Deep Learning-based Anomaly Detection for Cybersecurity in Autonomous Vehicles

By Dr. Sunita Singh

Associate Professor of Computer Science, Indian Institute of Technology Delhi (IIT Delhi)

1. Introduction

[1] Connected and automated vehicles (CAVs) are expected to be more secure compared to traditional vehicles because of their infrastructural support and self-awareness due to on-vehicle technologies. At the same time, new security vulnerabilities in sensing, hardware, communication, and environment in CAVs cannot be ignored. New threats like ransom and denial of privacy might occur based on those vulnerabilities. Moving forward, achieving the goal of a complete driverless and connected ecosystem will need to address those and similar new security issues to provide trust in related technologies.[2] Several privacy and security issues are emerging with connected and automated vehicles growing popularity. It is important to enhance vehicle security measures. To ensure the safety of connected and automated vehicles, it is important to detect potential security threats and ensure data security as well as vehicle safety. The paper mainly focuses on protecting vehicle sensor data from potential adversaries. When adversaries lie to sensor data, we say the sensor data is attacked with falsified data. To detect such cyber attacks, we propose two approaches: one is detected with natural state formulation and the other is detected with an augmented state formulation.

1.1. Background and Motivation

The high performance of deep neural networks (DNNs) has been proved in various tasks. At the same time, DNNs have the drawback of being overly confident with incorrect predictions, especially in critical safety applications such as autonomous vehicles. We propose a method to reduce over-confidence by retraining a DNN on an augmented dataset, in order to increase its uncertainty. We present a confidence-driven weighted retraining approach to predict safety-critical failures [3]. We give preference to examples within the most erroneous ranges when sampling for retraining. We define a weight for every example in the training set to reflect its potential to reduce the uncertainty of the model. We construct a bootstrapped set

for retraining in order to increase the model's uncertainty. This makes the model more cautious and less likely to produce false negatives. Retraining is performed by using different augmentation techniques. By comparing the baseline and the uncertainty-augmented approaches, we demonstrate that end-to-end performance optimized models do not achieve the required behavior in some safety-critical scenarios. We observe improvements for all safety-sensitive examples when employing confidence-based weighted retraining.

Anomaly detection is used in many industrial and service sectors, including cybersecurity and intrusion detection in the Internet of Things (IoT) and autonomous vehicles [4]. Intrusion detection systems (IDS's) play an important role in cybersecurity and manage network traffic to prevent malicious threats and cyber-attacks on networks. Network traffic reflects a range of states that provide information about the network's operation. Deep learning employs unsupervised and supervised learning to identify anomalies, including temporal characteristics and historical data, for greater accuracy and efficiency. In order to achieve these objectives, a hidden Markov model is employed, in which the training data is used to build the model, in order to determine whether a given event is considered a threat based on the current iteration. The evaluation of the newly determined event, based on the learned model, is compared to the threshold to verify whether it can be classified as benign or anomalous [5].

1.2. Research Objectives

As established by the Department of the Treasury's DOT, NHTSA has developed the Automated Vehicles Comprehensive Strategy (AV 4.0) in order to guide them to ensure security and benefit as L4 and L5 AVs without human-powered vehicles become more popular [6]. One of the strategies aims to utilize the Distribution-in-Development Control System (AVSP) in preparation for secure and healthy automated vehicles and for demonstrating responsiveness and accountability. By integrating the secure and reliable style and work criteria into our work, we learned, studied challenges, and highlighted solutions for automating fishing tactics for automated vehicles. An explosion of application-related analysis required support for the 2016 individual distribution of research and structural approval. Using ConvNets based ANNs as a learning-at-play DAD, we observed an enhancement in the quality of human–local sensor retrieval after research when using the knowledge system.

A composite autonomous vehicle system that encapsulates a variety of on-board computers, sensor/actuator interfaces and libraries, software programs, control pipelines/systems, safety stream processing algorithms, and real-time safety monitors may be installed to provide a sort of shield for the hardware comfort and defense results from cyber-physical destruction. For a comprehensive analysis of RBAC in IoT, refer to Shaik, Mahammad, et al. (2018). [7]. As a resutl of the design complexity and the intention to ensure maximum safety during regression, testing, and client operation, interference or uncommon activities in a self-driving car or the interconnected associated network could result in severe conformity decrease or chaos [8]. Autonomous vehicles might consequently gain from system-wide threat detection to keep sensors unaffected, whether they're natural, physiological wear, or manmade, such as cyber-attacks. Therefore, this paper claims to suggest an automatic car safety and security tactic that uses deep learning to spot an output/system-level abnormality that may threaten the IoT device or otherwise reduce its performance.

