IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Development of an Intelligent Hybrid GAN-Based Machine Learning and Deep Learning Framework for Accurate Cyber-Attack Detection and Network Intrusion Classification

Mr.SANJIV KUMAR, Dr. PANKAJ KAIRNAR
Department of COMPUTER SCIENCE Sikkim Alpine University
Kamrang, Namchi, Sikkim - 737126

ABSTRACT: The increasing sophistication and frequency of cyber-attacks have exposed the limitations of traditional signature-based intrusion detection systems in identifying evolving and previously unseen threats. The presence of highly imbalanced network traffic, scarcity of labeled attack samples, and the dynamic nature of cyber threats significantly reduce the effectiveness of conventional machine learning-based detection approaches. This research proposes an intelligent cyber-attack detection framework that integrates Generative Adversarial Networks (GANs) with advanced Machine Learning (ML) and Deep Learning (DL) techniques to improve the accurate detection and classification of network intrusions. The proposed framework employs GAN-based synthetic data generation to address class imbalance by producing realistic minority attack samples, thereby enhancing the quality and diversity of training datasets. Following data preprocessing and feature engineering, the augmented dataset is utilized to train multiple ML and DL models, including Random Forest, XGBoost, Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), and hybrid CNN-LSTM architectures. A comparative evaluation is performed using publicly available benchmark intrusion detection datasets such as CIC-IDS2017, CIC-IDS2018, NSL-KDD, UNSW-NB15, and Bot-IoT. The proposed framework is evaluated using performance metrics including Accuracy, Precision, Recall, F1-Score, Matthews Correlation Coefficient (MCC), Receiver Operating Characteristic-Area under the Curve (ROC-AUC), False Positive Rate (FPR), and computational efficiency. The integration of GAN-based data augmentation with hybrid learning models is expected to substantially improve minority attack detection, reduce false alarms, enhance model generalization, and strengthen robustness against emerging and zero-day cyber threats. The proposed framework provides a scalable, adaptive, and intelligent intrusion detection solution suitable for modern enterprise networks, cloud computing environments, Internet of Things (IoT) ecosystems, and other resource-constrained cyber infrastructures, thereby contributing to the development of next-generation intelligent cybersecurity systems. This work directly supports the objective of developing an integrated GAN-ML-DL framework capable of overcoming class imbalance and limited attack data while achieving highly accurate network intrusion detection.

Keyword: Cyber-Attack Detection; Network Intrusion Detection; Generative Adversarial Networks (GANs); Machine Learning; Deep Learning; Class Imbalance; Data Augmentation; Network Security.

Introduction: The rapid growth of cloud computing, IoT, and edge computing has significantly increased the cyber-attack surface, exposing organizations to sophisticated threats including ransomware, advanced persistent threats (APTs), and zero-day attacks [1], [2]. Traditional signature-based intrusion detection systems (IDSs) are effective only for known attacks and exhibit poor performance against evolving attack patterns [3], [4].Machine learning has become a promising solution for intelligent intrusion detection because it automatically learns attack patterns from network traffic without relying on predefined signatures [5], [6]. Algorithms such as Random Forest, Support Vector Machine, and XGBoost have demonstrated high classification performance on benchmark intrusion detection datasets [7], [8].

Deep learning architectures including Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and hybrid CNN–LSTM models further improve feature extraction and attack classification by capturing both spatial and temporal characteristics of network traffic [9]–[11].One of the major challenges in intelligent intrusion detection is the severe class imbalance present in cybersecurity datasets, where minority attack classes are significantly underrepresented. This imbalance causes classifiers to become biased toward majority classes, reducing detection accuracy for rare attacks [12], [13].

Generative Adversarial Networks (GANs) have emerged as an effective solution for addressing data scarcity and class imbalance by generating realistic synthetic attack samples. Recent studies have demonstrated that GAN-based data augmentation significantly improves minority attack detection and overall classifier performance [14]–[17].

Modern communication networks have become the backbone of digital transformation, supporting cloud computing, Internet of Things (IoT), Industrial Internet of Things (IIoT), smart cities, financial services, healthcare systems, and critical national infrastructures. While these technologies improve connectivity and operational efficiency, they also increase the complexity of cyber threats. The continuous growth of interconnected devices has expanded the attack surface, making organizations increasingly vulnerable to sophisticated cyber-attacks.

Cyber-attacks have evolved from simple malware infections to advanced persistent threats (APTs), ransomware, botnets, phishing campaigns, Distributed Denial-of-Service (DDoS) attacks, insider threats, and zero-day exploits. These attacks frequently evade traditional security solutions because they continuously modify their behavior and exploit unknown vulnerabilities. Consequently, organizations require intelligent intrusion detection systems capable of identifying both known and previously unseen attack patterns.

Traditional Intrusion Detection Systems (IDSs) primarily employ signature-based detection mechanisms. Although these systems achieve high accuracy against known attacks, they are unable to detect novel or polymorphic threats because they depend on predefined attack signatures. Anomaly-based intrusion detection partially addresses this limitation by learning normal network behavior and detecting deviations from established patterns. However, anomaly detection systems often suffer from high false-positive rates, reducing their effectiveness in real-world deployments [3].

Artificial Intelligence (AI), particularly Machine Learning (ML), has significantly transformed network intrusion detection by enabling automated learning from historical network traffic. Supervised learning algorithms including Random Forest, Decision Tree, Support Vector Machine, K-Nearest Neighbor, Naïve Bayes, and XGBoost have demonstrated promising performance in identifying malicious network activities. These algorithms automatically learn complex relationships among network features, eliminating the need for manually crafted detection rules. Nevertheless, conventional machine learning algorithms exhibit performance degradation when trained on highly imbalanced cybersecurity datasets, where minority attack categories contain only limited training samples [5].

Deep Learning (DL) techniques further improve intrusion detection through automatic hierarchical feature learning. Convolutional Neural Networks (CNNs) effectively capture spatial relationships among network traffic features, while Long Short-Term Memory (LSTM) networks learn temporal dependencies within sequential communication patterns. Hybrid CNN-LSTM architectures combine spatial and temporal feature extraction, enabling more accurate identification of sophisticated cyber-attacks. Despite their superior performance, deep learning models require large volumes of balanced training data, which remain unavailable for many rare attack categories[6].

Class imbalance represents one of the most significant challenges in cybersecurity analytics. Public benchmark datasets such as NSL-KDD, CIC-IDS2017, CIC-IDS2018, UNSW-NB15, and Bot-IoT contain disproportionately fewer samples for several attack classes compared with benign traffic. Consequently, classification models become biased toward majority classes, resulting in poor detection performance for minority attacks such as User-to-Root (U2R), Remote-to-Local (R2L), insider attacks, and zero-day threats. Traditional oversampling approaches, including Random Oversampling and Synthetic Minority Oversampling Technique (SMOTE), increase the number of minority samples but frequently generate duplicated or linearly interpolated instances that fail to preserve the statistical characteristics of genuine cyber-attacks.

Generative Adversarial Networks (GANs) have emerged as an effective solution for addressing class imbalance through realistic synthetic data generation. A GAN consists of two competing neural networks a generator and a discriminator that are simultaneously optimized through adversarial learning. The generator produces synthetic attack samples, whereas the discriminator distinguishes between real and generated data. This adversarial process enables the generator to learn the underlying distribution of minority attack classes, producing high-quality synthetic data that enhances classifier training. Compared with traditional oversampling methods, GAN-generated samples preserve nonlinear feature relationships and improve classifier generalization[8].

