Next-Gen Security: Enhanced DDoS Attack Detection for Autonomous Vehicles in 6G Networks

Sudha Anbalagan[1], Wajdi Alhakami[2], Mugundh Jambukeswaran Bhooma[3], Vijai Suria Marimuthu[4]
Kapal Dev[5], Gunasekaran Raja[6]
[1]Centre for Smart Grid Technologies, SCOPE, Vellore Institute of Technology, Chennai, India
[2]College of Computers and Information Technology, Taif University, Taif, Saudi Arabia
[3,4,6]NGNLab, Department of Computer Technology, Anna University, MIT Campus, Chennai, India
[5]Department of Computer Science, Munster Technological University, Ireland
[1]sudha.a@vit.ac.in, [2]whakami@tu.edu.sa, [3]mugundhjb@gmail.com, [4]vijaisuria04@gmail.com
[5]kapal.dev@ieee.org, [6]dr.r.gunasekaran@ieee.org

2024 IEEE 99th Vehicular Technology Conference (VTC2024-Spring) | 979-8-3503-8741-4/24/$31.00 ©2024 IEEE | DOI: 10.1109/VTC2024-SPRING62846.2024.10683431

***Abstract*—Autonomous Vehicles (AVs) have revolutionized transportation by utilizing 6G technologies such as automated driving assistance, navigation, connected intelligence, and independent decision-making. Yet, the increasing reliance on AVs exposes the Internet of Vehicles (IoV) to potential vulnerabilities, making it susceptible to cyber attacks. One prominent threat is Distributed Denial of Service (DDoS) attacks, which can significantly impact AVs' safety and operational integrity. DDoS attacks directly disrupt the fundamental functionality of AVs to make timely and informed decisions, potentially leading to accidents or system failures. Despite the existence of numerous systems for detecting DDoS attacks, their continuous evolution in various attack patterns poses a significant challenge for effective detection. This paper provides a vision of 6G Security by proposing an Advanced DDoS Attack Detection System (ADADS) to enhance the detection capabilities of DDoS attacks by employing a Hybrid Detection Model (HDM) and a Continuous Learning Model (CLM) to adapt the evolving patterns of DDoS attacks over time dynamically. The collaborative integration of these models leverages the overall efficiency of DDoS attack detection, delivering a robust and adaptive defense mechanism. The experimental findings reveal that the proposed ADADS achieves a remarkable accuracy 98.7% with rapid stabilization in a few iterations for the current 6G specifications and applications.**

***Index Terms*—Autonomous Vehicles, 6G Communications, DDoS Attack Detection, Hybrid Detection Model, Continuous Learning Model.**

I. Introduction

The advent of Autonomous Vehicles (AVs) holds immense potential to revolutionize the transportation landscape, fostering a future characterized by enhanced safety and reduced environmental footprint in the evolution of 6G [1]. AVs, relying on emerging technologies such as smart driving instruments and Machine Learning (ML) algorithms, demonstrate the capability to adeptly navigate their surroundings, thereby minimizing accidents and carbon emissions. The potential of achieving zero carbon emissions and enhanced security are the characteristics of 6G that need to be incorporated in the AVs [2]. This reliance on advanced technologies empowers these vehicles to make swift and precise decisions, consequently elevating safety standards in transportation [3]. However, while AVs demonstrate remarkable capabilities in navigating their surroundings, the integration of ML lacks standardized frameworks, posing challenges in ensuring reliability.

Concurrently, the evolution of AVs introduces a new landscape fraught with security vulnerabilities, particularly in the realm of data transmission across connected networks. This necessitates robust frameworks to fortify the security measures [4]. Cyber threats require attention, and a proposed framework aims to classify and enhance defense mechanisms. Nonetheless, complexities in AV cyber threats might be overlooked, necessitating real-world validation for the framework's effectiveness [5]. Despite the transformative potential of autonomous vehicles in diverse domains such as healthcare and transportation, incidents involving these vehicles have highlighted the susceptibility to errors and security breaches [6]. With the advancement of big data and communication technologies, techniques have gradually evolved to leverage Artificial Intelligence (AI) and ML to detect abnormalities in AV systems [7]. However, the lack of universally accepted frameworks for AI integration poses a significant challenge [8].

