Adaptive Intrusion Detection System for Controller Area Network (CAN) in Automotive Cybersecurity

Aniruddha Jaipurkar
Security Researcher

***Abstract*—The increasing complexity of automotive technology has led to the need for more effective ways to secure the communication between vehicles' controllers. This research paper presents an adaptive Intrusion Detection System (IDS) designed to protect the integrity of Controller Area Network (CAN) networks in modern vehicles. The study evaluates the performance of various machine learning and deep learning algorithms within the IDS framework. The goal is to understand the unique security challenges of CAN networks and develop effective intrusion detection measures. The research methodology involves collecting a representative dataset of CAN network traffic and extracting meaningful features using statistical, time-domain, and frequency-domain analysis. The dataset is then transformed into images for enhanced analysis and visualization. The comparative analysis of machine learning algorithms demonstrates that the CNN+GRU and DL-CNN models outperform others in terms of accuracy and robustness in detecting unauthorized access. These deep learning models excel at capturing complex temporal and pattern dependencies, critical for identifying abnormal activities in CAN networks. The findings contribute to the development of an adaptive IDS specifically tailored for CAN networks, addressing the security concerns of modern automotive systems. The utilization of image-based analysis techniques provides valuable insights into traffic patterns, aiding in effective intrusion detection. By leveraging machine learning algorithms, particularly the CNN+GRU and DL-CNN models, the IDS demonstrates improved performance in terms of accuracy and robustness. The research outcomes pave the way for advanced IDS development, enhancing the overall security of automotive systems.**

***Index Terms*—Intrusion detection system, In-vehicle, CAN, CNN, GRU.**

I. INTRODUCTION

The rapid emergence and evolution of new technologies in the automotive industry have led to significant changes in the way vehicles are built. As a result, the security of the various electronic systems in these vehicles has become a paramount concern[1], [2]. One of the most important factors that can be considered when it comes to protecting the integrity of the automotive cybersecurity ecosystem is the CAN, which is the central hub for communication between different vehicle control units. The CAN protocol, which was initially designed to allow in-vehicle communication between the various systems in a vehicle, has since become a standard in the automotive industry due to its simplicity and robustness. Unfortunately, the complexity and connectivity of modern vehicles have raised concerns about its security. Intruders can potentially take advantage of this to access the data or even harm the occupants[3], [4]. In response to the increasing number of security concerns in the automotive industry, the development of IDS has become a vital component of the cybersecurity framework for the automotive industry. An IDS is an essential part of a network's security framework, which aims to identify and respond to anomalous activities. In the case of the CAN network, an IDS can help prevent unauthorized access to the communication between the various control units[5]. The paper focuses on the development of an adaptive IDS that can be used for the CAN network in cars. It monitors the traffic in the network and analyzes the communication patterns to identify potential threats. The IDS can also respond to and detect anomalous activities in real time to protect the occupants and prevent unauthorized access to the vehicle's systems.
This study explores the use of machine learning methods in the development of an IDS framework. These techniques can learn patterns and identify anomalies in complex data sets, which makes them ideal for detecting intrusions in CAN networks. The study evaluates the performance of various ML-based

techniques, such as the ML-Logistic Regression and the Multilayer Perceptron, on the IDS for the CAN network. It also explores the applications of deep learning and the CNN+GRU in the domain. The research methodology for this study includes the preparation of a representative set of traffic patterns from the CAN network. This data allows the evaluation and training of the various machine learning techniques. In order to extract useful features from the data, various analysis methods, such as frequency- and time- domain analyses, are utilized.

The data collected from the CAN network is then transformed into images in order to improve its analysis and visualize it. This process can provide a more accurate representation of the traffic patterns' spectral and temporal characteristics. Different techniques, such as time-series mapping, heatmap analysis, and spectrograms, are utilized to convert the data into images, which helps in the subsequent evaluation and analysis. It is important to develop an adaptive IDS in order to protect the security of modern vehicles and prevent unauthorized access. This research utilizes image-based methods and learning algorithms to improve the detection of intrusions. The study's findings offer valuable insights into the creation of secure IDS frameworks for combating emerging threats in the automotive sector.

II. LITERATURE REVIEW

The goal of this literature review is to provide a comprehensive analysis of the various studies that are related to the cybersecurity of automotive systems. They cover topics such as intrusion detection and secure communication. The selected papers offer valuable insight into developing effective systems for the smooth operation of self-driving cars. Each study draws on a unique perspective and presents novel strategies to address the challenges of automotive systems. Alheeti et al.[6] explores the utilization of discriminant inference in detecting intrusions in the communication between self-driving cars and their networks. The researchers analyzed traffic patterns to identify potential threats and present the discriminant analysis's effectiveness. Jenefa et al.[7] proposed a secure communication method through an ID-based signature system. This article emphasizes the importance of such protocols in addressing the issue of tamper-proof messaging between vehicles in ad hoc networks.