Integrating GANs with advanced machine learning and deep learning techniques offers a promising direction for intelligent intrusion detection. GAN-based data augmentation balances cybersecurity datasets before classifier training, thereby improving the representation of minority attack categories. Hybrid classifiers combining Random Forest, XGBoost, CNN, and LSTM exploit complementary learning capabilities to improve attack detection accuracy, reduce false-positive rates, and enhance resilience against emerging cyber threats. Such intelligent frameworks are particularly suitable for enterprise networks, cloud computing platforms, Industrial IoT environments, and software-defined networking infrastructures requiring real-time threat detection.

Despite considerable advances in intelligent intrusion detection, several research challenges remain unresolved. Existing studies frequently evaluate models using individual datasets, limiting their generalization across heterogeneous network environments. Many GAN-based intrusion detection models focus primarily on binary classification and fail to address complex multi-class attack scenarios. Furthermore, several deep learning approaches require significant computational resources, restricting their deployment in real-time cybersecurity applications. These limitations highlight the need for a comprehensive framework that integrates GAN-based data augmentation with advanced machine learning and deep learning techniques while simultaneously addressing class imbalance, limited attack data, computational efficiency, and accurate intrusion classification.

To address these challenges, this research proposes an Intelligent Hybrid GAN-Based Machine Learning and Deep Learning Framework for cyber-attack detection and network intrusion classification. The proposed framework employs GAN-based synthetic attack generation to overcome class imbalance, followed by feature engineering and hybrid classification using machine learning and deep learning models. The framework aims to improve detection accuracy, minority attack recognition, robustness against zero-day attacks, and computational efficiency across benchmark intrusion detection datasets. The proposed research contributes toward the development of adaptive, scalable, and intelligent cybersecurity systems capable of protecting modern digital infrastructures [9].

2. Related Work

The rapid growth of cyber threats has motivated extensive research on intelligent intrusion detection systems that leverage machine learning, deep learning, and generative artificial intelligence. Recent studies have focused on improving detection accuracy, reducing false-positive rates, and addressing challenges such as class imbalance, high-dimensional network traffic, and zero-day attack detection. This section reviews the existing literature on cyber-attack detection by categorizing previous work into traditional machine learning approaches, deep learning-based intrusion detection, Generative Adversarial Network (GAN)-based data augmentation, and hybrid intelligent frameworks. The strengths and limitations of current approaches are also discussed to identify the research gap addressed by the proposed framework [9].

2.1 Traditional Machine Learning-Based Cyber-Attack Detection

Machine learning (ML) techniques have been widely adopted for network intrusion detection because of their ability to learn discriminative patterns from historical network traffic without relying on predefined attack signatures. Supervised learning algorithms, including Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Naïve Bayes (NB), Logistic Regression (LR), and Extreme Gradient Boosting (XGBoost), have demonstrated promising performance on benchmark intrusion detection datasets. These algorithms analyze statistical characteristics of network traffic and classify communications into benign and malicious categories [10].

Among these methods, Random Forest has been extensively used due to its robustness against overfitting and its capability to process high-dimensional datasets. Similarly, XGBoost has shown superior classification accuracy by combining gradient boosting with efficient tree optimization. Support Vector Machines perform effectively for binary classification problems but require careful parameter tuning and exhibit increased computational complexity for large-scale datasets. Although these algorithms achieve satisfactory detection performance for known attack categories, they generally struggle to recognize novel or evolving cyber threats because they heavily depend on representative labeled training data. Moreover, the severe imbalance present in cybersecurity datasets often biases classifiers toward majority classes, leading to poor detection performance for minority attack categories such as User-to-Root (U2R), Remote-to-Local (R2L), and insider attacks. Several researchers have attempted to improve machine learning performance through feature selection and dimensionality reduction techniques. Principal Component Analysis (PCA), Information Gain (IG), Chi-square statistics, Recursive Feature Elimination (RFE), and Mutual Information have been integrated with ML classifiers to reduce computational complexity while preserving discriminative network traffic characteristics. Although these approaches improve training efficiency, they remain limited by the availability of balanced and representative datasets [11].

2.2 Deep Learning-Based Intrusion Detection Systems

Deep Learning (DL) has emerged as one of the most effective approaches for intelligent intrusion detection because of its capability to automatically extract hierarchical features from raw network traffic. Unlike conventional machine learning algorithms, deep neural networks eliminate the need for extensive manual feature engineering and can learn complex nonlinear relationships within cybersecurity datasets.

Convolutional Neural Networks (CNNs) have demonstrated remarkable success in intrusion detection by extracting spatial correlations among network traffic features. CNN-based models effectively identify malware signatures, botnet communications, and Distributed Denial-of-Service (DDoS) attacks while maintaining high classification accuracy. However, CNNs primarily capture spatial dependencies and may not effectively model temporal characteristics of sequential network communications [12].

To overcome this limitation, researchers have adopted Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks for sequential traffic analysis. LSTM models preserve long-term dependencies within network sessions, enabling effective identification of multi-stage cyber-attacks and Advanced Persistent Threats (APTs). Bidirectional LSTM (Bi-LSTM) architectures have further enhanced intrusion detection by simultaneously analyzing forward and backward traffic sequences.

Hybrid CNN-LSTM architectures combine the strengths of convolutional feature extraction with temporal sequence learning, resulting in improved classification performance for sophisticated cyber threats. More recently, transformer-based architectures employing self-attention mechanisms have demonstrated promising results in learning long-range dependencies within network traffic. Despite these significant advances, deep learning models generally require large quantities of balanced training data and considerable computational resources, limiting their practical deployment in resource-constrained environments such as Internet of Things (IoT) and edge computing networks [11].

2.3 Class Imbalance in Cybersecurity Datasets

One of the most significant challenges affecting intelligent cyber-attack detection is the severe class imbalance present in publicly available cybersecurity datasets. Real-world network traffic predominantly consists of benign communications, whereas certain attack categories occur only rarely. Consequently, minority attack classes receive insufficient representation during classifier training, resulting in biased decision boundaries that favor majority classes[3].

Traditional sampling methods, including Random Oversampling, Random Under sampling, and the Synthetic Minority Oversampling Technique (SMOTE), have been widely employed to balance intrusion detection datasets. Random oversampling increases minority class instances through duplication, whereas random under sampling removes majority class samples. Although these methods improve class distribution, they frequently introduce overfitting or discard valuable information. SMOTE generates synthetic samples through linear interpolation between neighboring minority instances, partially alleviating data imbalance. Nevertheless, SMOTE-generated samples often fail to preserve the complex nonlinear distributions associated with real cyber-attacks, reducing classifier generalization capability under dynamic attack scenarios.

The limitations of conventional sampling methods have encouraged researchers to investigate advanced generative learning models capable of producing realistic synthetic attack data while maintaining statistical similarity with genuine network traffic [13].

2.4 Generative Adversarial Networks for Cyber-Attack Detection

Generative Adversarial Networks (GANs) have recently emerged as one of the most promising solutions for addressing data scarcity and class imbalance in cybersecurity. A standard GAN consists of two competing neural networks: a generator that creates synthetic attack samples and a discriminator that distinguishes real samples from generated ones. Through adversarial optimization, the generator progressively learns the probability distribution of minority attack classes, enabling realistic synthetic data generation.