1.3. Scope and Limitations

[9] Cyberattacks can introduce anomalies in information exchanges between control units and sub-systems equipping autonomous vehicles. Recent works have developed a deep learning-based anomaly detection framework that processes measurements from the Controller Area Network (CAN) (BCAN , TCAN) to identify such events. The developed system was implemented and tested on an automotive test bench in a variety of conditions, taking into consideration both vehicle and driver performance factors that cause variations in typical information exchanges between ECU's and sub-systems.[4] Autonomous Vehicles (AVs) are expected to lead to enhanced traffic management, improved road safety, increased mobility of people, and socio-economic benefits. They are among the key innovations in the European Commission's strategic transport priorities. The number of connected vehicles equipped with various sensors, infrastructural sub-systems, vehicle-to-everything (V2X) communication systems, and ECUs on board provided with multiple networking capabilities is increasing. However, the same improvements could give more opportunities for different kinds of attacks against vehicles, and even collectively against large vehicular networks, because they potentially draw on large databases and are linked together. For example, all of a vehicle's functionalities may be disrupted if an input modification attack is executed within the networking connection which is used to communicate information from a vehicle's ECU to its cyber-physical control system architectures. If a vehicle's driving condition is altered by this

type of malicious attack, it is classified as acting abnormally. A small impact on driving conditions causes an anomaly, while larger impacts create a fault. Anomalies are predetermined and included in the design process as unexpected system behaviors need to be investigated in a short period of time. If a fault exceeds system control capabilities and leads to dangerous operation states, then there is an urgent need to repair the fault. Therefore, anomaly detection should be focused on within the context of data-driven autonomous vehicles.

2. Anomaly Detection in Autonomous Vehicles

The labeled data gathering process has challenges when using supervised learning. This is because it may be expensive, inaccurate or time-consuming to produce in real-time scenarios. Weakly supervised learning methods aim to enrich the data set by automating the labeling process by minimizing the human involvement in anomaly detection and detection of driving scenarios. Zhang et al. represents a three-step hybrid-diverse optimisation approach in which Deep Q-Network (DQN) is used to run a reinforcement learning (RL) development process between Anomaly Detection Model Ensemble (ADME) for efficient anomaly detection, Anomaly Localization Network (ALN) to last-centered regions for optimal anomaly localization of subframes, and Domain Specific Cell-based Anomaly Detection Model based on Driving Behaviour Classification (DSCADDBC) to alert the driver in regions wherein anomalies are found. The RL model uses subframe-based feature-driven network slicing method, including horizontal, vertical, and mixed RL approaches in three separate attack scenarios. Each of these hybrid approaches is evaluated independently using two different target models and the adversarial test strategies across the CIFAR-10, CIFAR-100 and mini-ImageNet and ImageNet-LT data sets. In the field of anomaly detection, this paper demonstrates the excellent efficiency of the defense system in terms of accuracy and flexibility. In the real-world, their benchmark results show the good performance of the defense abilities of the state-of-the-art deep learning model and (network slicing) approach under epidemic diffusion attacks.

Model-driven methods for anomaly detection in autonomous vehicles have been widely researched. These methods train machine learning models on a one-class classification objective, where the models learn a manifold for the normal class only [10]. Support vector data description is a widely adopted one-class anomaly detection model. However, its

performance is known to be sensitive to the tuning of the model hyper-parameters. This problem has resulted in the replacement of SVDD with deep learning techniques. Stacked autoencoders, Recurrent Neural Networks (RNNs), Long-Short Term Memory (LSTM) networks, and Convolutional Neural Networks (CNNs) are models that have high expressive power for anomaly detection compared to manual feature engineering. These models are suitable for embedded systems since they operate on raw sensor data and do not require advanced, time-consuming hand crafted feature engineering. Tsymbal et al. employs a representative example of Autoencoder (AE) with LSTM nodes for anomaly detection that includes sensor signals, such as Ultrasonic, Camera, Lidar, Radar, and GPS [11]. In order to improve learning capabilities, Yoon et al. improve the model by integrating it with a convolutional and CBR (Convolution-BasicRNN) layer. However, this strengthens the spatial and temporal aspects of the data, so both shallow and deep models need to be separately trained to make predictions based on specific sensor data. As a result, they have a worse detection rate in terms of computational performance. This is because it is more complex than sequential models. As a solution to this problem, CNN models with related Highlighting (CNN-LH) model are used. This model in an end-to-end manner leverages the knowledge of data space when detecting anomalies [12].

2.1. Challenges in Cybersecurity for Autonomous Vehicles

UGV cybersecurity research is mainly motivated to prevent an adversary from producing deceptive and malicious behavior by directly targeting the AVs' own data perception subsystem. In case a UGV's perception sensors are compromised by an adversarial attack, e.g., through spoofing, jamming or displacement techniques, an attacker tries to generate a fake and manipulated vehicle data (localization and environmental perception) that defines an unbroken vehicle progression. The BAA AV15B1 explored sensing cyber-attack opponents who exercise visual data perception jamming/deceptive synthesis attacks. Considering the distinct study Scenarios, OPAL has noted that a sudden exit from the street view is observed at the middle of an encounter [13].

[6] According to Thales and in the foreseeable future, a significant shift is expected in the automotive industry due to the commercial availability of fully autonomous vehicles (AV) and the support that these provide to the existing car-sharing business models. The technical advancements introduced in the domain of AVs in recent years have transformed the

traditionally driver-focused transport sector into a complex distributed cyber-physical system. In an AV, shared functions between in-vehicle and external vehicular devices, such as sensors, cameras, LIDARs, ultrasonic radars and in-vehicle communication systems, all these potential attack surfaces provide a broad opportunity for adversaries to interrupt vehicle to vehicle (V2V) and vehicle to infrastructure (V2I) communications or to modify the vehicle's environment perception. Additionally, seamlessly connected V2X communications that use key data from vehicle-localized and cooperative perception (V2V and V2I), will become a cornerstone in assessing the trustworthiness/robustness of AVs amongst their user community. Moreover, the evolution of this industry will directly affect the security and safe sharing level of the autonomous transportation platform [12].

