Assessing Trust Level of a Driver-less Car Using Deep Learning

P Bhuvana Chandra
200070063
200070063@iitb.ac.in

S Venkata Sai Siddartha
200070074
200070074@iitb.ac.in

***Abstract*—The increasing adoption of driverless cars already providing a shift to move away from traditional transportation systems to automated ones in many industrial and commercial applications. Recent research has justified that driverless vehicles will considerably reduce traffic congestions, accidents, carbon emissions, and enhance the accessibility of driving to a wider cross-section of people and lifestyle choices. However, at present, people's main concerns are about its privacy and security. Since traditional protocol layers-based security mechanisms are not so effective for a distributed system, trust value-based security mechanisms, a type of pervasive security, are appearing as popular and promising techniques. In this paper, for the first time, we propose two deep learning-based models that measure the trustworthiness of a driverless car and its major On-Board Unit (OBU) components. The second model also determines its OBU components that were breached during the driving operation.**

***Index Terms*—Driverless car, trustworthiness measure, deep learning, intelligent transportation systems, On-Board Unit (OBU) components**

I. Introduction

While transportation systems play a critical role in the modern economy, the associated major problems affecting the world today are increased congestion and pollution, as well as the loss of life or incapacitating injury arising from fatal incidents. These problems directly impact national productivity and economy. Autonomous driving is viewed as one way of addressing these issues. Research has shown that autonomous vehicles can considerably reduce the congestion, accidents, and environmental pollution, and increase comfort and accessibility of car uses . As with other industrial revolutions prompted by Since the introduction of diesel engines, telephones, and electricity, some issues associated with driverless cars have arisen. For example, since the passengers' safety depends on autonomous driving, investigations were made to identify the key issues impacting the adoption of driverless cars. These studies clearly indicate the lack of people's trust in driverless cars because of concerns associated with privacy, security, expected performance and reliability.

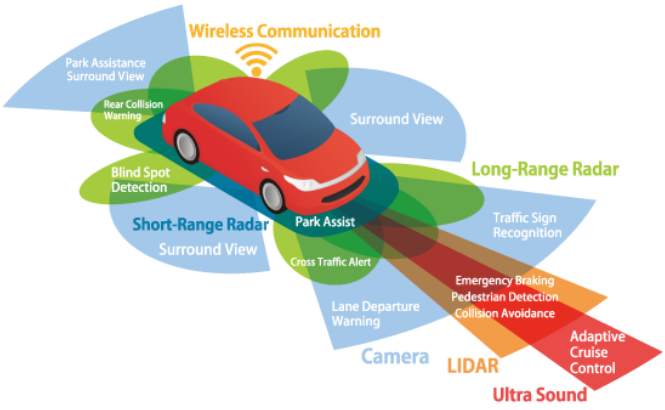

Fig. 1. Basic Overview of Autonomous Car Technology

Driverless cars are operating independently in a decentralized manner and thus less controlled by the traffic controller compared with the other infrastructure components like Roadside Units (RSUs). For these reasons, a driverless car is more exposed to the security threats than an RSU, demanding to leverage the trustworthiness value of an indi vidual driverless car for the proper functioning of an ITS. As a countermeasure, there are a few techniques available in the literature that assesses the trust level of a driverless car. These techniques can be categorized into three broad classes - (i) GPS location, (ii) safety or warning messages, and (iii) the combination of GPS location, safety message, and On-Board Unit (OBU) components-based techniques.

II. Assessing TrustLevel of Driver-less Car

A. Trustworthiness Measure of Driver-less Cars

Besides GPS, the OBU components of a driver-less car such as camera, LiDAR, radar, and acoustic Sensors play the main role in driving operations. There are different types of cameras, such as mid-range and long-range cameras, 3D video cameras, infrared cameras, and high-definition cameras. These cameras are mainly used to spot road signs (e.g., speed limit, right turn, left turn), hazards (e.g., incident sign, roadblock), pedestrians, and animals crossing in real-time. The modern cameras used in self-driving cars can produce up to 60 frames per second with millions of pixels per second.