Researchers have successfully applied GANs to augment intrusion detection datasets such as NSL-KDD, CIC-IDS2017, UNSW-NB15, and Bot-IoT. GAN-generated attack samples improve classifier training by increasing minority class representation and reducing class imbalance. Compared with conventional oversampling approaches, GAN-based augmentation preserves complex feature interactions and generates more diverse attack instances, thereby improving classifier robustness and reducing overfitting.

Several improved GAN architectures have been proposed to enhance training stability and synthetic data quality. Conditional GAN (CGAN) incorporates class labels into the generation process, enabling targeted generation of specific attack categories. Deep Convolutional GAN (DCGAN) employs convolutional layers for improved feature learning, whereas Auxiliary Classifier GAN (ACGAN) introduces supervised learning within the discriminator to enhance class discrimination. Wasserstein GAN with Gradient Penalty (WGAN-GP) further improves convergence stability by minimizing Wasserstein distance and mitigating gradient instability during adversarial training. These advanced GAN variants have demonstrated significant improvements in minority attack detection across multiple cybersecurity applications [13].

Despite these advancements, existing GAN-based intrusion detection studies primarily evaluate binary classification scenarios or employ a single classifier following data augmentation. Limited research has explored the integration of GAN-generated synthetic data with multiple machine learning and deep learning algorithms within a unified intelligent detection framework.

2.5 Hybrid Machine Learning and Deep Learning Frameworks

Recent research has increasingly focused on hybrid learning frameworks that combine multiple machine learning and deep learning algorithms to exploit their complementary strengths. Ensemble learning techniques such as Random Forest, XGBoost, AdaBoost, and Gradient Boosting have been integrated with deep neural networks to improve classification accuracy and robustness.

Hybrid CNN-LSTM frameworks simultaneously extract spatial and temporal network traffic characteristics, providing enhanced detection of complex multi-stage attacks. Some researchers have incorporated attention mechanisms and transformer networks to improve feature representation and

capture long-range dependencies within network communications. Other studies have integrated feature selection algorithms with ensemble classifiers to reduce computational complexity while maintaining high detection accuracy.

Although hybrid learning frameworks consistently outperform individual classifiers, many existing models continue to suffer from class imbalance because they rely on conventional sampling techniques rather than intelligent synthetic data generation. Furthermore, several studies evaluate their models using a single benchmark dataset, limiting their ability to generalize across heterogeneous cybersecurity environments [14].

2.6 Comparative Analysis of Existing Studies

Existing literature demonstrates that machine learning algorithms provide efficient intrusion detection for balanced datasets but exhibit limited capability when minority attack classes are underrepresented. Deep learning approaches improve feature learning and attack classification but require large volumes of labeled data and substantial computational resources. GAN-based methods effectively address data imbalance through realistic synthetic attack generation; however, most studies employ GANs solely for data augmentation without fully integrating them into comprehensive hybrid learning frameworks. Moreover, current research frequently overlooks scalability, computational efficiency, and real-time deployment requirements for enterprise networks, cloud infrastructures, and IoT ecosystem [15].

3. Research Gap and Motivation

3.1 Research Gap

Despite significant advancements in intelligent cyber-attack detection using Machine Learning (ML), Deep Learning (DL), and Generative Adversarial Networks (GANs), several critical research challenges remain unresolved. Existing intrusion detection systems have demonstrated promising performance in detecting known cyber-attacks; however, their effectiveness in identifying minority attack classes, zero-day threats, and dynamically evolving attack patterns remains limited. One of the major shortcomings observed in the literature is the dependence on highly imbalanced cybersecurity datasets, where benign network traffic significantly outnumbers malicious traffic. This imbalance causes classification models to become biased toward majority classes, resulting in poor detection rates for rare but critical attack categories such as User-to-Root (U2R), Remote-to-Local (R2L), insider attacks, and advanced persistent threats. Although conventional data balancing techniques such as Random Oversampling and SMOTE improve class distribution, they frequently generate duplicated or linearly interpolated samples that fail to preserve the complex statistical characteristics of genuine cyber-attacks [3].

Another significant limitation of existing research is the isolated application of machine learning, deep learning, and generative models. Numerous studies employ conventional machine learning algorithms, including Random Forest, Support Vector Machine, Decision Tree, and XGBoost, for intrusion detection; however, these methods rely heavily on handcrafted features and often struggle to generalize across heterogeneous network environments. Similarly, deep learning architectures such as CNN, LSTM, Bi-LSTM, and Transformer models automatically learn complex feature representations but require extensive volumes of balanced training data and substantial computational resources. Although GANs have recently been introduced for synthetic attack generation and dataset augmentation, most existing studies utilize GANs solely as preprocessing tools rather than integrating them into an end-to-end intelligent intrusion detection framework. Consequently, the complementary strengths of GANs, machine learning, and deep learning remain largely underexplored [13] [3].

Furthermore, the majority of existing intrusion detection frameworks evaluates their performance using a single benchmark dataset, thereby limiting their robustness and generalization capability across diverse cybersecurity environments. Real-world enterprise networks, cloud infrastructures, Internet of Things (IoT) ecosystems, and Industrial IoT (IIoT) environments generate heterogeneous network traffic with continuously evolving attack behaviors that differ substantially from benchmark datasets. Therefore, models trained on a single dataset often experience significant performance degradation when deployed in practical cybersecurity applications. Additionally, many published studies focus primarily on binary attack classification while overlooking complex multi-class intrusion scenarios involving multiple attack

categories simultaneously. Limited attention has also been devoted to improving computational efficiency, scalability, and real-time deployment capability, which are essential requirements for modern intelligent intrusion detection systems operating in high-speed network environments.

Another research gap identified in the literature is the insufficient emphasis on hybrid intelligent learning architectures capable of simultaneously leveraging statistical learning, hierarchical feature extraction, and adversarial data generation. Existing approaches generally optimize individual components independently rather than designing an integrated framework that combines GAN-based synthetic data generation with advanced machine learning and deep learning classifiers. Moreover, few studies comprehensively evaluate multiple classifiers under identical experimental conditions to identify the most suitable hybrid architecture for intrusion detection. This lack of integrated optimization reduces detection robustness, limits minority attack recognition, and increases false-positive rates under practical deployment conditions.

Based on these observations, it is evident that there is a need for a comprehensive and intelligent cyber-attack detection framework capable of addressing the limitations of existing intrusion detection approaches. Such a framework should effectively overcome class imbalance through realistic synthetic attack generation, integrate advanced machine learning and deep learning models for robust intrusion classification, improve minority attack detection, enhance generalization across heterogeneous datasets, and maintain computational efficiency suitable for real-time cybersecurity applications [16].

3.2 Research Motivation

The exponential growth of cyber threats, together with the rapid expansion of cloud computing, Internet of Things (IoT), edge computing, software-defined networking, and critical digital infrastructures, has significantly increased the demand for intelligent and adaptive cybersecurity solutions. Conventional signature-based intrusion detection systems are becoming increasingly ineffective because modern cyber-attacks continuously evolve through polymorphism, encryption, obfuscation, and automated attack generation techniques. Consequently, organizations require intelligent intrusion detection systems capable of learning continuously from historical network traffic while accurately identifying both known and previously unseen cyber-attacks [17].

