Hybrid Machine Learning And Deep Learning-Based Vpn Network Traffic Anomaly Detection System

Dr. K. Sangeetha[1], Mrs.N.Poornima[2]
[1]prof, Dept of Computer Science and Engineering
[2]Dept of Computer Science and Engineering
[1, 2] SNS College of Technology, Coimbatore, India

Abstract- *The rapid growth of internet usage and the widespread adoption of Virtual Private Networks (VPNs) have significantly enhanced secure communication, but have also introduced new challenges in detecting cyber threats hidden within encrypted traffic. Traditional intrusion detection systems often fail to identify such threats due to their reliance on signature-based methods and inability to analyze encrypted payloads effectively. This project presents a hybrid deep learning-based VPN traffic detection system that integrates Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks to accurately identify anomalous network behavior. The system utilizes structured network traffic data with features such as protocol type, service, flag status, traffic rates, and byte counts, which are preprocessed using label encoding and feature scaling techniques. In addition, conventional machine learning models such as Random Forest and Support Vector Machine (SVM) are implemented for performance comparison. The hybrid CNN–LSTM model captures both spatial and temporal patterns in network traffic, resulting in improved detection accuracy. A real-time web-based dashboard developed using Streamlit enables users to input parameters and visualize prediction results, including classification outcomes and confidence scores. An automated alert mechanism is also incorporated to notify users of suspicious activities, facilitating timely response to potential threats. Experimental results demonstrate that the proposed system outperforms traditional methods in terms of accuracy, precision, and recall, providing a scalable and efficient solution for real-time VPN traffic monitoring and cyber-attack detection in enterprise environments.*

Keywords: Virtual Private Network (VPN), Network Traffic Analysis, Anomaly Detection, Intrusion Detection System (IDS), Deep Learning, Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), Hybrid Model, Cybersecurity, Encrypted Traffic Analysis, Machine Learning, Random Forest, Support Vector Machine (SVM), Real-Time Monitoring.

I. INTRODUCTION

The rapid growth of internet usage and the widespread adoption of cloud-based services have significantly increased the demand for secure communication, leading to the extensive use of Virtual Private Networks (VPNs). VPNs provide confidentiality and data protection by encrypting network traffic, making them essential for modern enterprise and personal communication. However, this encryption also poses a major challenge for cybersecurity, as it can conceal malicious activities such as data exfiltration, unauthorized access, and distributed denial-of-service (DDoS) attacks. Traditional intrusion detection systems (IDS), which rely on signature-based or payload inspection techniques, often fail to effectively analyze encrypted VPN traffic, thereby limiting their ability to detect anomalies in real time.Recent advancements in Artificial Intelligence (AI) and deep learning have opened new possibilities for network traffic analysis and cybersecurity. Deep learning models are capable of automatically extracting complex and hidden patterns from large-scale data, making them highly suitable for detecting anomalies in encrypted traffic. In particular, Convolutional Neural Networks (CNN) are effective in capturing spatial relationships and feature patterns within network data, while Long Short-Term Memory (LSTM) networks are well-suited for modeling temporal dependencies and sequential behavior in traffic flows. The integration of these two models enables a more comprehensive analysis of network traffic by capturing both spatial and temporal characteristics.This project, titled "Hybrid Deep Learning-Based VPN Traffic Detection System," proposes an intelligent framework that combines CNN and LSTM architectures to detect anomalous behavior in VPN traffic. The system is designed to process structured network traffic data, learn complex traffic patterns, and accurately classify traffic as normal or anomalous. In addition to model development, a real-time web-based dashboard is implemented to allow users to input network parameters, visualize prediction results, and receive automated alerts for suspicious activities.