LiDAR measures the distances between the vehicle and the surrounding objects. It can send 400,000 pulses in a second and is used to create 3D maps of the surroundings of an autonomous car for better and safer driving performance as it can operate in severe conditions like low visibility, fog, and even complete darkness. LiDAR uses infrared light pulses to detect the objects' shapes and distances surrounding the car. The safe operation of autonomous cars also depends a lot on the radar. Two types (short and long-range) of radars are used in multiple applications: collision detection warning, lane mark detection, detecting fixed objects, and adaptive cruise control. Both short (24 GHz) and long (77 GHz) radars play an essential role in detecting the safe distance between two vehicles by detecting the movable objects and the distance between two objects.

To measure the trust level of non-trustworthy vehicles, we need to simulate the effect of one or more compromised components and then record the traffic conditions they impact. For this reason, our proposed trustworthiness measure comprises two main phases - (i) building a conceptual model and (ii) gathering data through simulation considering one or more compromised components. These two phases are described below.

B. Conceptual Model for Assessing Trust Using Deep Learning

There exist many deep learning algorithms. This paper uses Deep Neural networks (DNNs) in our proposed trustworthiness measure. This is because DNN learning is based on a feedforward neural network, which can represent a complex function more easily and does not create a cycle. Since traffic management is a complex system, in this work numerical values represent traffic conditions and specific traffic of a given short timeframe (e.g., an hour) does not generally form a cycle, these aspects justify the use of DNN in our proposed project.

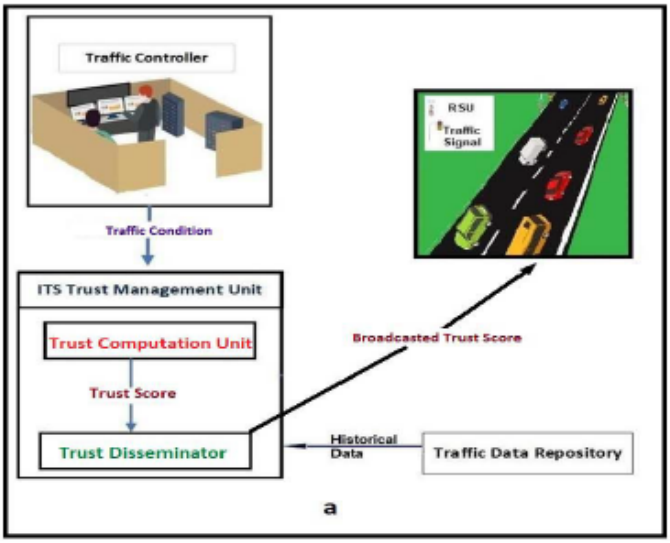

Fig. 2. Fundamental Components Of proposed technique

As shown in Fig. 2, the fundamental components associated with our proposed techniques are - (i) traffic controller, (ii) trust management scheme, and (iii) traffic infrastructure, including roads and highways. Our proposed trust management scheme receives traffic conditions (e.g., flow rate, vehicle speed, phase time, and distance between two vehicles) and historical traffic data from the traffic controller, calculates the trust score, and identifies whether a targetted car is compromised, and then which of its components are breached using two algorithms developed by deep learning techniques; and disseminates the trust score to the surrounding on-road vehicles and the traffic infrastructures.

III. Implimentation of the Model

A. Model Description

Here, we assume either central or regional traffic controllers will execute the trust management

scheme. Since regional traffic controllers are distributed across different geopolitical regions and coordinated by the central traffic controller, for obtaining the instantaneous traffic conditions of a specific area, computational and dissemination perspectives, regional traffic controllers are the good choice for running our proposed trust management scheme. We implement two models using deep learning - the first one (Model 1) assesses the trust level of a targeted driverless car and detects whether that car is trustworthy or not, while the second one (Model 2) determines which of its three components (e.g., Camera, LiDAR and radar) have been compromised, as well as their trust scores.