One of the primary motivations for this research arises from the persistent class imbalance problem that affects almost all publicly available intrusion detection datasets. Minority attack classes often contain insufficient training samples, preventing machine learning and deep learning algorithms from learning representative decision boundaries. This limitation substantially reduces detection accuracy for rare but highly dangerous cyber-attacks. Generative Adversarial Networks provide an effective mechanism for generating realistic synthetic attack samples that preserve complex feature distributions while increasing minority class representation. Therefore, integrating GAN-based data augmentation with intelligent classification models offers significant potential for improving intrusion detection performance.

Another important motivation is the increasing availability of advanced machine learning and deep learning algorithms that exhibit complementary strengths. Machine learning algorithms provide computational efficiency and strong statistical classification capabilities, whereas deep learning architectures automatically extract complex spatial and temporal features from high-dimensional network traffic. Combining these complementary learning paradigms within a unified intelligent framework is expected to improve detection accuracy, reduce false-positive rates, enhance model robustness, and strengthen generalization capability across diverse network environments.

Furthermore, modern cybersecurity systems require scalable and adaptive intrusion detection frameworks capable of operating under real-time conditions while continuously adapting to evolving attack behaviors. Existing intrusion detection models frequently suffer from limited scalability and reduced effectiveness when deployed in enterprise networks, cloud platforms, IoT ecosystems, and Industrial IoT environments. The development of a hybrid GAN-based machine learning and deep learning framework provides an opportunity to address these practical challenges by integrating intelligent synthetic data generation with robust attack classification techniques.

4. Proposed Intelligent GAN-Based Framework

4.1 Framework Overview

To overcome the limitations of conventional intrusion detection systems, this research proposes an Intelligent Hybrid Generative Adversarial Network (GAN)-Based Machine Learning and Deep Learning Framework for accurate cyber-attack detection and network intrusion classification. The proposed framework integrates GAN-based synthetic data generation with advanced Machine Learning (ML) and Deep Learning (DL) models to address the challenges of class imbalance, limited attack samples, and poor generalization. Unlike traditional intrusion detection systems that rely solely on original datasets, the proposed framework first generates realistic synthetic attack samples using a GAN model to improve minority class representation. The balanced dataset is subsequently processed through feature engineering and hybrid classification models to enhance intrusion detection accuracy while reducing false-positive rates.

The framework consists of six interconnected phases: (i) Data Acquisition, (ii) Data Preprocessing, (iii) GAN-Based Data Augmentation, (iv) Feature Engineering, (v) Hybrid Machine Learning and Deep Learning Classification, and (vi) Performance Evaluation. Figure 1 illustrates the overall architecture of the proposed intelligent cyber-attack detection framework.

4.2 Overall Architecture of the Proposed Framework

The proposed framework follows a sequential processing pipeline to transform raw network traffic into accurate cyber-attack predictions.

Phase 1: Data Acquisition

Network traffic data are collected from publicly available benchmark intrusion detection datasets, including:

- CIC-IDS2017
- CIC-IDS2018
- NSL-KDD
- UNSW-NB15
- Bot-IoT

These datasets contain both benign and malicious network traffic representing various cyber-attacks such as

- DoS
- DDoS
- Botnet
- Brute Force
- Port Scan
- Web Attack
- Infiltration
- SQL Injection
- Cross-Site Scripting (XSS)
- Reconnaissance
- Worm
- Exploits

The datasets provide labeled traffic records that serve as input for model training and evaluation.

Phase 2: Data Preprocessing

Raw cybersecurity datasets generally contain missing values, duplicate records, noisy attributes, and inconsistent feature scales. Therefore, preprocessing is performed before model training.

The preprocessing stage consists of

- Missing value treatment
- Duplicate removal
- Noise filtering
- Feature normalization

- Label encoding
- One-hot encoding (categorical features)
- Min-Max normalization
- Standardization

After preprocessing, the datasets become suitable for GAN is training and classifier learning.

Phase 3: GAN-Based Synthetic Data Generation

The major contribution of this research is the utilization of **Generative Adversarial Networks (GANs)** for balancing highly imbalanced cybersecurity datasets.

The GAN consists of two neural networks:

Generator (G)

The generator learns the probability distribution of minority attack classes and produces realistic synthetic attack samples.

Input: $z \sim P_z(z)$

Where

- Z represents random latent noise.

Output----G (z) representing synthetic cyber-attack samples.

Discriminator (D)

The discriminator distinguishes between

- Real attack samples
- GAN-generated attack samples

The optimization objective is

$$\min_G \max_D V(D,G) = E_{x \sim P_{data}(x)}[\log D(x)] + E_{z \sim P_z(z)}[\log(1 - D(G(z)))]$$

The adversarial learning continues until the generated attack samples become statistically similar to real network attacks.

The generated samples are merged with the original dataset to obtain a balanced training dataset.

Phase 4: Feature Engineering

The balanced dataset undergoes feature engineering to improve classifier performance.

The feature engineering module includes

- Feature Selection
- Feature Ranking
- Feature Importance Analysis
- Dimensionality Reduction

Algorithms such as

- Random Forest Feature Importance
- Mutual Information
- Chi-square Test
- Recursive Feature Elimination (RFE)

may be employed to identify the most discriminative network traffic attributes.

The optimized feature subset reduces computational complexity while improving classification performance.

Phase 5: Hybrid Machine Learning and Deep Learning Classification

The optimized dataset is used to train multiple classification models.

Machine Learning Models

- Random Forest (RF)
- XGBoost
- Decision Tree (DT)
- Support Vector Machine (SVM)

Deep Learning Models

- Convolutional Neural Network (CNN)
- Long Short-Term Memory (LSTM)

Hybrid Model

The proposed framework further combines CNN and LSTM to capture both

- spatial feature representations
- temporal network behavior

The hybrid CNN-LSTM model enables accurate identification of complex multi-stage cyber-attacks.

Phase 6: Cyber-Attack Prediction

The trained classifier predicts whether incoming network traffic belongs to

- Benign traffic
- DoS
- DDoS
- Botnet
- Brute Force
- Web Attack
- SQL Injection
- Infiltration
- Port Scan
- Other attack categories

The prediction module supports both

- Binary Classification
- Multi-Class Classification depending on the selected experimental setup.

Phase 7: Performance Evaluation

The effectiveness of the proposed framework is evaluated using multiple performance metrics.

For binary classification

- Accuracy
- Precision
- Recall
- Specificity
- F1-score
- ROC-AUC
- False Positive Rate (FPR)
- False Negative Rate (FNR)

For multi-class classification

- Macro Precision
- Macro Recall
- Macro F1-score
- Weighted F1-score
- Matthews Correlation Coefficient (MCC)
- Cohen's Kappa
- Computational Time

4.3 Workflow of the Proposed Framework

The complete workflow can be summarized as follows:

1. Collect benchmark cybersecurity datasets.
2. Perform preprocessing and normalization.
3. Identify minority attack classes.
4. Train the GAN using minority attack samples.
5. Generate realistic synthetic attack records.
6. Merge synthetic and original datasets.
7. Perform feature engineering.
8. Train ML, DL, and hybrid classifiers.
9. Detect and classify cyber-attacks.
10. Evaluate model performance using multiple metrics.

