machine learning," *Wireless Personal Communications*, Vol.**140**, No.**1**, pp.**739–775, 2025.**
[15] N. Mostafa, K. Metwally, and K. Badran, "Survey on SDN-based intrusion detection systems," in *Proc. 14th Int. Conf. Electrical Engineering* (ICEENG), May, pp.**317–322, 2024.**
[16] M. Chouikik, M. Ouaissa, M. Ouaissa, Z. Boulouard, and M. Kissi, "Detection and mitigation of DDoS attacks in SDN-based intrusion detection system," *Bulletin of Electrical Engineering and Informatics*, Vol.**13**, No.**4**, pp.**2750–2757, 2024.**
[17] A. Chetouane and K. Karoui, "New continual federated learning system for intrusion detection in SDN-based edge computing," *Concurrency and Computation: Practice and Experience*, Vol.**37**, No.**2**, **2025.**
[18] S. M'Rabet, H. Sahli, B. Yorobi, and M. Sayadi, "An inventive network intrusion detection system: Composite deep learning CNN-LSTM model," in *Proc. IEEE 6th Int. Conf. Image Processing, Applications and Systems* (IPAS), Jan., pp.**1–6, 2025.**
[19] U. K. Addanki, R. B. Devareddi, K. K. Kamarajugadda, M. Pavani, and P. Gera, "Machine-learning-powered intrusion detection system for Agriculture 4.0: Securing the next generation of farming," *International Journal of Innovative Research and Scientific Studies*, Vol.**8**, No.**1**, pp.**1964–1978, 2025.**
[20] V. Sharma, L. Kumar, and D. Srivastava, "Machine-learning-based prediction of users' involvement on social media," in *Advanced Applications of NLP and Deep Learning in Social Media Data*, Hershey, PA, USA: IGI Global, pp.**151–170, 2023.**
[21] V. Sharma, D. Srivastava, L. Kumar, M. Payal, and M. S. Adhikari, "Machine-learning-based image compression by reducing dimensionality," preprint, Feb. **2024.**
[22] R. Arthi, S. Krishnaveni, and S. Zeadally, "An intelligent SDN-IoT enabled intrusion detection system for healthcare systems using a hybrid deep learning and machine learning approach," *China Communications*, pp.**1–21, 2024.**

AUTHORS PROFILE

Pranjal Maurya received the Bachelor of Technology (B.Tech.) in Computer Science Engineering of Technology & Management and Master of Technology (M.Tech.) in Computer Science Engineering (CSE) from Madan Mohan Malaviya University of Technology. She is currently Ph.D. research Scholar in the Department of Computer Science, DDU Gorakhpur University. Her research interest includes WSN, Cloud Computing, IoT, Machine Learning and Deep Learning. She was previously working in Institute of Technology & Management as Assistant Professor for 1 years.

Sangeeta Devi received the Master of Computer Application (MCA) from IGNOU New Delhi and Master of Technology (M.Tech.) from AKTU Lucknow. She is currently Ph.D. research Scholar in the Department of Computer Science, DDU Gorakhpur University. Her research interest includes Data Science, WSN, IoT, Machine Learning and Deep Learning.

Dr. Upendra Nath Tripathi is currently Associate Professor in the Department of Computer Science, Deen Dayal Upadhyaya Gorakhpur University, Gorakhpur. He has 21 years of teaching and research experience. His areas of interests are Database, IoT, Machine Learning, Cloud Computing and Data Science.