Cyber assaults targeting AVs encompass a spectrum of methods, from malware injection to Distributed Denial of Service (DDoS) attacks. In recent years, strides in research have focused on devising techniques to identify and forestall these threats. Yet, detecting DDoS attacks in interconnected AV systems has received comparatively limited attention, which can endanger both the vehicles' functionality and passenger safety [9].

Therefore, while recent advancements in AV have brought about significant benefits, they have also raised important security considerations, particularly cybersecurity. Given the profound implications of DDoS attacks on AV safety and operations, it becomes imperative to formulate robust strategies for the identification and mitigation. In subsequent sections, we will delve into the urgency of detecting DDoS attacks within interconnected AV networks, scrutinizing recent advancements in the 6G research domain.

979-8-3503-8741-4/24/$31.00 ©2024 IEEE

Our study introduces the Advanced DDoS Attack Detection System (ADADS), an inventive approach that combines various Machine Learning algorithms within a hybrid model and integrates continuous learning capabilities for enhanced DDoS attack detection.

1) A comprehensive Hybrid Detection Model (HDM) has been developed to identify diverse forms of DDoS attack patterns by incorporating Random Forest (RF), K-Nearest Neighbor (KNN), Logistic Regression (LR), Support Vector Machine (SVM), and Naïve Bayes (NB) algorithms.
2) A novel Continuous Learning Model (CLM) is proposed to continuously evolve to new DDoS attack patterns utilizing HDM. This ensures that the CLM's detection capabilities remain up-to-date.

II. Related Works

Several existing works have focused on DADS using ensemble techniques. Q Liu et al. [1] combined XGBoost, RF, and Decision Tree (DT) to detect diverse external network attacks. This ensemble method, utilizing prediction confidence and majority voting, outperformed single attack models in detecting network intrusions, especially on the CICIDS2017 dataset. Z. Abdollahi Biron et al. [7] demonstrated a real-time scheme for diagnosing DoS cyber attacks in connected vehicles. It can potentially detect the occurrence of DoS and estimate its effect through a time delay in the information processing via a communication network. In [8], the authors developed a stacking ensemble framework for network intrusion detection with feature selection methods in Internet of Vehicles (IoVs) using tree-based ML models such as DT, RF, and XGBoost. Tree-based models combined using stacking improved accuracy, attack detection, and F1-score with longer execution time. This stacking ensemble model shows high accuracy on the Controller Area Network (CAN) intrusion and CICIDS2017 datasets.

A. Aljuhani [9] proposed a comprehensive DDoS defense system using ML/DL techniques in different environments. The aforementioned system classifies and analyzes the results based on various parameters, including the type of DDoS attack, the specific ML approach, the dataset utilized, and the evaluation metrics. G. Twardokus et al. highlighted vulnerabilities in Cellular Vehicle-to-Everything (C-V2X) technology, specifically in its physical and MAC layers. The paper depicted novel DoS attacks that exploit these weaknesses, raising concerns about the resilience of C-V2X against targeted attacks [10]. In [11], the authors developed a novel Intelligent Intrusion Detection System (IIDS) framework for efficient cyberattack detection in In-Vehicle networking and external vehicle networking. The IIDS framework employed a decentralized 5G-V2X network that enables several AVs to detect cyberattacks by transmitting alarm messages, resulting in fast and reliable intrusion detection achieving 98% accuracy.

A CNN-based IDS [12] is proposed for CAN bus systems, targeting common network attacks like DoS, Fuzzing, and Spoofing. Their approach efficiently selects optimal hyperparameter values such as detection accuracy, rate, false positives, false negatives, and F1-score to create an effective CNN-based IDS for the CAN bus. The usefulness of Recursive Feature Elimination was pointed out in [13] to enhance DADS by focusing on key features, reducing training and response times while notably improving detection accuracy.