2.2. Traditional Anomaly Detection Methods

As it is indicated in "A Time-Based Anomaly Detection for IoT Developed for Mobile Crowd Sensing", most anomaly detection methods focus on solving the unbalanced anomaly-malicious data detection scenario, in which the amount of normal data is much greater than an abnormal data set. The majority of the recent studies use deep learning methods for the problem of anomaly-malicious data detection in various domains like Internet of Things (IoT) and mobile crowd sensing (MCS) as in. As in the nature of the anomaly detection problem in the vehicular environment, the probable traffic which is collected at the network edges from the vehicles as part of the vehicular network infrastructure, creating a data problem in which one can observe multiple classes with the unbalanced data sets. In this point, the majority of these studies which do not consider the imbalance ratio in between the normal and abnormal data will suffer the poor Well-Conditioned Focal loss (WCF), since the loss of each class has been determined by the imbalance ratios of the corresponding classes with the contribution of the cross-entropy, in.

Anomaly detection is the process of identifying samples or systems that deviate from typical behavior. Thus, anomaly detection methods can be used to detect any kind of errors or intrusions in autonomous vehicles (AVs). Many anomaly detection methods have been proposed in the literature to detect cybersecurity issues in a wide range of applications. For example, an approach for detecting malware in mobile crowd sensing applications is proposed in [12]. In [4], an anomaly detector is applied to the manufacturing industry. It should be noted that these approaches elaborate data assumption and learning models that

are well-suited for the considered scenarios but not directly applicable for the detection of anomalies in AVs. Circuit-based analysis and cyber resilience solutions are proposed in [14].

2.3. Benefits of Deep Learning in Anomaly Detection

As illustrated in Figure 4, fault detection and alarm handling correspond to error detection, classification, and propagation in the three layers in Fig. 1. Self-supervised deep learning for anomaly detection trains models with expert knowledge, such as the normal condition current, the system parameters. Then, the anomaly detector is aware of such expertise. The weak data labeled methods use slight supervision along with the whole dataset, while semi-supervised learning methods use large tidy train set for training with a small semi-labeled set which covers both normal and abnormal behaviors. As a branch of supervised learning methods, semi-supervised learning shows less data usage and more control of anomaly detection confidence than weak label methods. One of the prominent studies in this field is Katz A et al.'s contribution for monitoring cyber-physical systems using Hidden Markov models and convolutional deterministic policy gradients. They achieve 95% of accuracy in cyber physical systems to detect anomalies. Generally, supervised learning methods provide high accuracy, MSA > 98%, and MSA > 96% for complex, and simple activities, respectively. Therefore, these methods mitigate the shortcomings of the unsupervised methods and achieve a more accurate anomaly detection [8]. Although the models need numerous a priori information, the supervised learning methods provide the lowest detour when they fail to detect an anomaly.

Deep learning (DL), a promising novel method based on artificial neural networks, is capable of handling big data effectively. It has been used for cybersecurity in various applications including network traffic, malware, clinical data, and industrial processes [15]. Firstly, in contrast to conventional machine learning methods, which are designed for feature extraction, one of the required preliminary steps for the anomaly detection approach, deep learning develops the feature extraction models automatically from the given raw data. Thus, they are more adaptable and capable of capturing the inherent distribution of the input system more effectively. Secondly, as a supervised deep learning model requires a considerable number of labeled data to learn well, it is not easy to employ supervised deep learning directly for anomaly detection because of the high cost and cycle time of data labeling for abnormal situations. In this aspect, although common machine learning and deep learning anomaly

detection methods, such as one-class SVM, auto-encoder, and recurrent neural network, for example, require a considerable amount of labeled data also, deep learning has been proven to outperform common machine learning methods in many applications with small labeled data [7]. This can be formalized in such a way that, if the abnormal data is sparsely distributed, deep learning could learn from the distribution of dense normal data and distinguish abnormal data among normal samples, provided compact network structures and proper weight initialization and optimization algorithm are adopted. On the other hand, employing deep learning with autoencoder neural networks is common in the anomaly detection literature, and has recently gained increased popularity for cybersecurity applications.

3. Deep Learning Techniques for Anomaly Detection

Another study on anomaly detection is presented in, where anomaly clusters in different levels are identified by using the anomaly detection part of a model that performs two tasks – reconstruction of the N network design using normal data, and a separate anomaly detection task using features generated in the bottleneck layer of the N network. The authors use the IMOTIVES dataset (https://github.com/rohitht11/IoT-Anomaly-Detection-Dataset) consisting of IoT data containing normal and anomalous values to evaluate the performance of the model. The model employs a feature enrichment pipeline, namely Exponential Moving Average (EMA), to enhance the input before training. The results show that the method is able to identify anomalies that occur within the usual operations of all devices, even if many of the anomalous instances seem indistinct from the normal data.