II. EXISTING SYSTEM

Current network intrusion detection systems face several challenges in effectively monitoring and securing Virtual Private Network (VPN) traffic. Many existing systems rely on traditional machine learning techniques or signature-based detection methods, which are limited in identifying sophisticated and evolving cyber threats. These systems often depend on predefined rules or known attack patterns, making them ineffective against new or unknown attacks. Additionally, due to the encrypted nature of VPN traffic, conventional methods struggle to inspect packet contents, leading to reduced detection accuracy.The lack of advanced feature extraction and intelligent analysis techniques hinders the ability of these systems to capture complex spatial and temporal patterns present in network traffic. As a result, important behavioral characteristics of anomalous traffic may go undetected. Furthermore, most existing systems are not designed to handle large-scale, high-dimensional data efficiently, which affects their scalability and real-time performance.Moreover, traditional intrusion detection systems often suffer from high false positive rates, where normal traffic is incorrectly classified as malicious, reducing trust in the system. They also lack interactive visualization tools and user-friendly interfaces, making it difficult for network administrators to monitor traffic patterns and respond quickly to potential threats.In summary, existing systems lack the intelligence, adaptability, and efficiency required to effectively detect anomalies in encrypted VPN traffic. This project aims to address these limitations by developing a hybrid deep learning-based VPN traffic detection system that improves accuracy, enables real-time monitoring, and enhances overall network security.

III. PROPOSED SYSTEM

The **"Hybrid Deep Learning-Based VPN Traffic Detection System"** is designed to address the limitations of existing intrusion detection systems by integrating advanced deep learning techniques to accurately identify anomalies in encrypted network traffic. The system provides an intelligent and scalable solution for real-time monitoring, classification, and alert generation of VPN traffic.

Key features of the proposed system include:

- **Hybrid CNN–LSTM Model:** The system utilizes a hybrid architecture combining Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks. CNN is used to extract spatial features and identify important traffic patterns, while LSTM captures temporal dependencies and sequential behavior in network traffic, enabling accurate anomaly detection.
- **Data Preprocessing and Feature Engineering:** The system performs data preprocessing techniques such as data cleaning, label encoding, and feature scaling to ensure high-quality input data. Feature selection is also applied to identify the most relevant attributes, improving model performance and reducing computational complexity.
- **Multi-Model Comparison:** In addition to the hybrid model, traditional machine learning algorithms such as Random Forest and Support Vector Machine (SVM) are implemented for comparison, allowing evaluation of performance improvements achieved by deep learning techniques.
- **Real-Time Traffic Classification:** The system classifies incoming VPN traffic as either normal or anomalous in real time. This enables continuous monitoring of network activity and quick identification of suspicious behavior.
- **Interactive Web-Based Dashboard:** A user-friendly web interface is developed using Streamlit, allowing users to input network parameters and visualize prediction results. The dashboard displays classification outcomes, confidence scores, and graphical representations of traffic analysis.
- **Automated Alert Mechanism:** The system includes an alert module that generates notifications when anomalous traffic is detected. This helps network administrators take immediate action to prevent potential cyber threats.
- **Scalability and Performance Optimization:** The proposed system is designed to handle large-scale network data efficiently. The hybrid deep learning model ensures improved accuracy, reduced false positives, and faster processing compared to traditional approaches.

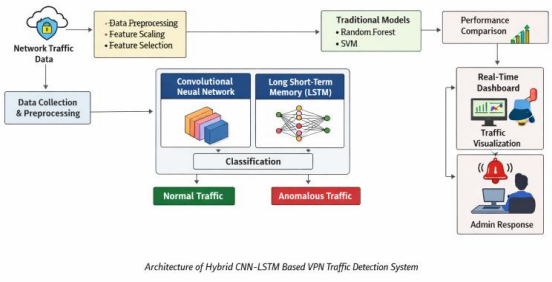


Figure 1: Architecture of Hybrid CNN-LSTM Based VPN Traffic Detection System

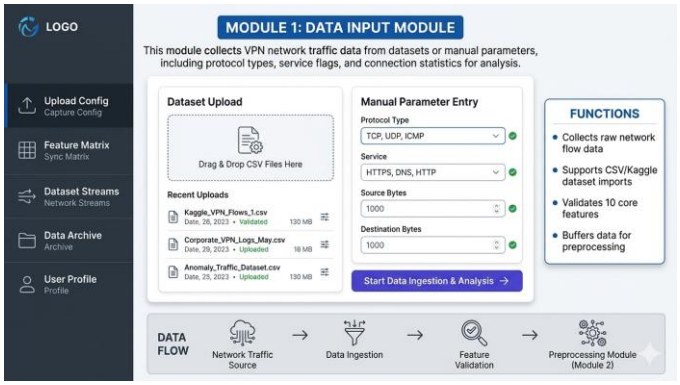
IV.METHODOLOGY

Development of the proposed Hybrid Machine Learning and Deep Learning–based VPN Network Traffic

