

Article

Performance Estimation in V2X Networks Using Deep Learning-Based M-Estimator Loss Functions in the Presence of Outliers

Ali R. Abdellah ^{1,2,*} , Abdullah Alshahrani ³, Ammar Muthanna ^{2,4,*}  and Andrey Koucheryavy ²

- ¹ Electronics and Communications Engineering, Electrical Engineering Department, Faculty of Engineering, Al-Azhar University, Qena 83513, Egypt
- ² Department of Communication Networks and Data Transmission, The Bonch-Bruевич Saint-Petersburg State University of Telecommunications, 193232 St. Petersburg, Russia; akouch@mail.ru
- ³ Department of Computer Science and Artificial Intelligence, College of Computer Science and Engineering, University of Jeddah, Jeddah 21493, Saudi Arabia; asalshahrani2@uj.edu.sa
- ⁴ Applied Probability and Informatics, Peoples' Friendship University of Russia (RUDN University), 117198 Moscow, Russia
- * Correspondence: alirefaee@azhar.edu.eg (A.R.A.); muthanna.asa@spbgut.ru (A.M.)

Abstract: Recently, 5G networks have emerged as a new technology that can control the advancement of telecommunication networks and transportation systems. Furthermore, 5G networks provide better network performance while reducing network traffic and complexity compared to current networks. Machine-learning techniques (ML) will help symmetric IoT applications become a significant new data source in the future. Symmetry is a widely studied pattern in various research areas, especially in wireless network traffic. The study of symmetric and asymmetric faults and outliers (anomalies) in network traffic is an important topic. Nowadays, deep learning (DL) is an advanced approach in challenging wireless networks such as network management and optimization, anomaly detection, predictive analysis, lifetime value prediction, etc. However, its performance depends on the efficiency of training samples. DL is designed to work with large datasets and uses complex algorithms to train the model. The occurrence of outliers in the raw data reduces the reliability of the training models. In this paper, the performance of Vehicle-to-Everything (V2X) traffic was estimated using the DL algorithm. A set of robust statistical estimators, called M-estimators, have been proposed as robust loss functions as an alternative to the traditional MSE loss function, to improve the training process and robustize DL in the presence of outliers. We demonstrate their robustness in the presence of outliers on V2X traffic datasets.

Keywords: 5G networks; V2X; deep learning; M-estimators; outliers



Citation: Abdellah, A.R.; Alshahrani, A.; Muthanna, A.; Koucheryavy, A. Performance Estimation in V2X Networks Using Deep Learning-Based M-Estimator Loss Functions in the Presence of Outliers. *Symmetry* **2021**, *13*, 2207. <https://doi.org/10.3390/sym13112207>

Academic Editors: Kuo-Hui Yeh and Tomohiro Inagaki

Received: 26 September 2021
Accepted: 15 November 2021
Published: 19 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The evolution of 5G networks is characterized by a multilayer nature, high complexity, low latency, large bandwidth, high capacity, and heterogeneity. Moreover, 5G networks need to maintain continuous connectivity to meet QoS requirements for many devices and must process a large amount of information about the natural environment. Artificial intelligence (AI) technologies in 5G networks achieve intelligent performance, realistic complex learning, organizational structure, and complex decision making due to their robust ability to analyze patterns, learn, progress, and intelligently detect. Recent developments in AI technologies have created new opportunities for intelligent transportation systems (ITS). Vehicle sensors have become more intelligent over time, allowing vehicles to assess their surroundings better. Advances in smart transport systems can improve the security and performance of vehicle communications at all levels. They can provide drivers with all kinds of information to avoid accidents, reduce traffic congestion, and reduce human error [1,2].

AI plays a critical role in 5G networks to extract and quickly gain insights from data. Machine learning (ML) can automatically detect patterns and discover outliers in information generated by sensors and smart devices. AI technologies extend ML strategies applied to smart IoT devices to make complex decisions based on pattern recognition, self-learning, self-healing, context awareness, and autonomous decision making. These will influence future dual digital models and continuous learning in autonomous vehicle applications, IoT, and predictive maintenance. ML has used techniques to evaluate the effectiveness of internal communications to significantly improve the wireless network performance and the ability of technology to process and learn from the massive amount of data collected through various methods, topologies, and mobility scenarios. The ML technology enables more flexible and reliable vehicular networks. It also provides access to information and a high level of understanding of wireless communication systems and proposes to manage the operational aspects of more heterogeneous networks. Adding ML to wireless networks supports opening up new possibilities for implementing and improving the networks and creating a reliable, intelligent, sustainable, inclusive, and robust network that operates without human intervention [3–6].

Symmetry is a widely studied pattern in various research areas. In computer science and telecommunication networks, symmetric network structures and symmetric algorithms are often studied. The study of symmetric and asymmetric faults and outliers (anomalies) in network traffic is also important. Moreover, many systems are symmetric because the data speed or volume is the same in both directions; improving QoS in the network depends on symmetry.

The technology of ML can be used in various symmetric IoT applications and makes it possible that it will be an essential source of information in the future. Therefore, the networking environment can access experimental symmetric data through different network devices, explore the current data, extract information, and make conscious decisions based on the available data. ML is an essential platform for realizing intelligent IoT applications. Wide-ranging, intelligent, ML-based IoT systems will be more efficient and effective if they exploit the properties of “symmetry” and “asymmetry.” This will be helpful in various IoT applications [7–10], for example:

- Network monitoring improves security and efficiency, where network monitoring that identifies and fixes problems in applications is much more valuable to an organization to prevent unwanted system failures.
- Detect intrusions, errors, and anomalies to monitor raw data, such as user records, devices, networks, and servers. Quickly detect attacks and unclear security risks.
- Improving QoS requirements can be achieved by monitoring the control and management of network resources, reducing interference such as packet loss, latency, and jitter in the network. Network resources are managed through QoS by providing advantages for particular classes of data on the network. Prediction of quality Characteristics such as throughput, resource allocation, or timing issues. Model enabled accuracy prediction is a kind of prediction that applies a model as input for the prediction. Enterprise networks should provide expected and measured services for video conferencing that use real-time audio and video communications. Both are delay-sensitive and bandwidth-intensive forms of communication. Enterprises use QoS to efficiently manage sensitive applications, such as real-time voice, video, and critical data, and to avoid degradation of QoS parameters. Enterprises can achieve QoS by using certain features, such as jitter buffers and bandwidth management. For many enterprises, QoS is part of SLA with their NSP to ensure a specific stage of network performance.
- Health applications. Recently, medical IoT systems have become one of the most essential modern medical advancements. This technique can reach a critical benefit by improving the distance control of healthcare. It can also support detecting medical problems rapidly, therefore protecting patients’ lives and health. Nevertheless, many networked medical devices in the IoT healthcare space have security flaws that make them susceptible to malicious threats. The above challenges can cause serious conse-

quences that influence patients' lives by disrupting medical equipment. Therefore, it is necessary to overcome these challenges to maintain the efficiency and accuracy of medical IoT systems. At the same time, the wide distribution of sensitive medical information in IoT healthcare systems leaves them vulnerable to complex attacks that aim at main security aspects such as privacy and safety. This will harm the reliability, acquisition, and widespread use of IoT healthcare systems.

- Smart cities are becoming more and more of a reality, thanks to the enormous technological research enabling the development of IoT, which offers a wide range of applications around different types of sensors. More sustainable, environmentally friendly, and economical smart cities and technologies are needed to cope with the growing population in cities. For smart-city management, numerous sensors, cameras, and actuators are installed everywhere. These sensors collect and send a bulk of data in actual time. The analysis and processing of the collected data should be almost instantaneous for efficient management of city operations. Additionally, for instant processing, high-speed internet connectivity is essential. With the advent of smart-city devices, internet-connected devices will transmit large amounts of data in real time. While this data contributes to the efficiency of city functions, it also poses serious security risks that cannot be ignored. Data from parking lots, security cameras, electric vehicle charging stations, and GPS systems contain citizens' confidential information. Not every networked device is already cyber-resistant. If it is, criminals can easily access the data and use it for illegal purposes. Therefore, governments and IT professionals should strengthen the security perimeters of smart devices and supporting infrastructure. Identifying and solving smart-city challenges is a collaborative approach. Governments and IT professionals, private organizations, and citizens should join together to work for a common goal—the success of the smart city
- Smart farming. With the world's population growing and climatic changes resulting in unpredictable weather in the world's food chains, the race for the sustainability of farming and the efficient use of dwindling resources such as water is a global challenge for countries worldwide. Smart farming uses sensors implanted in plants and fields to take measurements that assist make decisions and plant protection. Precision farming is an essential part of the smart-farming paradigm, in which sensors are implanted in plants to take a certain measure to allow targeted care measurement. Precision farming will be needed to ensure food protection in the future; therefore, it is essential for farming sustainability. The primary use cases of AI in IoT for farming are plant health and disease identification and data-based crop protection.

