# Multimodal-AI based Roadway Hazard Identification and Warning using Onboard Smartphones with Cloud-based Fusion

**Final Report**

By


Yunyi Jia
yunyij@clemson.edu
Clemson University

Gurcan Comert
Gurcan.Comert@Benedict.edu
Benedict College

Mayuresh Bhosale, Longxiang Guo, and Leo Hu
Clemson University


**July 2023**

**Center for Connected Multimodal Mobility ($C^2M^2$)**


*200 Lowry Hall, Clemson University*
*Clemson, SC 29634*

## DISCLAIMER

## ACKNOWLEDGMENT

*This study is partially supported by the Center for Connected Multimodal Mobility (C2M2) (USDOT Tier 1 University Transportation Center) headquartered at Clemson University, Clemson, SC. Any opinions, findings, conclusions, or recommendations expressed in this study are those of the authors and do not necessarily reflect the views of C2M2, and the official policy or position of the USDOT/OST-R, or any State or other entity, and the U.S. Government assumes no liability for the contents or use thereof.*

**Technical Report Documentation Page**

| 1. Report No. | 2. Government Accession No. | 3. Recipient's Catalog No. |
|---|---|---|
| **4. Title and Subtitle**<br>Multimodal-AI based Roadway Hazard Identification and Warning using Onboard Smartphones with Cloud-based Fusion | | **5. Report Date** |
| | | **6. Performing Organization Code** |
| **7. Author(s)**<br>Mayuresh Bhosale; ORCID: 0009-0008-0708-1559<br>Longxiang Guo; ORCID: 0000-0002-6889-6099<br>Gurcan Comert; ORCID: 0000-0002-2373-5013<br>Yunyi Jia; ORCID: 0000-0003-1334-1384<br>Leo Hu: ORCID: 0000-0003-2657-5382 | | **8. Performing Organization Report No.** |
| **9. Performing Organization Name and Address**<br>Clemson University, 4 Research Drive, Greenville, SC 29607<br>Benedict College, 1600 Harden St, Columbia, SC 29204 | | **10. Work Unit No.** |
| | | **11. Contract or Grant No.**<br>69A3551747117 |
| **12. Sponsoring Agency Name and Address**<br>Center for Connected Multimodal Mobility ($C^2M^2$)<br>Clemson University<br>200 Lowry Hall,<br>Clemson, SC 29634 | | **13. Type of Report and Period Covered**<br>Final Report July 2023 |
| | | **14. Sponsoring Agency Code** |
| **15. Supplementary Notes** | | |

**16. Abstract**

Road hazard is one of the significant causes of fatality in road accidents. Accurate estimation of road hazards can ensure safety and enhance the driving experience. Existing methods of road condition monitoring are time-consuming, expensive, inefficient, require much human effort, and need to be regularly updated. There is a requirement for a flexible, cost-effective, and efficient process to detect road conditions, especially road hazards. In this study, we present a new method to deal with road hazards using smartphones. Since most of the population drives cars with smartphones onboard, we aim to leverage this to detect road hazards in a more flexible, cost-effective, and efficient way. This study proposes a cloud-based deep-learning road hazard detection model based on a Long-Short Term Memory network (LSTM) to detect different types of road hazards from motion data. To address the issue of large data requests for deep learning, this study proposes to fuse both simulation data and experimental data for the learning. The proposed approaches are validated by experimental tests, and the results demonstrate the accuracy of road hazard detection based on cloud-based fusion.

| 17. Keywords | 18. Distribution Statement |
|---|---|
| road hazards, LSTM, Motion simulation, cloud-based fusion, Web User Interface | |

| 19. Security Classif. (of this report) | 20. Security Classif. (of this page) | 21. No. of Pages | 22. Price |
|---|---|---|---|
| Unclassified | Unclassified | 27 | NA |

# Table of Contents

## List of Tables



## List of Figures

# EXECUTIVE SUMMARY

Monitoring road conditions is necessary to ensure the safety of transportation systems. Surveyors are traditionally required to drive or walk along the roads in order to manually check for problems during monitoring and inspections of the state of the roads. Such procedures need a lot of labor from both people and machinery, but they hardly ever result in the provision of real-time information on road conditions. Current automated road condition monitoring methods typically call for specialized vehicles fitted with a particular set of sensors, along with the necessary processing and computing equipment. These methods also only employ a single vehicle, which typically still needs a surveyor to operate it, to carry out the detection on its own. Consequently, in this project, we created a far more affordable method of monitoring road conditions utilizing in-car smartphones from any user of a public vehicle. This method is cloud-based collaborative monitoring.

The acceleration signal, particularly the vertical acceleration, has a distinctive pattern in the trajectory when a vehicle travels over a certain kind of road problem. The broad outline of the acceleration wave is used to determine the type of road hazard. In our study, we trained an LSTM-based deep learning network to fully identify the different types of defects using acceleration data. For this study to reduce costs of data-collection, the motion data from BeamNG simulation platform is obtained to validate the LSTM Deep Learning model along with the experimental data collected from smartphones. Multiple detections for same locations can provide high confidence. The results of several smartphone detections from numerous vehicles were then combined to provide a comprehensive picture of the state of the roads. Cellular networks were used to transmit the vehicle GPS coordinates and data from the smartphone motion road hazard detection results to a cloud server. The top three types of damage inside a cluster were discovered to represent the state of the road at that place. All detection data were combined with the k-means clustering algorithm based on their GPS positions.

In this study, we achieved 89% testing accuracy combined to detect different classes of road hazards from the LSTM model. Data for this experiment was gathered on numerous roadways in South Carolina's Greenville, Spartanburg, Clemson, and Columbia. Moreover, very good Cloud-based fusion was achieved based on detections from various smartphones and vehicles. We developed a website that maps out the fusion results of found road damage. The website enables the responsible authorities to view the user-reported road damages using the mobile application we built.

# CHAPTER 1

## Introduction

Road hazards can result in significant injuries and fatalities, making them a major public health and safety concern worldwide. NHTSA estimated 42,915 fatalities in motor vehicle crashes in 2021, over a 10% increase from 2020[1]. One of the major causes of road accidents in the U.S. is due to road hazards. Road hazards such as potholes, roadwork, accident vehicles, dead animals, and other unexpected obstacles lead to fatal incidents every year. Being aware of such road hazards can contribute to a decrease in accidents and an increase in safety, comfort, and fuel economy. Traditionally used techniques for monitoring the road surface include surveying techniques and profilometer measurements. Surveying is the traditional technique to monitor road surface conditions where a technician walks down the road to assess road defects. Such a technique requires efforts by human inspection, is prone to human errors, has limited coverage, is time-consuming, and does not provide road defect data in real-time. Another method uses a profilometer to measure road surface profile, roughness, and other surface characteristics by laser non-contact profilometers or physical sensor contact-based profilometers. The profilometer equipment is expensive and requires trained professionals to operate it.

Smart detection and classification algorithms have been created to streamline monitoring procedures by reducing the need for specialized equipment. Many of those techniques rely on indications from the motion of the vehicle, like speed and acceleration. These recently developed methods for identifying road damage are far more practical than conventional approaches. However, the scope of their applications is still constrained by the fact that they still rely on specialized data-collection devices. So, continuously monitoring road degradation would be advantageous if an efficient and effective mechanism could be created. This would improve the transportation system in terms of driving comfort and safety. In recent years, the use of smartphones in scientific research has multiplied. Many individuals take their smartphones with them when they are driving. Smartphones have increased comput capability and a wide range of sensors, including accelerometers, gyroscopes, magnetometers, GPS, and cameras. Smartphones can be used as monitoring nodes thanks to these qualities.

