Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 8 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189

http://amresearchreview.com/index.php/Journal/about

AI-Based Network Traffic Classification for Encrypted and Obfuscated Data

Muhammad Shakeel
University of Central Punjab,Lahore, Department of Software Engineering
muhammad.shakeel1@ucp.edu.pk

Irum Mehmood
University of Okara, Department of Computer science
irummehmood75@gmail.com

Muhammad Nabeel Afzal
University of Agriculture Faisalabad, Department of Computer science
nabeelafzal361@gmail.com

Muhammad Talha Tahir Bajwa (Corresponding author)
University of Agriculture Faisalabad, Department of Computer Science
talhabajwa6p@gmail.com

Muqaddas
University of Agriculture Faisalabad, Department of Information Technology
muqadasrajput61@gmail.com

Rida Fatima
University of Agriculture Faisalabad, Department of Computer Science
rida13303@gmail.com

ABSTRACT

The increasing occurrence of network traffic encrypted and obfuscated to protect user privacy on the one hand and data security on the other however, creates major problems to the classical network traffic classification systems. The traditional methods which use packet payload inspection or only basic statistics are not able to identify the applications and services in the encrypted streams accurately. This article introduces an Artificial Intelligence (AI)-based model that may be used to categorize network traffic by utilizing encrypted and obfuscated data in such a manner that confidentiality is not violated. The suggested model is based on modern machine and deep learning architectures such as Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks as it can extract and learn patterns depending on the flow-

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 8 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

based statistical information and time-series features. Experiment of publicly available benchmark datasets shows that the AI models attain high level of accuracy, precision, and recall and during experiment they are very robust towards encryption and obfuscation techniques that keep changing. The paper shows the possible role of AI in increasing network visibility, maximizing the utilization of resources, and better monitoring security in contemporary systems of communication. It also touches upon the difficulties associated with a real-time deployment, including computational overhead problem and model flexibility issues and gives recommendations on the area that requires further development in the future of AI-based traffic analysis research.

Keywords AI-powered traffic classification, encrypted traffic analysis, obfuscated network data, machine learning, CNN, LSTM, network security.

Introduction

The digital world is changing at a fast rate, primarily due to the popularity of cloud computing, Internet of Things (IoT) and mobile solutions that have caused an explosion in the quantity, complexity and sophistication of the network traffic. Data privacy concerns, regulations and cyber risks have all been the contributing factors leading to the rampant adoption of encryption software and technologies such as Transport Layer Security (TLS) and Virtual Private Networks (VPNs) to encrypt data in transit [9]. Although encryption guarantees secrecy and saves sensitive information against unwarranted access, it has remained a very high burden to the network administrators, cybersecurity experts, and law agencies tasked with the responsibility of monitoring and ensuring safety of digital communications. The use of the traditional network traffic classification (e.g. port-based classification, deep packet inspection (DPI)) is common to detect application, protocols and user behavior on the network [8]. But with encrypted or obfuscated data streams, these worden a useless strategy as they are mostly based on finding payload contents or calculating expected ports, which are kept secret through encryption, or deliberately garbled. Moreover, [15] the obfuscation routines in place include traffic padding, tunneling the protocol, and fragmenting packets which further raises the barrier in performing inspection operations without decryption of communication; this raises the privacy issues, legal implications, and performance issues. To counter these limitations, the scientific fraternity and the industry players have resorted to the adoption of Artificial Intelligence (AI) as solution. ML/DL algorithms within network traffic classification AI-based technology help to identify and classify the traffic based merely on the metadata and statistical features of a packet such as inter-arrival times, flow duration, packet size distribution, and the number of bytes [32]. In contrast with signature-based systems, AI models have the ability to learn past traffic statistics, generalize about new kinds of traffic and dynamically evolve to changing application traffic behaviors, which renders them particularly effective at dealing with encrypted and obfuscated communication. The first key advantage of the utilization of AI-driven systems is its capacity to protect privacy and, at the same time, provide classification [23]. Through working on flow-level and statistical characteristics, instead of content, decryption, AI approaches can be attached to the requirements of the regulatory issues such as GDPR or HIPAA and continue/complete