In the computer science and symmetry domain [11], IoT can support various applications and services in different fields such as healthcare, home automation, security, and vehicles. The massive increase in data generated by vehicular networks makes vehicular communication vulnerable to attacks such as anomalies or outliers. These are critical factors that affect traffic flow and network security and impact the global economy. Robustness to outliers is a critical issue in modern machine-learning systems for many practical reasons, such as adversarial attacks, data corruption, etc. AI is a field of computer science that deals with developing intelligent machines, and complex and novel communication systems to maximize the chances of success. Nowadays, AI is a critical component to realize a vast amount of data collected by IoT devices. AI helps in analyzing the wireless network and data volume for data processing, detection, flow, accuracy, and timing reliability.

Robust intelligent techniques will be needed for adapting the network and managing resources of 5G for several facilities in multiple schemes. The ML is ideal for working on a 5G network because it requires a lot of data to make predictions and helps ensure that the 5G network can transmit a large amount of information. One of the best ML methods used for simulating nonlinear optimization functions is Deep Learning (DL). Recent advances in DL and ML techniques are promising for solving previously intractable and challenging problems. DL is a technique that uses the hierarchy of ANN to simulate ML activity during data processing in decision making, the human brain, and the nodal structures

of the nervous system. In the traditional method, the program analyzes the information linearly, while the functionality of a DL system allows the devices to process the data using a nonlinear method [1–16].

V2X technology enables information exchange between moving elements of the transportation network that can affect vehicles and the main components of V2X, such as V2V and V2I communication technologies, etc. The main benefits of the V2X network, include improved road safety, energy efficiency, transportation efficiency, and management of the road. V2X technology allows vehicles to communicate over a WANET using the IEEE 802.11p standard. WANET has several classes such as MANET, VANETs, FANETs, etc. The requirements of ad hoc networks are becoming more reliable and valuable for all their possible applications. However, the network requires establishing a mechanism to detect and mitigate security issues, as they are crucial in the network, and their failure leads to undesirable results. Wireless networks are typically vulnerable to interference caused by outliers or noisy data [17].

The presence of outliers or noise causes problems in the transport network as they affect the network's performance and lead to unsafe and unreliable connections. In recent research, various techniques have been used to reduce the impact of outliers in wireless networks. Among these techniques, the algorithms of ML are the most promising method that can be improved by reducing the effects of outliers in vehicular communications. The ML model can be adapted to the new network model depending on the data collection. Many researchers have been proposed robust learning algorithms to overcome the possibility of outliers in training data; these algorithms are insensitive to outliers and can detect unseen observations. However, in implementing the robust algorithms, ML requires close monitoring and the necessary measures to narrow down the scope, understanding, and threat model [18,19].

In many fields, such as industrial modeling, multilayer feedforward neural networks (MFNNs) or deep neural networks (DNNs) are used as approximators of nonlinear functions to provide advanced solutions for various applications. For training DNNs, a standard backpropagation algorithm based on minimizing the traditional MSE loss function is usually used. Unfortunately, this algorithm is prone and sensitive to outliers that distort the training data. As a result, the training data is contaminated with outliers, making the resulting model less reliable and potentially generating incorrect models, resulting in unacceptable performance.

Outliers are observations that deviate significantly from other values in a random sample of mass information. Examples of these outliers include the information population with gross errors caused by critical measurement errors, incorrect decimals, transcription errors, accidental scaling of a member from another population, rounding errors, grouping errors, and hypothetical deviation.

This article is about the performance estimation of V2X using DNNs. A set of robust statistical estimators, called M-estimators, have been proposed as robust loss functions to substitute the conventional MSE loss function, improve the training process, and robustize DL at the occurrence outliers in training data. The robustness in the presence of many outliers in V2X traffic datasets was illustrated in this work—finally, a comparative study with traditional and robust neural networks in terms of RMSE and MAPE.

The motivations behind this study include:

- Optimize quality of service (QoS) requirements and network monitoring to manage resources and ensure security.
- Monitor network availability and activity to identify and eliminate outliers (anomalies), including security and operational issues.
- One of the primary approaches to obtain a robust learning algorithm that is more robust to outliers is to replace the traditional loss function of the performance measure MSE with another robust function, to improve performance in the presence of outliers. In this approach, robustness against outliers has been satisfied by minimizing the effect of significant training errors due to outliers.

- The lack of accurate machine-learning analysis to achieve adequate performance.
 - The computational complexity of challenging problems in optimizing QoS measures.
- The research contributions have been summarized as follows:
- A novel DL algorithm has been proposed to estimate the performance of V2X using robust M-estimator loss functions instead of the standard MSE loss function.
 - A new comparison of the M-estimator approach and the conventional MSE loss function has been applied in terms of RMSE and MAPE, and under different sets of outliers on V2X traffic datasets used to verify the efficiency of the proposed method.
 - Finally, the results of the simulation-based tests show that:
 - ✓ The robust M-estimator loss functions have the best performance in all cases and outperform the conventional MSE loss function when data is clean or contains Gaussian noise or outliers.
 - ✓ When using noise-free data, the robust Fair loss function performs well and has the best performance compared to its peers.
 - ✓ When using data corrupted by Gaussian noise, the robust Cauchy loss function shows the best performance compared to the others.
 - ✓ Even on training data contaminated with outliers, the robust Fair loss function performs better than its competitors.

Outline of the article: The article is structured accordingly: the related literature review is presented in Section 2; the proposed work is explained in Section 3; the robust learning and outliers are introduced in Section 4; the basic concepts of M-estimator loss function are defined in Section 5; V2X simulation environment is introduced in Section 6; Deep Neural Network Learning is presented in Section 7; our theoretical results are illustrated in Section 8; and finally, in Section 9 we conclude.

2. Relevant Works

Many researchers focus on improving V2X communication performance using ML techniques to significantly improve traffic safety, energy efficiency, and traffic efficiency in establishing reliable communication. One of the primary approaches to obtain a robust learning algorithm that is more robust to outliers is to replace the loss function of the performance measure MSE with another robust function to enhance the training process and robustize DL in the presence of outliers. In this approach, robustness against outliers is satisfied by minimizing the effect of significant training errors due to outliers.

Moreover, there have been some previous efforts to develop robust learning algorithms to improve performance in the presence of outliers. Moreover, many researchers have focused on estimating and detecting outliers in wireless networks using ML techniques. In this work, we studied the performance of V2X traffic using DNNs based on M-estimator loss functions as a replacement for the standard MSE loss when the data contains outliers.

Abdellah et al. [20] presented robust ANN learning algorithms that enabled robust statistical methods, called M-estimators, as a robust loss function to substitute the conventional MSE loss function for studying VANET traffic performance. The research [21] addressed the estimation of energy for VANET performance using robust NN training algorithms and proposed robust M-estimator loss functions for the case of noise-free information. Liang et al. [22] investigated the applicability of machine learning to address problems in highly mobile vehicular communication and prompting ML to manage the resultant difficulties in vehicular networks. The research [23] presented a set of robust, statistical M-estimators as a replacement for the conventional MSE loss function using good, noise-free information. Furthermore, new transfer functions that depend on robust statistical m-estimators proposed alternatives to traditional transfer functions using a dataset containing outliers [24].