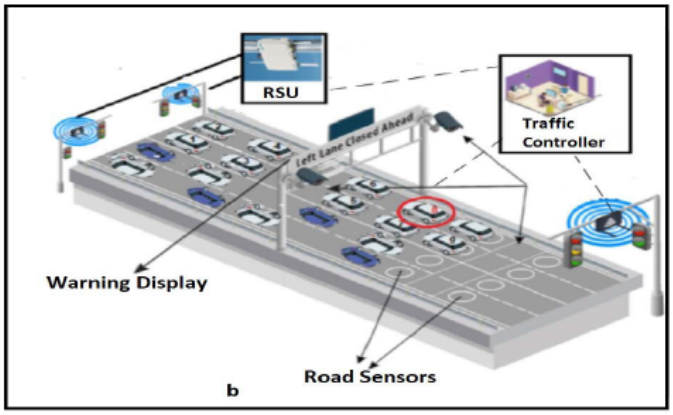

Fig. 3. traffic contextual scenario of a targeted vehicle indicated by a circle with its eight adjacent vehicles.

In this paper, for detecting the breach condition, we choose Camera, LiDAR, and radar as these are the most three important components are at least essential for the safe operation of a driverless car. Such identification requires leveraging the spatiotemporal situational information such as acceleration (deceleration), relative speed and distance of a targeted car considering its adjacent vehicles over a time window ΔT. This is because both the relative speed and distance of a compromised car with respect to the adjacent vehicles and their temporal changes are heavily affected.

We need to capture whether the behavior (e.g., acceleration/ deceleration, relative distance) of the targeted car resulted from current traffic conditions or the exhibited compromised behavior of the targeted car. To differentiate these two conditions, we need to consider the corresponding behavior of a sufficient number of cars in front and adjacent to the targeted one. The similar concept of using eight adjacent elements in calculating factors or measures is widely adopted in various domains, such as image processing and swam formation in robotics.

- Acceleration (deceleration) of a car i over time ΔT is computed as per the following equation,

$$a_i = \frac{\Delta V_i}{\Delta T} = \frac{V_i(T_1) - V_i(T_2)}{T_1 - T_2}$$

- The average relative speed of a car i with respect to its immediate preceding car j can be defined using

$$\bar{V_{ij}} = 0.5((V_i(T_1) - V_j(T_1)) + (V_i(T_2) - V_j(T_2)))$$

where j=x—x is i's preceding car for all $1 \leq x \leq 8$.

- The change of relative distance (position) of a car i with respect to the car j over time can be calculated using

$$\Delta X_{ij} = X_{ij}(T_1) - X_{ij}(T_2)$$

where $X_{ij}(T_1)$ denotes the relative distance between the cars i and j at time T_1

- Finally, the joint impact of a car's changing speed and the relative distance between two cars, i and j, on a lane can be captured using the following equation

$$\tau = (V_i(T_1) - V_i(T_2)) * (X_{ij}(T_1) - X_{ij}(T_2))$$

The mentioned four properties of the eight adjacent vehicles are considered as the feature set for the deep learning model (Model 2) to find the faulty sensors of the targeted car and their trustworthiness scores.

In the simulation, we monitored the impact on traffic conditions (vehicle speed, flow rate, phase time, and distance between vehicles) when one or more OBU components failed to work properly. Vehicle speed, flow rate, phase time, and distance between two cars for all nine cars were recorded periodically every 30 seconds. The simulation environment compromised one or more OBU components (e.g., LiDAR, radar, camera, acoustic sensor). For compromised situations, we created the following four different scenarios in the simulation:

- **Scenario 1**: We added an object on the street and monitored what happened if the camera,

LiDAR, or radar failed to detect the object. Different tests were done with the camera being unable to see the object. LiDAR could not detect the shape of the object, while the radar failed to assess the distance between the car and the object. We observed the change in traffic conditions if the camera failed to detect the object, but the LiDAR and radar were able to detect the shape and distance.

- **Scenario 2**: What happened if the camera fails to detect the speed limit was tested. We also investigated whether the LiDAR or radar works accurately when the camera cannot recognize the speed limit. When the camera misreads the speed limit. For example, the camera reads the speed limit as 90, where the actual speed limit is 60; if LiDAR or radar works fine, the car will not exceed the speed limit in case a safe distance is at risk. If LiDAR or radar fails along with the camera, there is a chance that the car may collide with the preceding vehicle.