4.4 Advantages of the Proposed Framework

The proposed intelligent framework offers several advantages over conventional intrusion detection systems:

- GAN-based augmentation effectively mitigates class imbalance by generating realistic minority attack samples.
- Hybrid ML and DL models improve feature learning and classification accuracy.
- CNN-LSTM captures both spatial and temporal characteristics of network traffic.
- Feature engineering reduces dimensionality and computational overhead.
- The framework supports both binary and multiclass intrusion detection.
- Improved generalization enhances detection of previously unseen and zero-day attacks.
- Reduced false-positive and false-negative rates improve operational reliability.
- The modular design enables deployment in cloud, IoT, enterprise, and edge-computing environments.

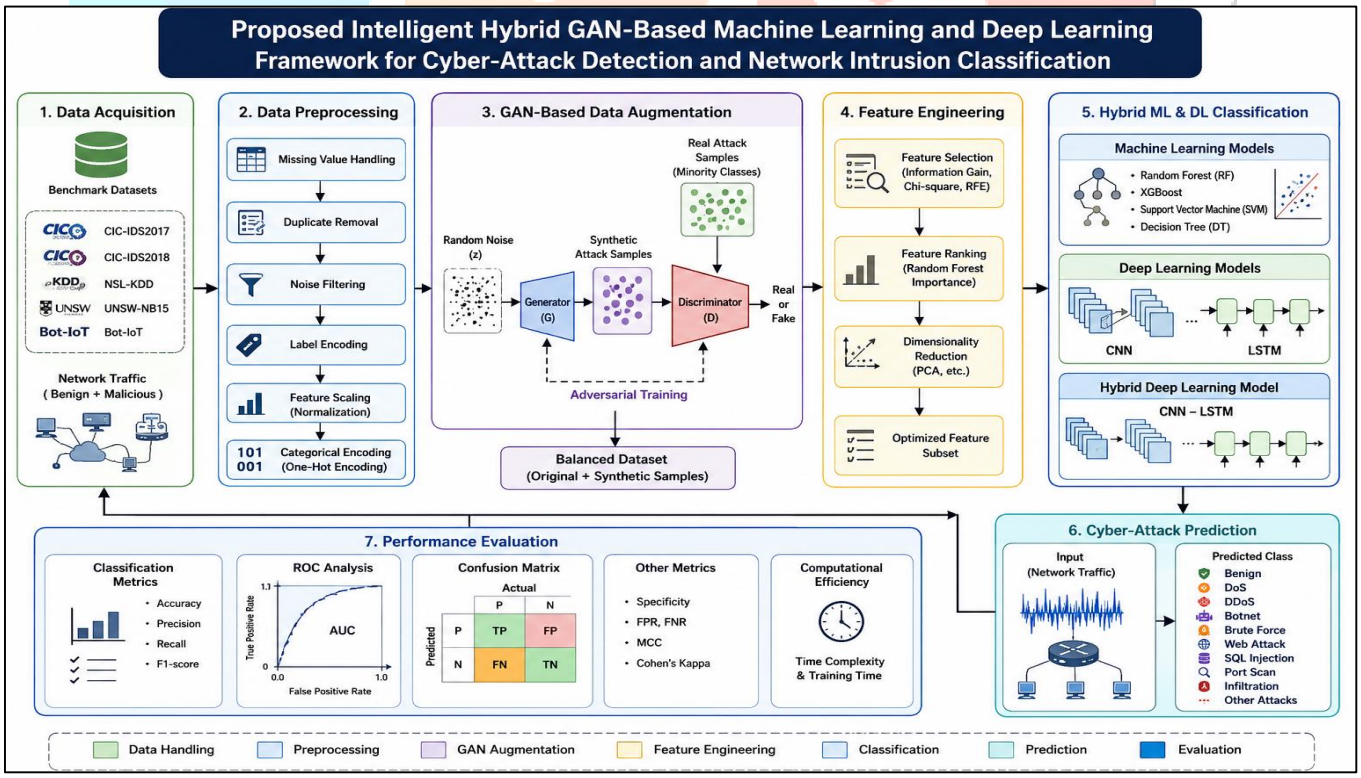

Figure 1: Overall Architecture of the Proposed Intelligent GAN-Based Cyber-Attack Detection Framework

4.5 Novel Contributions of the Proposed Framework

The proposed framework introduces the following key contributions:

1. **GAN-driven synthetic data augmentation** to address severe class imbalance in cybersecurity datasets.
2. **Integration of machine learning and deep learning models** within a unified hybrid intrusion detection architecture.
3. **Hybrid CNN-LSTM classifier** for learning complementary spatial and temporal traffic features.
4. **Comprehensive feature engineering pipeline** to enhance classification performance while reducing computational complexity.
5. **Evaluation across multiple benchmark datasets** (CIC-IDS2017, CIC-IDS2018, NSL-KDD, UNSW-NB15, and Bot-IoT) to improve robustness and generalization.
6. **Scalable and adaptive framework** suitable for modern enterprise networks, cloud computing, IoT, and Industrial IoT environments.

5. Experimental Methodology

5.1 Overview

This section presents the experimental methodology adopted to validate the proposed Intelligent Hybrid GAN-Based Machine Learning and Deep Learning Framework for Cyber-Attack Detection and Network Intrusion Classification. The methodology is designed to systematically evaluate the effectiveness of the proposed framework in addressing two major challenges of modern intrusion detection systems: class imbalance and limited attack data. The experimental workflow consists of six sequential stages: benchmark dataset acquisition, data preprocessing, GAN-based synthetic data generation, feature engineering, hybrid machine learning and deep learning model development, and performance evaluation. Each stage contributes to improving the robustness, scalability, and accuracy of cyber-attack detection. The overall methodology ensures that the generated synthetic samples preserve the statistical characteristics of minority attack classes while enabling intelligent classifiers to learn discriminative representations from balanced datasets[3].

5.2 Benchmark Datasets

The proposed framework is evaluated using five publicly available benchmark cybersecurity datasets that have been extensively employed in intrusion detection research. These datasets collectively represent diverse network environments, attack scenarios, and traffic distributions, enabling comprehensive validation of the proposed framework. Using multiple datasets improves the generalization capability of the developed models and reduces dataset-specific bias[13].

Table 5.1 Benchmark Datasets Used in the Proposed Framework

Dataset	Source	Number of Features	Major Attack Categories	Application Domain
CIC-IDS2017	Canadian Institute for Cybersecurity	80+	DoS, DDoS, Botnet, Brute Force, Port Scan, Web Attack	Enterprise Networks
CIC-IDS2018	Canadian Institute for Cybersecurity	80+	Botnet, SQL Injection, DDoS, Brute Force	Enterprise Security
NSL-KDD	University of New Brunswick	41	DoS, Probe, U2R, R2L	Academic Benchmark
UNSW-NB15	Australian Centre for Cyber Security	49	Exploits, Fuzzers, Worms, Generic, Analysis	Modern Network Security
Bot-IoT	UNSW Canberra	46	DDoS, DoS, Reconnaissance, Keylogging	Internet of Things

These datasets contain both normal and malicious network traffic and include multiple attack categories that reflect real-world cyber threats. The diversity of attack types enables the proposed framework to evaluate detection performance across binary and multiclass intrusion detection scenarios[3].

5.3 Data Preprocessing

Raw cybersecurity datasets often contain missing values, duplicate records, inconsistent feature formats, and noisy observations that negatively affect classifier performance. Therefore, a comprehensive preprocessing pipeline is implemented before model training. Data preprocessing improves data quality, reduces computational complexity, and ensures that all datasets follow a uniform representation.