Anomaly Detection System follows a structured design and implementation methodology. The process begins with analyzing system requirements to ensure that the data preprocessing modules, feature selection mechanisms, machine learning models, deep learning networks, and web application components operate efficiently under real-time conditions. Once the requirements are finalized, the design phase is carried out to define the overall system architecture, data flow, model integration, user interaction, and coordination between different modules. After completing the design stage, the system is developed as an intelligent network traffic analysis platform incorporating Python-based processing, data preprocessing techniques, machine learning algorithms, deep learning models, and Flask-based web deployment. All components of the system work together to achieve accurate classification of VPN network traffic as normal or anomalous. Functional testing is performed to verify correct data processing, feature extraction accuracy, model performance, prediction reliability, and system responsiveness.

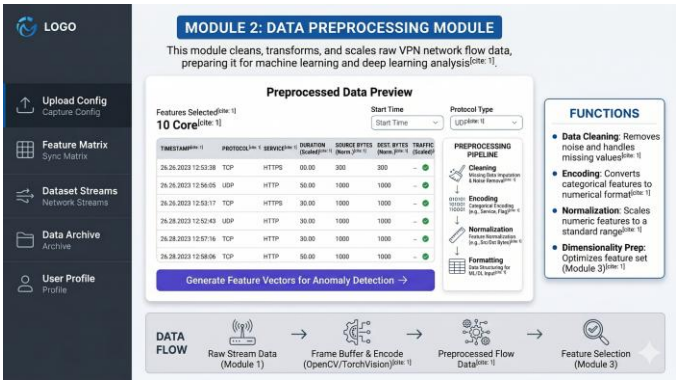
Data Input Module

The data input module is used to collect VPN network traffic data for analysis. The system accepts input from publicly available datasets or user-provided network parameters. The dataset consists of multiple features such as protocol type, service, flag status, source bytes, destination bytes, and connection statistics.



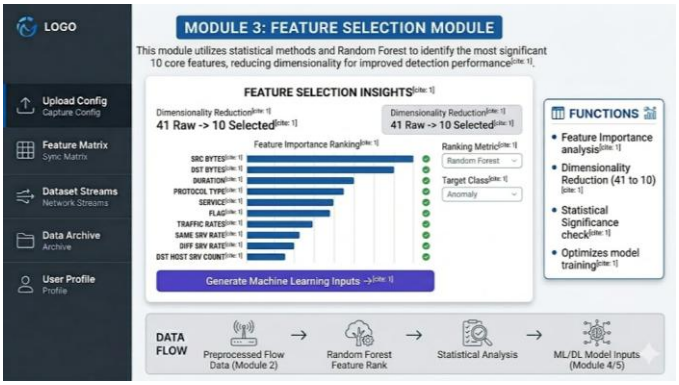
Data Preprocessing Module

The data input module is used to collect VPN network traffic data for analysis. The system accepts input from publicly available datasets or user-provided network parameters.



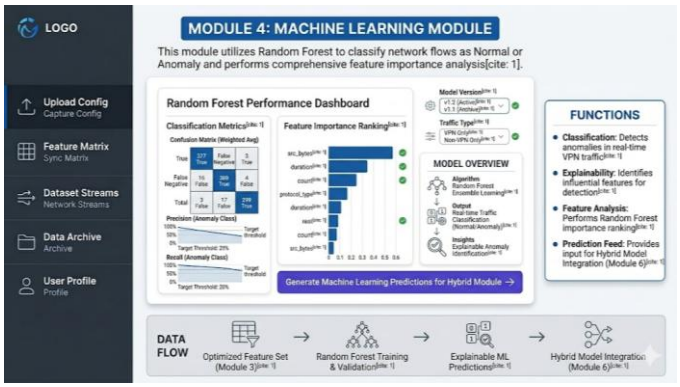
Feature Selection Module

The feature selection module identifies the most relevant attributes that significantly contribute to anomaly detection. Important features such as duration, protocol type, service, flag, source bytes, destination bytes, and traffic rates are selected based on their impact on classification performance. This module reduces dimensionality, improves model efficiency, and enhances prediction accuracy.