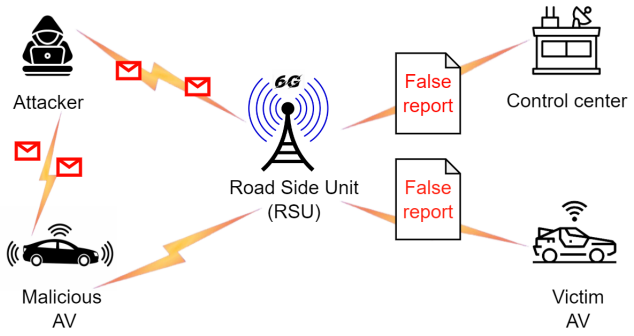

Fig. 1. Attack Scenario in AVs

In [14], the authors proposed an IDS incorporating Stochastic Gradient Descent for detecting intrusions in assistance with blockchain for an enhanced trust evaluation in a 5G-V2X IoV environment.

Although many of the related works achieve high performance in DDoS attack detection in IoV, they fail to identify and incorporate newly evolving attack patterns. Existing DADS frameworks can be improved using advanced algorithms and strategies. The DDoS attack scenario is depicted in Fig 1. It indicates a flood of messages the adversary is transmitting to attack the network. The proposed ADADS framework leverages DDoS attack detection using an HDM with ML algorithms such as NB, RF, LR, KNN, and SVM. Through CLM, it can effectively adapt to new attack patterns within a few iterations, resulting in high accuracy and model stabilization in 6G networks.

III. Advanced DDoS Attack Detection System

In the proposed ADADS system, as illustrated in Fig 2, the model commences by gathering data from the network traffic among autonomous vehicles. The CICDDoS2019 dataset [16], renowned for its comprehensive analysis of ML and DL models for DDoS attacks, was utilized for this purpose, providing real-time DDoS attacks from network traffic. Initially,

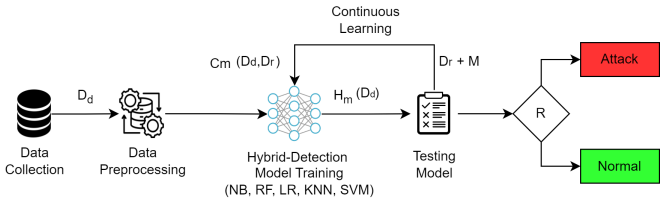

Fig. 2. Architecture diagram of proposed ADADS framework

Algorithm 1 Advanced DDoS Attack Detection System (ADADS) - HDM Training Phase

Require: AV dataset $\mathbf{D}_d$, Threshold limit ε
Ensure: Hybrid Detection Model $H_m(\mathbf{D}_d)$, Attack Classification $\mathbf{R}$, Accuracy $\mathring{A}$

for each $i = 1, 2, \ldots, n$ in AVs **do**
  Load and preprocess $\mathbf{D}_d$
  Split the dataset into training and test sets
  Calculate $Var(X)$ according to (2)
  Drop columns with zero variance
  for each model in H_m **do**
    Train each model
    Make prediction using each model
  end for
  Combine the predictions of each model
  Make predictions for the hybrid model using (5)
  for each preds in predictions **do**
    Find majority voting based on (4)
  end for
  Evaluate the hybrid model
  Calculate accuracy $\mathring{A}$ according to (1)
  Return $H_m(\mathbf{D}_d)$, $\mathbf{R}$, $\mathring{A}$
end for

Algorithm 2 Advanced DDoS Attack Detection System (ADADS) - CLM

Require: AV dataset $\mathbf{D}_d$, AV test data and evolved attack patterns $\mathbf{D}_r$, Threshold limit ε
Ensure: Continuous Trained Model $C_m(\mathbf{D}_d, \mathbf{D}_r)$, Attack Classification $\mathbf{R}$, Accuracy $\mathring{A}$