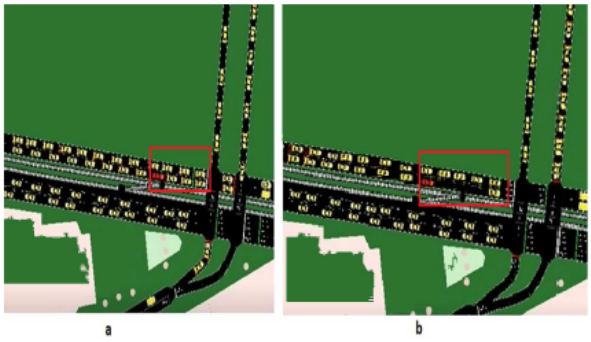

Fig. 4. Change of traffic condition due to camera fault in a simulation environment- (a) normal traffic condition and (b) traffic condition for the compromised camera. Normal and change traffic condition for the compromised camera are highlighted with a rectangle having red border. Note, the traveling direction of the targeted car (red car) is towards the intersection

- **Scenario 3**: We transmitted false signals to the LiDAR showing an obstacle on the road where there was no obstacle actuality. We monitored the impact if the LiDAR only falsely detected the object and also monitored what impact it had if the radar or camera failed at the sametime.
- **Scenario 4**: The radar and LiDAR's impact on wrongly detecting the gap between two vehicles or objects (e.g., lane divider) was investigated. In this case, two different outcomes may happen: (i) the observed vehicle could make a sudden stop or reduce speed falsely by sensing the car in front is too close, and (ii) the monitored vehicle could collide with the car in front by misjudging the safe distance between them.

Both of our proposed models use DNN based deep learning algorithm implemented using Keras libraries on Python. Also SVM and LSTM were implemented for comparision purpose.

B. Dataset

For this project, VicRoads' real traffic data available for five different intersections in Melbourne CBD were used to represent the original (uncompromised) traffic condition and develop the real traffic scenario in the simulation model. The intersections are- (i) Lonsdale and Russel street, (ii) Collins and Kings St, (iii) Elizabeth and Latrobe St, (iv) Collins and Swanston St, and (v) Flinders and Swanston St Melbourne. We selected the peak-time hourly average phase time, average vehicle speed, and flow rate of Mondays (Mondays with public holidays were not selected) of 2016, 2017, and 2018. The selected area (Melbourne CBD) has two peak periods creating traffic congestion.

The first one is the morning peak period (08:00 am - 09:30 am), while the second one is the evening peak period (04:30 pm - 06:00 pm). The morning peak period is crucial because most of the offices and schools start at 08:30 am - 09:00 am. Victoria has morning school speed zone time with reduced speed limit from 08:00 am to 09:30 am. The time window we selected is from 08:00 am to 10:00 am as any incident or disturbance to the normal traffic flow in the morning peak period is likely to make the most impact. We have gathered data from VicRoads from 2016-2018, natu rally these data contain diverse weather conditions and a wide spectrum of incidents and events occur on road. As stated before, the morning peak time is from 08:00-09:30 am, our data was collected from 08:00-10:00 am, that covers morning peak (08:00-09:30 am) and off-peak (09:30-10:00am) time. This gathered traffic data were embedded in our simulation

IV. Experimental Results

A. Experimental Setup

For both models, we implemented a DNN with three sequential hidden layers having 1024, 512, and 256 feed forward densely connected nodes, respectively. Model 1 which is a binary classifier uses traffic flow, vehicle speed and phase time as features, thus requiring three input nodes, and a single output node predicting whether a car is trustworthy. A total of 530 samples were used, drawn equally from both the classes. On the other hand, Model 2 computes a set of four features (details are provided in Section III-A) derived using Eq. (1)-(4) for each car in platoon of eight vehicles, thus, a total of 32 nodes are used in the input layer. The model employs three output nodes to predict the trustworthiness of individual sensors, namely cameras, LiDAR, and radar. A totalof 285 samples were used in this model, with 55percent positive class (trustworthy) samples for each binary classification task discussed below. For both models, the available dataset was split in the ratio of 80:20 to generate the training and test samples.