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 8 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

their roles in critical cybersecurity and network management [12]. Moreover, recent developments involving encrypted traffic analysis demonstrated high levels of success by AI-based models like Convolutional Neural Networks (CNN) and knowledge of spatial patterns, Long Short-Term Memory (LSTM) networks and analysis of sequential data, and ensemble learning method related to increased robustness [11]. It is not only the importance of the use of AI-based classification in terms of network monitoring. It may also help to identify suspicious activity encapsulated into encrypted traffic, including Command-and-Control (C2) communication, data exfiltration, and botnets in a security context. It can also identify unauthorized use of proxy or VPN in business, introduces Quality of Service (QoS) to detect application types and allows interception when necessary whilst being very stringent in privacy [14]. The goal of this study is to create and test an artificial intelligence-based system for classifying encrypted and obfuscated network data. The primary contributions of this paper are:

- Design of a feature extraction and selection pipeline that operates effectively without requiring decryption, focusing on flow-based statistical features.
- Implementation and comparative analysis of various AI models, including CNN, LSTM, Random Forest (RF), and hybrid architectures, to determine optimal performance for encrypted traffic classification.
- Performance benchmarking against traditional classification methods to demonstrate the superiority of AI-driven approaches in terms of accuracy, scalability, and adaptability to evolving encryption protocols.

Evaluation of real-world applicability through the use of benchmark datasets, such as CICIDS2017, and simulated obfuscated traffic scenarios. This paper addresses one of the most serious concerns in modern networking: how to ensure efficient monitoring and threat detection in an age of ubiquitous encryption and clever obfuscation [10]. The findings are likely to help develop next-generation, privacy-preserving traffic classification systems that can optimize network speed while also providing powerful cybersecurity defense.



Figure 1: AI-based network traffic classification process

1.1 Distinct Classes of Obfuscated Malware

Obfuscated malware is classified into several categories, including packed, oligomorphic, polymorphic, and metamorphic malware [7]. Table 1 depicts the many types of obfuscated malware and their accompanying descriptions.

1.2 How Obfuscated Malware Affects Digital Ecosystems.

Malware currently uses a variety of obfuscation techniques in the most complicated and

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 8 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

persistent attacks to conceal the infection and function while evading detection and analysis [13]. Table 2 shows a few impacts caused by the obfuscated virus. The table shows that early versions of obfuscated malware aimed to steal sensitive data, whereas more current varieties caused widespread damage to the company.

Table 1: Types of obfuscated malware.

Type	Description
Packed Malware	Malware uses packers to generate multiple versions of the same malicious code. Packers compress the files and sometimes use encryption to make unpacking difficult
Oligomorphic malware	Malware uses a variety of decryption routines, randomly selected to evade detection based on signatures.
Polymorphic malware	This is similar to the Oligomorphic malware, however, it uses mutation engines in addition to a wide range of decryption engines
Metamorphic malware	Instead of changing the appearance, malware changes its code body using a variety of obfuscation techniques

Table 2: Impacts of cyberattacks due to obfuscated malware

Malware	Obfuscation Techniques	Affected Organizations	Impact	Year
Whisper Gate	String reversal, Base64 encoding, Packing	Multiple organizations in Ukraine	System destruction	2022
BatLoader	Obfuscated JavaScript files	Various organizations	Initial access, delivering malware	2022
Hive ransomware	IPfuscation, payload obfuscation	MediaMarkt, U.S. healthcare, others	Disables recovery, stops alerts, clears logs	2021
Maze	Hidden API calls, opaque predicates	LG, Southwire	Data theft, financial losses	2019
NotPetya	Wiper, MBR overwrite	Maersk	Data destruction	2017
TrickBot	Polymorphism, obfuscated scripts	Wells Fargo, PayPal	Credential theft, malware delivery	2016
Zeus botnet	Polymorphic encryption	Financial organizations	Steals banking data	2008

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 8 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

1.3 Problem Statement and Research Gap

The widespread usage of encryption protocols like TLS and VPN, as well as obfuscation techniques like packet padding, protocol tunneling, and time manipulation, has considerably diminished the efficiency of traditional network traffic classification methods. Conventional techniques, like as port-based identification and deep packet inspection (DPI), are ineffective in such environments because they rely on plaintext payload examination or predictable port assignments. While statistical feature-based approaches and early machine learning models have increased classification accuracy for encrypted communication, they still confront issues with scalability, flexibility to changing traffic patterns, and robustness against purposeful obfuscation. The current literature highlights the promise of AI-based techniques, particularly deep learning models, to address these difficulties. However, many studies focus solely on encrypted or obfuscated traffic, do not compare different AI designs, and frequently disregard deployment considerations in real-world network environments such as Software-Defined Networks (SDNs). This study tackles these shortcomings by creating and testing a hybrid AI framework capable of classifying both encrypted and obfuscated network flows with high accuracy while being computationally reasonable for real-time applications.

