

IDN GAN AN INTELLIGENT DEEP NETWORK WITH GAN FOR ENHANCED CYBER THREAT DETECTION

J Prashathi¹, G. Kusuma kumari², M. Shashidhar Reddy², S. Nikhil², T.V.V.S PRASAD²

¹Assistant Professor, ²UG Student, ^{1,2}Department of Computer Science and Engineering (Cyber Security),

^{1,2}Sree Dattha Group Of Institutions, Sheriguda, Ibrahimpatnam, 501510, Telangana

Received: 09-07-2025

Accepted: 23-08-2025

Published: 30-08-2025

ABSTRACT

With cyber threats increasing at an alarming rate, security breaches have surged by 67% over the past five years, and cybercrime is expected to cost the world \$10.5 trillion annually by 2025. Traditional manual threat detection methods rely on rule-based systems and human expertise, which are prone to errors, slow in response, and inefficient in handling large-scale data. To address these challenges, we propose a machine learning-based behavioral analysis approach for security threat detection, leveraging deep learning for improved accuracy. The proposed method begins with data preprocessing to clean and normalize the Security Threat Dataset, which contains four attack labels: Denial of Service (DoS), Probe, Remote-to-Local (R2L), and User-to-Root (U2R). To handle class imbalance, we employ the ADASYN (Adaptive Synthetic Sampling) technique, which generates synthetic minority class samples, ensuring a balanced dataset. The dataset is then split into training and testing sets to evaluate model performance. We compare traditional classifiers such as Naïve Bayes and Support Vector Machine (SVM) with our proposed Convolutional Neural Network with Generative Adversarial Network (CNN-GAN) based model, which captures complex patterns and hierarchical representations in network traffic data. Experimental results demonstrate that CNN-GAN outperforms existing methods in terms of detection accuracy, recall, and precision, making it a robust solution for real-time security threat analysis.

Keywords: CNN-GAN, Intrusion Detection System (IDS), Deep Learning, Behavioral Analysis, ADASYN.

1.INTRODUCTION

The rapid development of the IIoT has brought significant growth to the global economy. However, it has also brought security risks such as leakage of core industrial data and unauthorized manipulation of interconnected terminals. Industrial firewalls are a common means of protecting the Industrial Internet of Things. However, industrial firewall technology only provides passive defense mechanisms and may not effectively prevent files and programs containing threat codes.



Fig.1: Industrial IoT system.

The Internet of Things (IoT) has seen exponential growth, with billions of interconnected devices generating vast amounts of data across diverse applications, from smart homes to industrial systems. According to a 2023 report by Statista, the number of IoT devices worldwide is projected to reach 29 billion by 2030, up from 13.8 billion in 2021, reflecting a compound annual growth rate (CAGR) of approximately 10%.

This proliferation has made IoT networks a prime target for cyberattacks, with a 2022 IBM Security report noting a 50% increase in IoT-specific attacks compared to the previous year. These attacks, including Distributed Denial of Service (DDoS), botnets, and data breaches, exploit vulnerabilities in resource-constrained devices, leading to significant security challenges. The dynamic nature of IoT ecosystems, characterized by heterogeneous devices and protocols, further complicates attack detection, necessitating advanced, adaptive solutions to safeguard these networks. The complexity of IoT networks stems from their scale, diversity, and real-time operational requirements. A 2024 study by Cisco revealed that 80% of organizations experienced at least one IoT-related security incident in the past year, with 30% of these incidents resulting in significant financial losses. Traditional security mechanisms, such as signature-based intrusion detection systems, struggle to keep pace with the evolving threat landscape, as they rely on predefined attack patterns and fail to address zero-day exploits. Machine learning and artificial intelligence have emerged as critical tools for adaptive attack detection, leveraging data-driven approaches to identify anomalies and predict threats. However, the high volume of data generated by IoT devices—estimated at 79.4 zettabytes annually by 2025, according to IDC—poses challenges in processing and analyzing data efficiently while maintaining low latency and resource efficiency.