The goal and contribution of this project are to create a practical method for employing in-vehicle smartphones to monitor road hazards, such as potholes and obstacles. The suggested system uses vehicle motion-based road hazard detection and classification techniques to increase detection accuracy. In order to create a comprehensive, precise, and all-encompassing road damage monitoring system, the technique also includes a cloud-based fusion algorithm to combine all road damage detection findings from various cars.

# CHAPTER 2

# Literature Review

## 2.1 Related Work

Previously, researchers devised different techniques for detecting road hazards smartly using different classification algorithms. Most of those attempts are based on signals from the motion of the vehicle, like speed and acceleration. For example, [2] used vehicle acceleration measurements to detect the road surface roughness and classify the road profile. To detect road damage levels, they proposed the estimation of power spectral density by using transfer functions and the relation between the road surface and vehicle acceleration. [3] proposed a low-cost inertial measurement unit (IMU) and a global positioning system (GPS) sensor in the vehicle to analyze the power spectral density and classify the road pavement roughness. [4] proposed an IMU-based distributed sensor network to estimate the road and traffic conditions. The use of vision-based techniques is another prominent trend in the current works. [5] proposed a CNN-based network to detect different types of road distress using information from a customized camera setup. [6] discussed the various pavement distress type segmentation algorithms. In this work [7], along with identification and classification, quantifying the type of distress was taken into consideration. [8] Talked about a deep, fully convolutional crack detection model (CrackPix) that used well-known image classification architectures for dense predictions by converting their fully connected layers into convolutional filters.

Although these recently developed methods for monitoring road conditions are far more practical than the conventional approaches, they still rely on specialized data collection devices, which restricts the scope of their uses. In order to improve the transportation system in terms of driving safety and comfort, it would be advantageous to continuously monitor road conditions if an effective and economical approach could be created. The use of smartphones has skyrocketed in recent years. Smartphones also feature increasing processing power and a wide range of sensors, including an accelerometer, gyroscope, magnetometer, GPS, and camera. Smartphones are the best devices for creating a mobile in-vehicle sensor network because of these features. Researchers have looked examined the potential of utilizing smartphones to monitor road conditions. [9] examined how deep learning techniques like convolutional neural network (CNN) and long short-term memory are used to detect road features like potholes or bumps from data collected from cyclists' iPhones (LSTM). [10] employed deep learning techniques and data from smartphones to recognize various types of road surfaces and potholes. In work [11], researchers examined what influences smartphone measurements in a moving vehicle the most and how that alters measurements of road roughness. Some cloud computing based has been used previously to monitor road conditions. In [15], researchers developed a cloud computing-based road condition monitoring technique using smartphone motion and vision data. In research [16], a similar cloud-based technique has been applied to detect road conditions with good precision in less time. [17] used a cloud-based system to alert the end users of road conditions. [18] uses vision-based road damage detection with faster R-CNN.

## 2.2 Challenges and Gaps

However, despite the recent efforts focusing on road conditions, road hazards have not been well studied yet, which could be caused by bad road conditions, such as potholes, and some road events, such as roadwork, accident vehicles, broken vehicles, dead animals, and other unexpected obstacles. In addition, sufficient road hazard data acquisition is another challenge in conducting such a study, especially for deep learning-based approaches, which usually require a large amount of data. Therefore, the objective and contribution of the study is to develop a cost-effective and data-abundant technique to detect road hazards using smartphones. The proposed method uses smartphone motion data with a deep learning network based on LSTM to estimate potential road hazards and leverage real-world and simulation vehicle data to generate sufficient data for the deep learning model.

# CHAPTER 3

# Data Collection and Analysis Methods

## 3.1 System Framework

When mounted in the vehicle, smartphones can reflect the road profile when driven over it. Smartphone sensors such as gyroscopes, accelerometers, g-force, and magnetometers can register the vehicle motion reflecting the road surface profile [11]. Additionally, Soft-body physics vehicle simulation platforms such as BeamNG Tech software have vehicle sensors such as g-force and accelerometer. These sensors are used to generate a large amount of road surface profiles. This study uses these motion data to classify road surfaces using Long Short-Term Memory (LSTM), a Recurrent Neural Network (RNN). This study categorizes the road hazard conditions into three major situations, including "No Hazard," "Road Defect Hazard" which represents hazards caused by road defects such as potholes and bumps the vehicles could go over, and "Road Event Hazard" which represents hazards caused by road events such as roadwork, accident vehicles, dead animals and other unexpected obstacles and the vehicles have to avoid or dodge. Figure 1 represents the system framework of the road hazard detection process.

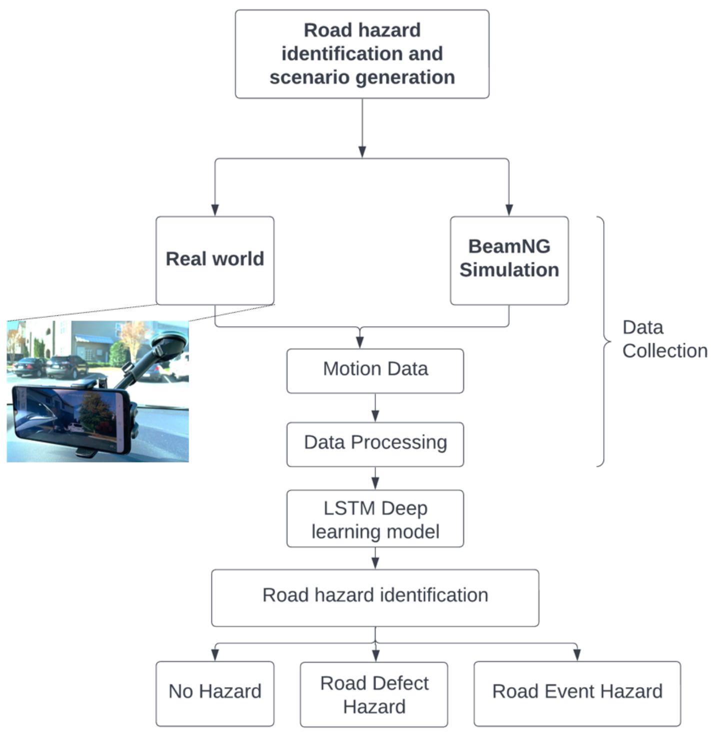



**Figure 1 Road hazard detection system framework**

## 3.2 Data Acquisition

Generally, the vehicle bump or pothole represents the road damage to the vehicle in the vertical direction, represented as vertical acceleration. During an obstacle avoidance course on the road, the vehicle sways in a lateral direction to generate lateral acceleration with some torsional acceleration. Therefore, it's crucial to find the relation between vertical vehicle acceleration, lateral acceleration, and the type of road damage. We gathered historical data on vehicle vertical, lateral, and torsional acceleration on various road damage sections to examine the type of road damage in simulation and real-world scenarios. For the real-world data, 3-axis acceleration data was measured from the smartphone mounted on the windshield or dashboard of the vehicle (refer to Figure 1). MATLAB application on an android smartphone was used to collect the motion data. For the simulation platform BeamNG Tech software, the g-force/acceleration sensor was used to measure the motion data.

Real-world data: The motion data from the real-world environment was collected over various roads of Greenville, Clemson, Spartanburg, and Columbia in South Carolina, USA. Multiple cars were used to collect the data from the MATLAB phone application. MATLAB android/iOS application has a sensor suite option that allows recording the data from different sensors such as acceleration, magnetic field, angular velocity, orientation, and GPS position. For our research, motion data with a sampling rate of 100 hertz was considered to classify different road hazards. Data was collected at various speeds, placements, road inclination, and damages on the roads. For road defect hazards, acceleration data from potholes and bumps with varying widths and depths/heights were recorded. For road event hazards, acceleration data was collected considering obstacles of various sizes ahead of a vehicle. For this, the driver would steer the vehicle to avoid hitting obstacles. These obstacles could generally be a dead animal, accidental vehicles, roadwork equipment, or any unexpected objects that the vehicle should avoid hitting. To recreate all common scenarios for machine models to detect the hazards in real-time, as many diverse data points as possible were collected with varying vehicle speeds, hazard sizes, different locations, and multiple smartphones.