The DL algorithm has addressed detecting an anomaly in 5G networks regarding network latency [25]. In [26], the robustness of different ANNs training algorithms has been investigated using the robust M-estimators as a loss function in order to robustize learning

in the presence of outliers. Maimó et al. [27] studied DL's performance for abnormality discovery in 5G networks. Reddy et al. [28] examined anomaly detection by DNN in tracking IoT traffic for future smart-city applications.

Abdellah et al. [29] proposed an approach based on a robust estimation technique called M-estimator as a robust loss function to improve the robustness of the ANN training procedure. Rusiecki [30] introduced a novel robust performance measure depending on SoftMax Loss for DNN training in the presence of noisy data and showed its robustness for different sets of contaminated data. Zahra et al. [31] presented robust NN classifiers using a novel robust loss function known as M-estimator in the presence of corrupted data. Alpha et al. [32] combined the advantages of the assumed τ -estimates of the Nonlinear Regression Model (NRM) and the Backpropagation Algorithm depending on M-estimators as a robust loss function, in order to create the TAO robust learning algorithm that handles the problems of model fitting with outliers.

Moreover, Noor-A-Rahim et al., 2021 and Liu et al., 2020 have presented 5G and 6G technology components for V2X communication [33,34]. In particular, in [33], the authors discussed the evolution from the existing cellular V2X technology to New Radio-V2X, focusing on the main features and functions of the physical layer, sidelink communication, and its resource allocation, precise positioning techniques, security, privacy mechanisms, and architecture flexibility. On the other hand, several critical technologies from different areas, such as new tools, algorithms, and system structures, are presented [34].

3. Proposed Work

In this proposed work, the performance of the V2X network using a deep neural network (DNN) was estimated with new robust loss functions, a set of robust statistical techniques called M-estimators. The robustness of DNNs for different levels of outliers on V2X datasets was demonstrated.

The training datasets were created from the V2X network. The V2X network was simulated using MATLAB software. The collected data was analyzed and processed before the training phase. After loading the dataset as input to the network, the dataset was divided into two subsets: Input (I) and Output (O) columns. Then, it was also divided into training, testing, and validation subsets. The normalization of the input data must be in the interval $[-1, 1]$, compatible with the actual maximum or minimum values. The ANN was generally trained using a standard backpropagation (BP) learning algorithm. This technique is the most popular supervised learning algorithm, where I/O pairs are provided to the network, and weights are adapted to minimize the error between the network's actual and estimated outputs using a loss function based on delta rules. During training, the training model was adjusted, the gradient of the loss function was calculated, and then the network weights and biases were updated in response to the gradients. This process was repeated until the minimum output error was reached.

In the next stage, the test network requires test groups to assess the estimated model performance. The initial training stage generally reduces the validation error as an error of the training set. However, when the network overfits the training, the validation error always begins to increase. In this case, the network parameters were stored with the minimum error of the validation set.

However, the traditional BP algorithm is not robust when data contains outliers that can contaminate the training data. To overcome the potential of outliers, a robust BP algorithm was used in this work. A robust performance measurement function, i.e., the well-known M-estimator loss function, was proposed as a replacement for the conventional MSE loss function to improve the robustness of DNNs and obtain the optimal performance on contaminated data in V2X datasets. DNNs can estimate the optimal performance of a V2X network based on the collected V2X dataset; V2X throughput was used as input and packet loss rate as output (desired output). A comparative study between robust and traditional DNN performance using RMSE and MAPE shows that the proposed algorithm

provides outstanding results for the desired application. Figure 1 illustrates the flowchart of the proposed DNN for estimating V2X performance.

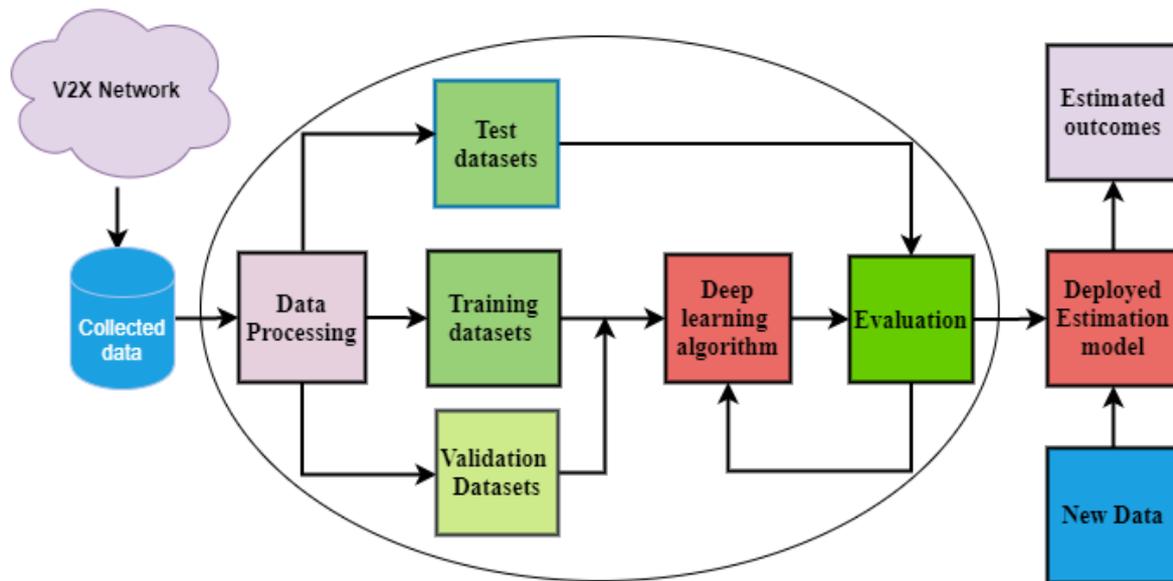


Figure 1. The flowchart of the suggested DNN for estimating V2X performance.

4. Robust Learning and Outliers

In statistics, outliers are observations that deviate dramatically from other data points. Outliers can result from measurement error, rounding error, human error, long-tailed noise distribution, etc. Outliers are a general feature of several realistic datasets; their likelihood in raw data varies between 1–10% [35,36]. However, it is hard to identify outliers even though there are many of them in actual data, and it is not easy to detect such multiple outliers; moreover, it is sometimes impossible to tell whether a point is an outlier or not. Furthermore, defining outliers in multivariate or methodological data can be difficult or even impossible, and testing for such various deviations is fully involved and sometimes leads to extensive computations. There is usually no need to disregard these observations when implementing a data-driven model, only to reduce their impact on the design parameters.

Training DNNs with the traditional backpropagation learning algorithm minimizes the training data's traditional performance measurement function (MSE loss function). However, this conventional backpropagation learning algorithm is sensitive to outliers that can distort the training data. It can be considered optimal only for noise-free data or for data corrupted by errors from the Gauss distribution when the mean is zero [35,36]. When gross errors or outliers contaminate the data, it leads to an unreliable method. To overcome the problem of outliers, several robust statistical estimators that are not severely affected by deviations from observations, such as M-estimators [20,21,23,24,26,29,31], R-estimators, L-estimators, and LTS [30], and LMedS [37], have been used in recent research papers. In this paper, we address the robust training of DNN based on using M-estimators as robust loss functions to substitute the conventional MSE loss function, which loses robustness in the presence of outliers in the V2X dataset.

5. M-Estimators Loss Function

M-estimators play a vital role in NN society and constitute a broad group of estimators; maximum likelihood estimation and nonlinear least-squares methods are special cases. M-estimators are an effective mechanism in robust statistics that is resistant to outliers. M-estimators are robust statistical estimators usually used as an alternative to the least-squares estimator (LS) when the data contain outliers or extreme observations or do not conform to

a normal distribution. The M-estimator is advantageous when the data have outliers and are contaminated by a single outlier (or an error with a strong tail). Even a single outlier can destroy the least-squares estimator and lead to its breakdown. Therefore, there are two options: 1. eliminate outliers that give a bad result; 2. use the robust M-estimators.

Recently, several researchers have proposed M-estimators as loss functions to improve the learning process of NNs. The M-estimators use multiple loss functions that are less increasing than the LS estimators. The M-estimator restricts the response once the fitting error surpasses a threshold, i.e., the fitting error is not zero. Hence, the M-estimator loss function has greater reliability, robustness, and flexibility when the data contains outliers than the conventional MSE loss function, which is highly susceptible to outliers; thus, this set of estimators should be a suitable replacement for the MSE loss function outliers.