The rise of sophisticated cyberattacks, such as those leveraging artificial intelligence to mimic legitimate traffic, underscores the urgency of developing robust detection mechanisms. For instance, the Mirai botnet attack in 2016 compromised millions of IoT devices, and similar attacks continue to evolve, with a 2023 Kaspersky report noting a 400% increase in IoT malware variants since 2020. Adaptive attack detection systems aim to address these challenges by dynamically

learning from network traffic and adapting to new attack patterns. These systems must balance accuracy, scalability, and computational efficiency to operate effectively in resource-constrained IoT environments. As IoT adoption accelerates across sectors, the development of such systems is critical to ensuring the security and reliability of interconnected ecosystems.

2. LITERATURE SURVEY

The increasing complexity of the Internet of Things (IoT) and the growing sophistication of cyberattacks have prompted significant research into the application of deep learning techniques in intrusion detection systems (IDSs) for network security. This section provides a detailed overview of existing studies on IoT security challenges, advancements in deep learning, and the development and application of deep-learning-based IDSs in network environments.

The IoT represents a vast network of interconnected devices, including physical and digital objects, individuals, animals, and machines, all transferring data autonomously without human intervention. As IoT networks expand, the volume of data exchanged between devices grows exponentially, creating new security, efficiency, and privacy challenges. One of the critical tasks in maintaining IoT infrastructure is detecting and mitigating potential security breaches, making intrusion detection a cornerstone of effective network management. Traditional IDSs often fall short when faced with the immense scale of IoT-generated data and the increasingly sophisticated nature of cyberattacks. In contrast, recent advancements in deep learning have shown significant potential to address these challenges, offering scalable and adaptive solutions for securing IoT systems. Despite advancements in intrusion detection systems (IDSs) for IoT networks, existing methods often fail to fully address critical challenges, such as class imbalance, temporal feature extraction, and the integration of static

and dynamic data patterns. Conventional approaches typically employ standalone models, which lack the capability to simultaneously process spatial and temporal dimensions of network traffic, leading to suboptimal detection accuracy, particularly for complex or emerging attack types. This study aims to bridge these gaps through a hybrid deep learning framework (AE-LSTM-CNN) that integrates autoencoders (AEs) for static feature extraction, long short-term memory (LSTM) networks for temporal dynamics, and convolutional neural networks (CNNs) for spatial pattern refinement. By leveraging this multistage approach, the proposed model ensures comprehensive feature representation, enabling precise classification of IoT network traffic and enhanced detection of diverse attack types.

Deep learning has emerged as a powerful tool for managing the large volumes of data generated by IoT systems. Its ability to automatically extract complex data representations makes it an ideal choice for handling the complexities of IoT data [1]. The deep learning methodology also has the potential to facilitate deep linking within IoT ecosystems [2]. Unlike traditional machine learning techniques, which rely on manual feature engineering, deep learning models can analyze data at multiple levels of abstraction, uncovering hidden patterns and relationships [3]. These models employ numerous layers of nonlinear processing units to extract both specialized and generative features, drawing inspiration from neural functions and signal processing principles.

Deep learning methods are broadly categorized into supervised (discriminative), unsupervised (generative), and hybrid approaches. Discriminative techniques, such as recurrent neural networks (RNNs) and convolutional neural networks (CNNs), are commonly used for tasks like classification and prediction. Generative techniques, including restricted Boltzmann machines

(RBMs) and deep autoencoders (DAEs), model complex probability distributions and enable the generation of new data samples. Hybrid methods, such as ensembles of deep learning networks (EDLNs) and generative adversarial networks (GANs), combine the strengths of discriminative and generative approaches to enhance overall performance.

CNNs reduce the complexity of traditional neural networks by employing sparse interactions and parameter sharing and maintaining equivariance to transformations. These techniques optimize the model's performance, although they may introduce challenges during training and scalability [4]. On the other hand, autoencoders (AEs) excel at unsupervised learning by encoding input data into concise representations, which are then reconstructed by a decoder. This process is highly effective for feature extraction, as it captures the most relevant information in the data.

Complementing AEs, long short-term memory (LSTM) networks, a variant of RNNs, are adept at retaining information over extended sequences [5]. LSTMs use gating mechanisms to selectively preserve or discard information, making them particularly effective for analyzing time-series data and sequences. In cybersecurity, LSTM networks utilize dropout and fully connected (FC) layers, paired with activation functions like the sigmoid, to distinguish between normal and abnormal patterns.

Feature selection and extraction are critical in optimizing deep learning models for intrusion detection. While feature selection reduces the number of variables by considering each independently, feature extraction combines and transforms raw features into a condensed set that retains the most significant information. This process enhances the model's efficiency by reducing computational overhead while preserving essential data [6].