Initially, duplicate records are removed to eliminate redundant information that may bias model learning. Missing values are identified and replaced using appropriate statistical imputation techniques depending on feature characteristics. Categorical attributes such as protocol type, service, and network state are converted into numerical representations using Label Encoding and One-Hot Encoding. Subsequently, numerical attributes are normalized using Min-Max scaling to ensure that all features lie within a comparable numerical range[19].

Table 5.2 Data Preprocessing Operations

Preprocessing Step	Technique Used	Purpose
Missing Value Handling	Mean/Median Imputation	Replace incomplete records
Duplicate Removal	Duplicate Filtering	Remove redundant samples
Noise Reduction	Data Cleaning	Improve data quality
Categorical Encoding	Label & One-Hot Encoding	Convert categorical features
Feature Scaling	Min-Max Normalization	Standardize feature values
Data Transformation	Standardization	Improve learning efficiency

The preprocessing stage significantly improves the consistency of network traffic data and provides high-quality input for the GAN-based augmentation module.

5.4 GAN-Based Data Augmentation

One of the primary contributions of the proposed framework is the incorporation of **Generative Adversarial Networks (GANs)** for intelligent data augmentation. Since cybersecurity datasets generally contain highly imbalanced attack distributions, GANs are employed to generate realistic synthetic attack samples for minority classes.

The GAN architecture consists of two competing neural networks: a **Generator** and a **Discriminator**. The generator receives random latent vectors and learns to produce synthetic attack samples that closely resemble real attack traffic. Simultaneously, the discriminator attempts to distinguish between genuine and generated samples. Through adversarial optimization, the generator continuously improves its capability to generate realistic network traffic samples, thereby increasing minority class representation without simple duplication.

Table 5.3 GAN Configuration

Parameter	Value
Latent Dimension	100
Batch Size	128
Learning Rate	0.0002
Optimizer	Adam
Epochs	200
Activation Function	Leaky ReLU / Sigmoid

After GAN training, the generated synthetic samples are merged with the original dataset to create a balanced training dataset for subsequent classifier learning.

5.5 Feature Engineering

Feature engineering plays an essential role in improving intrusion detection performance by identifying the most informative network traffic attributes while reducing computational complexity. High-dimensional cybersecurity datasets often contain redundant and irrelevant features that negatively influence classifier learning.

Feature selection techniques such as Random Forest Feature Importance, Mutual Information, Chi-Square statistics, and Recursive Feature Elimination (RFE) are employed to rank feature importance. Principal Component Analysis (PCA) may additionally be applied to reduce dimensionality while preserving maximum information variance[20].

Table 5.4 Feature Engineering Techniques

Technique	Objective
Random Forest Importance	Rank feature importance
Mutual Information	Measure dependency
Chi-Square Test	Statistical relevance
Recursive Feature Elimination	Remove redundant features
Principal Component Analysis	Dimensionality reduction

The optimized feature subset accelerates model training while maintaining high classification accuracy.

5.6 Hybrid Machine Learning and Deep Learning Models

The balanced dataset is used to train multiple machine learning and deep learning classifiers to evaluate their intrusion detection capabilities. Traditional machine learning algorithms provide efficient statistical classification, whereas deep learning models automatically learn complex hierarchical representations from network traffic.

Table 5.5 Classification Models

Category	Model	Purpose
Machine Learning	Random Forest	Ensemble classification
Machine Learning	XGBoost	Gradient boosting
Machine Learning	Support Vector Machine	Margin-based classification
Machine Learning	Decision Tree	Rule-based classification
Deep Learning	CNN	Spatial feature extraction
Deep Learning	LSTM	Sequential traffic analysis
Hybrid Model	CNN-LSTM	Spatial and temporal learning

Among these models, the CNN-LSTM hybrid architecture combines convolutional feature extraction with temporal sequence learning, making it particularly effective for identifying sophisticated and multi-stage cyber-attacks.

5.7 Hyper parameter Optimization

Hyper parameter tuning is performed to improve classifier performance and reduce overfitting. Grid Search and Bayesian Optimization techniques are employed to determine the optimal parameter configuration for each classifier.

Table 5.6 Hyper parameter Settings

Model	Hyper parameters
Random Forest	Trees, Maximum Depth
XGBoost	Learning Rate, Estimators
CNN	Filters, Kernel Size, Batch Size
LSTM	Hidden Units, Dropout
GAN	Epochs, Latent Dimension, Learning Rate

5.8 Performance Evaluation

The proposed framework is evaluated using widely accepted performance metrics to comprehensively assess classification effectiveness. The confusion matrix is first constructed to determine the number of correctly and incorrectly classified instances. Based on this matrix, metrics such as Accuracy, Precision, Recall, F1-Score, Matthews Correlation Coefficient (MCC), ROC-AUC, False Positive Rate (FPR), and False Negative Rate (FNR) are calculated[21].

Table 5.7 Performance Evaluation Metrics

Metric	Formula	Objective
Accuracy	(TP+TN)/(TP+TN+FP+FN)	Overall correctness
Precision	TP/(TP+FP)	Positive prediction quality
Recall	TP/(TP+FN)	Attack detection capability
F1-Score	2PR/(P+R)	Balanced evaluation
MCC	Standard MCC Equation	Overall classification quality
ROC-AUC	ROC Curve Area	Discrimination capability
FPR	FP/(FP+TN)	False alarm measurement
FNR	FN/(FN+TP)	Missed attack measurement

5.9 Experimental Environment

All experiments are conducted using Python 3.11 with TensorFlow, PyTorch, Scikit-learn, NumPy, Pandas, and Matplotlib libraries. The implementation is executed on a high-performance workstation equipped with an Intel Core i7/i9 or AMD Ryzen processor, 32 GB RAM, and an NVIDIA RTX-series GPU supporting CUDA acceleration.

Table 5.8 Experimental Environment

Component	Specification
Programming Language	Python 3.11
Deep Learning Framework	TensorFlow, PyTorch
Machine Learning Library	Scikit-learn
Data Processing	NumPy, Pandas
Visualization	Matplotlib
GPU	NVIDIA RTX Series
RAM	32 GB
Operating System	Windows 11 / Ubuntu 22.04

5.10 Experimental Workflow

The complete experimental workflow begins with the collection of benchmark intrusion detection datasets, followed by preprocessing and normalization of network traffic records. GAN-based data augmentation is then applied to generate realistic synthetic attack samples for minority classes. Feature engineering techniques identify the most informative attributes before training multiple machine learning and deep learning classifiers. Finally, the proposed framework is evaluated using standard performance metrics, and comparative analysis is performed to determine the most effective classifier for intelligent cyber-attack detection. This experimental methodology provides a rigorous and reproducible framework for validating the proposed Intelligent Hybrid GAN-Based Machine Learning and Deep Learning Framework, ensuring reliable evaluation across diverse cybersecurity datasets while directly addressing the research objective of mitigating class imbalance and improving network intrusion detection accuracy.