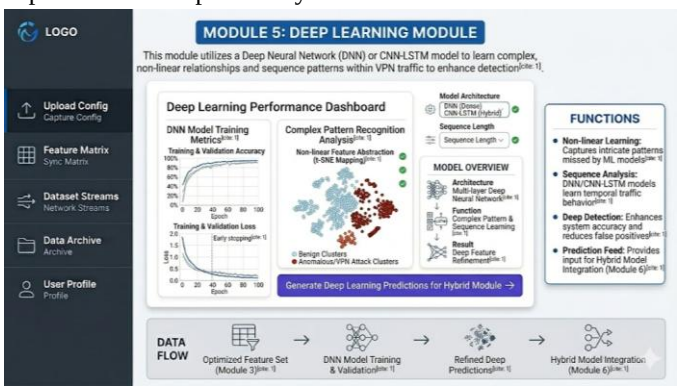
Machine Learning Module

The machine learning module implements algorithms such as Random Forest to perform initial classification and feature importance analysis. The model is trained on the processed dataset to identify patterns in normal and anomalous traffic. Random Forest provides robustness, interpretability, and efficient handling of large datasets, making it suitable for initial prediction and feature evaluation.



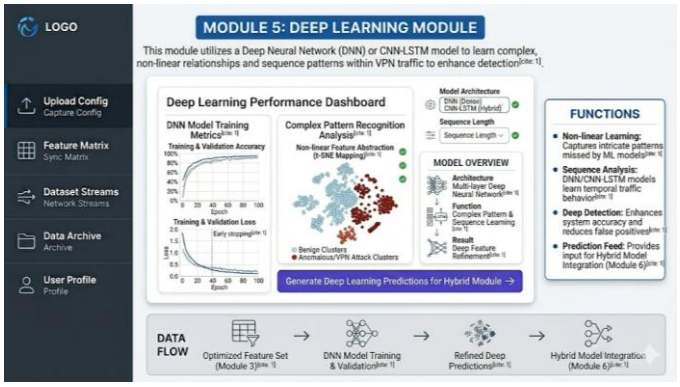
Deep Learning Module

The deep learning module utilizes a Deep Neural Network (DNN) to learn complex and non-linear relationships within the network traffic data. The model consists of multiple layers with activation functions that enable high-level feature extraction. This module enhances the system's ability to detect sophisticated and previously unseen anomalies in VPN traffic.



Hybrid Model Integration Module

The hybrid model integration module combines the outputs of the machine learning and deep learning models to produce a final prediction. Techniques such as ensemble learning or weighted averaging are used to integrate the results.



