العدد الرابع والثلاثون
تاريخ الإصدار: 2 – آب – 2021 م
www.ajsp.net
ISSN: 2663-5798

"Measure the Difference between Current and Expected Performance using Machine Learning (For Autonomous Driving Systems)"

Researcher:

Eng. Ahmad Abdullah Alghamdi

MSc Computer Security and Forensics and Ph.D. student in Computer Science

King Abdulaziz University

Abstract:

Measuring the performance of the program is one of the most important stages of testing, as we first test the program to detect problems that may occur to the user in the future, and these problems are the difference between the performance of the program and what is required of it, and we search for the difference and from it the program is evaluated and to measure this difference we need to measure Performance with measurable and comparable measures The problem with this is that we need to measure something qualitative, which is the quality of the program with a quantitative thing, which is its standards, and if the quality is calculated by reducing the difference between what is expected and what is in place, the problem arises in how to define what is expected, especially since the subject goes beyond defining what It is required because the program can have achieved all the points required of it in the current circumstances, but what happens if the circumstances and capabilities change in the future, of course, there are many methods and models that help in evaluating the program's performance in the future through different tests, but it is not possible to measure all possibilities as well. Using the human element in conducting this research increases the time and cost, especially if it deals with a large number of possibilities and special circumstances

Therefore, in this research, we use more effective methods to detect future problems of the system through the use of machine learning to reach the best performance, speed and accuracy as it is less in cost and time.

I. Introduction

We do not live in an ideal world free of errors and there is no program or development flawless of course the program is tested in earlier stages and make sure that there are no problems or explicit errors as it is assumed that it passed a performance review and ensure that it achieved what is required of it and also that there are no errors in Operating Here we search for problems that may occur in the future or appear in certain circumstances, they must be monitored in order to either solve them or alert the user not to use the program in these circumstances and all that has been discovered the largest number of problems and solved as the quality of the program increased and to reach the best level and discovery .

The largest amount of expected problems should have been to determine the testing methodology and repeat the process of measuring performance through different circumstances. This problem appears in most large programs, especially those that rely on data processing. When testing this type of program, it is difficult to apply all the cases that the program can go through and therefore it has been used Machine learning to ensure that all cases are covered to obtain a clear picture of the project in any situation and the probability of losing any activity is low and this study clearly defines the methodology for testing the program and measuring the current performance difference with the performance What is expected in all circumstances, quality assurance and access to all work requirements and in order for the system's evaluation to appear clearly to the program's accomplishments, as the machine learning system is designed to train itself and benefit in evaluation, research and discovery of problems as in the system because it is automatic and does not depend on human intervention that guarantees the application of standards With minimal error, the Autonomous Driving Program has been identified as a practical example to demonstrate the problem and the solution.

In this research we will deal with the dimensions of the problem and its treatment. In the second chapter, we will review some studies that dealt with the idea of integrating and using machine learning in the program's testing system, what it succeeded in achieving and what it could not implement. The third chapter will review the program's test measures and their types and review a set of testing strategies used Which needs to be followed through the machine learning system, and in the fourth chapter we will review the system in question, self-driving where we use it to implement the idea of research and present the problem and the solution in the fifth chapter we will review the methodology in which we review the problem and treatment and how we used machine learning and how this technology manages and implements the aforementioned test processes Finally, the results of the research and our findings will be discussed.

II. Related work

Several previous studies have dealt with the issue of using performance testing and evaluation through machine learning technology, especially since the autonomous driving system in and of itself works through it and during continuous research and access to several studies that directly benefit the subject of the research. We will deal with several researches as shown below.