Initialize $\mathbf{D}_d$ and $\mathbf{D}_r$
for each $i = 1, 2, \ldots, n$ in AVs **do**
  while $\mathbf{D}_r$!= null **do**
    Initialize numpy array
    Load $\mathbf{D}_r$ and combine with $\mathbf{D}_d$ and $\mathbf{M}$
    Perform $\mathbf{D}_c$ transformation
    Split the dataset into training and test sets
    Make prediction according to (5)
    for $\mathbf{R}$ == true **do**
      Determine the pattern and threshold limit ε
      Update the model $C_m(\mathbf{D}_d, \mathbf{D}_r)$
    end for
    Calculate accuracy $\mathring{A}$ according to (1)
    Save the model at checkpoint
    Load the new model
    Return $C_m(\mathbf{D}_d, \mathbf{D}_r)$, $\mathbf{R}$, $\mathring{A}$
  end while
end for

categorical data is encoded using label encoding, and one-hot encoding is applied to the `'Protocol'` data due to its absence of natural order, resulting in binary columns for each protocol type. One-hot encoding was crucial to prevent the model from interpreting categorical variables with any inherent order or significance. After these procedures, columns with zero variance were eliminated. Subsequently, the preprocessed data was split into a 3:1 ratio for training and testing.

A. Hybrid Model

Table I depicts the analysis of five distinct ML models, incorporating them into a hybrid model to further enhance the accuracy in detecting DDoS among AVs. This hybrid ensemble technique outperforms NB, LR, RF, SVM, and KNN using a voting mechanism. It stands out for its robustness and reduced risk of overfitting. Additionally, Algorithm 1 outlines the HDM Training Phase, detailing the steps involved in training the hybrid detection model. Thus, the proposed ADADS system approach provides more interpretable results by capitalizing on the unique interpretations of various algorithms.

1) Random Forest (RF): Analyzes traffic patterns to spot anomalies caused by DDoS attacks. It's adept at handling complex data and flagging outliers that might signify malicious activities.
2) K-Nearest Neighbors (KNN): Identifies unusual patterns in incoming data by comparing them to historical data. Any sudden deviations raise red flags for further investigation.
3) Logistic Regression (LR): Assesses the probability of incoming data packets being part of a DDoS attack based on various features, contributing to the overall probability assessment of an attack.
4) Support Vector Machines (SVM): Classifies incoming data as normal or malicious by finding the best separation between different classes. It's efficient in detecting complex patterns in high-dimensional spaces.
5) Naive Bayes (NB): Complements other algorithms by calculating probabilities of various attack scenarios based on observed data. It assumes feature independence, offering an additional perspective to the analysis.

Consider an ensemble of ML models $M = \{m_1, m_2, ..., m_n\}$ tasked with classifying objects into distinct categories $C = \{c_1, c_2, ..., c_k\}$, each model is assigned a weight $W = \{w_1, w_2, ..., w_n\}$ based on its performance metrics, such as accuracy and mean average error rate.

Accuracy represents the ratio of correctly predicted instances to the total instances in the dataset for classifier 'i'.

$$\text{Acc}_i = \frac{\mathbf{T} - \mathbf{F}}{\mathbf{T}} \times 100\% \tag{1}$$

where T is the total predictions and F is the misclassified instances.

Variance of a dataset (X) is obtained by calculating the average squared difference between each data point X_i and the mean $\bar{X}$.

$$Var(X) = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 \tag{2}$$

where n is the number of data points.