Simulation data: BeamNG is a high-fidelity soft-body vehicle simulation software that is authentic and provides real-like vehicle behavior [14]. In this study, BeamNG Tech software is used to generate the required road hazard data and the data generation framework can be seen in figure 2. The custom environment was built with varying road hazard dimensions and scenarios. Data was collected from multiple software-inbuilt vehicles, such as Ibishu Pessima, a mid-sized sedan, and Mazda CX7, a large-size vehicle. The motion data was recorded by the G-Force/Acceleration sensor mounted on the simulation vehicle and accessed through BeamNGpy (Python-based API). Indentations were made in the asphalt road with constant width of 0.5 meters and depth of pothole ranging from 0.1 meters to 0.5 meters (refer to Fig 3). Multiple vehicles traveling at speeds ranging from 10 miles to 55 miles per hour went over generated potholes to generate the motion data for potholes. For obstacle avoidance(dodge) road hazards, Stanley control was used to steer the simulation vehicle tracking the pre-defined vehicle path. Scenarios were built to track the vehicle's motion, avoiding different sizes of obstacles, with speeds ranging from 10 to 55 miles per hour. A Proportional, Integral, and Derivative (PID) control was used for throttle commands with smoothening to avoid jarring vehicle motion for both potholes and dodging road hazards. The motion at 100-hertz frequency was collected in a text file and later used in the deep learning model. Considering the vehicle frame axes, the X-axis represents the lateral motion, Y-axis represents the torsional motion, and Z-axis represents the vehicle's vertical motion for road.

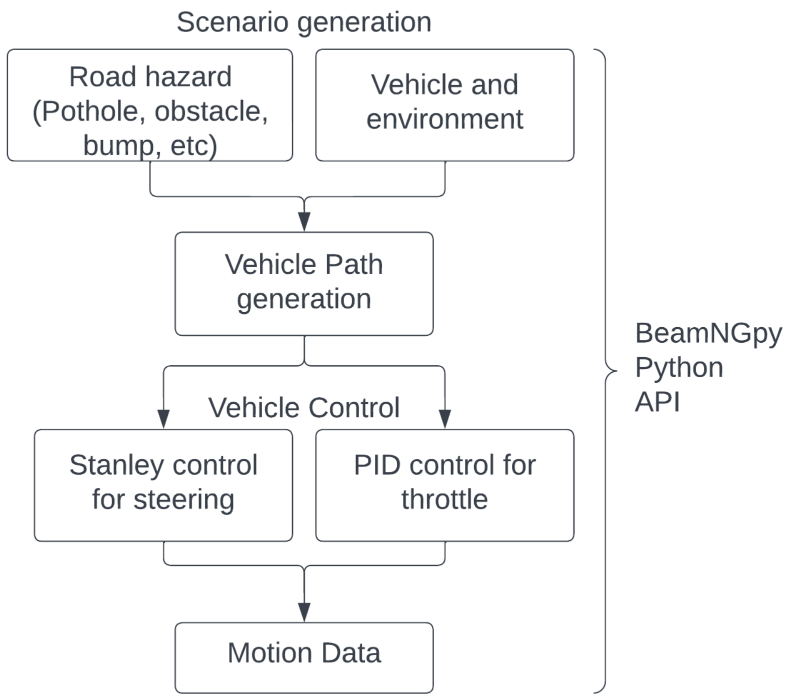



**Figure 2 BeamNG simulation motion data generation framework**

**Figure 3 BeamNG simulation environment and one example of potholes**

### 3.3 Data Processing

The accelerometer sensor in the smartphone and BeamNG provide motion data over time but has noise associated with it. To accurately identify the type of road hazard, data should be free of noise and needs to be filtered. This study explored two different filtering techniques to eliminate the high-frequency noise: Kalman filtering and low-pass filtering. For Kalman filtering, different measurement noise covariance and dynamic noise covariance matrix was tuned to eliminate the high-frequency noise [19]. For low pass filtering, band pass frequency and sampling rate were tuned to eliminate the high-frequency noise [20].

### 3.3 Deep Learning based Road Hazard Detection Model

The LSTM network was chosen because it features feedback connections in contrast to traditional feed-forward neural networks. Both individual data points (like photos) and complete data sequences, such as time series motion data, can be processed [13] [15]. Since a single motion data point cannot identify a certain road surface condition, this specialization is crucial for our strategy. Only when the data are contextualized do they become meaningful. Motion data obtained from both simulation and the real world is input for the LSTM network. The suggested solution uses a stacking architecture of two completely connected layers and two LSTM layers. The proposed architecture is depicted in Fig 4; 80 hidden units are the input size for the first completely connected layer, followed by another hidden layer. This is followed by two LSTM layers that each have 64 units layered on top of one another. The final cell's output is extracted, and the SoftMax function is applied. This is the likelihood that each class will exist, according to the model. A cell, an input gate, an output gate, and a forget gate make up an LSTM unit in Fig. 4. The three gates control the flow of information into and out of the cell, and the cell remembers values across arbitrary time intervals. The first gate, referred to as the forget gate, chooses which portion of the cell state data should be deleted. The second input gate chooses the data that will be included in the cell state. Last but not least, the output gate produces output data based on cell status providing a classification of the road hazard and corresponding probabilities.

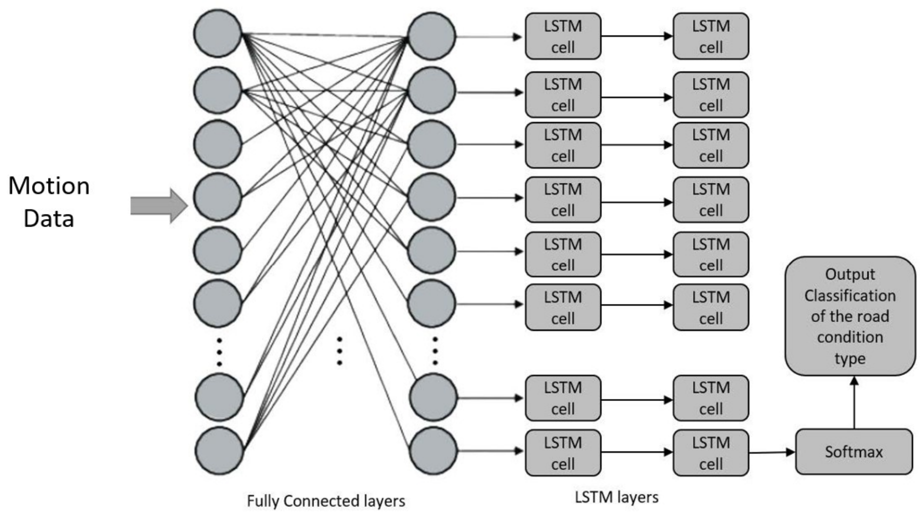



**Figure 4 LSTM architecture for deep learning-based road hazard detection**

### 3.3 Heterogeneous Training Methods

Simulations do not represent the real world exactly. The LSTM model was tested on three heterogeneous training methods to validate the accuracy of simulation and real-world data for road hazard detection. The three types of tests were as follows:

#### 3.3.1 Test 1 (Simulation Only)

In this case, the road hazard data for training the LSTM model was only simulation data, and the testing data was also only simulation data. Total 2008 simulation data points were split into 1236 for training and 772 for testing.