The first study [3]

In this study we explain to you the most prominent problems and the way to solve them as it dealt with verifying the inputs and ascertaining the values coming from the various sensing tools and this is through a mathematical assessment based on a set of mathematical models and the use of machine learning to reach good results and the most prominent contribution to this study As follows:

- Significantly reduce the effort in the treatment process
- Provide high safety in using the system by verifying the received input
- Provide clear means and direct tools through which the consumption of processing tools is reduced

Among the most prominent dilemmas that have been noticed from our point of view, which may be an obstacle to implementation and highlighting them:

- Expensive and intended by the large number of procedures and treatments, which need several stages to apply, and also requires human resources to conduct the registration and analysis process
- Difficult to implement, because it requires several stages of treatment to reach the desired mathematical model
- Despite its reliance on DNN to reduce problems, the use of mathematical models remains a major obstacle in the application of this study.

The second study [4]:

This study is based on solving the problem of the absence of Global Navigation Satellite System (GNSS) signals coming from the satellite and what affects the system of using autonomous mobile vehicles, as there may be a loss of signal during traffic in different environmental and climatic conditions such as canyons, bridges and tunnels. It has used the learning system. The automated Input Delay Neural Network (IDNN) and among the most prominent features of this study:

- The ability to learn error deviation because it is more powerful and fast mathematically
- We can complete the missing data from the satellites
- The study focused on difficult cases such as brakes, turns, sharp and successive turns, and speed, and reached good results that reached 89.55%

Among the most prominent defects from our point of view in this study are as follows:

- Its reliance on fixed models that it designed and tested
- · You have not developed flexible variables that can be used for a huge number of tests, which is a big obstacle. Despite adding difficult cases, there are many conditions and cases that the study failed to address and test.

The third study [4]

This study focused on studying the durability and safety characteristics of all sensor systems by inserting them into the system based on Deep neural networks (DNNs) technology. The most prominent contribution of this study is

- Contributing to finding solutions to safety dilemmas when the system receives incorrect or confused data by ensuring the context of the inputs by relying on images and equations that can be compared when exposed to specific events or conditions that may affect the data received from the system's sensors or satellites.

Among the most prominent dilemmas, despite the success that he achieved from the use of DNN in determining access to numbers and simple variables, the study is weak in covering non-local aspects of work and its connection with external systems such as variables or the system’s association with a global review system.

ISSN: 2663-5798
www.ajsp.net

The Fourth Study [5]

This study deals with the use of machine learning in predicting the expected paths of pedestrians and crowd control, as it focused on the idea of using a machine learning system in the autonomous driving exercise to understand and predict pedestrian movement to avoid problems and partially succeeded in reaching a repetitive model through which we can add movement and anticipate reaction and from it increased. The power of the analysis system

However, the study had some flaws, including that it did not address the testing of unique scenarios or unexpected circumstances, as it relied on developing the awareness system for self-driving and did not focus on unexpected situations such as sudden, irresponsible human behavior. Solutions were not included or addressed, their limits, and training on what Will be with her

III. Program test metrics

Program testing metrics are the quantitative measures used to assess the progress, quality and productivity of the program testing process, the aim of which is to improve the effectiveness and efficiency of the testing process and help in making better decisions. This is done by providing reliable data for the testing process and we are quantitatively determined with what the system has of a certain value, for example. Compare the vehicle's weekly mileage with the ideal mileage recommended by the manufacturer.

The metrics work to improve the efficiency and effectiveness of the program testing process, whether it is material, capacity, amount, dimensions, and size, we simply cannot evaluate and improve what we cannot measure as we benefit from measurement in making decisions and predicting faults and understanding capabilities and capacity and how we can improve it and the ability to determine the type of technology or Units that the program can deal with. The measurement process is divided into three stages [1].

- Measuring the performance of phases and processes to understand and evaluate the performance of each process determining the program development life cycle
- Measuring the level of the product where the quality and efficiency of the product can be assessed
- Measuring the feasibility of the project as the efficiency of the work team and the testing tools used are measured

Of course, determining the required and useful test metrics is one of the most important steps, as you must define the goal of the measurement, what we will benefit from and the cost that we will pay, and compare it with the return from it. During the application of the equations from which the test report comes out, where the percentage of completion of the test and the extent of coverage of the program are shown

Standards differ according to the type of program used and can be classified as follows [1]:

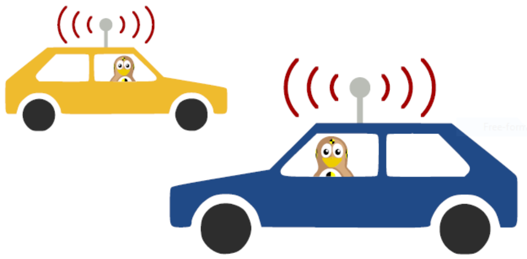
- Quality
- Priority
- Status
- Severity
- Implementation

In order for the measurements and tests to be carried out with high efficiency, we design scenarios for the tests in order to ensure that all the tests are covered and determine what each test will do. It also shows the cost of the test and the return on it and the ability to organize and arrange the work. And the schedule of it and the possibility of studying the overall performance of the program. The scenario will include a set of tests such as

- Positive tests, which are carried out during the provision of a set of reliable data for the required performance, compare this data with the data resulting from the program’s performance, and the more identical, the better the quality of the program
- Negative tests, which is the provision of invalid or correct data as input, and the performance of the program is verified upon receiving it and ensuring that the program is not disrupted and stable despite this data
- Limit tests, which is the introduction of a set of data and values greater than the specified limits, whether in size or number of operations, and evaluating the performance of the program in dealing with these circumstances

IV. Autonomous driving

Self-driving work requires four jobs [8]:



- Localization
- perception
- planning
- system management

A. Localization

It is considered one of the most important functions to assist in self-driving because it locates the unit and is used to set this function.
GPS devices are widely used to identify the location, but it cannot be used directly as exceptions can appear in reading due to problems with satellite communication and after conflict in some places. Integrate GPS data with the vehicle's various sensors as it gives speed, time and directions data and is matched with digital maps to provide accurate and reliable location information.

B. perception

It is a service that provides information that has been sensed to the surrounding environment through vision, sensation and measurement systems, which can monitor neighboring units, classify their type and determine their nature, and use a set of functions to identify the fixed and moving elements around the car and in the end sends the data obtained from all the measurement units and cameras attached to it.

C. planning

It is used to provide alternatives and options under different situations and divides this function into three steps

1. Orientation: It directs the unit after it gets the location and interface and uses digital maps to reach the fast, safe and best path according to the recorded data
2. Behavioral thinking: It uses knowledge and perception information to implement automatic movement during the course of the car, avoiding collisions and static and moving obstacles.
3. Driving: It uses the data provided by digital maps to determine the speed allowed by the traffic rules and regulations that must be adhered to

D. system management

It is a function of the general leadership of the unit, of which there is self or manual administration, or there is a system for calculating safety limits and temporary suspensions that work in three sections

1. The driving position, which is a means of determining whether it is self-driving, manual, or manual with control, and whether it is possible to switch between them
2. Fault management, which is a review of the driving method, determination of the size of the failure, and the possibility of intervention to control and adjust by changing the driving position or stopping the unit
3. Work interface, which is the functions used for driving, whether it is self-driving or manual

V. Methodology

After reviewing many studies and research on measuring the performance of autonomous driving and giving the best results after testing it, despite the tremendous growth of autonomous driving technology and what awaits it in the future, it faces an obstacle in the signal reception process that directly affects the transmission of data across the network as the system relies on transmission A huge group of data in the form of messages received across layers between management and self-driving, which in turn depend on analysis and results by the administration and access decisions necessary to contribute to the distribution of the load and accelerate the movement and the problem appears directly when the communication between the administration and self-driving is interrupted when exposure to natural disasters or traffic By tunneling or passing through neighborhoods used for insulators, saving energy that causes poor reception of signals, either