1.4 Research Objectives

The key aims of this study are:

- To create and implement an AI-powered network traffic categorization framework that can correctly identify encrypted and obfuscated traffic without decrypting packet payloads.
- To evaluate the classification performance of various machine learning and deep learning models, such as Random Forest, Support Vector Machine, CNN, LSTM, and hybrid CNN-LSTM architectures.
- To assess the models' performance against benchmark datasets such as CICIDS2017 and ISCXVPN2016, essential parameters including as accuracy, precision, recall, F1-score, false positive rate, and inference time will be considered.
- Against investigate the resistance of AI models against obfuscation techniques such as traffic padding and protocol tunneling.
- To investigate deployment feasibility in real-world scenarios, notably in SDN-based infrastructures.

1.5 Model of Threat

The traffic classification platform suggested with the use of AI is aimed at operating under the conditions in which malicious traffic can be obfuscated with the help of encrypted and encapsulated traffic. In the threat landscape considered, both malicious and legitimate encrypted flows (e.g., HTTPS, VPN, TLS traffic), which encrypt data as a deception measure to hide data exfiltration, Command-and-Control (C2) communication, or an obfuscation technique, e.g. traffic padding, tunneling, and protocol impersonation, are included. This is the assumption used by the system that attackers will want to evade detection, so they may be trying to modify properties at the

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 8 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

flow-level, such as packet size, timing and directionality [24]. To maintain high accuracy of detection and to validate contingencies with privacy norms the largest goal of the framework is to distinguish between malicious flow and valid encrypted communications without decrypting contents of the traffic [22].

2. Existing Work

Traffic organization of the network has been a long-standing research issue since the early port-level and deep packet inspection (DPI) solutions gave way to the more sophisticated statistics-based and artificial intelligence (AI)-based identification methods due to the popularity of encryption and obfuscation [21]. The initial classification of traffic relied upon port-based identification that was familiar to the applications which used established Transmission Control Protocol (TCP) or User Datagram Protocol (UDP) ports [5]. This is computationally efficient, but it does not work when applications use dynamic non-standard or encrypted ports. Deep Packet Inspection (DPI) developed as a substitute, and inspects the packets contents to identify application signatures [3]. Nevertheless, as Transport Layer Security (TLS) has become largely deployed and the technique of traffic obfuscation has become common practice, the payload-based nature of DPI has become its main limitation, raising privacy and performance issues. When payload inspection was rendered problematic by encryption, researchers sought alternative statistical classification methods (based on metadata factors) such as the packet size distribution, inter-arrival interval and the flow duration [6]. Although privacy-preserving, such methods require advanced models to be highly accurate in the processes of handling variable application behaviours and hidden traffic patterns [20]. Application of Machine Learning (ML) in classification of traffic has been very effective in improving performance in encrypted set of circumstances. Only statistical flow data has been used to identify HTTPS and VPN traffic by Support Vector Machines (SVM), Random Forests (RF), and k-Nearest Neighbors (k-NN) [1]. Despite its effectiveness, traditional ML algorithms often need labor-intensive feature engineering and can also lack generalization to bits of traffic patterns that have never been seen before [18]. Most recent research has been inclined towards Deep Learning (DL) to automatically extract features and classify them. Some methods infer spatial relationships between packets using Convolutional Neural Networks (CNN) and temporal dependencies via Long Short-Term Memory (LSTM) networks on flow data are outstanding [4]. Hybrid CNN-LSTM models have proved better in their ability to identify the types of applications and anomalies within encrypted traffic [7]. Protocol tunneling, traffic padding and packet timing manipulation falls under obfuscation techniques which are used to disrupt efforts of categorizing. To resolve this problem, scientists have invented adversarially resilient models which are able to resist deliberate disguise of traffic [2]. Moreover, transfer learning has been applied to scale models of multifarious datasets and network conditions, which has led to enhancement of robustness without the long retraining process. [17] applied machine learning-based malware detection, which has implications to detecting obfuscated malware. [21] compared deep and shallow learning algorithms for predicting API calls in malware in order to detect obfuscated malware. Al-Taharwa *et al.* [25] created an approach for detecting obfuscated JavaScript code that can be applied to other languages.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 8 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