Lokman et al.[8] reviewed the various intrusion detection techniques for the CAN bus systems of cars. Their findings offer a comprehensive overview of the available techniques and their applications. This review serves as a valuable reference material for researchers in this field. Song et al.[9] present a novel method for detecting in-vehicle network threats through deep learning. They use this approach to develop an efficient and robust method for protecting the car's systems. Serinelli et al[10]. used open-source data to analyze the implementation of IDSs for zero-day and classic attack detection. Their findings support the understanding of the vulnerabilities and performance of IDSs in the networks of cars.

Various authors[11]–[15] covers different topics such as network-based intrusion detection, identity-based authentication schemes, vehicle re-identification, and model-based IDS designs. These provide a variety of perspectives on how to address the security concerns of automotive systems. The studies presented in this literature review provide valuable insight into the various aspects of automotive cybersecurity. They also highlight the use of blockchain and deep learning in system design, as well as the communication between cars and their networks. The findings and their methodologies can help develop secure and robust vehicle systems, which would ensure the safety and longevity of transportation.

III. METHODOLOGY

i. Dataset - The "CAN Intrusion Dataset" from “OCSLab” is a valuable source of information for researchers in the field of detecting and monitoring intrusions into a network's control area[16]. It includes a variety of real-world traffic patterns, including abnormal and normal communication. This dataset allows practitioners and researchers to develop and evaluate techniques for detecting intrusions into an organization's control area using real CAN network scenarios. Its availability provides them with a significant advantage in developing effective and efficient intrusion detection systems for automotive cybersecurity. Sample dataset shown in figure-1

	CAN ID	DATA[0]	DATA[1]	DATA[2]	DATA[3]	DATA[4]	DATA[5]	DATA[6]	DATA[7]	Label
0	1201	41	39	39	35	0	0	0	154	R
1	809	64	187	127	20	17	32	0	20	R
2	1349	216	0	0	136	0	0	0	0	R
3	1201	41	39	39	35	0	0	0	154	R
4	2	0	0	0	0	0	3	2	228	R
...	...	...	...	...	...	...	...	...	...	...
818435	848	5	32	52	104	117	0	0	12	R
818436	1088	255	0	0	0	255	134	9	0	R
818437	848	5	32	100	104	117	0	0	92	R
818438	1349	216	90	0	137	0	0	0	0	R
818439	790	5	33	48	10	33	30	0	111	R

818440 rows × 10 columns

Figure 1 Sample dataset

ii. Pre-processing

a. Drop Labels: In the pre-processing phase, the process of dropping labels involves taking out the target variable from the dataset. This step can be performed when there are features or columns that represent the prediction or outcome target. In some cases, it is necessary to remove the target because it is not relevant to the analysis.

b. Standard Scaler: In data pre-processing, standard scaling is a technique that involves transforming a dataset's features. This process can be beneficial when there are multiple units or scales in the data, as it allows them to be brought to a similar scale. For machine learning systems that are sensitive to the input features' scale, standard scaling can be beneficial. It can help them perform better by making the features comparable and provide better convergence.

iii. Data Transformation

a. Transform all features into the scale of [0, 1]: It is usually necessary to change the features' scale to a specific range in order to maintain consistency throughout the transformation process. Doing so will ensure that the minimum and maximum values of the feature remain consistent.

b. Multiply the feature values by 255 to transform them into the scale of [0, 255]: To prepare the data for presentation, the feature values must be multiplied by 255. This method will change the features' intensities into the range of colors that can be used in image composition. For instance, the minimum intensity of the black features is 0, while the maximum intensity of the white is 255. By doing this, the transformed features can fall within the [0, 255) range.

c. Reshape: In order to convert tabular data into an image, the process of reshaping it involves arranging the various features into a framework that fits the image's dimensions. This can be done in the context of creating a grid, which will depend on the requirements of the image.