The dropout rates for the hidden layers were chosen as 0.5, 0.4, and 0.3, respectively. The 'Adam' optimizer was used with a learning rate of 0.01, and 'binary cross-entropy' was chosen as the loss function in training the models. ReLU (rectified linear unit) and 'sigmoid' activation functions were utilized at the hidden layers and output layer, respectively. Finally, the DNN was trained for 200 epochs with a batch size of 5 in each epoch. The binary outputs from output layer nodes were used as the class labels to determine the accuracy and related performance metrics, whereas their actual outputs (prediction probabilities) were used as the trustworthiness scores of the vehicles (Model 1) and of their three OBU components individually (Model 2). Note that in Model 2, each output separately determines the trust score of camera, LiDAR and radar in the range of 0-1, the use of binary crossentropy ensures that. The performance metrics used for model evaluation in this work such as precision, recall, and area under the ROC curve are widely used in literature.

B. Performance and Results of the model

For measuring the capability of our models (the true positive rate or recall) in distinguishing between the classes, Receiver Operating Char acteristic (ROC) curves are generated and AUC is measured. The precision-recall (PR) curves are also generated to show the accuracy of our model in terms of precision for the different levels of recall. The results obtained are then compared with two other machine learning models, namely Support Vector Machine (SVM) with polynomial kernel and Long Short- Term Memory (LSTM) neural network model with 100 hidden layer units, to further substantiate our models' performance in achieving higher precision over different recall values. The results are illustrated in Fig. 5,6 for Model 1 and following figures for Model 2

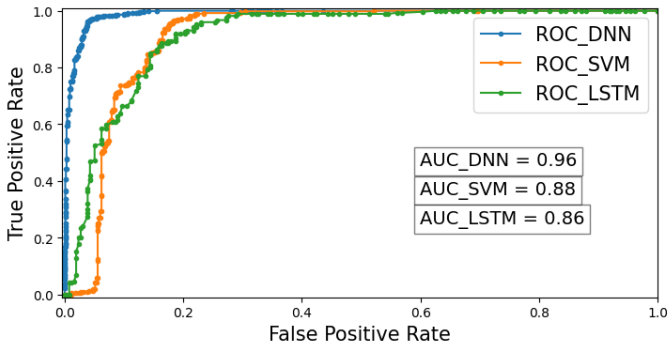

Fig. 5. The performance for classifying trustworthy and not-trustworthy vehicles - ROC curve with ROC-AUC score

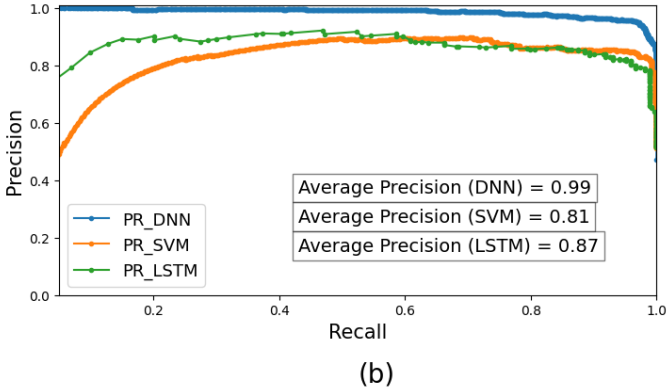

Fig. 6. The performance for classifying trustworthy and not-trustworthy vehicles - Precision-Recall curve with average precision score.