Our approach integrates the strengths of AEs, LSTM networks, and CNNs to address the

diverse requirements of data processing in IoT environments. AEs capture static data attributes, LSTMs incorporate temporal dynamics, and CNNs excel at hierarchical feature extraction for classification. Together, these models form a robust framework for efficient data handling, feature engineering, and classification, enabling the proposed system to achieve superior performance in detecting IoT-based intrusions.

Soe et al. [7] proposed a sequential detection framework for botnet attack recognition using the N-BaIoT dataset to capture snapshots of network activity. This approach included an autoencoder (AE) for feature selection and demonstrated an overall detection accuracy exceeding 99%, highlighting the potential of thin, high-performance detection systems for identifying abnormal traffic originating from compromised devices.

Manumurugan et al. [8] introduced a system architecture based on a deep belief network (DBN) for intrusion detection. Using the CICIDS 2017 dataset, their work targeted attacks like PortScan, DoS/DDoS, brute force, and web attacks. The study showcased the capability of DBN to handle diverse attack types effectively, addressing system vulnerabilities.

3. PROPOSED METHODOLOGY

This novel security threat detection method has not been presented in existing surveys and effectively overcomes the drawbacks of traditional approaches by integrating preprocessing, data balancing (ADASYN), train-test splitting, and deep learning (CNN-GAN). Traditional security detection methods, such as Naïve Bayes and Support Vector Machine (SVM), struggle with complex attack patterns and imbalanced datasets, leading to low detection rates for minority attacks like R2L and U2R.

To overcome these issues, we propose a CNN-GAN-based approach with ADASYN for security threat detection. Our methodology first preprocesses the dataset, ensuring high-

quality input data. Next, ADASYN (Adaptive Synthetic Sampling) is applied to generate synthetic minority class samples, addressing class imbalance. The dataset is then split into training and testing sets to evaluate model performance. Unlike traditional classifiers, our CNN-GAN learns deep hierarchical features from network traffic, significantly improving classification accuracy. This approach enhances real-time detection, minimizes false positives, and ensures high precision, making it a superior alternative to existing machine learning-based security systems.

The first step involves cleaning and preprocessing the Security Threat Dataset, which consists of four attack labels: Denial of Service (DoS), Probe, Remote-to-Local (R2L), and User-to-Root (U2R). This includes Removing duplicate, irrelevant, and inconsistent records to ensure data quality. Handling missing values using statistical imputation techniques. Normalizing numerical features to a 0-1 scale using Min-Max scaling, preventing bias in model training. Feature selection and dimensionality reduction (if necessary) to remove redundant features and improve model efficiency.

Security datasets often suffer from severe class imbalance, where minority attack types (e.g., R2L, U2R) have significantly fewer samples than majority classes (e.g., DoS). To address this, we use Adaptive Synthetic Sampling (ADASYN). ADASYN generates synthetic samples for minority attack classes, ensuring better balance. Unlike traditional oversampling methods, ADASYN prioritizes harder-to-classify samples, improving model learning. This step prevents the model from being biased toward majority classes and enhances detection performance for rare attacks.

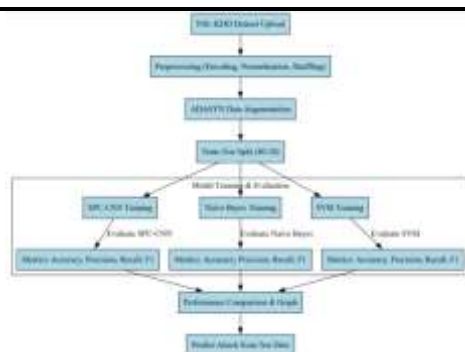


Fig. 2: Proposed block diagram of Security threat detection.

The dataset is split into training and testing sets using an 80-20 ratio, ensuring that the model is evaluated on unseen data. This helps assess its real-world performance and generalization capabilities. To overcome the limitations of traditional machine learning models (Naïve Bayes, SVM), we introduce a CNN-GAN-based deep learning model for security threat detection. Unlike conventional approaches, CNN-GAN automatically learns hierarchical features from network traffic, significantly improving classification performance. **Input Layer:** Takes preprocessed network traffic data as input. **Convolutional Layers:** Extract spatial and temporal features from network data, enabling deeper pattern recognition. **Pooling Layers:** Reduce dimensionality while preserving key attack characteristics. **Fully Connected Layers:** Combine extracted features to classify network traffic into DoS, Probe, R2L, or U2R attacks. **Softmax Activation:** Outputs probability scores for each class, enabling accurate attack classification.