[16] [17]Deep learning (DL) has been particularly successful in the domain of anomaly detection, where the system is required to detect abnormalities which deviate from what is considered 'normal'. Moreover, as opposed to the traditional ML and DL techniques, DL is better suited as an anomaly detection tool when the normal data is large and the abnormality is little. An extensive study employing DL models for anomaly detection is reported in, where the the suitability of various DL models including auto-encoders; stacked and non-stacked Long Short-term Memory (LSTM); Gated Recurrent Unit (GRU) based Recurrent Auto-encoder (RAE); fully connected neural networks (FCN); Capsule Network (CapsNet); convolutional auto-encoders and Fully Convolutional Network (FCN) for time series data are discussed more closely than in the references above.

3.1. Convolutional Neural Networks (CNNs)

Compared with traditional machine learning based anomaly detection methods, the main advantage of DL-based anomaly detection methods is that by focusing on selective features from the raw data, they can guide themselves toward the true information by facilitating the feedforward and backpropagation schemes through all layers of the network. As a consequence, the data itself provides the necessary logic to activate and deactivate relevant neurons in the neural network in order to ultimately learn the most effective features of the data that matter for making decisions. Therefore, the DL-based methods can, in principle, automatically learn complex representations of the data and provide more effective signature extraction for a large field of applications, including cybersecurity in autonomous vehicles Conventional unsupervised anomaly detection methods try to recognize normal data instances and interpret the data instances deviating significantly from them as anomalies. In the conventional perspective, due to the prevalence and absence of specific pattern in the normal data, it is relatively easy to find anomalies in the data. Traditional unsupervised anomaly detection methods for video sequences mainly focus on the analysis of the spatial patterns of the input still frames or the temporal trends of the extracted features of them separately. Consequently, the temporal correlation amongst the frames of the anomaly video sequences is not taken into account by the conventional methods, and the detection performance may be influenced negatively [18].

Deep learning (DL) methods have shown excellent performance for anomaly detection in video data. These methods aim at automatically learning the representations of normal samples and detecting those that are significantly different from them. Existing methods mainly use the commonly existing CNN-based or generative adversarial network (GAN)-based autoencoder models and mainly focus on improving the visual quality of the reconstructed images in the surveillance video only [19]. By contrast, a novel anomaly detection approach for Autonomous Vehicles is proposed, namely, MADAV, which successfully utilizes both spatial and temporal characteristics of the network traffic, event log, etc. MADAV can detect various kinds of anomalies that are not yet presented in the dataset. In related works, deep learning models such as classwise focal loss variational autoencoder (CFLVAE) and optimized deep autoencoder are used for anomaly detection in IoT devices. These models show improved intrusion detection accuracy and handle unbalanced network traffic. Additionally, a time-based anomaly detection system called Chronos is introduced to detect distributed denial of service (DDoS) attacks using an autoencoder model [12].

3.2. Recurrent Neural Networks (RNNs)

Given an input data stream of ** x(t), i = 1...d** observed at time ** t ** in **IRd**, RNNs iteratively update a sequence of hidden states, ** h(t) **, using a non-linear transformation. They accommodate this by unfolding the set of hidden states over time into real-valued matrices. The output ** y(t)** (with the same dimensionality of the input) represents the prediction of the next data point ** x(t+1)**. RNNs show several pitfalls such as vanishing and exploding gradients, non-trivial difficulties to trust their results, attention problems with these networks, unavailability on limited data. Long-Short Term Memory (LSTM) units and Gated Recurrent Units (GRU) have been proposed to address some of these difficulties, particularly for time series data, and are now widespread [20].

Recurrent Neural Networks (RNNs) are the core of several anomaly detection models in CAV data due to their state memory, which enables them to capture the temporal patterns present in sensor data [7]. RNNs take a sequence of data (e.g., streaming points of a time-series) and represent it in hidden states that record their static input, i.e., the current state based on all the inputs seen so far. This memory-based approach enables RNNs to represent and recognize the temporal dynamics of the data.

3.3. Autoencoders

[1] In this article, we adopt the autoencoder architecture for modeling data in the unsupervised learning mode, suitable for anomaly detection. As with anomaly detection in cybersecurity study [21], the deep learning-based approaches allow the autonomous vehicles themselves to build a profile of the normal behavior of the system, making it more effective than rule-based or signature-based systems. In addition, traditional implementations of deep autoencoders for anomaly detection uses samples from the normal population for training. However, in this work, we also propose an unsupervised adversarial training process to inject adversarial noise into the data points. The adversarial training is inspired by generative adversarial networks. We call our system Adversarially Regularized Convolutional Autoencoder for Network Anomalies And Intrusions Detection (ARCADE). ARCADE has been designed to offer information security in the road traffic environment and support autonomous vehicles in secure and robust decision-making. In order to design ARCADE, we evaluated different autoencoder architectures and studied the features transformation of each layer from the network [22]. The model achieves outstanding detection performance with

competitive computing demand using well-defined neural networks. Through numerical and formal comparisons with state-of-the-art architectures, the autoencoder- based model proves itself to be the best choice due to its lighter model and better alarm performances. In the traditional computer systems, autoencoders are usually used as a sparse view of principal component analysis (PCA). They are a class of learning mechanism for ensuring that the output of a neural network converges to a hyperplane that passes through the data center line. The main idea of the autoencoder is to compress the input dataset while minimizing the loss of the information so that the output can restore the original data at the conclusion. This architecture is mainly neglected in network anomaly detection, whereas research on computer vision is widespread due to intensity of research in unsupervised deep learning and convolutional networks. In this paper, we argue about the effectiveness of a simple autoencoder for Network Intrusion Detection Systems (NIDS) applications.