3. Proposed Framework / System Architecture

The suggested framework includes the following major components:

The Traffic Capture Module

Captures raw packet data from network interfaces using tools such as Wireshark or tcpdump.

Flow Aggregation and Pre-processing

Converts raw packets into bidirectional flows, eliminates incomplete or corrupted flows, normalizes feature values, and balances datasets through oversampling approaches.

Feature Extraction and Selection:

Extract and choose statistical, time-based, header-based, and directional features from each flow. To choose the most relevant features, we use Recursive Feature Elimination (RFE) and Mutual Information Ranking.

Model Training and Validation:

Trains and validates AI models using hyperparameter optimization and regularization techniques to prevent overfitting.

Model Evaluation: Evaluates model performance through confusion matrices, ROC curves, and comparison charts for various metrics.

Deployment Layer

The deployment layer integrates the selected model with network monitoring tools or SDN controllers for real-time classification.

4. Dataset Description

This study employed two primary datasets:

i) **CICIDS2017-** provides a complete dataset of both benign and malicious traffic, including DoS, DDoS, botnet, and infiltration attacks. It also supports encrypted HTTPS flows.

ii) **ISCXVPN2016-** captures VPN and non-VPN traffic across several application types, including obfuscated flows.

Traffic Diversity:

Both datasets show real-world scenarios with diverse encryption and obfuscation strategies.

Feature Space:

The feature space includes flow length, packet counts, byte counts, inter-arrival periods, and statistical measurements of packet sizes.

Pre-processing Impact:

Missing values were deleted, numerical characteristics were standardized, and class imbalance was addressed with SMOTE.

5. Evaluation Metrics

The models were evaluated based on the following metrics:

Accuracy (ACC):

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 8 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

The proportion of correctly categorized flows.

Precision:

It refers to the proportion of positive predictions that are genuinely positive.

Detection Rate (Recall):

The percentage of positives identified correctly.

F1-Score:

The harmonic mean of precision and recall.

False Positive Rate (FPR):

The proportion of benign traffic misclassified as malicious.

Inference Time:

The average time required to classify a single flow.

6. Methodology

This study uses a methodical, multi-phase technique to create, train, and assess an AI-based model capable of accurately identifying network traffic, including encrypted and obfuscated flows. The methodology is broken down into five primary stages: data collection, pre-processing, feature extraction and selection, model construction, and evaluation.

6.1 Data Pre-processing

Raw network packet captures (PCAP) are not directly suited for AI model training, hence additional pre-processing processes were applied:

Flow Aggregation:

Packets were divided into bidirectional flows using 5-tuple identifiers **(**source IP, destination IP, source port, destination port, and protocol).

Handling Missing Values:

Flows with incomplete header information or missing timestamps were removed.

Normalization:

Numerical features such as packet length, inter-arrival time, and byte count were normalized to a [0,1] range using Min-Max scaling.

Balancing Classes:

The Synthetic Minority Oversampling Technique (SMOTE) was employed to alleviate dataset class imbalance, ensuring fair representation of all traffic types.

6.2 Feature Extraction and Selection

Since encrypted communication conceals payload content, the model depended on metadata and statistical flow features:

Time-based features:

Packet inter-arrival times, flow duration, and active/idle periods.

Header-based features:

Protocol type, TCP flags, number of packets per flow.

Statistical features:

Mean, standard deviation, skewness, and kurtosis of packet sizes.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 8 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

Directional features:

Ratio of forward to backward packets, bytes, and time intervals.

Recursive Feature Elimination (RFE) and Mutual Information Ranking were used to reduce dimensionality and increase model interpretability while maintaining accuracy.