V. EXPERIMENTAL RESULTS

The experimental results of the proposed Hybrid Machine Learning and Deep Learning–based VPN Network Traffic Anomaly Detection System demonstrate that the developed model achieves improved classification accuracy, reduced false positive rates, and efficient real-time prediction performance. During testing, the data input module successfully handled both dataset-based inputs and user-provided network parameters without significant delay. The backend processing pipeline effectively managed data preprocessing, feature extraction, and model execution, ensuring smooth data flow across all modules for reliable system operation.The data preprocessing and feature selection modules were able to clean, transform, and optimize the dataset efficiently. Categorical features were encoded accurately, and numerical features were normalized to improve model performance. The feature selection process successfully identified the most relevant attributes contributing to anomaly detection, which reduced computational complexity and enhanced overall prediction accuracy.The machine learning module, based on the Random Forest algorithm, performed well in identifying important features and providing initial classification results. It demonstrated strong performance in handling structured network data and provided interpretable predictions. The deep learning module, implemented using a Deep Neural Network (DNN), effectively captured complex non-linear relationships within the data and improved the detection of hidden and unknown anomalies.The hybrid model integration module combined the outputs of the Random Forest and Deep Neural Network models to produce final predictions. This integration significantly improved system performance compared to standalone models. The hybrid approach reduced misclassification of normal traffic and enhanced detection capability for anomalous traffic, resulting in higher accuracy and better generalization across different network conditions.The model training and evaluation module validated the system using performance metrics such as accuracy, precision, recall, and F1-score. The results indicated that the hybrid model outperformed individual machine learning and deep learning models in terms of prediction reliability and consistency.The Flask web application module successfully enabled real-time interaction with the system. Users were able to input network traffic parameters through the web interface and receive instant classification results. The frontend module provided a responsive and user-friendly interface, ensuring ease of use for both technical and non-technical users.The result and visualization module clearly displayed the classification output as either normal or anomalous. In addition, the precaution and suggestion module provided useful recommendations to mitigate potential

network threats, such as monitoring suspicious activity and applying security controls.Overall, the system demonstrated improved detection accuracy, reduced false positives, efficient real-time performance, and enhanced usability compared to traditional intrusion detection methods. The integration of machine learning and deep learning techniques proved to be effective in analyzing encrypted VPN traffic and identifying anomalies with high reliability.

VI. CONCLUSION

The proposed Hybrid Machine Learning and Deep Learning–based VPN Network Traffic Anomaly Detection System provides a practical and efficient solution for improving network security and anomaly detection using advanced artificial intelligence techniques. The integration of Random Forest for feature selection and initial classification with Deep Neural Networks for complex pattern recognition enables the system to effectively analyze encrypted VPN traffic and identify anomalous behavior. This improves detection accuracy, enhances system reliability, and reduces dependency on traditional rule-based intrusion detection methods when compared with conventional approaches.The system ensures effective coordination between data input processing, preprocessing, feature selection, machine learning, deep learning, and web application modules through the use of Python-based implementation, statistical data analysis, and hybrid model integration techniques. In addition, the system enables structured network traffic analysis, which supports better understanding and monitoring of network behavior.The incorporation of a hybrid ML–DL framework further enhances the system by enabling accurate detection of both known and unknown anomalies while maintaining low false positive rates. The deployment of the model through a Flask-based web application provides real-time prediction capability and user-friendly interaction, making the system suitable for practical applications.Ultimately, the proposed system offers a cost-effective, intelligent, and scalable solution suitable for network administrators, cybersecurity professionals, and organizations, while contributing to improved detection of malicious activities and strengthening security in modern network environments.

VII. FUTURE WORKS

Future development of the **"Hybrid Deep Learning-Based VPN Traffic Detection System"** will focus on several key areas:

- **Enhancements to Deep Learning Models:** Future work will explore advanced deep learning architectures such as Attention-based models, Transformers, and Autoencoders to further improve detection accuracy and reduce false positives. Optimization techniques will also be applied to enhance training speed and model efficiency.
- **Real-Time Traffic Monitoring and Deployment:** The system will be extended to support real-time network traffic capture and live deployment in enterprise environments. Integration with network monitoring tools and intrusion detection systems will enable continuous and automated threat detection.
- **Multi-Class Attack Classification:** The current system can be enhanced to classify different types of cyber-attacks (e.g., DDoS, phishing, brute force) instead of binary classification (normal vs anomalous), providing more detailed insights into network threats.
- **Integration with Cloud and Edge Computing:** Future improvements will include deploying the system on cloud and edge platforms to handle large-scale network traffic efficiently and reduce latency in detection.
- **Incorporation of Advanced Feature Engineering:** Additional feature extraction techniques and dimensionality reduction methods will be explored to improve model performance and adaptability to evolving traffic patterns.
- **Improved Visualization and User Interface:** The web-based dashboard will be enhanced with advanced visualization tools, real-time analytics, and customizable alerts to provide better user experience and decision-making support.
- **Scalability and Big Data Handling:** The system will be optimized to process large volumes of high-speed network traffic using distributed computing frameworks such as Apache Spark or Hadoop.
- **Security and Privacy Enhancements:** Future work will focus on incorporating advanced security mechanisms, including secure data transmission, encryption techniques, and privacy-preserving machine learning to protect sensitive network information.
- **Adaptability to Emerging Threats:** Continuous learning mechanisms and model updates will be implemented to ensure the system can adapt to new and evolving cyber threats in dynamic network environments.