ISSN: 2663-5798

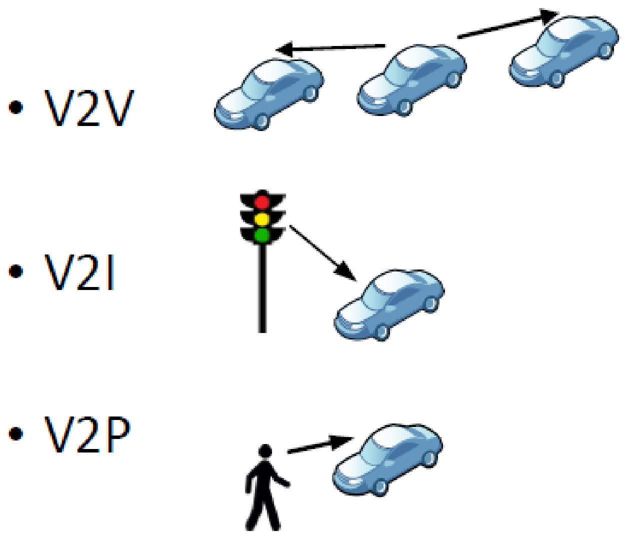
from the car or the administration side, which causes the inability to analyze data, test them and obtain the correct results, and a conflict will also appear in car decisions, which causes conflicting results and decisions that affect the movement of vehicles directly. It has a great reluctance to use and safety self-driving, which results in difficult testing and prevention S system performance as we cannot simulate these signal outage problems and situations.

From this research we will review the process of integrating treatment and testing into the machine learning system and it can manage the testing process and design tests for many dilemmas where it repeats the measurement process with modification in the simulation elements and from it the system can test many cases and conditions that the self-driving can go through and from which the system can discover problems The future that can happen and teach self-driving how to deal with it and will use the neural network technology because of its flexibility and the ability to invent different cases and conditions

And the development of a test system based on the use of high-resolution three-dimensional maps by integrating digital maps with the input of amplifiers and simulating sensors in order to carry out the testing process through the technology of deep neural networks that are fed with a set of different scenarios to test the autonomous system, and we will use the available panoramic images for free. On google street view and digital maps from the open street map website and what it provides in terms of data and specifications of the roads in terms of their dimensions and the size of the permissible speed. Through machine learning, we use inputs to design different scenarios in which to display the many situations that an autonomous driving program could go through

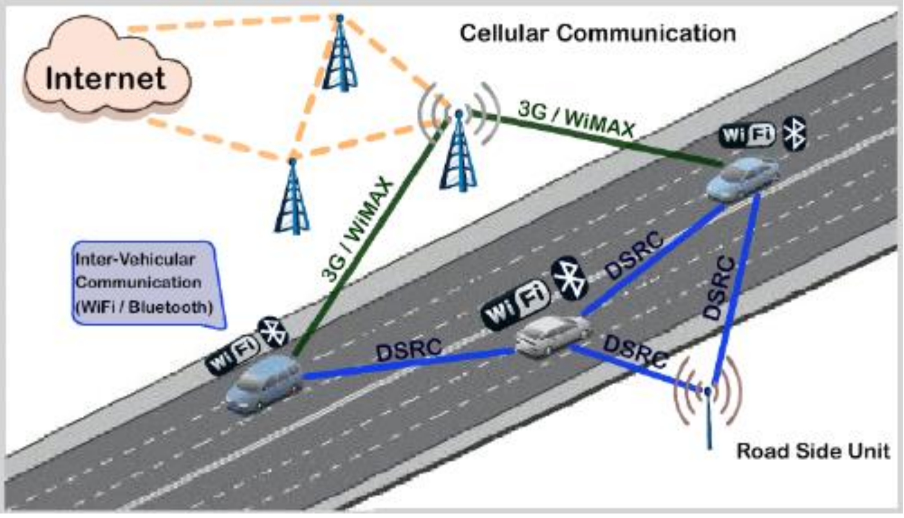
The form consists of :

The data level [8] It is responsible for data collection, transmission and processing and is represented in vehicles and signals. This level collects data and encapsulates packets and redirects them to the management level. This level is supported by different types of communication V2V & V2i & I2I.