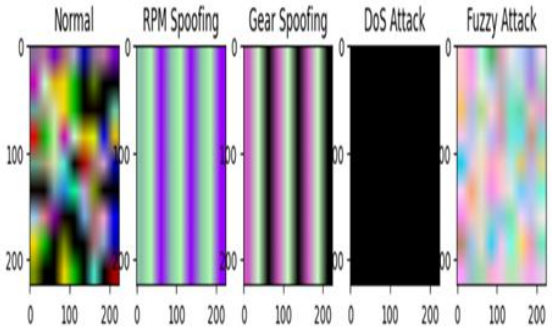

Figure 2 Image after converting from CSV to image file

iv. Load images and labels

a. Convert to Grayscale: When converting images to grayscale, the data they are loaded with are usually only shades of gray. This eliminates the need for additional information in the colors in the original files, and it simplifies the data. This method is particularly useful when the color information in images isn't important for a particular analysis. It can also be utilized with algorithms that are made to operate with grayscale images.

b. Resize the Image: In image resizing, the dimensions of the pictures are adjusted to a specific size so that they can be used in a certain manner. Doing so helps ensure that they are compatible with the dataset and uniformity across all the images. It can also be used to fit pictures into a predefined size. Choosing the right technique for image resizing can help prevent data loss and ensure the preservation of crucial features.

c. Flatten the Image: A flattening process involves converting an image's two-dimensional structure into a single vector. This process can be performed by reshaping the array into a single column or row. It is usually done when the modeling or analysis techniques that need the data

in a flattened format are required. Flattened images can be processed by algorithms that prefer to receive vector input instead of the matrix-like structure. This process is particularly beneficial when working with machine learning systems that need to be equipped with flattened input, like neural networks.

IV. RESULTS AND OUTPUTS

The performance of different machine learning algorithms in detecting intrusions within the Controller Area Network (CAN) system was evaluated as shown in figure-3,4,5. The results showed that the CNN + GRU algorithm achieved the highest accuracy of 99.59%, indicating its superior performance in capturing complex temporal and pattern dependencies. The CNN algorithm followed closely with an accuracy of 95.74%, showcasing its effectiveness in analyzing visual patterns and extracting features. The Logistic Regression algorithm achieved an accuracy of 94.28%, while the Stochastic Gradient Descent and Multi-Layer Perceptron algorithms achieved accuracies of 91.44% and 89.27% respectively. These findings highlight the robustness and effectiveness of deep learning-based approaches, particularly the CNN + GRU model, in intrusion detection for the CAN network, emphasizing the importance of leveraging advanced machine learning techniques for automotive cybersecurity.

a. Machine Learning

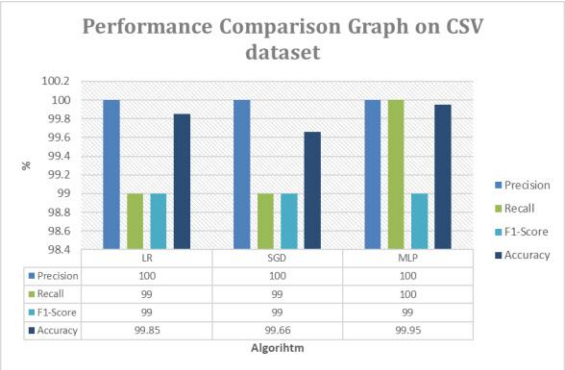

Figure 3 Performance analysis – ML using CSV file

b. Deep Learning

Figure 4 Performance analysis - DL using Image file

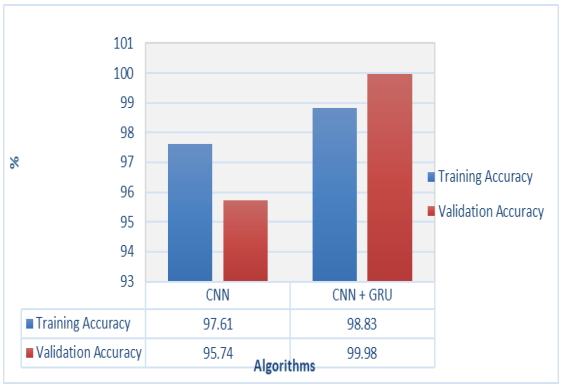

Figure 5 Training and Validation Accuracy Comparison Graph of Deep Learning Algorithms

V. CONCLUSION AND FUTURE SCOPE

The goal of this study was to develop an adaptive IDS for the CAN in automotive cybersecurity. It evaluated various deep learning and machine learning algorithms. The CNN + GRU algorithm was the most accurate and robust in terms of its accuracy and robustness. The findings of the study highlight the importance of utilizing advanced techniques for detecting intrusions in the CAN network, particularly due to its complexity and increasing number of automotive components. Furthermore, it shows the transformation of data into images for enhanced traffic analysis and visualization. The study's future aims include developing an adaptive IDS that can be used to identify and prevent unauthorized access to the CAN network. It will also explore the use of deep learning and machine-learning algorithms to analyze and improve the effectiveness of the detection process. In addition, the use of diverse datasets, such as real-world ones, can help improve the accuracy of the intrusion detection. In addition, the use of anomaly

scoring and other techniques for detecting anomalies may be integrated into the IDS' capabilities. Researchers may also develop countermeasures to enhance the security of the CAN network. The study's findings provide valuable information on how to secure the CAN network's integrity. By utilizing cutting-edge machine learning methods and visual analysis, the researchers were able to attain an important insight into how to prevent unauthorized access to modern vehicles. The study's future directions will allow for the development of more comprehensive and robust IDSs, which can help safeguard the systems of modern cars.