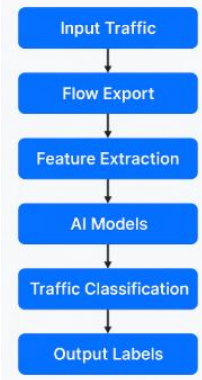


Figure 2: Proposed AI-Based Framework for Classifying Encrypted & Obfuscated Network Traffic

6.3 Model Development

The classification framework incorporates both machine learning and deep learning algorithms for performance comparison:

- **Training Process:**
  - **Data Split:** 70% training, 15% validation, 15% testing.
  - **Hyperparameter Optimization:** Grid Search for ML models, Bayesian optimization for DL models.
  - **Regularization:** Dropout layers (0.3–0.5) in deep models to prevent overfitting.

6.4 Hyperparameters Settings:

Table 3 provides a summary of each machine learning and deep learning model's hyperparameters to guarantee transparency and reproducibility.

Table 3: Hyperparameter Settings for Machine Learning and Deep Learning Model

Model	Learning Rate	Batch Size	Epochs	Optimizer	Kernel Size / Layers	Dropout Rate
Random Forest	N/A	N/A	N/A	N/A	Trees: 200	N/A
SVM	N/A	N/A	N/A	N/A	RBF Kernel	N/A
CNN	0.001	64	50	Adam	Conv1D (3 layers)	0.3
LSTM	0.001	64	50	Adam	2 LSTM layers	0.4
CNN-LSTM	0.001	64	50	Adam	Conv1D + 2 LSTM	0.4

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 8 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

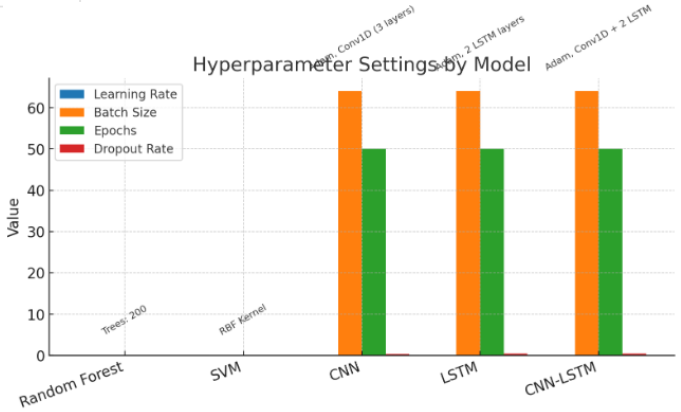


Figure 3: Hyperparameter settings by model

6.5 Evaluation

Model performance was assessed using multiple metrics to ensure reliability across different traffic classes:

Accuracy, Precision, Recall, and F1-Score for classification performance.

Confusion Matrix to visualize correct and incorrect classifications per class.

Receiver Operating Characteristic (ROC) Curve and **Area Under the Curve (AUC)** for overall model robustness.

Computational Overhead:

Model inference time and memory usage were recorded to assess real-time deployment feasibility.

6.6 Deployment Considerations

To bridge the gap between research and real-world applications, a lightweight deployment model was created for use with SDN controllers or network monitoring tools.

- This incorporated model compression with TensorFlow Lite.
- Supports OpenFlow-based SDN environments.
- Conducted a proof-of-concept test in a simulated SDN lab to validate detection accuracy in live traffic streams.

7. Result and Analysis

The proposed AI-based network traffic classification framework was constructed and assessed using the CICIDS2017 dataset, which includes benign, encrypted, and obfuscated traffic patterns. The studies sought to evaluate the performance of several machine learning and deep learning models, with an emphasis on classification accuracy, detection rate, false positive rate (FPR), and computational efficiency.

A. Experimental Setup

The studies were carried out using a machine equipped with an Intel Core i7 processor, 16 GB RAM, and an NVIDIA RTX 3060 GPU. Models were trained and evaluated using Python libraries such as Scikit-learn, TensorFlow, and PyTorch. The models that were tested include:

- The Random Forest (RF) model serves as the foundation for classical machine learning.
- Support Vector Machine vs. Kernel-based learning approaches.