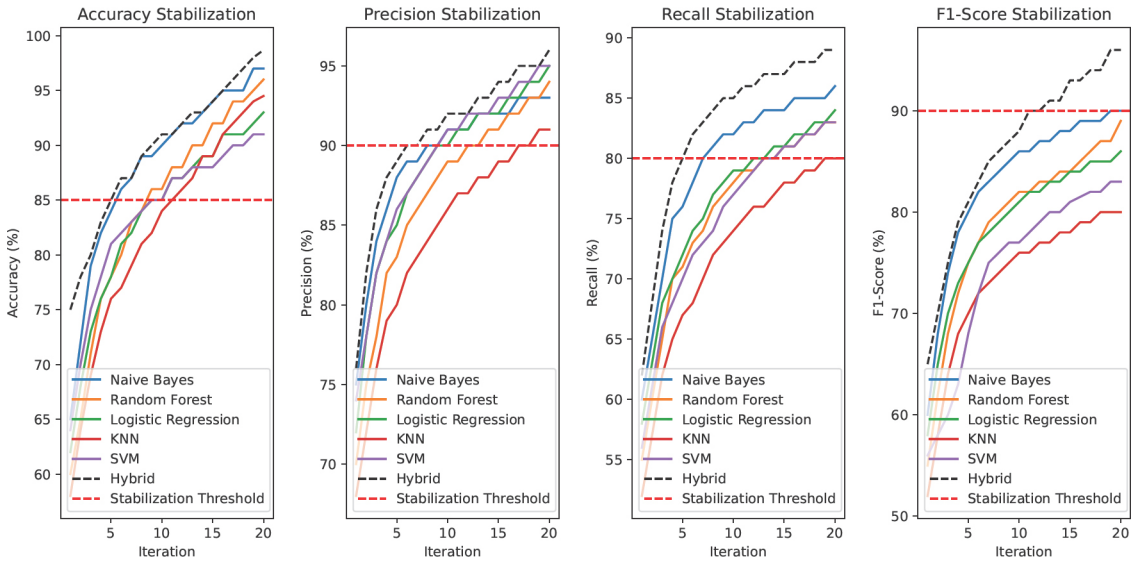

Fig. 3. Results of ADADS with Continuous Learning Model

Mean Average Error (MAE) quantifies the average error between predictions and actual values for classifier 'i'.

$$\text{MAE}_i = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{3}$$

The term w_i refers to the weight and is a coefficient assigned to classifier 'i', indicating its influence or importance in the weighted voting scheme.

$$w_i = \frac{\text{Acc}_i - \text{MAE}_i}{\sum_{i=1}^{n}(\text{Acc}_i - \text{MAE}_i)} \tag{4}$$

Weighted Voting Prediction calculates the final prediction by considering the weighted contribution of each classifier. The individual prediction of the i-th classifier is represented as C_i.

$$\text{Weighted Voting Prediction} = \frac{\sum_{i=1}^{n} C_i \cdot w_i}{\sum_{i=1}^{n} w_i} \tag{5}$$

The weights adjust the impact of each classifier's prediction on the final decision. $\hat{y}_\kappa$ is the predicted probability for class κ. The same can be represented as below,

$$\hat{y}_\kappa =_\kappa \sum_{i=1}^{M} \omega_i \Theta_i(y = \kappa|x) \tag{6}$$

B. Continuous Learning Model

Various ML models exhibit differing performance across various attack patterns. Therefore, our ADADS aims to propose a CLM for AV data, addressing varying model performance across attack patterns as detailed in Algorithm 2. The dataset, comprising both the test dataset and evolved attack patterns of DDoS attacks, is denoted as $\mathbf{D}_r$. The misclassified instances $\mathbf{F}$ from previous iterations are identified and denoted as $\mathbf{M}$. For a training iteration $(\mathbf{D}_r, \mathbf{M})$, $\mathbf{R}$ is the attack classification that determines whether the vehicle is attacked or not. The proposed CLM is constructed using a Hybrid Detection algorithm and combined dataset. The combined dataset $\mathbf{D}_c$ can be represented as