#### 3.3.1 Test 2 (Simulation and real separate)

In this case, the road hazard data for training the LSTM model was only simulation data, and the testing data was only real-world data. This test was performed to check the accuracy and correlation of simulation and real-world tests. A total of 2758 data points was split into 2008-only simulation data for training and 750 real-world data for testing.

#### 3.3.1 Test 3 (Simulation and real mixed)

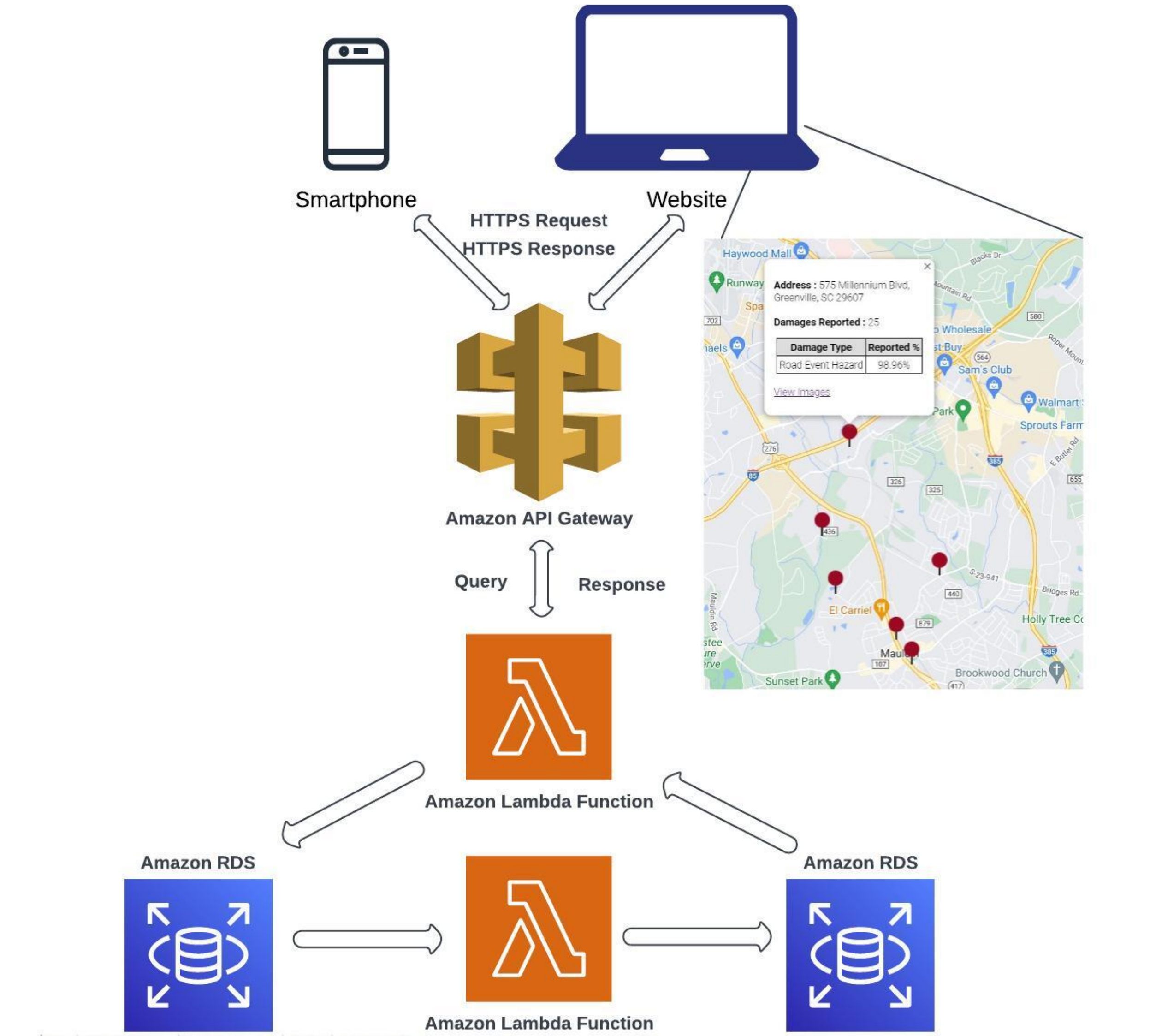
In this case, the road hazard data for training the LSTM model was simulation data and real-world data mixed, whereas the testing data was only real-world data. A total of 2758 data points were split into 2211 mixed data for training and 547 real-world data for testing.

### 3.3 Cloud-based fusion

The motion-based road hazard data points recorded by multiple vehicles at multiple locations are used to generate more reliable and accurate detection. For cloud-based fusion, the type of hazard, GPS coordinates of the road hazard detection, and confidence are utilized. The recorded hazards are stored in a database through a Relational Database Service (RDS) of the Amazon Web Services (AWS) cloud. The RDS has the type of hazard, GPS coordinates, and confidence ratings. For multiple readings of road hazards detected by different vehicles at the same location, clustering is required to report the damages and is achieved by cloud-based fusion. AWS Lambda function posts the data from the AWS API gateway to the relational database. Raw road hazard detection data is populated and clustered to provide optimized results. Further, the road hazard data is posted to the website through AWS lambda and the API gateway function. Figure 5 gives an overview of optimized clustering and cloud-based fusion approach. Due to various noises and environmental factors, road hazard detection at the same location will have slightly varying GPS coordinates. To consolidate these detections, k-means clustering, an unsupervised, non-deterministic, and iterative algorithm, is implemented. This clustering method is proven effective in obtaining accurate results and is used in many practical applications. At first, K centers are randomly chosen by the algorithm, and each data point is assigned to the nearest k center calculated by using the Euclidean distance. This generates the initial k clusters. After assigning data points to k centers, the algorithm recalculates new centers by averaging the data points assigned to the initial centers. These new centers are then recalculated and reassigned until the criterion function is minimum or if the algorithm is looped a certain number of times. This study considers only latitudes and longitudes from the data points collected. The initial value of k is considered 1; consequently, k-means clustering is performed on the data set. The value of k increases until the average Within Cluster Sum of Squares (WCSS) of previous k and current k is less than or equal to 0.001, thus obtaining a final and optimum number of k clusters. WCSS is the average squared distance from every point inside a cluster to the centroid of the cluster. Each cluster represents data points that are nearest to one another and have a very high probability of being nearby or at the same location. Each cluster is allotted with a centroid latitude and longitude value of that cluster, type of hazard, average confidence of all the data points in that cluster for each hazard type, the total count of the data points in a cluster, and a cluster ID containing top three hazard types based on decreasing average confidences, as shown in Figure 5. A web UI displays the cluster information on a map with its address, hazard type, total damages reported, and respective confidence.