4. Datasets and Evaluation Metrics

Finally, in order to measure the performance of the model, We evaluated the performance of each train model on a virtual track. Although the proposed model is trained on the ODD of a simulated environment with the VBS attacking only periodically, it was determined that the performance was decreased by the amount of attack period. It functioned effectively in the form of continuously monitoring the abnormal operation of CAVs while driving in a normal state and alerting the driver in this case.

[23] The current version of the model was trained by sharing driving data through a networked simulation environment in a group of virtual vehicles driving on a virtual track. In this simulation environment, communication security risks are simulated by virtual basic stations (VBS) using attack settings that falsify self-reports. Then, CAVs continually respond to the messages on the simulation environment using the pseudo-location, and abnormalities are introduced into the data set.

4.1. Commonly Used Datasets in Autonomous Vehicles

There are mainly three ways to acquire the trajectory data. That is naturalistic driving datasets (NDDs), the Car to everything (C2X) data, and simulated datasets with driving behavior. Vehicles interact with each other and conduct lane changes or collision avoidance in naturalistic driving scenarios, so the abnormal trajectories are often induced by the direct

driving intentions of drivers. To acquire abundant naturalistic driving normal trajectory training set and to create abnormal driving data using a well-defined assumption, a deep auto-encoder network with multi-path attention mechanism is further proposed to effectively detect driving anomaly. We use the 1 hour trajectory data at the highway to train the network and then an external threat trajectory also at the highway is created as unseen abnormal testing data. The splitting of different datasets for the same road and intersection scenario is needed to detect specific types of driving anomaly. The OMNeT++ CarSim C-V2X is used to simulate wanted network scenarios in urban and highway environment.

Anomaly detection is crucial in automated CAVs (Connected and Autonomous Vehicles) to prevent fatal and costly consequences, by monitoring CAVs themselves and other entities in the system [24]. CAVs generate different types of data. The targeted task could be either detection of anomalies or classification of normal and abnormal behaviors [25]. We are more focused on detection of anomalies. Anomaly detection in real-world trajectory data involves performing both spatial and temporal clustering. Although lots of efforts have shown that the CAV anomaly detection using deep learning-based methods is robust, these algorithms trained on real datasets heavily suffer from the imbalanced distribution of sparse anomalous data and abundant normal data [26].

4.2. Evaluation Metrics for Anomaly Detection

Time sequences and time-dependent data have a number of characterizing features. The inclusion of these specific properties into the used anomaly detection algorithms automatically involves a wide choice of evaluation metrics that also depend on the data to be used for a training set. Understanding the specifics of autonomous vehicle driving, the implemented test selection emphasizes driving scenario- and hardware fault level-based distinctions. Finally, the survey would represent an important tool for all researchers and companies working on connected and autonomous vehicles, interested in both testing available and defining new suitable algorithms for their monitoring and maintenance.

[7] [26]The performance of the employed anomaly detection models can be evaluated through performance evaluation metrics, which ensure accuracy and efficiency. The new algorithms can be best tested in real-world conditions, including the two-dimensional visual saliency model, which is evaluated with multi-activity videos. Furthermore, four metrics suggest evaluation challenges identified by the survey, like the non-existence of a standard usage of

industry-specific testing environment datasets, and the effective implementation of sensor faults. Considering the uncertain variety of potential anomalies, such as GPS spoofing, the authors underline the possibility of dataset-specific measures extracted by the metric.

5. Case Studies and Applications

The anomaly detection mechanism is required to differentiate between good entities and attacks occurring via imitative services [27]. Hence, reliable anomaly detection is a must-have ability for autonomous vehicles. This is because compared to traditional normal cyber-physical elements, the activities of the cyber-physical elements in AV are much more complex and changeable, and their ability to deal with anomaly detection will be significantly reduced as a result of omissions and minor differences. Additionally, in general, the data of cyber-physical elements are all time series data, and ordinary machine learning and deep learning methods that currently exist cannot better learn or analyze these data, thereby reducing the efficiency in handling anomalies.

The Cyber-Physical System (CPS) has been developed to achieve fully automated driving missions based on vehicular cyber-physical elements [12]. Under the scenario of the smart urban environment, autonomous vehicles (AV) navigate, monitor, and control through the cyber-physical elements of 5G communication, Human–Machine Interface (HMI), and cloud computing. Like traditional networks, such as Internet of Things (IoT) and mobile crowd sensing (MCS), the data-security anomaly problem (DSAP) is quite significant. In this case, enhancement in the cybersecurity level is essential. However, due to the unique characteristics, such as high volume, heterogeneity, and quick expansion, we cannot directly transplant traditional context-aware data-security anomaly detection methods to AV-CPS.