REFERENCES

[1] E. U. H. Qazi, M. H. Faheem, and T. Zia, "HDLNIDS: Hybrid Deep-Learning-Based Network Intrusion Detection System," *Applied Sciences*, vol. 13, no. 8, 2023.

[2] M. A. S. Sandila, S. Sultan, Z. U. Hassan, and A. Ali, "A Hybrid Deep Learning Approach for Intrusion Detection in Network Traffic Using Convolutional and Recurrent

Neural Networks," *Spectrum of Engineering Sciences*, vol. 3, no. 7, 2025.

[3] H. Rana, F. Zainab, F. Raoof, and A. Zahoor, "Prediction of Network Intrusion Using CNN-LSTM: Hybrid Deep Learning Approach," *KIET Journal of Computing and Information Sciences*, vol. 7, no. 2, 2025.

[4] I. Izhar, A. Abdullah, M. Z. Hussain, and M. Z. Hasan, "Enhancing IoT/IIoT Intrusion Detection Using Hybrid CNN-LSTM Models," *Spectrum of Engineering Sciences*, vol. 3, no. 10, 2025.

[5] C. Zhang, J. Li, N. Wang, and D. Zhang, "Intrusion Detection Method Based on Transformer and CNN-BiLSTM in IoT," *Sensors*, vol. 25, no. 9, 2025.

[6] A. Gueriani, H. Kheddar, and A. C. Mazari, "Adaptive Cyber-Attack Detection Using Attention-Based LSTM-CNN Models," *arXiv preprint*, 2025.

[7] M. A. Talukder et al., "A Dependable Hybrid Machine Learning Model for Network Intrusion Detection," *arXiv preprint*, 2022.

[8] R. Nazre et al., "Temporal Convolutional Network-Based Approach for Network Intrusion Detection," *arXiv preprint*, Dec. 2024.

[9] S. A. Ahmed et al., "Enhancing Cloud Data Center Security through Deep Learning Models," *Engineering, Technology & Applied Science Research*, vol. 15, no. 1, 2025.

[10] V. P. M. Priya and S. Soumya, "Advancements in Anomaly Detection Techniques in Network Traffic," *Journal of Scientific Research and Technology*, vol. 2, no. 6, 2024.

[11] M. Cai, "Network Traffic Anomaly Detection Method Based on Python," Academic Journal of Science and Technology, vol. 4, no. 1, 2024.

[12] K.Lu, "Network Anomaly Traffic Analysis," *Academic Journal of Science and Technology*, 2025.

[13] N. Vinisha and S. G. Krishna, "Detecting Network Traffic Anomalies with Machine Learning," *IJRASET*, 2025.

[14] B. Zhang et al., "Hybrid CNN-LSTM-GRU Model for Intrusion Detection Using NSL-KDD Dataset," *Utilitas Mathematica*, vol. 122, no. 2, 2025.

[15] J. Li, Q. Du, and F. Huang, "Intrusion Detection Technology Based on CNN-SaLSTM," in *Proc. WCNA 2021*, Springer, 2022.

[16] W. Zhang and J. P. Lazaro, "Survey on Network Security Traffic Analysis and Anomaly Detection Techniques," *IJETAA*, 2024.

[17] M. Vignesh and S. Shanthini, "Machine Learning-Driven Detection of Encrypted VPN Traffic," *IJIREEICE*, 2025.

[18] "Anomaly Detection in Encrypted Network Traffic Using Self-Supervised Learning," *Scientific Reports*, 2024.

[19] "Intrusion Detection Based on Transformer and CNN-BiLSTM," *Sensors*, 2025.

[20] "Data Mining for Anomaly Detection in Network Traffic," *IJSRST*, 2025.