**Algorithm 1** Clustering Algorithm

```
for k in range(k_initial,k_maximum+1) do
    kmeans = k_clusters.fit(locations)
    centroids = kmeans.random_centers
    predict = kmeans.centroids
    for i in range(number of locations) do
        centroids = kmeans.random_centers
        WCSS = WCSS + (locations(i) - current_center(i))^2
        if WCSS < 0.001 & k > 1 then
            return WCSS, centroid, k, predict
        end if
    end for
end for
return WCSS, centroid, k, predict
```




| id | latitude | longitude | type | confidence |
|---|---|---|---|---|
| 1 | 34.78126134 | -82.35030519 | 3 | 95.2914 |
| 2 | 34.75609949 | -82.37012646 | 10 | 94.0998 |
| 3 | 34.75185708 | -82.36962732 | 2 | 84.5459 |
| 4 | 34.92708759 | -81.90524608 | 3 | 90.1694 |
| 5 | 34.83482296 | -82.44577792 | 2 | 90.9333 |
| 6 | 34.77335331 | -82.36957711 | 3 | 81.769 |

Raw Data Points

| id | clusterId | latitude | longitude | type | confidence | total_count |
|---|---|---|---|---|---|---|
| 1 | 0 | 34.80031413 | -82.32742422 | 11 | 86.4124 | 25 |
| 2 | 1 | 34.76942187 | -82.39536847 | 12 | 99.9052 | 25 |
| 3 | 2 | 34.77672644 | -82.30748234 | 12 | 92.4708 | 25 |
| 4 | 2 | 34.77672644 | -82.30748234 | 11 | 86.5116 | 25 |

Clustered Data

**Figure 5 Cloud-based fusion approach**

# CHAPTER 4

# Experimental Results

## 4.1 Experimental Data representation and Data processing

This section represents data information introduced in section 3.1.1 and data processing explained in section 3.1.2.

### 4.1.1 Data Representation

Motion data from BeamNG simulation and real-world smartphones were collected at 100 Hertz frequency. Figure 6 and 7 represent the motion of a vehicle for road event hazard and road defect hazard, respectively. As observed for the road event hazard, the vehicle experiences high lateral acceleration from yaw movement when trying to avoid obstacles, accidents, etc., on the road. For a road defect hazard, vertical vehicle acceleration has a spike corresponding to pitch movement when going over potholes, bumps, etc., on the road. The data distribution used for this study is presented in Table 1. A total of 2758 motion data points from simulation and real-world tests were collected for three road conditions and were identified as classes with no hazard, road event hazard, and road defect hazard. Out of these, 2008 data points were from BeamNG simulation software, and 750 data points were from real-world testing. These data points were labeled as 0 for no hazard, 1 for a road defect hazard, and 2 for a road event hazard associated with a respective time stamp. All the data processing, namely labeling, and filtering, was performed on MATLAB R2021a software. For the cloud-based fusion experiment, a total of 250 road hazard detections were used to evaluate the approach from 5 different locations, each of road event hazard and road defect hazard. The data were collected from 5 different smartphones mounted in 5 different vehicles. The GPS coordinates of each detection were recorded to be utilized for k-means clustering and cloud-based fusion.

**Table 1 Motion data distribution**

| Data Type | Simulation | Real-world | Total |
|---|---|---|---|
| No hazard | 1091 | 94 | 1185 |
| Road defect hazard | 217 | 278 | 495 |
| Road event hazard | 700 | 378 | 1078 |
| Total | 2008 | 750 | 2758 |

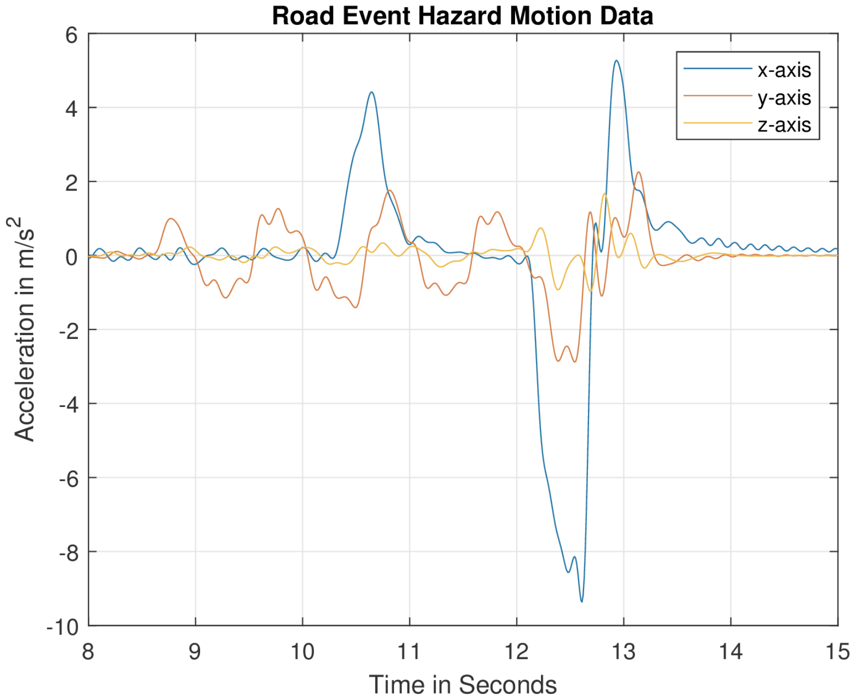



**Figure 6 Vehicle motion data for road event hazard**

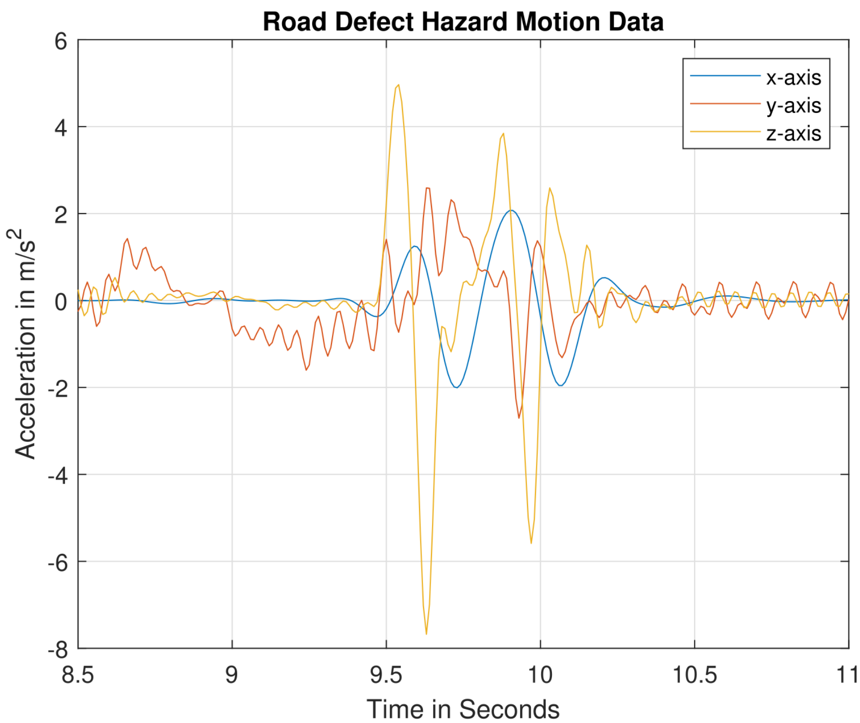



**Figure 7 Vehicle motion data for road defect hazard**

### 4.1.2 Data Processing

Figure 8 represents the lateral acceleration plot over time in the road event hazard scenario with and without filtering. As observed, unfiltered data from simulation and smartphone sensors consist of high-frequency noise. Kalman-filtered motion data reduces noise with a bit of delay in tracking the unfiltered motion data. Low-pass filtered data eliminates the high-frequency noise with good tracking of unfiltered motion data.