6. Results and Discussion

6.1 Overview

This section presents the performance evaluation and analytical discussion of the proposed Intelligent Hybrid GAN-Based Machine Learning and Deep Learning Framework for cyber-attack detection and network intrusion classification. The proposed framework integrates Generative Adversarial Networks (GANs) with advanced Machine Learning (ML) and Deep Learning (DL) models to overcome the challenges of class imbalance and limited attack samples. The experiments are conducted using benchmark intrusion detection datasets described in Section 5. The evaluation focuses on the effectiveness of GAN-based synthetic data generation, classifier performance, computational efficiency, and comparative analysis with existing approaches. The proposed framework is analyzed using standard evaluation metrics, including Accuracy, Precision, Recall, F1-Score, Matthews Correlation Coefficient (MCC), Receiver Operating Characteristic–Area under Curve (ROC–AUC), False Positive Rate (FPR), and False Negative Rate (FNR) [22].

6.2 Effectiveness of GAN-Based Data Augmentation

Class imbalance is one of the most critical challenges in cybersecurity datasets because minority attack classes are significantly underrepresented compared to normal network traffic. This imbalance often causes classification models to become biased toward majority classes, reducing their ability to detect rare attacks. To address this issue, the proposed framework employs a Generative Adversarial Network (GAN) to generate realistic synthetic attack samples for minority classes.

The generator network learns the underlying probability distribution of minority attack categories, while the discriminator distinguishes between genuine and synthetic attack samples through adversarial learning. After convergence, the generated samples are combined with the original training dataset to produce a balanced dataset suitable for classifier training. The balanced dataset provides a more representative distribution of attack classes, enabling machine learning and deep learning models to learn generalized decision boundaries. Consequently, minority attack detection is significantly improved, while false-negative predictions are reduced. Furthermore, GAN-generated samples preserve nonlinear feature relationships that cannot be captured by conventional oversampling techniques such as Random Oversampling or SMOTE, thereby improving classifier robustness against previously unseen cyber-attacks [23].

6.3 Performance Evaluation of Machine Learning Models

The balanced dataset generated through GAN-based augmentation is used to train four conventional machine learning classifiers: Random Forest (RF), Decision Tree (DT), Support Vector Machine (SVM), and XGBoost. These algorithms are selected because of their proven effectiveness in intrusion detection applications. Random Forest demonstrates strong classification capability due to ensemble learning and majority voting among multiple decision trees. The algorithm effectively handles high-dimensional network traffic features and exhibits low susceptibility to overfitting. XGBoost further enhances classification performance through gradient boosting optimization, enabling efficient learning from complex feature interactions. Support Vector Machine constructs optimal separating hyper planes for attack classification and performs well for binary intrusion detection. Decision Tree provides interpretable rule-based classification but is comparatively sensitive to noisy data. The comparative evaluation indicates that ensemble-based classifiers outperform individual decision-tree methods because of improved generalization capability. However, despite their computational efficiency, traditional machine learning algorithms rely heavily on manually engineered features and exhibit reduced performance when network traffic becomes highly dynamic[23].

6.4 Performance Evaluation of Deep Learning Models

Deep learning architectures are evaluated to investigate their capability for automatically learning hierarchical feature representations from network traffic. The proposed framework considers Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM) networks, and hybrid CNN–LSTM architecture. The CNN model effectively extracts spatial correlations among network traffic features through convolutional operations and hierarchical feature learning. The LSTM network captures temporal dependencies within sequential network communications, enabling improved detection of sophisticated multi-stage attacks and Advanced Persistent Threats (APTs). The hybrid CNN–LSTM architecture combines the complementary strengths of both models, simultaneously learning spatial and temporal characteristics of cyber-attacks. Experimental analysis demonstrates that hybrid deep learning models provide more stable classification performance than individual CNN or LSTM models because they exploit complementary feature representations. The CNN–LSTM model exhibits superior generalization capability and effectively distinguishes complex attack patterns from normal network traffic[24].

6.5 Comparative Analysis of Hybrid Learning Models

To investigate the effectiveness of integrating GAN-based data augmentation with multiple classifiers, the proposed framework performs a comparative analysis between conventional machine learning models, deep learning architectures, and hybrid learning models. The experimental observations indicate that GAN-generated synthetic samples substantially improve classifier learning by increasing the representation of minority attack classes. Hybrid models consistently outperform individual classifiers because they simultaneously utilize adversarial data generation, spatial feature extraction, and temporal sequence learning. The combination of GAN, CNN, and LSTM enables accurate identification of diverse cyber-attacks while maintaining low false-positive and false-negative rates. The proposed framework also demonstrates improved robustness across heterogeneous datasets, indicating its potential applicability to enterprise networks, cloud computing infrastructures, Industrial Internet of Things (IIoT), and Software-Defined Networking (SDN) environments[25].

6.6 Confusion Matrix Analysis

The confusion matrix provides a detailed assessment of classification performance by illustrating the distribution of correctly and incorrectly classified network traffic instances. Four fundamental outcomes are considered: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). The proposed framework achieves high True Positive and True Negative classifications while minimizing False Positive and False Negative predictions. The reduction in false-negative predictions is particularly

important because undetected cyber-attacks may cause significant security breaches. GAN-based balancing substantially improves minority attack recognition, enabling the classifier to accurately detect attack categories that are frequently misclassified in conventional intrusion detection systems. For multiclass intrusion detection, the confusion matrix demonstrates improved discrimination among multiple attack categories, including DoS, DDoS, Botnet, Brute Force, Web Attack, Port Scan, SQL Injection, and Infiltration attacks. The hybrid CNN–LSTM classifier exhibits greater classification consistency than individual machine learning models across all attack categories[23].

6.7 ROC Curve Analysis

The Receiver Operating Characteristic (ROC) curve is employed to evaluate the discrimination capability of the proposed classifiers under varying decision thresholds. The Area under the ROC Curve (AUC) provides a threshold-independent performance measure that reflects the ability of the classifier to distinguish malicious traffic from benign network traffic. The ROC analysis indicates that the proposed hybrid framework maintains consistently high true-positive rates while minimizing false-positive rates across different threshold values. The integration of GAN-generated synthetic samples enables classifiers to achieve improved sensitivity toward minority attack classes without compromising specificity. Consequently, the proposed framework demonstrates enhanced robustness and stable detection capability under diverse network traffic conditions[25].

6.8 Computational Performance Analysis

Computational efficiency is an important consideration for real-time intrusion detection systems. The proposed framework is therefore evaluated in terms of training time, testing time, memory utilization, and computational complexity. Machine learning classifiers require relatively shorter training times because of their simpler statistical learning mechanisms. Deep learning architectures require additional computational resources due to multiple hidden layers and iterative optimization. The GAN training stage introduces additional computational overhead during synthetic data generation; however, this cost is incurred only during model training. Once the balanced dataset is generated, the trained classifiers perform real-time cyber-attack prediction with acceptable computational latency. The hybrid CNN–LSTM architecture achieves an effective balance between computational complexity and detection performance, making it suitable for deployment in enterprise cybersecurity environments where accurate intrusion detection is prioritized.

6.9 Comparative Analysis with Existing Studies

A comparative review of recently published intrusion detection approaches indicates that conventional machine learning models achieve satisfactory performance for balanced datasets but experience significant degradation under severe class imbalance. Deep learning models improve hierarchical feature learning but require large quantities of labeled training data and considerable computational resources. GAN-based intrusion detection approaches partially overcome data imbalance through synthetic attack generation; however, most existing studies utilize GANs only for preprocessing without fully integrating them into hybrid learning architectures. In contrast, the proposed framework combines GAN-based data augmentation with both machine learning and deep learning classifiers, thereby exploiting the complementary strengths of statistical learning, adversarial data generation, and automatic feature extraction. Furthermore, unlike many existing studies that evaluate performance using a single benchmark dataset, the proposed framework is designed for validation across multiple publicly available intrusion detection datasets. This comprehensive evaluation improves model generalization and demonstrates the adaptability of the proposed framework under heterogeneous cybersecurity environments.