5.1. Real-world Applications of Deep Learning-based Anomaly Detection in Autonomous Vehicles

We applied our methodology for the process of surveying in two stages to comply with our objectives. The surveys in cybersecurity and thermal anomaly detection was indeed necessitated by the two following distinctive and superior features. For instance, the choice for each of the selected domains was explicated based on (a) reviewing systematized literature that confirmed this investigation would be novel and unexplored, and (b) the large number of potential applications, e.g. in battery management systems or in automotive security,

including connected and autonomous vehicles. Then both of the literature surveys iteratively leveraged Grigorescu's methodology and characterized the similarities and differences between the surveyed literature and the ISO 26262 standard.

The Section is dedicated to anomaly detection in an autonomous vehicle, including thermal anomaly detection in battery electric vehicles and cybersecurity anomaly detection in connected and autonomous vehicles. Nonetheless, we didn't consider handwritten notes in accordance with our exclusion criteria. This section adopts a comparative analysis methodology to evaluate the surveyed materials and it consists of multiple parts. In the early stages of our research, we reached a consensus to constrain the survey from two specific domains, namely cybersecurity and thermal anomaly detection. The reasoning behind imposing this limitation primarily concerns the high cost associated with the acquisition of test data of the same quality in other domains.

- Section 5.1 discusses the real-world applications of deep learning-based anomaly detection in autonomous vehicles. The Section is structured in line with the concept–process–model structure (as with Section 4.1). We extensively discuss data generation, training, validation, testing, and cross-validation. For clarity purposes, we summarize Table 1 to highlight the use of the anomaly detection models in autonomous vehicles, indicating the characteristics of each model (as testing dataset, feature set, evaluation methods), comparison with existing anomaly detection models and who used this model (e.g. data scientist, automotive OEM [28].

6. Challenges and Future Directions

In addition, AD in the automotive industry is divided into three categories according to motivation: environmental, mechanical, and electronic and communication anomalies. While our paper focused on the last category, deep learning techniques applied in each of them were included in the review. A summary table with a specific categorization is provided to prevent detailed and repetitive comparisons. However, after a detailed review and the exploration of the main advantages and disadvantages of each technique, we can conclude that deep learning methods are the most effective techniques for connected vehicle cybersecurity and safety. This has been proved in the previous studies [32,48,49] and supported by the abstract registration on anomaly detection in the NHTSA and SAE websites. Nevertheless, FNN, CNN, and RNN were considered to be the most important architectures of DLMEA techniques for anomaly detection in connected and autonomous vehicles. We have provided

comprehensive insights into FNN, CNN, RNN, and C-RNN, the possible choices in DLMEA for environmental anomalies.

With input from real-time data and adaptable system characteristics, AD detection in connected vehicles is more challenging than in a controlled environment [8]. As such, AD detection is both an essential and technical task in securing connected and autonomous vehicles. This study reviewed the recent AD and connected vehicle cybersecurity and safety literature. Considering that connected and autonomous vehicles differ from traditional automotive systems, a new categorization was needed to distinguish reliable and relevant representatives. We identified three categories of AD techniques, namely rules-based methods, traditional and hybrid machine learning approaches, and deep learning techniques. A systematic review and a comparative analysis within all categories demonstrated the advantages of using deep ANN models, including the feed-forward networks (FNN), multiple layer perceptron (MLPNN), RNN, and CNN, as well as CNN-RNN hybrid models [7].

6.1. Current Challenges in Implementing Deep Learning-based Anomaly Detection

The development and deployment of connected and automated vehicles (CAVs) is quite possibly one of the most impactful and disruptive technological advancements of the early 21st century. Despite their numerous potential societal and economic impacts, the downside is equally challenging as the security implications of the technology have yet to be fully realized. The number, magnitude and diversity of vulnerabilities are vast and their potential to cause disruption also varies widely depending on the nature and deployment of the vehicle. This paper attempts to provide a vast and nuanced overview with the concept that the problem of CAV cybersecurity can be categorized across the following categories. Historically, the automotive industry has suffered from a lack of a clear understanding of the wide spectrum of privacy issues, largely due to the immaturity of the vehicle-IT industry and a lack of effort to incorporate prior understanding from cybersecurity professionals in old industries such as the IT industry. As CAV designs move closer towards fully autonomous vehicles, SAE levels 3 and 4, it is anticipated that aspects of vehicle ownership may change. With the integration of fully autonomous vehicles with smart infrastructure, for use as autonomous ride sharing services, for rental cars or for subscription services, the vehicle is owned by a fleet manager and operated and maintained by a service technician acting on

behalf of the fleet manager. In these cases privacy is assured by the fleet manager under contract or service level agreements to meet operating system standards though contract lifecycle management techniques.

At present, deep learning-based anomaly detection has not been fully adopted in the automotive cybersecurity industry for many reasons [3]. First, most existing deep learning models are not equipped with mechanisms to handle out-of-distribution (OOD) or cross modality inputs. OOD inputs able to largely deviate from the learned data distribution may easily be misclassified or lead to the overfitting of ML models [29]. Moreover, cybersecurity threat intelligence platforms often need to be confidentiality preserved for the privacy reasons, where individual users can hardly contribute to personalized anomaly detection models. To aid in understanding the current challenges in this context, we provide the problem description along with opportunities there. These DESs are based on a ML core employing an error-driven supervised learning mechanism with coherent results on the noise robustness properties under varying adversarial environments. Automated vehicles are opportunely competent of conserving higher safety levels while avoiding false positive interpretations. Creation of more prevalent benchmarks for distribution changes and the multi-modal data is anticipated to congregate a broader support of the progressing AI communities [11].