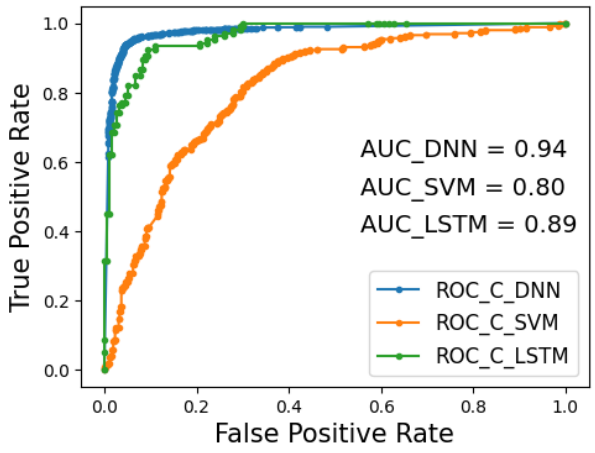

Fig. 7. ROC Curve and ROC-AUC score for the individual OBU components' trustworthiness classification task for - Camera

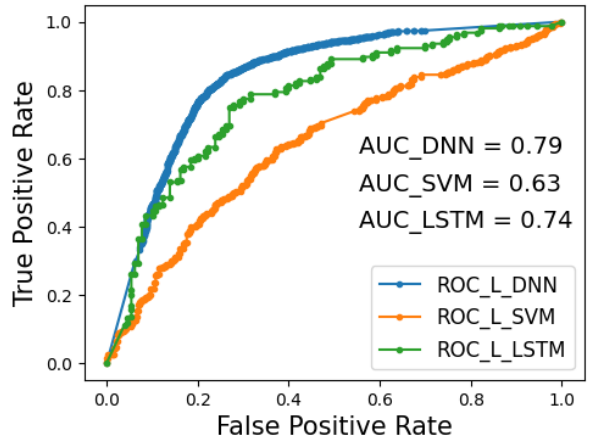

Fig. 8. ROC Curve and ROC-AUC score for the individual OBU components' trustworthiness classification task for - LiDAR

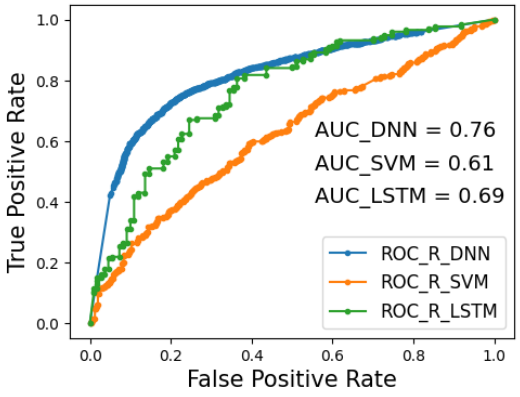

Fig. 9. ROC Curve and ROC-AUC score for the individual OBU components' trustworthiness classification task for - Radar

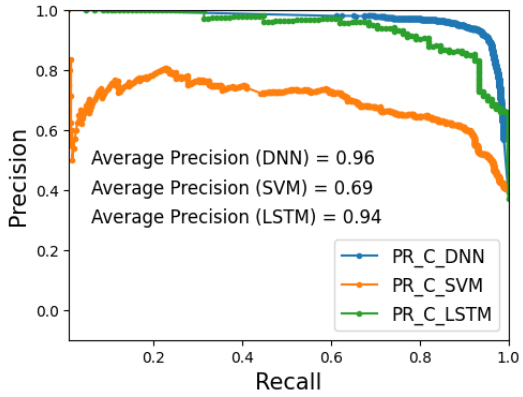

Fig. 10. Precision-Recall curve and average precision score for the individual OBU components' trustworthiness classification task for - Camera

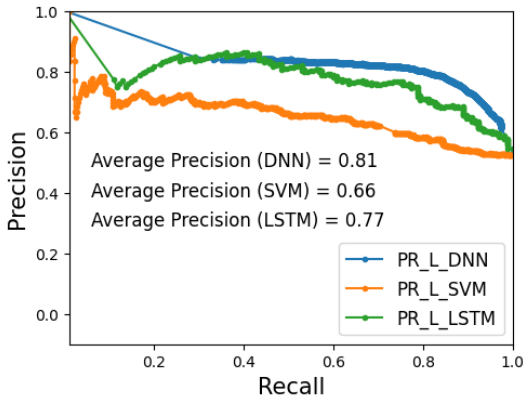

Fig. 11. Precision-Recall curve and average precision score for the individual OBU components' trustworthiness classification task for - LiDAR