In this work, the traditional MSE loss function has been replaced by M-estimator loss functions to improve the training process and increase the learning of NNs. Assume that e_i is the residual of the i th data point that is the difference between the observed and estimated values. The least-squares method attempts to reduce the error $\sum_i e_i^2$ that is unstable with outliers. Outlier data has extreme values that can spoil the estimation of the parameters. The M-estimators attempt to minimize the effects of outliers by exchanging the squared residuals e_i^2 with various fitting error functions, resulting in:

$$\min \sum_i \rho(e_i) \quad (1)$$

- $\rho(e)$ is a positive-definite symmetric function, where $\rho(e) = \rho(-e)$ for all e .
- $\rho(e) \geq 0$ for all e and has a unique minimum at $e = 0$.
- $\rho(e)$ increases when e increases from 0 but does not become too large when e increases and is selected to be less increasing than square.

Table 1 illustrates the used M-estimators, and their influence function can be represented as follows [36,38].

Table 1. Commonly used M-estimators.

Type	$P(e)$	$\Psi(e)$	$\omega(e)$
L2	$e^2/2$	e	1
Fair	$c^2 \left[\frac{ e }{c} - \log \left(1 + \frac{ e }{c} \right) \right]$	$\frac{e}{1+ e /c}$	$\frac{1}{1+ e /c}$
Cauchy	$\frac{c^2}{2} \log \left(1 + (e/c)^2 \right)$	$\frac{e}{1+(e/c)^2}$	$\frac{1}{1+(e/c)^2}$

6. V2X Simulation Environment

In this section, V2X in a smart city has simulated using MATLAB environment. First, the mobility map for V2X has been created. The AODV routing protocol was prepared in a virtual mobility map to study and evaluate its work. The road network is developed when creating mobility maps containing basic entities: simulation domain, connection point (node), and RSU (roadside unit). The simulation domain needs to define the connection points that move in arbitrary directions to implement AODV. Implementing the AODV routing protocol requires a maximum city size and many nodes; you also need to install multiple RSUs. If the simulation domain is larger than the simulation time, this will be automatically expanded and assumed that the simulation domain is 100×100 for the x - y axis. In the mobility model, the desired connection points in the boundary can move in any direction along fixed routes.

Figure 2 shows that the dots represent the individual connection points and RSU locations by the ID numbers assigned to them by the network design and configuration. The beginning and end of this model are indicated sequentially by the connection point numbers 20 and 70. This unit causes the arrangement of multiple connection points; the flow of connection points assigns groups of connection-point activities that flow over the

simulation and the turn ratio that determines the probability of routes at every intersection. The simulation module visualizes the network architecture and selects the start and end times of the simulation. The position of the RSU in the simulation map helps connection points that move randomly to communicate with other connection points far away from other connection point areas. The RSU allows for connecting and exchanging messages with moving vehicles, including safety alerts and traffic information, and they can also connect directly or through multihop connections. Figure 2 illustrates the V2X simulation in a smart city.

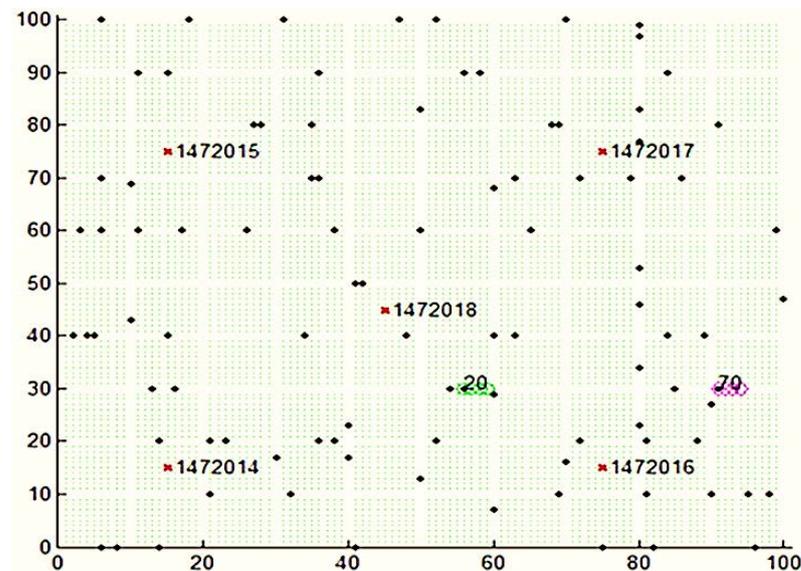


Figure 2. The V2X simulation in a smart city.

7. Deep Neural Network Learning

Deep learning is an ML method that depends on the architecture and function of artificial neural networks (ANNs). For this reason, DL, also known as DNN, consists of multiple hidden layers. DL has usually been implemented to deal with massive amounts of data and complex algorithms to train the model. Every layer has neurons connected to the neurons of the previous layer by a set of weights. As shown in Figure 3, there are many different layers, but the most common is the dense layer, which connects all modules directly to each neuron in the previous layer. The neurons in each subsequent layer can represent increasingly complex aspects of the original input by stacking layers.

The DNN consists of many units, so the output of one unit becomes the input for the following units in the network, as shown in Figure 3. The first and last layers are the input and output layers, respectively, and the last layer is the output layer. The input layer contains several neurons corresponding to the input samples from the V2X dataset. In this work, V2X performance has been estimated for a particular traffic dataset; the output layer also contained a single neuron, considered the output or estimated output value since there was only one output.

During DNN training, the weights are initialized randomly. Then, sample data is added one by one. The learning rule has been applied to adjust the connections' weights considering the input data and the expected result. The efficiency of DNN, i.e., the accuracy and the results, largely depends on the data examples used in training. An extensive set of data examples with high content variability is the key to more accurate inference results. When the training depends on similar or repetitive data, the DNN will not analyze the basic information differently from the example data. In this case, the DNN will retrain.

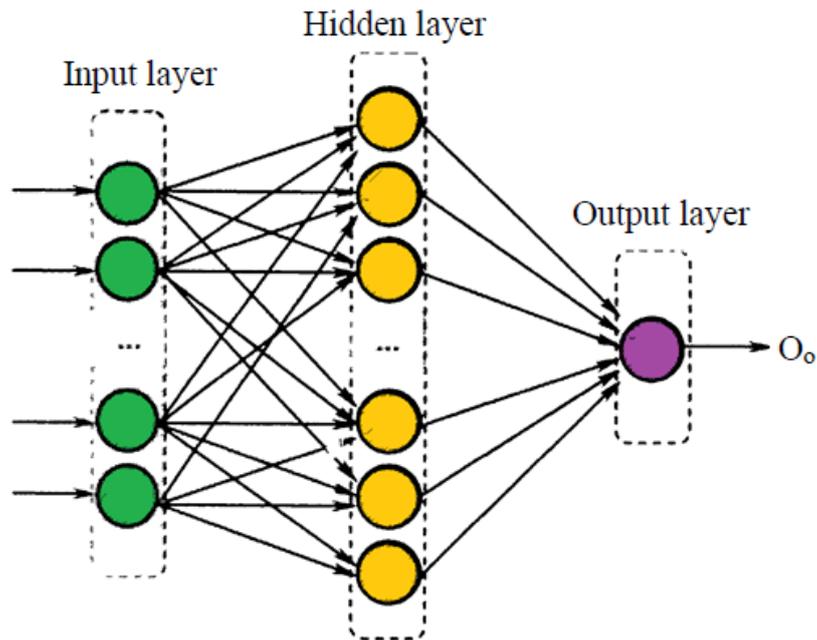


Figure 3. Deep neural network architecture.

In this work, the architecture of DNN consists of three layers with a hidden layer containing 100 hidden neurons, as shown in Figure 3. The most appropriate architectural parameters (e.g., batch size, epochs, activation function, loss function, learning rate, and performance goal) have been selected for the proposed model DL.