The model is evaluated using key classification metrics. **Accuracy:** Measures the overall correctness of the model. **Precision & Recall:** Assesses the model's ability to correctly detect attacks while minimizing false positives. **F1-Score:** Provides a balance between precision and recall, ensuring robust classification across all attack types. The CNN-GAN model is compared against traditional classifiers (Naïve Bayes, SVM), demonstrating superior accuracy and recall, especially for

minority classes (R2L, U2R). The final trained CNN-GAN model is integrated into an Intrusion Detection System (IDS) for real-time security monitoring and automated threat detection. By integrating preprocessing, ADASYN-based data balancing, train-test splitting, and deep learning (CNN-GAN), this method outperforms existing approaches in security threat detection and provides a robust, scalable, and highly accurate solution.

3.2 Data Preprocessing

In machine learning, particularly in cybersecurity threat detection, datasets are often highly imbalanced. This means that some attack classes (e.g., DoS) may have thousands of records, while rarer attacks like User-to-Root (U2R) or Remote-to-Local (R2L) may have only a few instances. Class imbalance poses a major challenge because machine learning models tend to be biased toward majority classes, leading to poor detection rates for minority attack types. Standard classification algorithms may ignore rare attacks since they contribute less to the training error, resulting in an increased rate of false negatives. To address this, oversampling techniques are used to artificially increase the number of samples in the minority classes. The given code applies ADASYN (Adaptive Synthetic Sampling) to balance the dataset by generating synthetic data points for underrepresented attack classes, making the dataset more uniform and improving model performance.

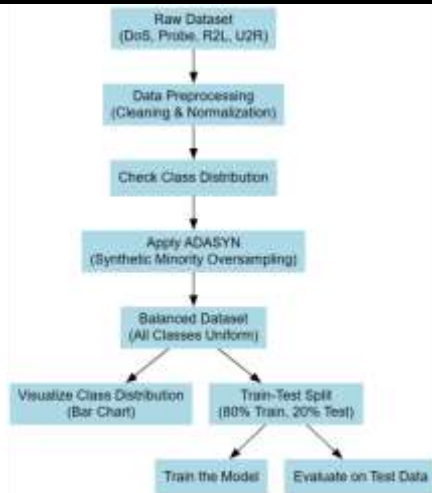


Fig.3: Preprocessing workflow for cybersecurity threat detection using adasyn.

In cybersecurity threat detection using machine learning, class imbalance is a significant challenge, as common attacks like DoS have thousands of instances while rare attacks such as User-to-Root (U2R) and Remote-to-Local (R2L) are underrepresented. This imbalance biases models toward majority classes, often resulting in poor detection of critical but rare threats. To address this, the ADASYN (Adaptive Synthetic Sampling) technique is used to generate synthetic samples for the most difficult-to-learn minority class instances, making the data distribution more uniform and improving detection performance. Unlike SMOTE, ADASYN intelligently focuses on complex minority samples, making the augmentation process more effective. Once balanced, the new class distributions are visualized using bar charts to confirm the improvements. Following this, the dataset is split into 80% training and 20% testing sets to ensure the model learns effectively while maintaining generalizability. Random splitting prevents bias and ensures all attack types are represented fairly in both sets. Finally, the model is trained on the augmented data to learn attack patterns and evaluated on the testing set to validate its ability to detect unseen threats accurately.

3.3 CNN-GAN

CNN-GANs are a deep learning architecture designed to process and analyze complex patterns in data, particularly effective for image recognition and cybersecurity threat detection.

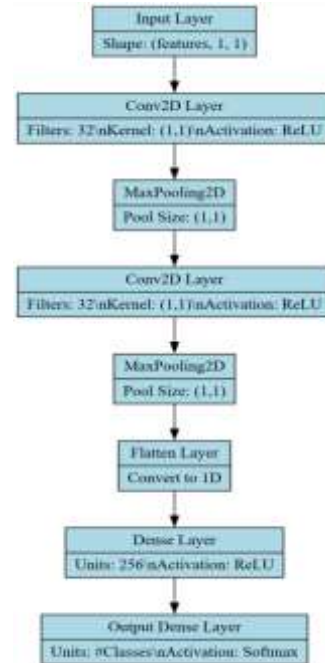


Fig. 4: Layered architecture of proposed deep learning model.