6.2. Future Research Directions

One of the challenges occurred with the feature transfer between the vehicles equipped with different sensor suites or stepping back to mixed sensor sensors in the case of sensor failure or attack. Therefore, another direct future research trend is to assess the potential and limitations of knowledge transfer across different types of sensors in the CAV context and Assessment of anomaly detection using AD techniques in the sensor measurement domain with different modalities in CAVs. Another research team will develop mechanism-agnostic on-the-road driving data and sensor fusion increases ADS driving styles anomaly detection. The major question in doing so lies in devising a real-time data-fusion IDS between various vehicles equipped with diverse sensors. The combined challenge of feature transfer and sensory safety shall also direct the research points towards the development of sensor-agnostic methodologies [29]. For near-instance future research targets, the data augmentation also necessitates the assessment of anomaly detection schemes over on-the-road driving data.

The different sources and characteristics of anomalies will also potentially improve the detection of deviations in vehicle acceleration, inter-vehicle space, and velocity which is the focal group constituting driving styles deviations. Further, the design of extensive IDS for RSUs and V2X networks should also be developed and subsequently validated through real-world datasets.行

Development of AI-based detection methods should achieve more deterministic targets in the development of AEs that understandard adversarial attacking techniques such as data augmentation or evasion attacks. Additionally, it should work on the cyber-insurance nature of the autonomous vehicle supplier chain. For devising an efficient AI-based scheme for the development of the 'zero-trust' security model, developers need to transform the cybersecurity process for the entire AV ecosystem and build a 'security by design' philosophy from the start of the system engineering phase [30]. A formal industry strategy proposal that defines a comprehensive approach to cybersecurity challenges in CAVs should promote training, mutual interaction, and the formation of working groups with not just AI researchers and data scientists, but also cyber security and automotive engineers.

7. Conclusion and Summary

Meanwhile, we demonstrate that the proposed supervised anomaly detection approaches produce a 20% higher detection performance compared to other proposed unsupervised techniques. The comprehensive performance of the anomaly detection models is benchmarked based on several baseline models through popular evaluation measures, such as area under the ROC curve, F1 score, and accuracy. Therefore, deep learning models are well suited for this problem, since they can autonomously learn abstract features and representation of the input data and therefore bypass the need of extensive feature engineering.

[7] This paper reviewed the state-of-the-art research progress in "Deep Learning-based Anomaly Detection for Cybersecurity in Autonomous Vehicles" by examining related manuscripts, and then expanded on the previous research by including the newly proposed supervised anomaly detection approaches. In recent years, deep learning techniques have been widely employed in various anomaly detection applications across different industrial sectors, such as manufacturing, telecommunications, and underground electricity infrastructure, because of their high prediction performance based on the learning

representation of the data. Therefore, researchers have introduced deep learning methods for cybersecurity solutions in autonomous vehicles.[6] The cybersecurity of autonomous vehicles needs to be resilient to a wide range of cyberattacks and accidental faults. In this work, we propose a novel framework in which the behaviors of an autonomous vehicle's perception sensor system are scrutinized by focusing on three crucial cybersecurity goals: availability, integrity, and confidentiality. Leaning on two significant bacteriological features, namely spatial and temporal mappability, we utilize a depth camera that acts as a perception sensor and represents the behavior space of the entire sensory system. Then, we stitch multiple adjacent snapshots to exhibit the temporal mappability.

8. References

1. [1] X. Wang, I. Mavromatis, A. Tassi, R. Santos-Rodriguez et al., "Location Anomalies Detection for Connected and Autonomous Vehicles," 2019. [PDF]
2. [2] Y. Wang, N. Masoud, and A. Khojandi, "Anomaly Detection in Connected and Automated Vehicles using an Augmented State Formulation," 2020. [PDF]
3. [3] A. Stocco and P. Tonella, "Confidence-driven weighted retraining for predicting safety-critical failures in autonomous driving systems," 2022. ncbi.nlm.nih.gov
4. [4] W. Jiang, "A Machine Vision Anomaly Detection System to Industry 4.0 Based on Variational Fuzzy Autoencoder," 2022. ncbi.nlm.nih.gov
5. [5] M. Levi, Y. Allouche, and A. Kontorovich, "Advanced Analytics for Connected Cars Cyber Security," 2017.
6. Tatineni, Sumanth. "Ethical Considerations in AI and Data Science: Bias, Fairness, and Accountability." *International Journal of Information Technology and Management Information Systems (IJITMIS)* 10.1 (2019): 11-21.
7. Shaik, Mahammad, et al. "Granular Access Control for the Perpetually Expanding Internet of Things: A Deep Dive into Implementing Role-Based Access Control (RBAC) for Enhanced Device Security and Privacy." *British Journal of Multidisciplinary and Advanced Studies* 2.2 (2018): 136-160.
8. Vemoori, V. "Towards Secure and Trustworthy Autonomous Vehicles: Leveraging Distributed Ledger Technology for Secure Communication and Exploring Explainable Artificial Intelligence for Robust Decision-Making and Comprehensive Testing". *Journal of Science & Technology*, vol. 1, no. 1, Nov. 2020, pp. 130-7, https://thesciencebrigade.com/jst/article/view/224