The model was trained using both the traditional loss-function MSE and the robust loss-function M-estimator. For all neurons in the hidden layer, the activation function Tansig was chosen, and Purlin was selected for the neuron in the output layer. We chose a batch size of 32, a number of epochs of 1000, a learning rate of 1×10^{-3} , and a performance goal (minimum loss) of 1×10^{-3} . The model DL was trained several times with different configurations. We created different combinations of the input dataset with minor changes to the network parameters and ran them with all possible topologies. However, this is only the initial phase; once the input dataset is prepared, only this dataset is used, and the number of data type values is fixed.

8. Simulation Results

In this section, the V2X traffic dataset was collected and processed after simulating the V2X environment. The dataset was split into 70% training, 15% validation, and 15% testing. A DNN model was built to estimate the performance of the V2X traffic. DNN trained with the robust BP algorithm uses both M-estimators as robust loss functions and the traditional MSE loss function. All the above loss functions try to find the optimal V2X performance in the presence of outliers—comparing the performance of robust and traditional DNNs using RMSE and MAPE for each model.

The performance of DNNs was examined for the percentage of outliers in three cases as follows:

- (1) Set A: The DNN is trained with noise-free, high-quality, clean data.
- (2) Set B: The network is trained with perfect data contaminated with slight Gaussian noise (GN): $G2 \sim N(0, 0.1)$.
- (3) Set C: DNN is trained with data contaminated with GN, $G2 \sim N(0, 0.1)$, in addition to very good, randomized outliers of the form:

$$H_1 \sim N(+15, 2), H_2 \sim N(-20, 3), H_3 \sim N(+30, 1.5), H_4 \sim N(-12, 3).$$

The training data used in this case is as follows:

$$\text{Data} = (1 - \varepsilon)G_2 + \varepsilon(H_1 + H_2 + H_3 + H_4)$$

Data are error distribution, and $G_2, H_1, H_2, H_3,$ and H_4 are probability distributions occurring on probability $1 - \varepsilon$ and ε , respectively, where ε is fixed and H random. In this case, the outliers were included in the training data with a percentage $\varepsilon = 10\%$ of the data. We assigned the outliers randomly to the desired percentage of data (percent outliers) (25% of this percentage will have outlier H_1 type, the other 25% will have H_2 type, and so on . . . H_3 and H_4). The dataset used for this experiment was generated from the V2X network. The dataset is then contaminated in the x - y axis by Gaussian noise with a mean of zero and a standard deviation of 0.1, $G_2 \sim N(0, 0.1)$. A variable percentage, ε , of data was randomly selected and then replaced with probability, ε , by background noise uniformly distributed in the specific range [24,26,29,31,32].

In this work, a DNN architecture is MFNN consisting of three layers with one hidden layer containing 100 hidden neurons. The network was trained using the robust BP algorithm based on the robust loss functions mentioned earlier. The training function is Traincgf, the max epoch for training is 1000 epochs. Moreover, the performance goal (min loss) is 1×10^{-3} . The normalization of the input data must be in the interval $[-1, 1]$, which is compatible with the actual maximum to minimum values. DNN can estimate the optimal performance of a V2X network based on the collected V2X dataset; it used V2X throughput as input and packet loss rate as output (desired output). The goal is to develop a robust DNN that can estimate V2X performance when the data contains outliers. We conducted a comparative study between robust and traditional DNN performances regarding RMSE and MAPE, to prove which algorithm gives excellent results for the considered application.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2} \quad (2)$$

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{x_i - \hat{x}_i}{x_i} \right| \quad (3)$$

where n is the number of data points, x_i is the observed value, and \hat{x}_i is the estimated value. The subscript i denotes the corresponding individual values of the observed and estimated values.

Table 2 shows the performance values for networks trained with MSE, Fair, and Cauchy loss functions for estimating V2X traffic.

Table 2. Performance scores for networks trained with MSE, Fair, and Cauchy loss functions for estimating V2X traffic.

Loss Function	Set A		Set B		Set C	
	RMSE	MAPE%	RMSE	MAPE%	RMSE	MAPE%
MSE	0.0156	1.3	0.1786	10.6	0.5179	16.7
Cauchy	0.0194	1.5	0.0630	5.4	0.0923	9.8
Fair	0.0132	1.1	0.0740	6.2	0.0756	8.6

Table 2 shows the performance predicted by robust and traditional DNN loss functions for robust estimation of V2X traffic in the presence of outliers. The performance has been estimated using RMSE and MAPE in three cases according to the percentage of outliers.

The loss function Fair has the best performance, as shown by the tabulated results, compared to its competitors in RMSE and MAPE in the case of set A with noise-free data. The maximum average performance improvement, in this case, is 0.4%. Moreover, the traditional MSE loss function has approximately the same performance as the Fair loss

function, and the maximum average improvement is 0.2% in this case. On the other hand, the Cauchy loss function provides the lowest performance compared to the other functions.

Looking at the tabulated results using set B, it is found that the robust Cauchy loss function has the best performance compared to the others in both RMSE and MAPE. The maximum average of performance improvement, in this case, is 5.2%. Additionally, the robust Fair loss function has semi-equal performance to the robust Cauchy loss function, and the maximum average of performance improvement is 4.4%. On the other hand, the traditional MSE loss function performs poorly compared to the others.

As the table shows in the case of set C with contaminated data, the robust neural networks are insensitive to outliers. The robust loss function Fair has the best performance compared to the others. The average performance improvement, in this case, is 8.1%. In addition, the robust Cauchy loss function performs approximately the same as the robust Fair loss function, and the average performance improvement is 6.9% in this case. However, the traditional MSE loss function offers poor performance compared to its peers.

Figure 4 shows, using Set A with high-quality noise-free data, that all models trained with robust and traditional loss functions have approximately equivalent responses to the ideal model. We noted that the estimated model gradually decreases with an increase in the input pattern.

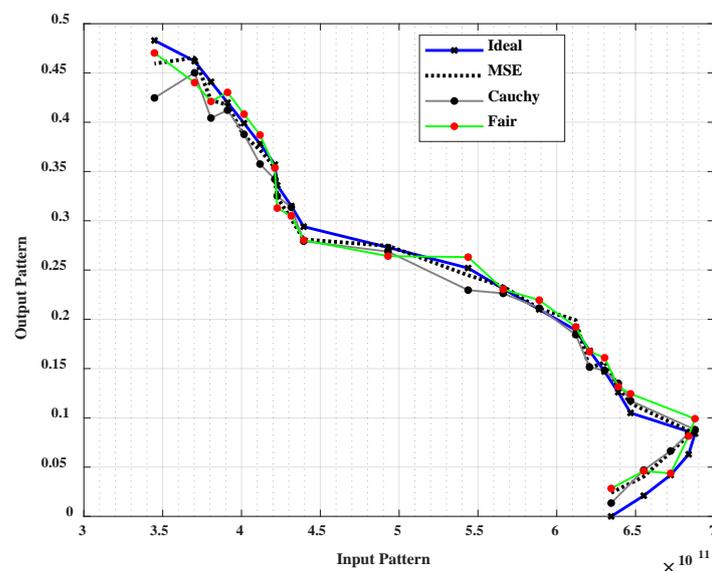


Figure 4. The predicted models in the robust loss functions Cauchy and Fair and the traditional loss function MSE in Set A.

Figure 5 illustrates the validation loss of each model over the number of epochs during DNN training in three cases according to the loss function used—MSE, Cauchy, and Fair, respectively. It was found that the error gradually decreases after more training epochs in all cases. The best validation performance is 0.015838 at the 63rd epoch, 0.0012387 at the 45th epoch, and 0.018973 at the 43rd epoch of the training network using the traditional loss functions MSE, Cauchy, and Fair, respectively.