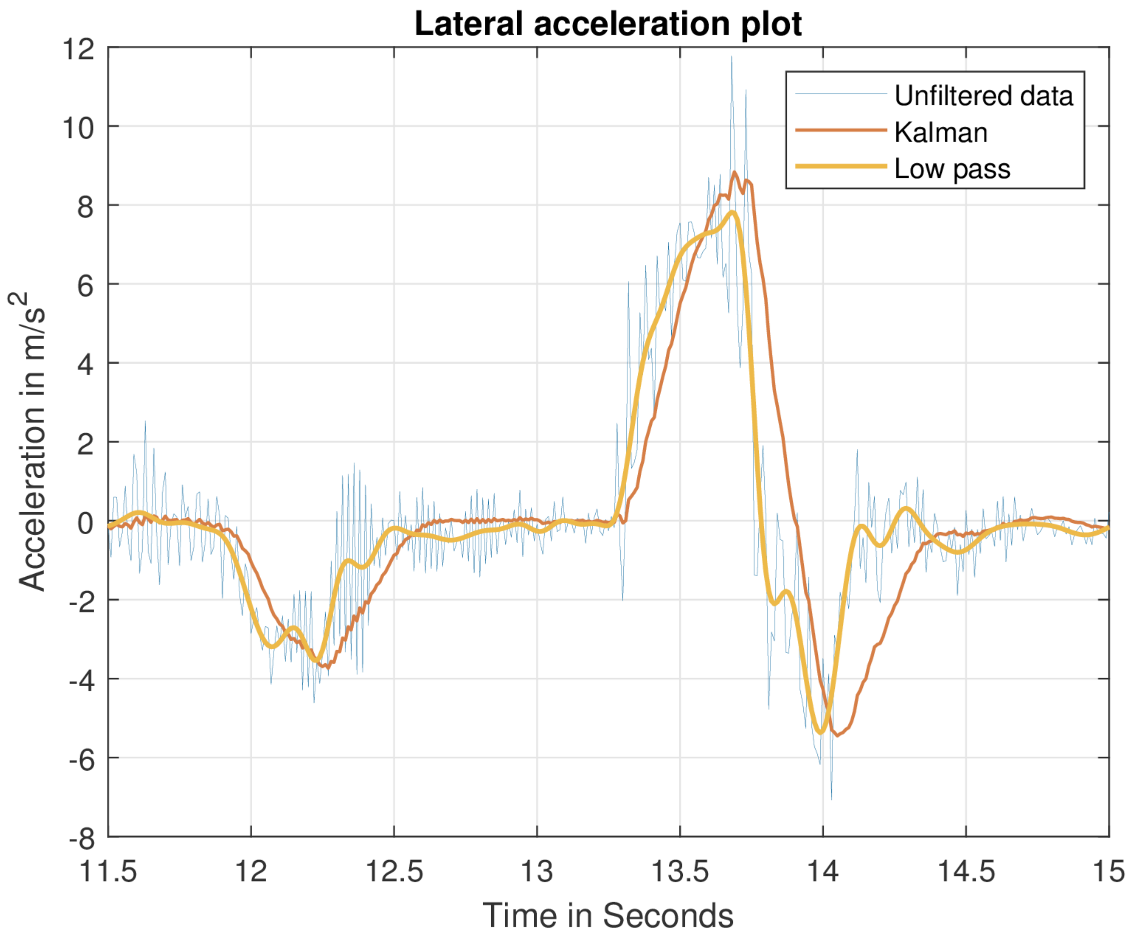



**Figure 8 Lateral acceleration motion data with and without filtering**

### 4.2 Experimental Results and Analysis

The LSTM model was trained and tested on the Google Colab platform with Tensor- Flow2/Keras deep learning library. The tuned parameters of the model have mentioned in table 2. These tuned parameters are the same for all three heterogeneous training methods mentioned in section 2.5. Three features are no hazard, road defect hazard, and road event hazard, respectively. For model training, the time series labeled motion data are combined together, and 80-time step data are fed into the LSTM model. Tables 3 to 5 represent the LSTM training and testing accuracies for three tests performed, which are explained in section 3.2.1.

**Table 2 LSTM tuning parameters**

| Parameter Name | Parameter Used |
|---|---|
| Number of features | 3 |
| Number of time steps | 80 |
| Number of training epoch | 15 |
| Optimizer | Adam |
| Batch size | 512 |
| Learning rate | 0.0025 |
| Loss regularization | L2 loss 0.0015 |

**Table 3 LSTM accuracy results for only simulation data - Test 1**

| Training data | Testing data | Filter type | Training accuracy | Testing accuracy |
|---|---|---|---|---|

| | | | | |
|---|---|---|---|---|
| Simulation | Simulation | Kalman | 95.5% | 97.8% |
| Simulation | Simulation | Low-pass | 94.5% | 97.3% |

### 4.2.1 Model training results

Figures 9 to 11 show training accuracies over epochs for three test cases. It can be observed that the training process was terminated when the model's performance on the validation dataset and testing dataset almost ceased to increase and began to vary from one another. This shows that the model was trained prior to the emergence of an overfitting problem. The trained LSTM result for test 1, consisting of only motion data from the simulation, gives a training accuracy of 95.5%. Test 2, which consists of motion data from the simulation, gives a training accuracy of 96.1%, and test 3, which consists of motion data from the simulation and real-world mixed, gives a training accuracy of 95.6% to classify road hazards. For all the tests, the loss has a good learning rate with less decay and decreases with increasing training accuracy until the model is overtrained. Both the Kalman-filtered data and the low-pass filtered data for each test provide very similar training accuracies suggesting these techniques to be viable for noise reduction of sensor-based time series motion data.