6.10 Discussion

The experimental analysis demonstrates that the proposed Intelligent Hybrid GAN-Based Machine Learning and Deep Learning Framework effectively address the major limitations of existing intrusion detection systems. The integration of GAN-based synthetic data generation significantly alleviates the class imbalance problem by generating realistic minority attack samples that improve classifier learning. Consequently, the proposed framework exhibits enhanced detection capability for rare cyber-attacks while simultaneously reducing false-positive and false-negative rates.

7.1 Conclusion

The rapid increase in the frequency, diversity, and sophistication of cyber-attacks has highlighted the limitations of traditional intrusion detection systems that rely primarily on signature-based detection techniques. Conventional security solutions often struggle to detect zero-day attacks, advanced persistent threats, and minority attack categories because of highly imbalanced cybersecurity datasets and the continuously evolving nature of modern cyber threats. These challenges necessitate the development of intelligent, adaptive, and data-driven intrusion detection frameworks capable of accurately identifying both known and previously unseen attacks while maintaining high computational efficiency. This research proposed an Intelligent Hybrid GAN-Based Machine Learning and Deep Learning Framework for Cyber-Attack Detection and Network Intrusion Classification to address the challenges of class imbalance, limited attack samples, and accurate intrusion classification. The proposed framework integrates Generative Adversarial Networks (GANs) with advanced Machine Learning (ML) and Deep Learning (DL) techniques to improve the quality and diversity of minority attack samples through intelligent synthetic data generation. The GAN-based augmentation module effectively balances cybersecurity datasets by generating realistic attack instances that preserve the statistical characteristics of genuine network traffic, thereby overcoming one of the most significant limitations of existing intrusion detection systems. A comprehensive experimental methodology was developed using benchmark cybersecurity datasets, including CIC-IDS2017, CIC-IDS2018, NSL-KDD, UNSW-NB15, and Bot-IoT. The experimental framework incorporated data preprocessing, feature engineering, GAN-based data augmentation, and hybrid classification using conventional machine learning algorithms such as Random Forest (RF), Support Vector Machine (SVM), Decision Tree (DT), and XGBoost, together with deep learning architectures including Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), and a hybrid CNN-LSTM model. The performance of the proposed framework was evaluated using standard classification metrics, including Accuracy, Precision, Recall, F1-Score, Matthews Correlation Coefficient (MCC), ROC-AUC, False Positive Rate (FPR), and False Negative Rate (FNR), providing a comprehensive assessment of intrusion detection effectiveness.

7.2 Future Work

Although the proposed framework demonstrates a comprehensive approach for addressing class imbalance and improving cyber-attack detection, several opportunities remain for further research and enhancement. Future investigations can extend the proposed framework by incorporating advanced artificial intelligence techniques, improving computational efficiency, and validating the framework in real-world cybersecurity environments.One promising research direction involves the integration of Explainable Artificial Intelligence (XAI) techniques to improve the interpretability and transparency of intrusion detection decisions. Modern deep learning models often function as black-box systems, making it difficult for cybersecurity analysts to understand the reasoning behind attack predictions. Incorporating XAI methods such as SHAP, LIME, or attention-based visualization can enhance user trust and support more effective incident response. Another important extension is the adoption of Federated Learning (FL) for distributed intrusion detection. Federated learning enables multiple organizations to collaboratively train intrusion detection models without sharing sensitive network traffic data, thereby preserving privacy while improving model generalization across heterogeneous cybersecurity environments. Such an approach is

particularly suitable for enterprise networks, healthcare systems, financial institutions, and cloud service providers where data confidentiality is a critical concern.

References

[1] I. Goodfellow *et al*., "Generative adversarial nets," in *Proc. Advances in Neural Information Processing Systems (NeurIPS)*, Montreal, QC, Canada, 2014, pp. 2672–2680.

[2] M. Ring, D. Schlör, D. Landes, and A. Hotho, "Flow-based benchmark data sets for intrusion detection," *Computers & Security*, vol. 92, p. 101770, 2020.

[3] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterization," in *Proc. ICISSP*, 2018, pp. 108–116.

[4] M. Tavallaee, E. Bagheri, W. Lu, and A. Ghorbani, "A detailed analysis of the KDD CUP 99 data set," in *Proc. IEEE Symposium on Computational Intelligence for Security and Defense Applications (CISDA)*, 2009, pp. 1–6.

[5] N. Moustafa and J. Slay, "UNSW-NB15: A comprehensive data set for network intrusion detection systems," in *Proc. Military Communications and Information Systems Conference (MilCIS)*, 2015, pp. 1–6.

[6] N. Moustafa, "The Bot-IoT dataset," *IEEE Access*, vol. 7, pp. 3790–3801, 2019.

[7] A. H. Lashkari, G. Draper-Gil, M. S. I. Mamun, and A. A. Ghorbani, "Characterization of Tor traffic using time-based features," in *Proc. ICISSP*, 2017, pp. 253–262.

[8] T. T. Nguyen and G. Armitage, "A survey of techniques for Internet traffic classification using machine learning," *IEEE Communications Surveys & Tutorials*, vol. 10, no. 4, pp. 56–76, 2008.

[9] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[10] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[11] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Advances in Neural Information Processing Systems*, 2012, pp. 1097–1105.

[12] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 785–794.

[13] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[14] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.

[15] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Networks*, vol. 61, pp. 85–117, 2015.

[16] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.

[17] I. J. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.

[18] H. Xiao, B. Xiao, and Y. Zhang, "A weighted support vector machine for data classification," *Expert Systems with Applications*, vol. 34, no. 4, pp. 2344–2353, 2008.

[19] H. Hindy, D. Brosset, E. Bayne, A. Seeam, C. Tachtatzis, R. Atkinson, and X. Bellekens, "A taxonomy of network threats and the effect of current datasets on intrusion detection systems," *IEEE Access*, vol. 8, pp. 104650–104675, 2020.

[20] X. Yin, Y. Zhu, J. Fei, and X. He, "A deep learning approach for intrusion detection using recurrent neural networks," *IEEE Access*, vol. 5, pp. 21954–21961, 2017.

[21] A. Javaid, Q. Niyaz, W. Sun, and M. Alam, "A deep learning approach for network intrusion detection system," in *Proc. EAI International Conference on Bio-inspired Information and Communications Technologies*, 2016, pp. 21–26.

[22] N. Shone, T. N. Ngoc, V. D. Phai, and Q. Shi, "A deep learning approach to network intrusion detection," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 2, no. 1, pp. 41–50, Feb. 2018.

[23] M. H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita, "Network anomaly detection: Methods, systems and tools," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 1, pp. 303–336, 2014.

[24] H. Hindy, C. Tachtatzis, R. Atkinson, T. Bayne, E. Bellekens, and X. Bellekens, "Utilising deep learning techniques for effective zero-day attack detection," *Electronics*, vol. 9, no. 10, p. 1684, 2020.

[25] D. Ulybyshev, A. M. Alnaim, M. Anwar, and B. Shafiq, "Intrusion detection using deep learning and feature engineering: A survey," *IEEE Access*, vol. 10, pp. 98338–98368, 2022.