$$\mathbf{D}_c = \mathbf{D}_d \cup \mathbf{D}_r \cup \mathbf{M} \tag{7}$$

$\mathbf{D}_c$ is preprocessed and trained on the HDM to obtain the refined model parameters. During this training, new attack patterns are detected, and the model $C_m(\mathbf{D}_d, \mathbf{D}_r)$ updates its decision boundaries (Δ) and threshold limit ε accordingly. The updated model is then saved as a checkpoint using the *dump* function and stored with the filename `'model_checkpoint.joblib'`. The refined model is then employed for continuous learning in the next iteration, incorporating $\mathbf{D}_r$ and adjusting its detection capabilities based on the misclassified instances of attacks $\mathbf{M}$. The process continues until no new datasets remain, yielding improved accuracy with each subsequent repetition. This iterative process ensures the CLM adapts to evolving attack patterns, maintaining optimal performance over time.

TABLE I
RESULTS OF ADADS BEFORE CLM

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
NB	0.9435	0.9504	0.95	0.94
RF	0.9705	0.9845	0.9783	0.9764
LR	0.9605	0.9902	0.9666	0.9566
KNN	0.9820	0.9875	0.9982	0.9781
SVM	0.9715	0.9789	0.9765	0.9806
ADADS	**0.9656**	**0.9783**	**0.9769**	**0.9663**

IV. Results & Discussion

The primary focus is identifying DDoS attacks in AVs using ML techniques to analyze network traffic patterns and identify anomalies. This research introduced an innovative hybrid model that integrates five supervised ML algorithms into an ensemble learning through a soft voting mechanism. This approach capitalizes on the strengths of individual algorithms to enhance overall performance. Experimental results demonstrate that the hybrid model surpasses individual algorithms' accuracy and detection rate.

The study tabulated the results of the proposed system's DDoS attack detection using various models, presenting metrics such as accuracy, recall, precision, and F1-score in Table I. The weighted soft voting scheme emerged as the optimal choice among the ensemble voting mechanisms, yielding hybrid results. Notably, the accuracy achieved 96.56%, while precision and recall attained close to 98%. This improvement was attributed to dataset standardization, proper feature selection, and labeling. This research surpassed previous combinations such as KNN, SVM, and a hybrid model of RF and NB [15] on the NSL-KDD dataset.

The outcomes from ADADS, integrating the CLM, are visually represented in Fig 3. They depict the performance metrics of the proposed system, including accuracy, recall, precision, and F1-Score. The results consistently improve with each iteration, reaching a peak accuracy of 98.7% after 20 iterations when tested with evolved unseen datasets. The CLM achieves higher accuracy and superior performance compared to other models, ensuring swift response in Internet of Vehicles (IoVs). Continuous learning with future advancements notably boosts the model's accuracy and performance with each iterative update. However, this enhancement is accompanied by increased latency. To address this, iterations were capped at the maximum threshold, ceasing further model training.

V. Conclusion & Future Works

Highlighting the pressing need to safeguard AVs from DDoS attacks within the IoVs environment, the proposed ADADS showcasing remarkable effectiveness, achieving greater accuracy with minimal iterations. ADADS demonstrates resilience and adaptability against evolving attack patterns in 6G networks, offering a robust defense mechanism crucial for maintaining safety and operational integrity amidst evolving cyber threats. Future works could focus on identifying any types of attack patterns, such as spoofing and replay attacks, resulting in a zero-day vulnerability solution.

References

[1] Q. Liu, W. Bao and Q. Liu, "Research on Vehicular External Network Intrusion Detection System Based on Ensemble Learning," 2023 IEEE 98th Vehicular Technology Conference (VTC2023-Fall), Hong Kong, Hong Kong, 2023, pp. 1-6.

[2] Hongzhi Guo, Xiaoyi Zhou, Jiajia Liu, Yanning Zhang, "Vehicular intelligence in 6G: Networking, communications, and computing," Vehicular Communications, Volume 33, 2022, 100399, ISSN 2214-2096.