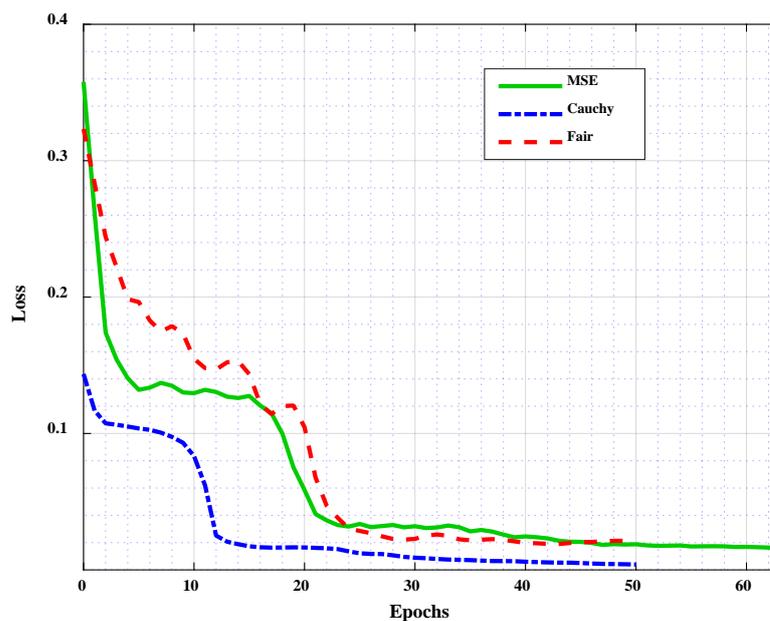


Figure 5. The best validation performance of DNN training in the case of Set A.

In Figure 6, it can be seen that in the case of using set B on the data corrupted with Gaussian noise, the predicted models using Cauchy and Fair robust loss functions have responses that are semi-equal and close to the ideal model, except that the model predicted using the conventional MSE loss function deviates from the ideal model. The performance of the training, in this case, is shown in Figure 7. We noticed that the loss decreases as the number of epochs increases. The best validation performance using the traditional MSE loss function is 0.06399 at the 46th epoch. The best validation performance in the case of robust Fair loss function is 0.07479 at the 44th epoch. Moreover, the best validation performance using the robust Cauchy loss function is 0.0040613 at the 54th epoch.

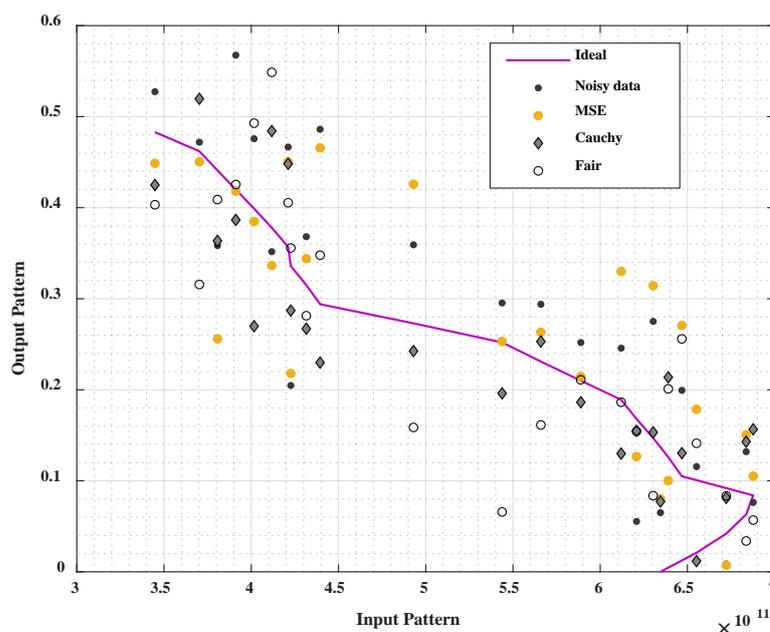


Figure 6. The predicted models in robust loss functions Cauchy and Fair and traditional loss function MSE in set B.

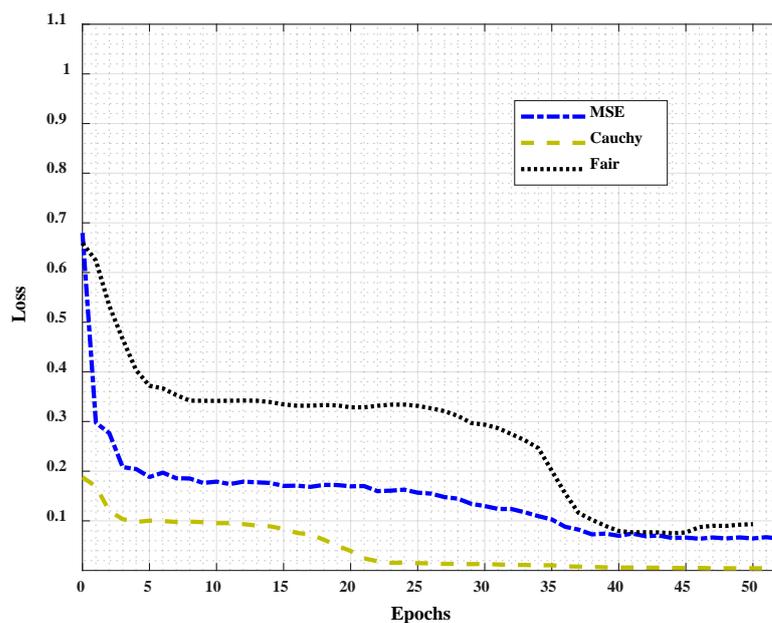


Figure 7. The best validation performance of DNN training in the case of set B.

As illustrated in Figure 8, when using set C with data corrupted with outliers, the robust neural networks are immune to outliers and produce approximately ideal results. Otherwise, the model generated by traditional neural networks (MSE) has been strongly affected by the outliers.

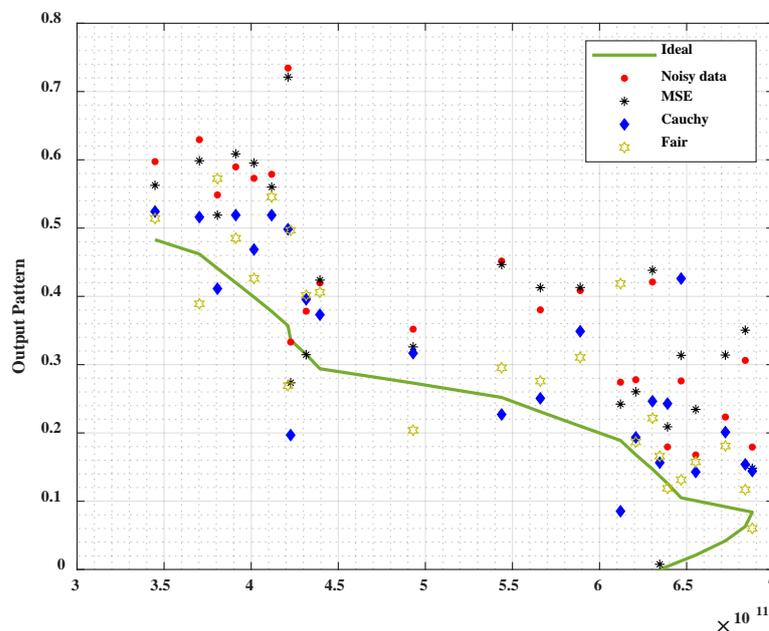


Figure 8. Predicted models in the robust loss functions Cauchy and Fair and the traditional loss function MSE in the set C case.

The performance of the training, in this case, is shown in Figure 9. We noted that the loss decreases with the increasing number of epochs. The best validation performance when using the traditional MSE loss function is 0.069745 at the 65th epoch; the best validation performance when using the robust, fair loss function is 0.047096 at the 50th epoch. Moreover, the best validation performance when using the robust Cauchy loss function is 0.0027879 at the 71st epoch.