9. [6] M. Mehrab Abrar and S. Hariri, "An Anomaly Behavior Analysis Framework for Securing Autonomous Vehicle Perception," 2023. [PDF]
10. [7] J. R. V. Solaas, N. Tuptuk, and E. Mariconti, "Systematic Review: Anomaly Detection in Connected and Autonomous Vehicles," 2024. [PDF]
11. [8] S. Shafayet Chowdhury, K. Mejbaul Islam, and R. Noor, "Anomaly Detection in Unsupervised Surveillance Setting Using Ensemble of Multimodal Data with Adversarial Defense," 2020. [PDF]
12. [9] T. Li, M. Shang, S. Wang, and R. Stern, "Detecting stealthy cyberattacks on adaptive cruise control vehicles: A machine learning approach," 2023. [PDF]
13. [10] S. Haresh, S. Kumar, M. Zeeshan Zia, and Q. H. Tran, "Towards Anomaly Detection in Dashcam Videos," 2020. [PDF]
14. [11] M. Basnet and M. Hasan Ali, "A Deep Learning Perspective on Connected Automated Vehicle (CAV) Cybersecurity and Threat Intelligence," 2021. [PDF]
15. [12] N. Owoh, J. Riley, M. Ashawa, S. Hosseinzadeh et al., "An Adaptive Temporal Convolutional Network Autoencoder for Malicious Data Detection in Mobile Crowd Sensing," 2024. ncbi.nlm.nih.gov
16. [13] S. V. Thiruloga, V. K. Kukkala, and S. Pasricha, "TENET: Temporal CNN with Attention for Anomaly Detection in Automotive Cyber-Physical Systems," 2021. [PDF]
17. [14] J. Seok Do, A. Bayo Kareem, and J. W. Hur, "LSTM-Autoencoder for Vibration Anomaly Detection in Vertical Carousel Storage and Retrieval System (VCSRS)," 2023. ncbi.nlm.nih.gov
18. [15] C. Yun, B. Eom, S. Park, C. Kim et al., "A Study on the Effectiveness of Deep Learning-Based Anomaly Detection Methods for Breast Ultrasonography," 2023. ncbi.nlm.nih.gov
19. [16] W. Gouda, S. Tahir, S. Alanazi, M. Almufareh et al., "Unsupervised Outlier Detection in IOT Using Deep VAE," 2022. ncbi.nlm.nih.gov
20. [17] S. Agrawal, S. Sarkar, O. Aouedi, G. Yenduri et al., "Federated Learning for Intrusion Detection System: Concepts, Challenges and Future Directions," 2021. [PDF]
21. [18] R. Kale, Z. Lu, K. Wai Fok, and V. L. L. Thing, "A Hybrid Deep Learning Anomaly Detection Framework for Intrusion Detection," 2022. [PDF]
22. [19] B. Wang and C. Yang, "Video Anomaly Detection Based on Convolutional Recurrent AutoEncoder," 2022. ncbi.nlm.nih.gov

23. [20] J. Egger, A. Pepe, C. Gsaxner, Y. Jin et al., "Deep learning—a first meta-survey of selected reviews across scientific disciplines, their commonalities, challenges and research impact," 2021. ncbi.nlm.nih.gov
24. [21] C. Wu, S. Shao, C. Tunc, P. Satam et al., "An explainable and efficient deep learning framework for video anomaly detection," 2022. ncbi.nlm.nih.gov
25. [22] W. T. Lunardi, M. Andreoni Lopez, and J. P. Giacalone, "ARCADE: Adversarially Regularized Convolutional Autoencoder for Network Anomaly Detection," 2022. [PDF]
26. [23] L. P. Yuan, E. Choo, T. Yu, I. Khalil et al., "Time-Window Group-Correlation Support vs. Individual Features: A Detection of Abnormal Users," 2020. [PDF]
27. [24] P. J. Gunkel, C. S. Fischer, and P. Isserstedt, "Quarks and light (pseudo-)scalar mesons at finite chemical potential," 2019. [PDF]
28. [25] J. Wiederer, J. Schmidt, U. Kressel, K. Dietmayer et al., "A Benchmark for Unsupervised Anomaly Detection in Multi-Agent Trajectories," 2022. [PDF]
29. [26] M. Bikandi, G. Velez, N. Aginako, and I. Irigoien, "Synthetic outlier generation for anomaly detection in autonomous driving," 2023. [PDF]
30. [27] M. Ayalew Belay, S. Stenen Blakseth, A. Rasheed, and P. Salvo Rossi, "Unsupervised Anomaly Detection for IoT-Based Multivariate Time Series: Existing Solutions, Performance Analysis and Future Directions," 2023. ncbi.nlm.nih.gov
31. [28] H. Cao, W. Zou, Y. Wang, T. Song et al., "Emerging Threats in Deep Learning-Based Autonomous Driving: A Comprehensive Survey," 2022. [PDF]
32. [29] D. Bogdoll, M. Nitsche, and J. Marius Zöllner, "Anomaly Detection in Autonomous Driving: A Survey," 2022. [PDF]
33. [30] V. Kumar Kukkala, S. Vignesh Thiruloga, and S. Pasricha, "Roadmap for Cybersecurity in Autonomous Vehicles," 2022. [PDF]