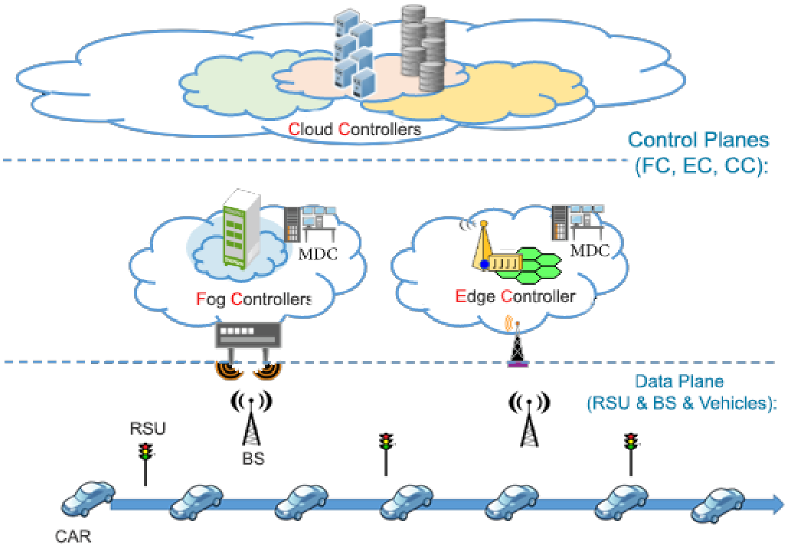
It uses different connection types to transfer data such as

- DSRC
- WIFI
- LTE



The level of management [8], which is responsible for allocating potential resources based on the services they provide, which makes the decision to direct specific traffic. There are three levels of management.

- Cloud level control
- Edge level control
- fog level control



Self-driving training

In order to be able to train autonomous driving on how to behave and work in different conditions, you need to test the driving style by simulating the environment surrounding the cars and entering all the data you need to perceive the surrounding environment in order to be able to plan the path and control. We have used high-resolution three-dimensional maps and simulated sensor data where we do By feeding the perception of self-driving with data obtained from the deep neural network, we prepared it with data and an approach to reality. This is the cheapest and best way to test the performance of a self-driving program [2].

We first use network simulation tools and we used sumo simulation, which can create a model that contains the necessary data for the network to be tested, as it provides a simulation of the car, its specifications, the road, the number of lanes and areas, which you pass using a set of images used from Google, from which we can simulate the real world and then enter Scenarios for data or not sending one of the sensor data.

Driving begins by discovering if there is a specific path to the next point and from it we make prediction through the simulation system where we enter the current address of the panorama image and what is in line with the road and then we optimize the choice of the directed address as there should not be any fixed object in front of the panorama (edges of the road , Endings). Classified vehicle angles are calculated based on the OSM lane characteristics and the turns applied in the GSV panorama. Note that real posters are not necessarily accurate because the calculation is based on one assumption that the vehicle that collected the GSV panoramic photos is traveling in a straight direction in most lanes. However, we found that our model predictions regarding the address angle using the generated grid showed a good level of effectiveness.

As for dealing with different intersections, it is one of the most common scenarios in urban driving environment where the human driver needs to select the high-level command in order to follow the planned driving route. We refer to the following features of the road to deal with an intersection: Is there a forward intersection? And the distance to the intersection. We follow a similar procedure for mapping intersections as they first locate them on the OSM by finding the common vertices and then mapping them using panoramic GSV images. The closest intersection distance is used from the vehicle's current address. We use GSV images and ground truth computed images to train our model to predict distance but the forward inter-section? It is a binary classification problem. After comparing and evaluating manually with the pictures taken from the samples, we chose a certain distance of 30 meters and less, such as the intersection in front of us. Likewise, we choose that intersection distance over 100 meters as a false prediction in order to provide clear distinction for the classification task.

Obviously, determining or measuring the distance between an intersection when approaching one of them however, the error of estimation grows for our model when predicting a scene at the intersection. The visual inputs at intersections are usually not as organized as sections of a public road. The open view of unstructured terrain confuses the neural network-based model because only one snapshot of the image is presented. We believe that prediction results can be improved using a memory-based model [2].