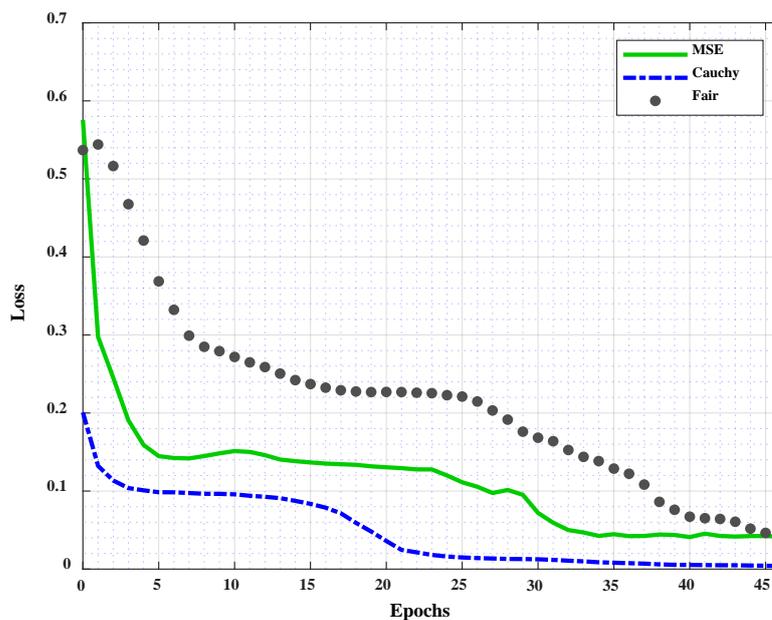


Figure 9. The best validation performance of DNN training in the case of set C.

In Figure 10, the relationship between the performance represented by RMSE and the percentage of outliers is depicted. It can be observed that RMSE increases with an increasing percentage of outliers, which is significant when using the conventional MSE loss function.

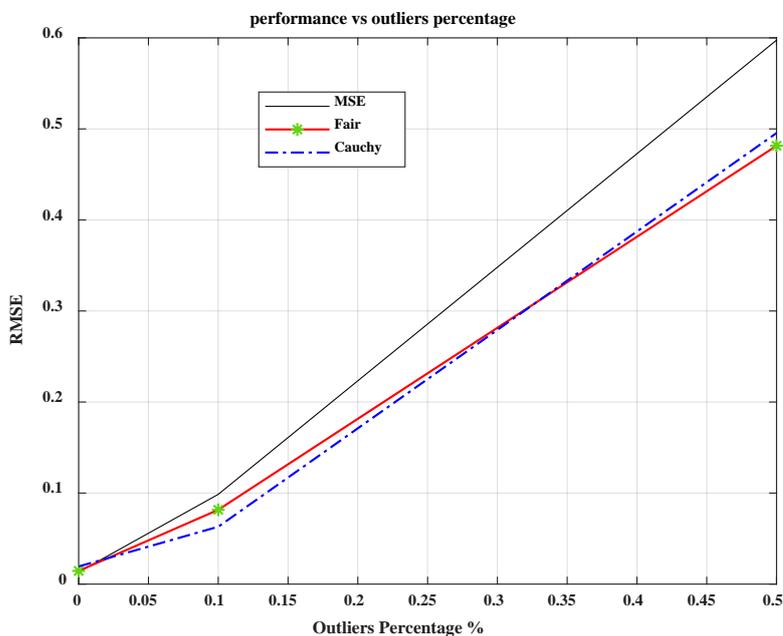


Figure 10. Performance vs. outliers percentage.

9. Conclusions

In this work, V2X performance was estimated using DNN learning based on robust M-estimators as robust loss functions for performance measurement (a robust statistical technique), in order to replace the traditional MSE loss function in the presence of outliers in V2X datasets. The simulation results show that the deep-learning approach with M-estimator loss functions provides excellent results for the V2X network in the presence of outliers. The proposed robust M-estimator loss functions were compared with the conventional MSE loss function for different percentages of outliers. The results show

that the robust M-estimator loss functions have the best performance in all cases and outperform the conventional MSE loss function. When using noise-free data, the robust Fair loss function performs well and has the best performance compared to its peers. When using data corrupted by Gaussian noise, the robust Cauchy loss function shows the best performance compared to the others. Even on training data contaminated with outliers, the robust Fair loss function performs better than its competitors. Based on the above results, the robust M-estimator loss functions are a promising solution for the highly corrupted datasets in wireless networks.

Author Contributions: Conceptualization, A.R.A.; methodology, A.R.A.; software, A.R.A. and A.K.; validation, A.M. and A.R.A.; formal analysis, A.R.A.; investigation, A.M.; resources, A.K.; data curation, A.R.A. and A.K.; writing—original draft preparation, A.R.A. and A.M.; writing—review and editing A.A.; visualization, A.R.A. and A.K.; supervision, A.K.; project administration, A.K.; funding acquisition, A.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by the Applied Scientific Research under the SPbSUT state assignment 2021.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data are contained within the article and/or available from the corresponding author upon reasonable request.

Acknowledgments: This research is based on the Applied Scientific Research under the SPbSUT state assignment 2021. The researcher (Ali R. Abdellah) is funded by a scholarship (Ph.D.) under the Joint Executive Program between the Arab Republic of Egypt and the Russian Federation.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

M-estimators	Maximum-likelihood estimators
LMedS	Least Median of Squares
L-estimators	Linear combination of order statistics
R-estimators	Estimates based on rank transformations
MSE	Mean square error
DNN	Deep neural networks
DL	Deep Learning
BP	Backpropagation
RMSE	Root Mean square error
MAPE	Mean absolute percentage of error
LTS	Least Trimmed Squares
LMLS	Least mean log square
QoS	Quality of service
ITS	Intelligent Transportation Systems
AI	Artificial Intelligence
IoT	Internet of Things
ML	Machine Learning
ANN	Artificial Neural Networks
MFNNs	Multilayer feedforward neural networks
V2X	Vehicle to Everything
WANET	Wireless ad hoc network
MANET	Mobile ad hoc network
VANETs	Vehicular ad hoc networks
FANETs	Flying ad hoc networks
LMLS	The least mean log squares
Traincgf	Conjugate gradient backpropagation with Fletcher-Reeves updates
SLA	Service level agreement
NSP	Network service provider