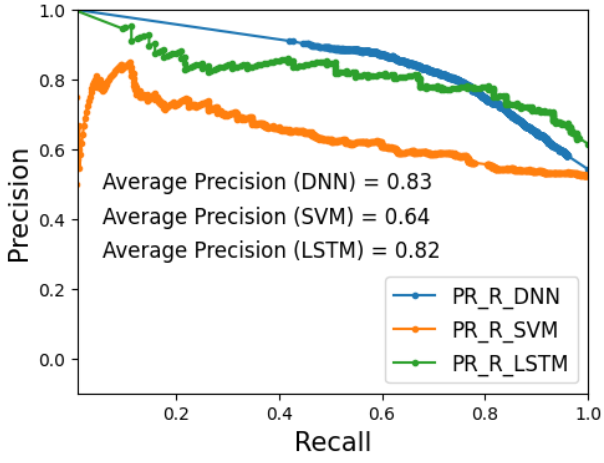

Fig. 12. Precision-Recall curve and average precision score for the individual OBU components' trustworthiness classification task for - Radar

Finally, the F1-scores were computed for the models to determine the balance between the precision and recall that the models can achieve. F1 score for Model 1 is reported as 0.96, whereas those for the SVM and LSTM models are 0.89 and 0.87, respectively. F1 scores obtained from Model 2, SVM, and LSTM for the three OBU components (camera, LiDAR, radar) are (0.93, 0.81, 0.78), (0.75, 0.64, 0.63) and (0.86, 0.76, 0.74), respectively. The F1 score for detecting a faulty camera is very high, whereas the other two components also yielded promising results.

V. Conclusion

The traffic system is a decentralized, complex, and dynamic entity, demanding a machine-learning model to encapsulate its unpredictable processes. In assessing trust levels for driverless cars, we introduce two models utilizing the DNN (deep neural network) algorithm. The first model determines the trustworthiness of a specific car, accompanied by its evaluated trust level. Simultaneously, the second model identifies compromised elements among LiDAR, camera, and radar, calculating their trust values. Employing real traffic data from VicRoads for normal conditions and simulation data from a SUMO-developed model for compromised scenarios, our results demonstrate the accurate assessment of trustworthy scores by our proposed models. Furthermore, the effectiveness of deep learning models can be enhanced with more extensive training datasets. As smart city initiatives advance, the autonomous collection of significant traffic and on-road event data through Intelligent Transportation Systems (ITS) at various driving conditions and locations holds promise for refining and advancing our model.

VI. References

1) G. Karmakar, A. Chowdhury, R. Das, J. Kamruzzaman and S. Islam, "Assessing Trust Level of a Driverless Car Using Deep Learning," in IEEE Transactions on Intelligent Transportation Systems, vol. 22, no. 7, pp. 4457-4466, July 2021, doi: 10.1109/TITS.2021.3059261.
2) K. E. Brown and R. Dodder, "Energy and emissions implications of automated vehicles in the U.S. energy system", Transp. Res. D Transp. Environ., vol. 77, pp. 132-147, Dec. 2019.
3) A. Chowdhury, G. Karmakar, J. Kamruzzaman and S. Islam, "Trustworthiness of self-driving vehicles for intelligent transportation systems in industry applications", IEEE Trans. Ind. Informat., vol. 17, no. 2, pp. 961-970, Feb. 2021.
4) K. Kaur and G. Rampersad, "Trust in driverless cars: Investigating key factors influencing the adoption of driverless cars", J. Eng. Technol. Manage., vol. 48, pp. 87-96, Apr. 2018.
5) J. Cui, L. S. Liew, G. Sabaliauskaite and F. Zhou, "A review on safety failures security attacks and available countermeasures for autonomous vehicles", Ad Hoc Netw., vol. 90, Jul. 2019.
6) Reference Notebook Model Code