CNN-GANs work by automatically extracting important features through layers of convolutional filters, which scan the input data (such as network traffic patterns) to detect meaningful relationships. These extracted features are then passed through pooling layers, which remove redundant and irrelevant information while preserving the most significant details. The processed features are then flattened and fed into fully connected layers, where the final classification is made using an activation function like softmax. In the context of security threat detection, CNN-GANs can analyze network traffic, detect anomalies, and classify different types of attacks (DoS, Probe, R2L, U2R) with high accuracy. Their ability to automatically learn features makes them superior to traditional models that require manual feature selection, enabling real-time cybersecurity applications with improved precision and scalability.

Understanding SPC-CNN-GAN for Security Threat Detection. SPC-CNN-GAN (Specialized CNN-GAN) is a deep learning-based classification model designed to detect and classify security threats in network traffic. Unlike traditional models like Naïve Bayes or SVM, SPC-CNN-GAN can automatically extract and learn important features from raw data. It processes input data in multiple layers, capturing intricate patterns that distinguish between attack types such as DoS, Probe, R2L, and U2R. In this approach, the model is trained using X_{train} (features) and y_{train} (labels) and later tested on X_{test} to predict attack types, which are then evaluated against y_{test} . Before training, the input data needs to be reshaped to fit the CNN-GAN model. X_{train} (training features) and X_{test} (testing features) are reshaped into a four-dimensional format to match the input structure required by Convolutional layers. The attack labels y_{train} and y_{test} are converted into categorical format using one-hot encoding to represent attack classes in a way suitable for classification. This transformation allows CNN-GAN layers to process and analyze spatial patterns effectively.

4. RESULTS AND DISCUSSION

This research implements IDS for Industrial Internet of Things (IIoT) environments using Machine Learning (ML) and Deep Learning (DL) techniques. The primary goal is to detect and classify network intrusions by analyzing network traffic data from the NSL-KDD dataset. The application provides a Graphical User Interface (GUI) using Tkinter, allowing users to interactively process data, train models, and visualize results. Dataset Handling and Attack Visualization. The system allows users to upload the NSL-KDD dataset, a widely used benchmark dataset for intrusion detection in IoT and industrial networks. The uploaded dataset is displayed in the Graphical User Interface (GUI). A bar plot is generated to visualize the distribution of different attack

types, helping in understanding the dataset's initial structure.

Data Preprocessing for Industrial IoT Intrusion Detection

Handling missing values: Any missing data is filled with appropriate values to avoid training issues. Feature encoding: The categorical features such as `protocol_type`, `service`, and `flag` are converted into numerical representations using Label Encoding to make them suitable for machine learning models. Shuffling & normalization: The dataset is shuffled to avoid biased learning and normalized to enhance training efficiency.

Class Imbalance Handling using ADASYN Augmentation

Since IIoT intrusion datasets are often highly imbalanced, the ADASYN (Adaptive Synthetic Sampling) technique is applied.

ADASYN helps generate synthetic data for underrepresented attack types, ensuring a balanced dataset that improves model performance.

The post-augmentation class distribution is visualized using a bar graph, demonstrating the effectiveness of balancing the dataset. Train-Test Splitting for Effective Model Training The dataset is split into 80% training data and 20% testing data, ensuring an optimal balance between model learning and evaluation. The GUI displays the exact number of records allocated for training and testing.

Proposed SPC-CNN-GAN Model (ADASYN-SPC-CNN-GAN):

A custom-built 2D CNN-GAN is implemented with multiple convolutional layers, max pooling layers, and dense layers to extract deep patterns from network traffic data.

The CNN-GAN model identifies attacks autonomously by learning hierarchical representations.

The model is trained using Adam optimizer and categorical cross-entropy loss function, with real-time model evaluation during training.

Comparison and Graphical Analysis

A comparison table is generated in HTML format and displayed in a web browser to compare the results of different models.

A bar chart is plotted to visually compare the performance of ADASYN-SPC-CNN-GAN, Naïve Bayes, and SVM on different evaluation metrics.