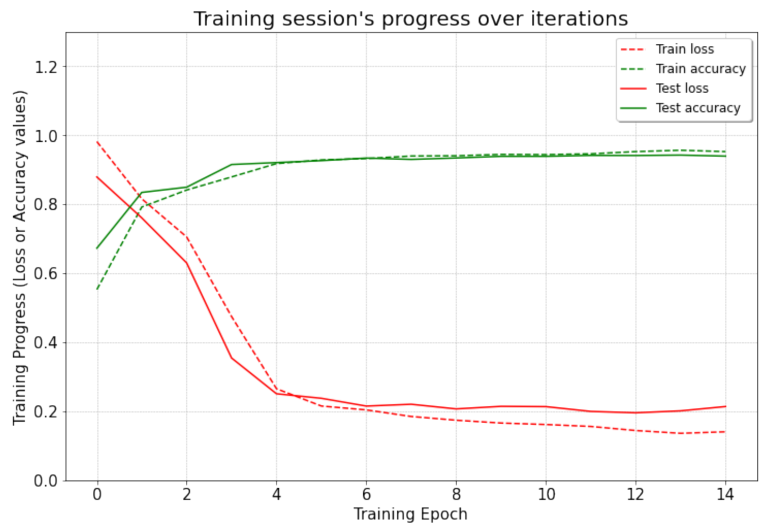



**Figure 9 LSTM training accuracy and loss for simulation data only - Test 1**

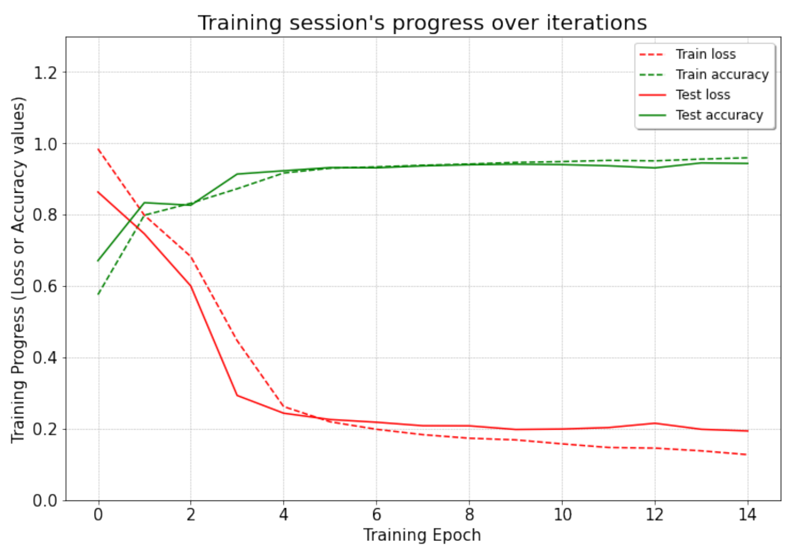



**Figure 10 LSTM training accuracy and loss for simulation and real separated data - Test 2**

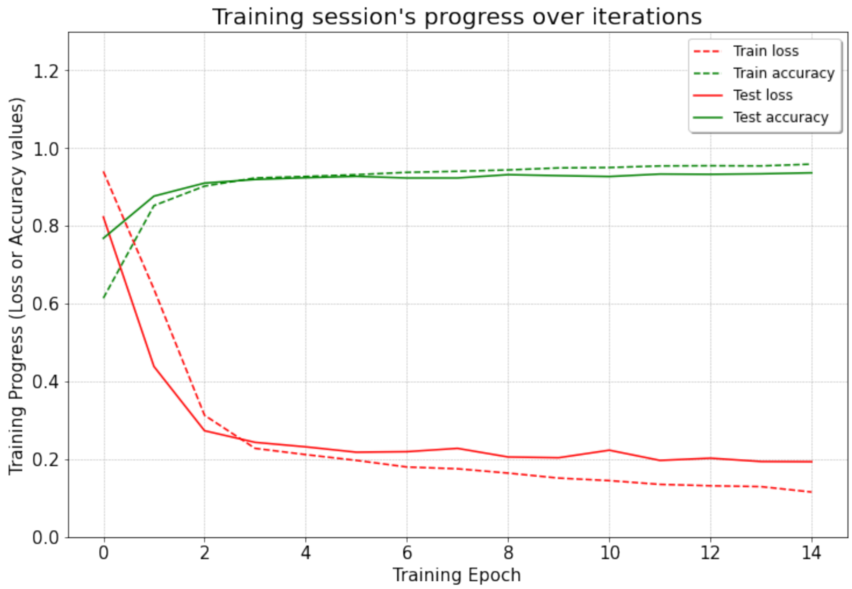



**Figure 11 LSTM training accuracy and loss for simulation and real mixed data - Test 3**

### 4.2.2 Model testing results

Figures 12 to 14 show the confusion matrix of the model on testing data set only. Tables 3 to 5 provide results for three different tests with two types of data filtering techniques, namely the Kalman and low-pass filters. For test 1, where training and testing motion data is only from simulation, Kalman filtered data gave a training accuracy of 95.5% and testing accuracy of 97.8%. Whereas, Low-pass filtered motion data gave a training accuracy of 94.5% and a testing accuracy of 97.3%. The confusion matrix for road hazard class results can be seen in figure 11. As observed, all three classes provide an accuracy of over 96%. This suggests that the simulation-only training and testing data set is very similar and identification of road hazard patterns by the LSTM model is accurate.

For test 2, where training motion data is only simulation and testing motion data is only real, Kalman-filtered data gave a training accuracy of 96.1% and a testing accuracy of 75.6%. Whereas, Low-pass filtered data gave a training accuracy of 96.1% and a testing accuracy of 79.7%. The confusion matrix for road hazard class results can be seen in figure 13. As observed, the identification of road event hazard is 99%. In contrast, road defect hazard and no hazard have 60% accurate detection. Lower model testing accuracy corresponds to a slightly low correlation between the simulation and the real world but is sufficient to validate the LSTM model. When observed, the road event hazard motion data from the simulation and the real world is highly similar, thus resulting in high detection accuracy. To improve the accuracy of other classes, more test scenarios with varying potholes and multiple vehicles can be produced and validated. For test 3, where training motion data is mixed with simulation and real-world, and testing motion data is only real-world, Kalman-filtered data gave a training accuracy of 95.6% and a testing accuracy of 89.6%. Whereas low-pass filtered data gave a training accuracy of 94.6% and a testing accuracy of 89.0%. The confusion matrix for road hazard class results can be seen in figure 14. As observed, the identification of road defect hazard has a high accuracy of 83%, and the accuracy for road event hazards is very promising as well. For the hazard road, the accuracy is 64%. The increased test accuracy the test 2 corresponds to the introduction of a small amount of real-world motion data to train the model. This suggests that the real-world motion data for pothole and undamaged classes is slightly different than the simulation motion data. The accuracy further can be improved by training the model with more test and training motion data.

Overall, the LSTM model performed better than the deep learning models in research [7] to [11] with a similar objective. The simulation-based motion data took less time to generate with almost no cost involved. Moreover, the LSTM model trained with simulation and real motion data and tested with real motion data provided good road hazard detection accuracy proving the legitimacy of the usage of simulation motion data. Although the real-world motion data is just a small portion of all the motion data collected, the trained model with most simulation motion data provides good performance in real-data testing. Filtering techniques proved effective in reducing noise and accurately detecting road hazards. These filtering techniques can further be implemented in real-time.

### 4.2.3 Cloud-based fusion results

The LSTM training model used simulation and real-world data for cloud-based fusion. Table 6 shows the results obtained from cloud-based fusion approach. The k-means clustering algorithm deployed on AWS lambda clustered 250 individual data points into 9 clusters, each with a centroid latitude and longitude. K-means algorithm also segregates the clusters by different types, with ’11’ being road defect hazard and ’12’ being road event hazard.

The accuracy before fusion is more than 92% of all the data points, indicating good performance by the LSTM model. Out of 25 total detections for each cluster for the same hazard type, almost more than 23 were true detections. Thus, accuracy after cloud-based fusion is very good for road hazard detection. Figure 15 shows a screen capture of the web page UI displaying clustering results with damage types and their confidence with their address. The website allows relevant authorities to see the road damages people have reported using a mobile application.

**Table 4 LSTM accuracy results for simulation and real separate data - Test 2**

| Training data | Testing data | Filter type | Training accuracy | Testing accuracy |
|---|---|---|---|---|
| Simulation | Real | Kalman | 96.1% | 75.6% |
| Simulation | Real | Low-pass | 95.2% | 79.7% |

**Table 5 LSTM accuracy results for simulation and real mixed data - Test 3**

| Training data | Testing data | Filter type | Training accuracy | Testing accuracy |
|---|---|---|---|---|
| Simulation + real | Real | Kalman | 95.6% | 89.6% |
| Simulation + real | Real | Low-pass | 94.6% | 89.0% |