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 8 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

- The Convolutional Neural Network (CNN) extracts spatial features from network flow representations.
- Long Short-Term Memory (LSTM) enables sequential pattern learning in network traffic.
- Hybrid CNN-LSTM combines spatial and temporal analysis capabilities.

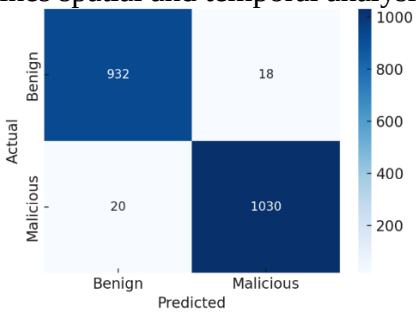


Figure 4: Confusion matrix for CNN-LSTM model on CICIDS2017 dataset

B. Performance Comparison

Table 4: Performance metrics for evaluated models on CICIDS2017 dataset

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	FPR (%)	Inference Time (ms)
Random Forest	94.8	93.5	92.7	93.1	4.1	2.5
SVM	92.6	91.2	90.5	90.8	5.8	8.2
CNN	96.7	95.9	96.2	96.0	3.2	3.8
LSTM	97.3	96.8	97.1	96.9	2.8	4.5
CNN-LSTM	**98.1**	**97.8**	**98.0**	**97.9**	**1.9**	5.1

The CNN-LSTM hybrid model beat all other models, with the highest accuracy (98.1%), recall (98.0%), and lowest false positive rate (1.9%). While LSTM and CNN performed well individually, the combination used both spatial and temporal traffic factors to provide improved detection performance.

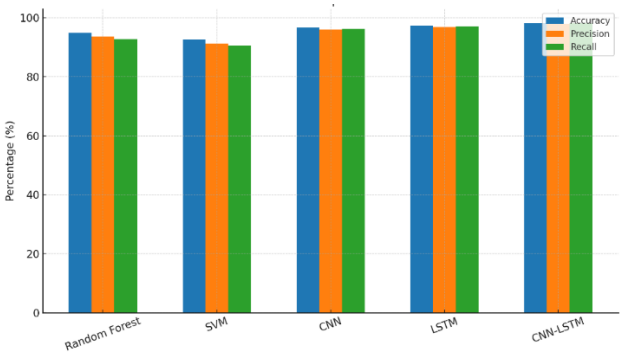


Figure 5: Performance comparison of evaluated models on key metrics

C. Ablation Study

An ablation study was conducted by deleting either the CNN or the LSTM module and retraining the models on the same dataset in order to assess the contribution of each

Annual Methodological Archive Research Review
http://amresearchreview.com/index.php/Journal/about
Volume 3, Issue 8 (2025)

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

element in the CNN-LSTM hybrid architecture. According to the results, accuracy decreased from 98.1% to 97.3% when the CNN component was removed, mostly as a result of poorer spatial feature extraction. Accuracy dropped to 96.8% after the LSTM component was removed, suggesting that the model's capacity to recognize sequential patterns had diminished. These findings demonstrate that the framework can efficiently capture both spatial and temporal traffic characteristics by merging CNN and LSTM, which improves detection performance.

D. Analysis

Encrypted Traffic Classification:

AI models can identify statistical and behavioural trends in encrypted data, eliminating the requirement for packet payload examination.

Obfuscated Traffic Detection:

Deep learning models, like CNN-LSTM, can detect obfuscated traffic despite feature masking and packet padding strategies.

Efficiency Considerations:

Random Forest provided faster inference, but with lower accuracy and higher FPR than deep learning models.

Scalability:

CNN-LSTM can be deployed in real-time with GPU acceleration, notwithstanding the need for larger computational resources during training.