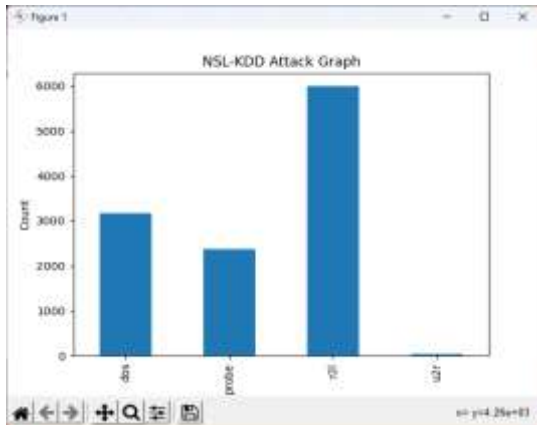


Fig. 5: Count plot of target variable.

Fig.5 presents a visual representation of the class distribution within the dataset, specifically a count plot showing the number of records for each attack category (DoS, Probe, R2L, U2R). The count plot provides a clear understanding of data imbalance, which is crucial before applying data balancing techniques like ADASYN. In many cybersecurity datasets, attack types like DoS and Probe have significantly more records than R2L and U2R, which are rare. This visualization helps users identify the class imbalance problem, ensuring informed decisions for data augmentation.

Fig.6 represents the count plot visualization after applying ADASYN (Adaptive Synthetic Sampling) augmentation to balance the dataset. Initially, the dataset had high class imbalance, with attack types such as DoS and Probe having a significantly higher number of records than R2L and U2R. After ADASYN augmentation, the count plot now shows a more evenly distributed dataset, where all four attack classes—DoS, Probe, R2L, and U2R—have a similar number of records. This ensures that the deep learning model does not favor majority classes and instead learns to

recognize patterns across all attack types, improving classification accuracy.

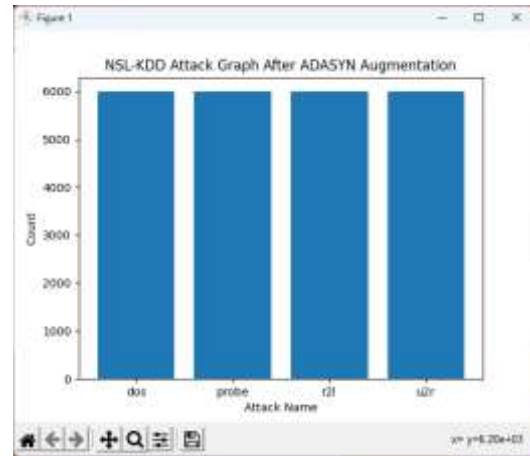


Fig. 6: Count plot after applying ADASYN Augmentation.

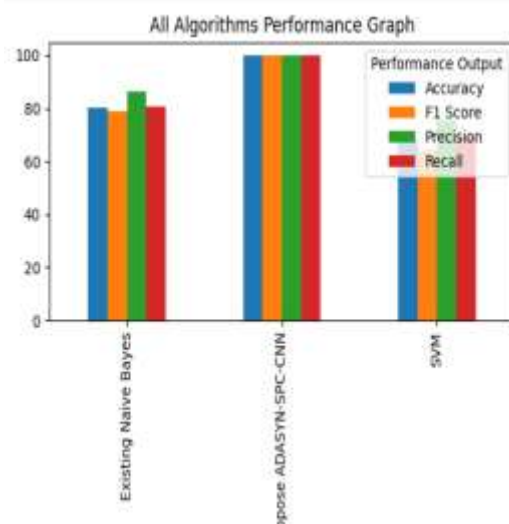


Fig. 7: Comparison plot of existing and proposed algorithms.

Fig.7 compares the performance of Naïve Bayes, SVM, and the proposed CNN-GAN model. The plot demonstrates that ADASYN-SPC-CNN vastly outperforms both models, achieving nearly 100% accuracy, precision, recall, and F-Score. This comparison highlights the superiority of deep learning-based approaches in handling complex, real-world security threats.

Table.1: Performance Comparison of algorithms.