REFERENCES

[1] D. Kosmanos et al., "A novel Intrusion Detection System against spoofing attacks in connected Electric Vehicles," Array, vol. 5, no. October 2019, p. 100013, 2020, doi: 10.1016/j.array.2019.100013.

[2] H.Alqahtani and G. Kumar, "A deep learning-based intrusion detection system for in-vehicle networks," Comput. Electr. Eng., vol. 104, p. 108447, 2022, doi: https://doi.org/10.1016/j.compeleceng.2022.108447.

[3] P.Wei, B. Wang, X. Dai, L. Li, and F. He, "A novel intrusion detection model for the CAN bus packet of in-vehicle network based on attention mechanism and autoencoder," Digit. Commun. Networks, vol. 9, no. 1, pp. 14–21, 2023, doi: 10.1016/j.dcan.2022.04.021.

[4] H. Taslimasa, S. Dadkhah, E. C. P. Neto, P. Xiong, S. Ray, and A. A. Ghorbani, "Security issues in Internet of Vehicles (IoV): A comprehensive survey," Internet of Things, vol. 22, p. 100809, 2023, doi: https://doi.org/10.1016/j.iot.2023.100809.

[5] M. Chen and M. Yan, "How to protect smart and autonomous vehicles from stealth viruses and worms," ISA Trans., 2023, doi: https://doi.org/10.1016/j.isatra.2023.04.019.

[6] K. M. A. Alheeti, A. Gruebler, and K. McDonald-Maier, "Using discriminant analysis to detect intrusions in external communication for self-driving vehicles," Digit. Commun. Networks, vol. 3, no. 3, pp. 180–187, 2017, doi: 10.1016/j.dcan.2017.03.001.

[7] J. Jenefa and E. A. Mary Anita, "Secure Vehicular Communication Using ID Based Signature Scheme," Wirel. Pers. Commun., vol. 98, no. 1, pp. 1383–1411, 2018, doi: 10.1007/s11277-017-4923-7.

[8] S. F. Lokman, A. T. Othman, and M. H. Abu-Bakar, "Intrusion detection system for automotive Controller Area Network (CAN) bus system: a review," Eurasip J. Wirel. Commun. Netw., vol. 2019, no. 1, 2019, doi: 10.1186/s13638-019-1484-3.

[9] H. M. Song, J. Woo, and H. K. Kim, "In-vehicle network intrusion detection using deep convolutional neural network," Veh. Commun., vol. 21, p. 100198, 2020, doi: https://doi.org/10.1016/j.vehcom.2019.100198.

[10] B. M. Serinelli, A. Collen, and N. A. Nijdam, "On the analysis of open source datasets: Validating IDS implementation for well-known and zero day attack detection," Procedia Comput. Sci., vol. 191, pp. 192–199, 2021, doi: 10.1016/j.procs.2021.07.024.

[11] J. Jenefa and E. A. Mary Anita, "An Enhanced Secure Authentication Scheme for Vehicular Ad Hoc Networks Without Pairings," Wirel. Pers. Commun., vol. 106, no. 2, pp. 535–554, 2019, doi: 10.1007/s11277-019-06178-4.

[12] S. Rajasekaran, A. Maria, F. Al-Turjman, C. Altrjman, and L. Mostarda, "ABRIS: Anonymous blockchain based revocable and integrity preservation scheme for vehicle to grid network," Energy Reports, vol. 8, pp. 9331–9343, 2022, doi: 10.1016/j.egyr.2022.07.064.

[13] M.-H. Monzer, K. Beydoun, A. Ghaith, and J.-M. Flaus, “Model-based IDS design for ICSs,” Reliab. Eng. Syst. Saf., vol. 225, p. 108571, 2022, doi: https://doi.org/10.1016/j.ress.2022.108571

[14] Y. Xu, L. Rong, X. Zhou, X. Pan, and X. Liu, “Joint Multiple Fine-grained feature for Vehicle Re-Identification,” Array, vol. 14, no. April, p. 100152, 2022, doi: 10.1016/j.array.2022.100152.

[15] T. Thilagam and R. Aruna, “Intrusion detection for network based cloud computing by custom RC-NN and optimization,” ICT Express, vol. 7, no. 4, pp. 512–520, 2021, doi: 10.1016/j.icte.2021.04.006.

[16] H. K. Kim, “HCRL - Car-Hacking Dataset.” [Online]. Available: https://ocslab.hksecurity.net/Datasets/car-hacking-dataset.