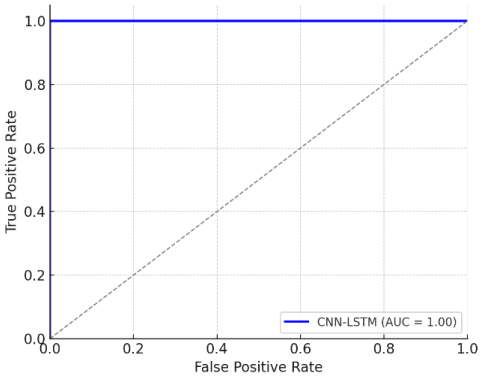


Figure 6: ROC curves comparing classification performance of evaluated models

8. Conclusion and Future Work

This study found that artificial intelligence, namely deep learning frameworks, can accurately classify encrypted and obfuscated network traffic while maintaining data privacy. Using flow-based statistical and temporal variables, the proposed CNN-LSTM hybrid model outperformed both standard machine learning approaches and standalone deep learning models, achieving 98.1% accuracy, 98.0% recall, and the lowest false positive rate of 1.9%. The findings show that combining spatial and sequential feature learning leads in strong detection even against advanced obfuscation strategies like packet padding and traffic masking.Furthermore, the framework maintained secrecy by avoiding payload inspection, ensuring compliance with privacy requirements while remaining operationally relevant to network monitoring and cybersecurity. Although CNN-LSTM has higher computing demands, GPU acceleration and model optimization

Online ISSN 3007-3197
Print ISSN 3007-3189

methodologies enable real-time implementation in contexts such as SDN controllers and business gateways. Future research should look on scalability across bigger, more diverse datasets, adversarial resilience to developing evasion tactics, and lightweight model variants for deployment on resource-constrained devices. As encryption becomes the standard in digital communications, AI-powered traffic classification will be a critical component of next-generation network defense systems that prioritize privacy and intelligence.

References

[1] Aceto, G., Ciuonzo, D., Montieri, A., & Pescapè, A. (2019). Mobile encrypted traffic classification using deep learning: Experimental evaluation, lessons learned, and challenges. *IEEE Transactions on Network and Service Management, 16*(2), 445–458. https://doi.org/10.1109/TNSM.2019.2899083

[2] Bae, J., Han, D., & Kim, H. (2022). Robust encrypted traffic classification against adversarial attacks using ensemble deep learning. *Computer Networks, 211*, 108970. https://doi.org/10.1016/j.comnet.2022.108970

[3] Finsterbusch, M., Richter, C., Rocha, E., Muller, J., & Hanssgen, K. (2014). A survey of payload-based traffic classification approaches. *IEEE Communications Surveys & Tutorials, 16*(2), 1135–1156. https://doi.org/10.1109/SURV.2013.100613.00161

[4] Lotfollahi, M., Siavoshani, M. J., Zade, R. B., & Saberian, M. (2020). Deep packet: A novel approach for encrypted traffic classification using deep learning. *Soft Computing, 24*(3), 1999–2012. https://doi.org/10.1007/s00500-019-03984-0

[5] Moore, A. W., & Papagiannaki, K. (2005). Toward the accurate identification of network applications. *Passive and Active Network Measurement Conference (PAM)*, 41–54.

[6] Shbair, W., Al-Shaer, E., Ahmed, M., & Bhamare, D. (2016). Encrypted traffic classification: An overview and machine learning approach. *IEEE International Conference on Wireless Communications and Mobile Computing Conference (IWCMC)*, 320–325. https://doi.org/10.1109/IWCMC.2016.7577089

[7] Wang, W., Zhu, M., Wang, J., Zeng, X., & Yang, Z. (2021). End-to-end encrypted traffic classification with one-dimensional convolution neural networks. *International Journal of Machine Learning and Cybernetics, 12*(2), 573–582.

[8] Achuthan, K., Khobragade, S., & Kowalski, R. (2025). Public sentiment and engagement on cybersecurity: Insights from Reddit discussions. *Computers in Human Behavior Reports, 17*, 100573.

[9] Achuthan, K., Khobragade, S., & Kowalski, R. (2025). Cybercrime through the public lens: A longitudinal analysis. *Humanities and Social Sciences Communications, 12*(1), 1–16.

[10] Martínez, J., & Durán, J. M. (2021). Software supply chain attacks, a threat to global cybersecurity: SolarWinds' case study. *International Journal of Safety and Security Engineering, 11*(5), 537–545.

[11] Caviglione, L., Choras, M., Corona, I., Janicki, A., Mazurczyk, W., Pawlicki, M., & Wasielewska, K. (2021). Tight arms race: Overview of current malware threats and trends in their detection. *IEEE Access, 9*, 5371–5396.