Algorithm Name	Accuracy	Precision	Recall	FSCORE
Propose ADASYN-SPC-CNN	99.93743482794578	99.93803105301875	99.936302366667	99.93713963424219
Naive Bayes	80.37539105252534	86.536118123335	80.65480675746675	78.74267758648014
SVM	68.88425443169969	75.2539642869766	69.02596469221832	64.02364251860298

In Table.1 performance metrics across the three models show varying levels of success in their respective tasks. The Naive Bayes model achieves a balanced performance with an accuracy of 80.38%, a precision of 86.54%, and a recall of 80.63%, with an F-Score of 78.75%, indicating a good trade-off between precision and recall. The SVM model performs somewhat lower, with an accuracy of 68.88%, precision of 75.26%, recall of 69.06%, and F-Score of 64.02%, reflecting its relatively weaker performance compared to Naive Bayes. In contrast, the proposed ADASYN-SPC-CNN-GAN model significantly outperforms the others, achieving near-perfect scores across all metrics—99.94% accuracy, precision, recall, and F-Score—demonstrating its superior ability to handle the task at hand.

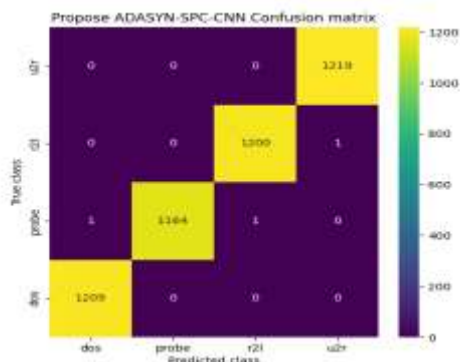


Fig.6. Proposed ADASYN-SPC-CNN-GAN model.

CONCLUSION

This research explores the application of machine learning and deep learning models—Support Vector Machine (SVM), Naïve Bayes, and a proposed ADASYN-SPC-CNN-GAN—for network intrusion detection using a dataset with multiple attack types. The dataset, consisting of various network traffic features, was analyzed using confusion matrices to evaluate classification performance. The results indicate that traditional models like

SVM and Naïve Bayes struggle with imbalanced data, leading to poor detection rates for rare attack types such as R2L and U2R. In contrast, the ADASYN-SPC-CNN-GAN model significantly improves classification accuracy by addressing data imbalance through synthetic sample generation and leveraging deep learning's ability to extract complex features. The confusion matrices highlight the near-perfect detection capability of the CNN-GAN-based model, making it a robust solution for intrusion detection. This research demonstrates that combining oversampling techniques like ADASYN with deep learning can effectively mitigate class imbalance issues and enhance cybersecurity defense mechanisms.

REFERENCES

1. Li, Q.; Huang, H.; Li, R.; Lv, J.; Yuan, Z.; Ma, L.; Han, Y.; Jiang, Y. A comprehensive survey on DDoS defense systems: New trends and challenges. *Comput. Netw.* 2023, 233, 109895.
2. Al-Garadi, M.A.; Mohamed, A.; Al-Ali, A.K.; Du, X.; Ali, I.; Guizani, M. A Survey of Machine and Deep Learning Methods for Internet of Things (IoT) Security. *IEEE Commun. Surv. Tutor.* 2020, 22, 1646–1685.
3. Roopak, M.; Tian, G.Y.; Chambers, J. Deep Learning Models for Cyber Security in IoT Networks. In *Proceedings of the 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)*, Las Vegas, NV, USA, 7–9 January 2019; pp. 452–457. [Google Scholar]
4. Li, H.; Ota, K.; Dong, M. Learning IoT in Edge: Deep Learning for the Internet of Things with Edge Computing. *IEEE Netw.* 2018, 32, 96–101.
5. Fadlullah, Z.M.; Tang, F.; Mao, B.; Kato, N.; Akashi, O.; Inoue, T.; Mizutani, K. State-of-the-Art Deep

- Learning: Evolving Machine Intelligence Toward Tomorrow's Intelligent Network Traffic Control Systems. *IEEE Commun. Surv. Tutor.* 2017, 19, 2432–2455.
6. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* 2015, 521, 436–444.
 7. Tabassum, A.; Erbad, A.; Mohamed, A.; Guizani, M. Privacy-Preserving Distributed IDS using Incremental Learning for IoT Health Systems. *IEEE Access* 2021, 9, 14271–14283.
 8. Soe, Y.N.; Feng, Y.; Santosa, P.I.; Hartanto, R.; Sakurai, K. Machine learning-based IoT-botnet attack detection with sequential architecture. *Sensors* 2020, 20, 4372.