It is somewhat important for autonomous driving to define a multi-lane driving context especially in urban driving or highway driving in order to advance path planning and maneuver driving control. The number of lanes property determines the number of lanes in the current driving route. Where only one-way roads are included in the training data due to the inconsistency of two-way roads when looking in the driving direction.

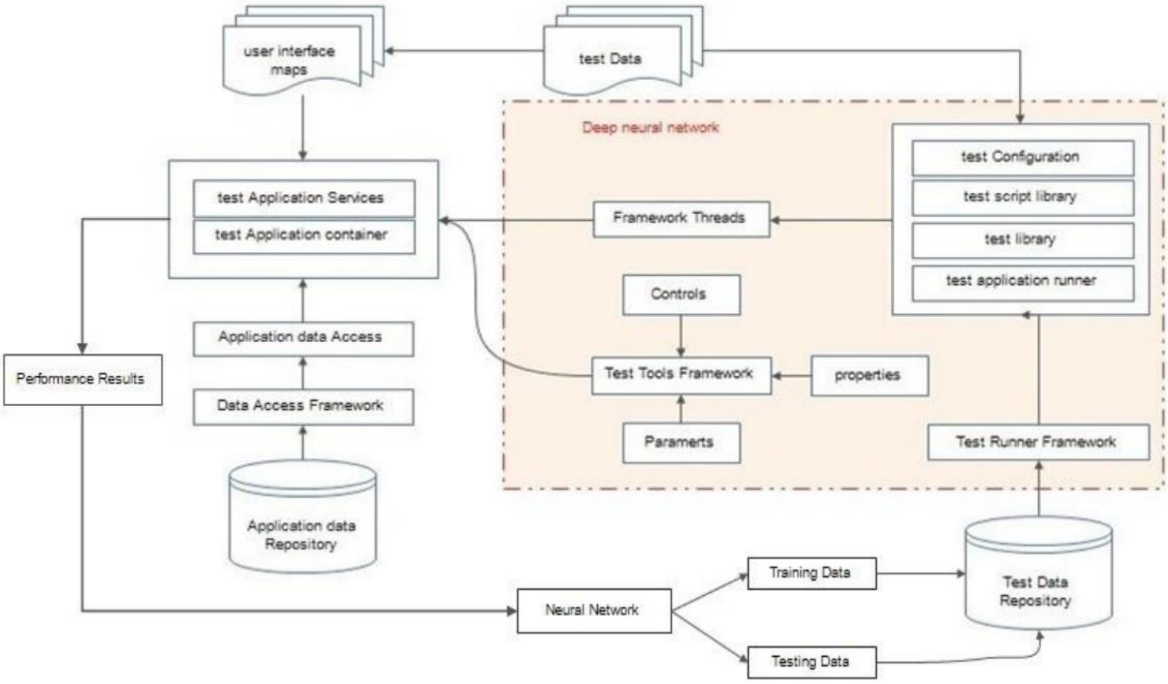
A side issue of lane counting is identifying potential roadside bike lanes. If there is a bike lane based on the given panoramic photo.

Design the training form

Initially we follow the paths specified by the map simulator as we adjust the parameters to find the most efficient structure and produce the best accuracy for validating training and we ended up with a network of five wrapping layers and three layers that are completely connected. We found that 3 different scenarios used for each wrapping layer. And we also applied padding and maximum clustering to preserve the input size and spatial resolution, respectively, through the warp seams. The structure of the output layer depends on whether the model is intended for regression or binary classification. For regression, we used a single core output layer without an activation function.

Before training, the images were normalized by changing the pixel values of float and dip by 255. The random images were increased by applying a 22 ° rotation, shifting width and height by 0.2, shearing 0.2, and zooming by a factor of 0.2. Images were trained in batch size 32 and 50 bags. The steps for each period for both training and validation are based on the total number of images. The corresponding images and tags were also randomly mixed during the training phase. Their model was pre-tested on a database of places while our model used a random configuration without pre-trained weights. The model tends to converge after "a few thousand iterations". Instead, our model used only 50 epochs and fewer than two hundred steps per era since we are using a much smaller data set.