References

1. Morocho-Cayamcela, M.E.; Haeyoung, L.; Wansu, L. Machine Learning for 5G/B5G Mobile and Wireless Communications: Potential, Limitations, and Future Directions. *IEEE Access* **2019**, *7*, 137184–137206. [\[CrossRef\]](#)
2. Sumalee, A.; Ho, H. Smarter and more connected: Future intelligent transportation system. *IATSS Res.* **2018**, *42*, 67–71. [\[CrossRef\]](#)
3. Sarker, I.H. Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Comput. Sci.* **2021**, *2*, 160. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Sun, Y.; Peng, M.; Zhou, Y.; Huang, Y.; Mao, S. Application of Machine Learning in Wireless Networks: Key Techniques and Open Issues. *IEEE Commun. Surv. Tutor.* **2019**, *21*, 3072–3108. [\[CrossRef\]](#)
5. Ali, E.S.; Hasan, M.K.; Hassan, R.; Saeed, R.A.; Hassan, M.B.; Islam, S.; Nafi, N.S.; Bevinakoppa, S. Machine Learning Technologies for Secure Vehicular Communication in Internet of Vehicles: Recent Advances and Applications. *Secur. Commun. Netw.* **2021**, *2021*, 8868355. [\[CrossRef\]](#)
6. Tong, W.; Hussain, A.; Bo, W.X.; Maharjan, S. Artificial Intelligence for Vehicle-to-Everything: A Survey. *IEEE Access* **2019**, *7*, 10823–10843. [\[CrossRef\]](#)
7. Alsharif, M.H.; Kelechi, A.H.; Yahya, K.; Chaudhry, S.A. Machine Learning Algorithms for Smart Data Analysis in Internet of Things Environment: Taxonomies and Research Trends. *Symmetry* **2020**, *12*, 88. [\[CrossRef\]](#)
8. Jagannath, J.; Polosky, N.; Jagannath, A.; Restuccia, F.; Melodia, T. Machine learning for wireless communications in the Internet of Things: A comprehensive survey. *Ad Hoc Networks* **2019**, *93*, 101913. [\[CrossRef\]](#)
9. Zhu, J.; Xu, W. Real-Time Data Filling and Automatic Retrieval Algorithm of Road Traffic Based on Deep-Learning Method. *Symmetry* **2020**, *13*, 1. [\[CrossRef\]](#)
10. Hassan, R.; Qamar, F.; Hasan, M.K.; Aman, A.H.M.; Ahmed, A.S. Internet of Things and Its Applications: A Comprehensive Survey. *Symmetry* **2020**, *12*, 1674. [\[CrossRef\]](#)
11. Martin, T.; Geneiatakis, D.; Kounelis, I.; Kerckhof, S.; Fovino, I.N. Towards a Formal IoT Security Model. *Symmetry* **2020**, *12*, 1305. [\[CrossRef\]](#)
12. Khedkar, S.P.; Canessane, R.A.; Najafi, M.L. Prediction of Traffic Generated by IoT Devices Using Statistical Learning Time Series Algorithms. *Wirel. Commun. Mob. Comput.* **2021**, *2021*, 5366222. [\[CrossRef\]](#)
13. Boutaba, R.; Salahuddin, M.A.; Limam, N.; Ayoubi, S.; Shahriar, N.; Estrada-Solano, F.; Caicedo, O.M. A comprehensive survey on machine learning for networking: Evolution, applications and research opportunities. *J. Internet Serv. Appl.* **2018**, *9*, 16. [\[CrossRef\]](#)
14. Kaur, J.; Khan, M.A.; Iftikhar, M.; Imran, M.; Haq, Q.E.U. Machine Learning Techniques for 5G and Beyond. *IEEE Access* **2021**, *9*, 23472–23488. [\[CrossRef\]](#)
15. Singh, D.P.; Sharma, D. Traffic Prediction Using Machine Learning and IoT. In *Integration of Cloud Computing with Internet of Things*; Wiley: Hoboken, NJ, USA, 2021; pp. 111–129. [\[CrossRef\]](#)
16. Cayamcela, M.E.M.; Lim, W. Artificial Intelligence in 5G Technology: A Survey. In Proceedings of the 2018 International Conference on Information and Communication Technology Convergence (ICTC), Jeju, Korea, 17–19 October 2018.
17. Arena, F.; Pau, G. An Overview of Vehicular Communications. *Futur. Internet* **2019**, *11*, 27. [\[CrossRef\]](#)
18. Zhang, Y.; Meratnia, N.; Havinga, P. Outlier Detection Techniques for Wireless Sensor Networks: A Survey. *IEEE Commun. Surv. Tutor.* **2010**, *12*, 159–170. [\[CrossRef\]](#)
19. Rassam, M.; Zainal, A.; Maarof, M.A. Advancements of Data Anomaly Detection Research in Wireless Sensor Networks: A Survey and Open Issues. *Sensors* **2013**, *13*, 10087–10122. [\[CrossRef\]](#)
20. Abdellah, A.R.; Muthanna, A.; Koucheryavy, A. Robust Estimation of VANET Performance-Based Robust Neural Networks Learning. In *Internet of Things, Smart Spaces, and Next Generation Networks and Systems*; NEW2AN 2019, ruSMART 2019. Lecture Notes in Computer Science; Galinina, O., Andreev, S., Balandin, S., Koucheryavy, Y., Eds.; Springer: Cham, Switzerland, 2019; Volume 11660. [\[CrossRef\]](#)
21. Abdellah, A.R.; Muthanna, A.; Koucheryavy, A. Energy Estimation for VANET Performance Based Robust Neural Networks Learning. In *Distributed Computer and Communication Networks*; DCCN Communications in Computer and Information Science; Vishnevskiy, V., Samouylov, K., Kozyrev, D., Eds.; Springer: Cham, Switzerland, 2019; Volume 1141. [\[CrossRef\]](#)
22. Liang, L.; Ye, H.; Li, G. Toward Intelligent Vehicular Networks: A Machine Learning Framework. *IEEE Internet Things J.* **2019**, *6*, 124–135. [\[CrossRef\]](#)
23. Zahra, M.M.; Essai, M.H.; Abd Ellah, A.R. Performance Functions Alternatives of Mse for Neural Networks Learning. *Int. J. Eng. Res. Technol. (IJERT)* **2014**, *3*, 967–970.
24. Essai, M.H.; Abd Ellah, A.R. M-Estimators Based Activation Functions for Robust Neural Network Learning. In Proceedings of the IEEE 10th International Computer Engineering Conference (ICENCO2014), Cairo, Egypt, 29–30 December 2014; pp. 76–81.
25. Doan, M.; Zhang, Z. Deep Learning in 5G Wireless Networks—Anomaly Detections. In Proceedings of the 29th Wireless and Optical Communications Conference (WOCC2020), Newark, NJ, USA, 1–2 May 2020; pp. 1–6. [\[CrossRef\]](#)
26. Abd Ellah, A.R.; Essai, M.H.; Yahya, A. Comparison of Different Backpropagation Training Algorithms Using Robust M-Estimators Performance Functions. In Proceedings of the IEEE 2015 Tenth International Conference on Computer Engineering & Systems (ICCES), Cairo, Egypt, 23–24 December 2015; pp. 384–388.

27. Maimo, L.F.; Clemente, F.J.G.; Gil Perez, M.; Perez, G.M. On the performance of a deep learning-based anomaly detection system for 5G networks. In Proceedings of the 2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), San Francisco, CA, USA, 4–8 August 2017; Institute of Electrical and Electronics Engineers (IEEE): Piscataway, NJ, USA, 2017; pp. 1–8.
28. Reddy, D.K.; Behera, H.S.; Nayak, J.; Vijayakumar, P.; Naik, B.; Singh, P.K. Deep neural network based anomaly detection in Internet of Things network traffic tracking for the applications of future smart cities. *Trans. Emerg. Telecommun. Technol.* **2020**, *32*, 2161–3915. [[CrossRef](#)]
29. Ellah, A.R.A.; Essai, M.H.; Yahya, A. Robust Backpropagation Learning Algorithm Study for Feed Forward Neural Networks. Master's Thesis, Al-Azhar University, Cairo Governorate, Egypt, 2016.
30. Rusiecki, A. Trimmed Robust Loss Function for Training Deep Neural Networks with Label Noise. In *Artificial Intelligence and Soft Computing. ICAISC 2019; Lecture Notes in Computer Science*; Rutkowski, L., Scherer, R., Korytkowski, M., Pedrycz, W., Tadeusiewicz, R., Zurada, J., Eds.; Springer: Cham, Switzerland, 2019; Volume 11508. [[CrossRef](#)]
31. Mohamed, M.Z.; Mohamed, H.; Essai Ali, R.; Abd, E. Robust Neural Network Classifier. *Int. J. Eng. Dev. Res. (IJEDR)* **2013**, *1*, 326–331.
32. Pernia-Espinoza, A.; Ordieres-Meré, J.; Martínez-De-Pisón, F.; González-Marcos, A. TAO-robust backpropagation learning algorithm. *Neural Netw.* **2005**, *18*, 191–204. [[CrossRef](#)] [[PubMed](#)]
33. Bagheri, H.; Rahim, N.A.; Liu, Z.; Lee, H.; Pesch, D.; Moessner, K.; Xiao, P. 5G NR-V2X: Toward Connected and Cooperative Autonomous Driving. *IEEE Commun. Stand. Mag.* **2021**, *5*, 48–54. [[CrossRef](#)]
34. Liu, Z.; Lee, H.; Khyam, M.O.; He, J.; Pesch, D.; Moessner, K.; Poor, H.V. 6g for vehicle-to-everything (v2x) communications: Enabling technologies, challenges, and opportunities. *arXiv* **2020**, arXiv:2012.07753.
35. Peter, J.H.; Elvezio, M.R. *Robust Statistics*, 2nd ed.; John Wiley and Sons: New York, NY, USA, 2009.
36. Peter, J.R.; Annick, M.L. *Robust Regression and Outlier Detection*; John Wiley and Sons: New York, NY, USA, 2005.
37. Rusiecki, A.; Kordos, M.; Kamiński, T.; Greń, K. Training Neural Networks on Noisy Data. In *Artificial Intelligence and Soft Computing. ICAISC 2014; Lecture Notes in Computer Science*; Rutkowski, L., Korytkowski, M., Scherer, R., Tadeusiewicz, R., Zadeh, L.A., Zurada, J.M., Eds.; Springer: Cham, Switzerland, 2014; Volume 8467. [[CrossRef](#)]
38. Zhang, Z. Parameter estimation techniques: A tutorial with application to conic fitting. *Image Vis. Comput.* **1997**, *15*, 59–76. [[CrossRef](#)]