[3] L. Shangguan et al., "Dynamic Watermarking for Cybersecurity of Autonomous Vehicles," in IEEE Transactions on Industrial Electronics, vol. 70, no. 11, pp. 11735-11743, Nov. 2023.

[4] S. Ghosh et al., "An Integrated Approach of Threat Analysis for Autonomous Vehicles Perception System," in IEEE Access, vol. 11, pp. 14752-14777, 2023.

[5] A. Slyamkhanov, Z. Bozbayev, A. Alzhanova, A. Pan, T. Ramazan and L. Rzayeva, "Usage of Machine Learning in DDOS Attack Detection", 2023 10th International Conference on Wireless Networks and Mobile Communications (WINCOM), Istanbul, Turkiye, 2023, pp. 1-6.

[6] A. Chattopadhyay et al., "Autonomous Vehicle: Security by Design," in IEEE Transactions on Intelligent Transportation Systems, vol. 22, no. 11, pp. 7015-7029, Nov. 2021.

[7] Z. Abdollahi Biron, S. Dey and P. Pisu, "Real-Time Detection and Estimation of Denial of Service Attack in Connected Vehicle Systems," in IEEE Transactions on Intelligent Transportation Systems, vol. 19, no. 12, pp. 3893-3902, Dec. 2018.

[8] L. Yang, A. Moubayed, I. Hamieh and A. Shami, "Tree-Based Intelligent Intrusion Detection System in Internet of Vehicles," 2019 IEEE Global Communications Conference (GLOBECOM), Waikoloa, HI, USA, 2019, pp. 1-6.

[9] A. Aljuhani, "Machine Learning Approaches for Combating Distributed Denial of Service Attacks in Modern Networking Environments," in IEEE Access, vol. 9, pp. 42236-42264, 2021.

[10] G. Twardokus and H. Rahbari, "Vehicle-to-Nothing? Securing C-V2X Against Protocol-Aware DoS Attacks," IEEE INFOCOM 2022 - IEEE Conference on Computer Communications, London, United Kingdom, 2022, pp. 1629-1638.

[11] S. Anbalagan, G. Raja, S. Gurumoorthy, R. D. Suresh and K. Dev, "IIDS: Intelligent Intrusion Detection System for Sustainable Development in Autonomous Vehicles," in IEEE Transactions on Intelligent Transportation Systems, vol. 24, no. 12, pp. 15866-15875, Dec. 2023.

[12] M. Delwar Hossain, H. Inoue, H. Ochiai, D. Fall and Y. Kadobayashi, "An Effective In-Vehicle CAN Bus Intrusion Detection System Using CNN Deep Learning Approach," GLOBECOM 2020 - 2020 IEEE Global Communications Conference, Taipei, Taiwan, 2020, pp. 1-6.

[13] A. El-Ghamry and M. Elhoseny, "Detecting distributed DoS attacks in autonomous vehicles external environment using machine learning techniques," The 3rd International Conference on Distributed Sensing and Intelligent Systems (ICDSIS 2022), Hybrid Conference, Sharjah, United Arab Emirates, 2022, pp. 292-308.

[14] S. Anbalagan, G. Raja, S. Gurumoorthy, D. S. R and K. Ayyakannu, "Blockchain Assisted Hybrid Intrusion Detection System in Autonomous Vehicles for Industry 5.0," in IEEE Transactions on Consumer Electronics.

[15] T. K. Mohd, S. Majumdar, A. Mathur and A. Y. Javaid, "Simulation and Analysis of DDoS Attack on Connected Autonomous Vehicular Network using OMNET++," 2018 9th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), New York, NY, USA, 2018, pp. 502-508.

[16] Iman Sharafaldin, Arash Habibi Lashkari, Saqib Hakak, and Ali A. Ghorbani, "Developing Realistic Distributed Denial of Service (DDoS) Attack Dataset and Taxonomy," IEEE 53rd International Carnahan Conference on Security Technology, Chennai, India, 2019.