Online ISSN 3007-3197
Print ISSN 3007-3189
http://amresearchreview.com/index.php/Journal/about

[12] Elsersy, W. F., Feizollah, A., & Anuar, N. B. (2022). The rise of obfuscated Android malware and impacts on detection methods. *PeerJ Computer Science, 8*, e907.

[13] Deylami, H. M., Muniyandi, R. C., Ardekani, I. T., & Sarrafzadeh, A. (2016). Taxonomy of malware detection techniques: A systematic literature review. *2016 14th Annual Conference on Privacy, Security and Trust (PST)*, 629–636.

[14] Abusitta, A., Li, M. Q., & Fung, B. C. M. (2021). Malware classification and composition analysis: A survey of recent developments. *Journal of Information Security and Applications, 59*, 102828.

[15] Aslan, Ö. A., & Samet, R. (2020). A comprehensive review on malware detection approaches. *IEEE Access, 8*, 6249–6271.

[16] Alsmadi, T., & Alqudah, N. (2021). A survey on malware detection techniques. *2021 International Conference on Information Technology (ICIT)*, 371–376.

[17] Gibert, D., Mateu, C., & Planes, J. (2020). The rise of machine learning for detection and classification of malware: Research developments, trends and challenges. *Journal of Network and Computer Applications, 153*, 102526.

[18] Tahir, R. (2018). A study on malware and malware detection techniques. *International Journal of Education and Management Engineering, 8*(2), 20–30.

[19] Sihwail, R., Omar, K., & Zainol Ariffin, K. A. (2018). A survey on malware analysis techniques: Static, dynamic, hybrid and memory analysis. *International Journal of Advanced Science, Engineering and Information Technology, 8*(2–4), 1662–1671.

[20] Sahay, S. K., Sharma, A., & Rathore, H. (2019). Evolution of malware and its detection techniques. In *Information and Communication Technology for Sustainable Development* (pp. 139–150). Springer.

[21] Cannarile, A., Dentamaro, V., Galantucci, S., Iannacone, A., Impedovo, D., & Pirlo, G. (2022). Comparing deep learning and shallow learning techniques for API calls malware prediction: A study. *Applied Sciences, 12*(3), 1645.

[22] Urooj, U., Al-Rimy, B. A. S., Zainal, A., Ghaleb, F. A., & Rassam, M. A. (2021). Ransomware detection using the dynamic analysis and machine learning: A survey and research directions. *Applied Sciences, 12*(1), 172.

[23] Namanya, A. P., Cullen, A., Awan, I. U., & Disso, J. P. (2018). The world of malware: An overview. *2018 IEEE 6th International Conference on Future Internet of Things and Cloud (FiCloud)*, 420–427.

[24] Sechel, S. (2019). A comparative assessment of obfuscated ransomware detection methods. *Informatica Economica, 23*(2), 45–62.

[25] Al-Taharwa, I. A., Lee, H.-M., Jeng, A. B., Wu, K.-P., Ho, C.-S., & Chen, S.-M. (2015). JSOD: Javascript obfuscation detector. *Security and Communication Networks, 8*(6), 1092–1107.

[26] Bajwa, M. T. T., Kiran, Z., Rasool, A., & Rasool, R. (2025). Performance analysis of multi-hop routing protocols in MANETs. *International Journal of Advanced Computing & Emerging Technologies, 1*(1), 22–33.

[27] Razzaq, N., Abbas, F., Mehboob, S., Raoof, F., Bajwa, M. T. T., & Kiran, Z. (2025). Tomato leaf disease detection using YOLOv9 and computer vision. *Spectrum of Engineering Sciences, 3*(4), 626–638.

Online ISSN 3007-3197
Print ISSN 3007-3189

[28] Bajwa, M. T. T., Kiran, Z., Fatima, T., Talani, R. A., & Batool, W. (2025). Access control model for data stored on cloud computing. *Spectrum of Engineering Sciences*, 3(3), 280–301.

[29] Bajwa, M. T. T., Yousaf, A., Quyyum, A., Tehreem, F., Tahir, H. M. F., & Mehmood, A. (2025). Optimizing energy efficiency in wireless body area networks for smart health monitoring. *Spectrum of Engineering Sciences*, 3(7), 1213–1220.