Models are evaluated on the same images and compared to expected results. We focused on classification tasks as we found it difficult for humans to usefully measure angles or distances from low-quality images. Thus, we did not take into account the distance to the intersection and the heading angles were inferred to the binary classification by asking humans to predict whether the image showed negative rotation (left turn with respect to the road) or positive rotation (turn right with respect to the road).



To see how well our model succeeded in generalizing to data collected in different geographic locations, we took advantage of panoramic GSV images and crossed validation by comparing prediction on GSV images with a model that was trained on the dataset and prediction on GSV images and then used neural network models For address, intersection distance, number of lanes trained, it appears that increasing the data applied to the training data may not be strong enough for driving scene training since it is relatively difficult to increase buildings and other features (such as trees, grass, and sidewalks) near the road. We made sure that the test images were never used during training and validation of the models. However, the same set of test data was preserved.

VI. Conclusion

In this research we have dealt with the problem of testing and reviewing the performance of systems that rely on processing big data, as we have corrected that there is a problem in testing these programs because they are exposed to many different scenarios in the future that are difficult to anticipate in the traditional way. Evaluate the performance of the program with each of the scenarios of special events that may occur with the expected performance of it, which we seek to be better by testing the system, returning to training it for appropriate performance, and then re-experimenting again until the program passes all the scenarios with efficiency and serious results.

To show the idea, we have presented a clear practical example of this type of program, which is the self-driving program, and what challenges this program may go through resulting from its exposure to different circumstances, whether from interruption of communication with direct management or that one of its systems entered error data resulting from either an error reading or A defect in the unit or a delay in transmitting the reading, and from these cases the extent of the damage that results from the inability of the program to deal with these exceptional circumstances appears.

Here, the problem appears if the company focuses on testing the program in all scenarios first, it will not be able to implement all the scenarios presented, and it has also increased the duration of training and the cost and required more human resources to reach satisfactory results. It has been proven that when using the deep neural network system it can develop test models and is able Implementing many scenarios, making measurements on them, then returning to the program with the results to correct the concepts and methods of dealing and return again for the application to reach the best performance and save a lot of time, money and the need for dense human resources.

VII. REFERENCES

[1] J. Janai, F. G¨uney, A. Behl, and A. Geiger, "Computer vision for autonomous vehicles: Problems, datasets and state-of-the-art," arXiv preprint rXiv:1704.05519, 2017.
[2] W. Schwarting, J. Alonso-Mora, and D. Rus, "Planning and decision-making for autonomous vehicles," Annual Review of Control, Robotics, and Autonomous Systems, vol. 1, pp. 187–210, 2018.
[3] Omer Achrack, Ouriel Barzilay, Raizy Kellerman, "**Multi-Loss Sub Ensembles for Accurate Classification with Uncertainty Estimation**"
[4] Uche Onyekpe, Vasile Palade, Stratis Kanarachos,
"Learning to Localise Automated Vehicles in Challenging Environments using Inertial Navigation Systems (INS),"
[5] Chih-Hong Cheng, Rongjie Yan, "**Continuous Safety Verification of Neural Networks**"
[6] Simone Zamboni, Zekarias Tilahun Kefato, Sarunas Girdzijauskas, Noren Christoffer,
"Pedestrian Trajectory Prediction with Convolutional Neural Networks"
[7] P. Gaussier, S. Moga, M. Quoy, and J.-P. Banquet, "From perceptionaction loops to imitation processes: A bottom-up approach of learning by imitation," Applied Artificial Intelligence, vol. 12, no. 7-8, pp. 701–727,1998.
[8] M. Bojarski, D. Del Testa, D. Dworakowski, B. Firner, B. Flepp,P. Goyal, L. D. Jackel, M. Monfort, U. Muller, J. Zhang et al., "End to end learning for self-driving cars," arXiv preprint arXiv:1604.07316,2016.