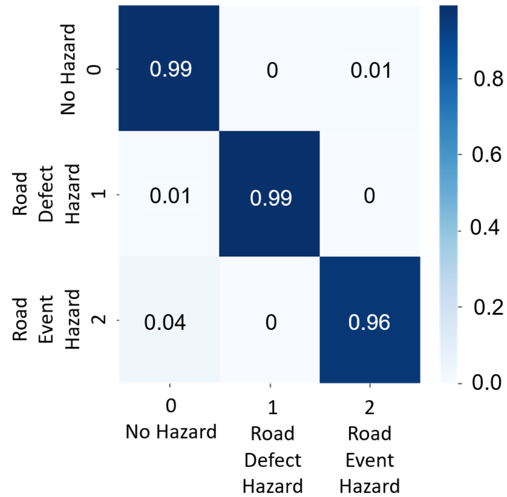



**Figure 12 Confusion matrix for Test 1 (Simulation only) - with Kalman filtered data**

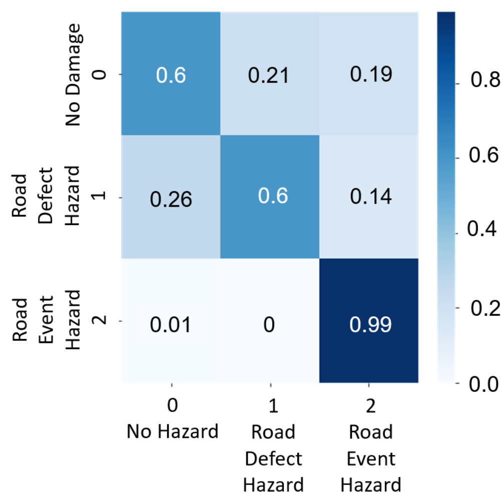



**Figure 13 Confusion matrix for Test 2 (simulation and real separate) - with low-pass filtered data**

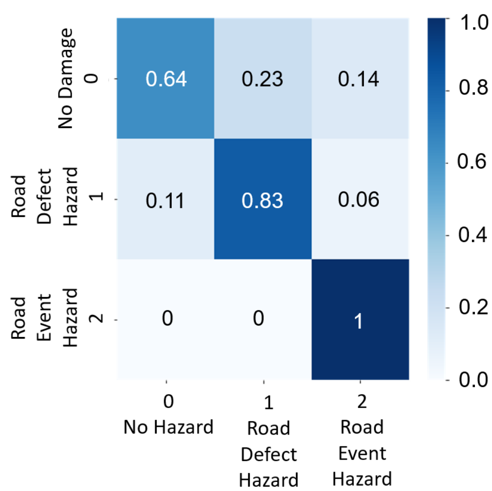



**Figure 14 Confusion matrix for Test 3 (simulation and real mixed) - with low-pass filtered data**

**Table 6 K-means clustering and cloud based fusion results**

| Cluster ID | Latitude | Longitude | Type | Total count | True count | Accuracy |
|---|---|---|---|---|---|---|
| 0 | 34.8003 | -82.3274 | 11 | 25 | 23 | 92 |
| 1 | 34.7694 | -82.3954 | 12 | 25 | 25 | 100 |
| 2 | 34.7767 | -82.3075 | 12 | 25 | 24 | 96 |
| 2 | 34.7767 | -82.3075 | 11 | 25 | 23 | 92 |
| 3 | 34.7344 | -82.3744 | 11 | 25 | 20 | 80 |
| 4 | 34.7521 | -82.2983 | 12 | 25 | 23 | 92 |
| 5 | 34.7930 | -82.3013 | 11 | 25 | 24 | 96 |
| 6 | 34.8166 | -82.3215 | 12 | 25 | 25 | 100 |
| 7 | 34.7896 | -82.3245 | 12 | 25 | 25 | 100 |
| 8 | 34.7813 | -82.3109 | 11 | 25 | 24 | 96 |

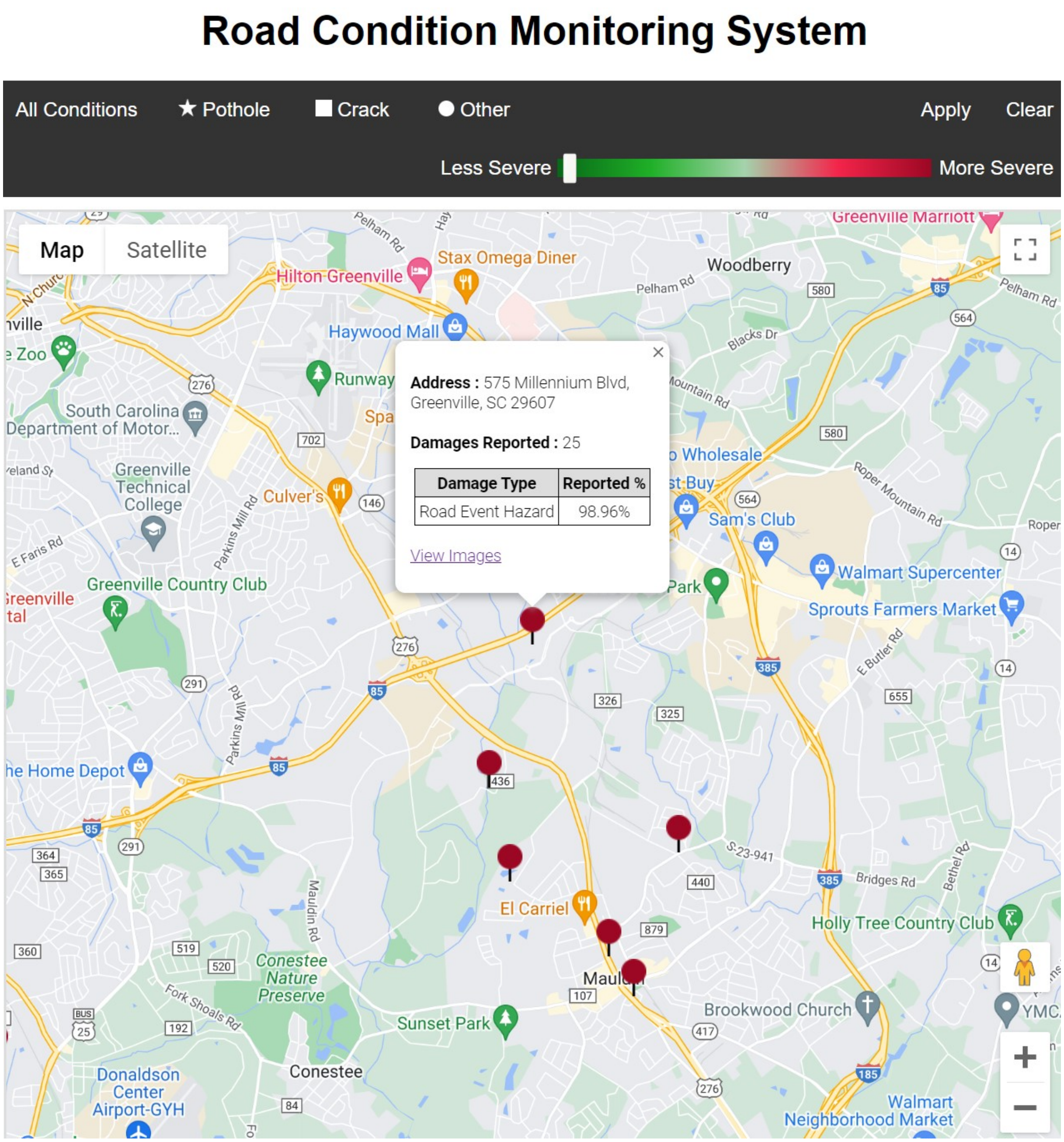



**Figure 15 Road hazard representation on web UI**

### 4.3 Conclusion

This study is on the problem of monitoring road surfaces for various road hazards detected. The method uses vehicle motion data from a simulation platform and the real world collected from smartphones. A deep learning-based LSTM technique was trained for this task. A soft-body physics-based simulation platform was explored to provide real-like vehicle behavior. Further validating its usage to train the deep learning models with abundant data, less time, and lower costs. The performance of the proposed method was proven in the simulation platform and real-world experiments. Cloud-based fusion techniques provided more accurate results allowing to monitor of road hazards with more reliability. However, the following areas of the work could still use improvement: we can add more information to the dataset currently used to train the deep-learning models. This study involves only motion-based data. In the future, we will combine this with vision-based road hazard detection to provide better results. The current work does not provide the severity of the road hazards. In our future work, we will provide metrics to include the severity index for various road hazards. The current work involves gathering and testing the data on the deep learning model.

## REFERENCES

[1] https://www.nhtsa.gov/press-releases/early-estimate-2021-traffic-fatalities

[2] A. Gonzalez, E.J. O'brien, Y.-Y. Li and K. Cashell (2008) The use of vehicle acceleration measurements to estimate road roughness, Vehicle System Dynamics, 46:6, 483-499, DOI: 10.1080/00423110701485050

[3] Kongyang Chen, Mingming Lu, Xiaopeng Fan, Mingming Wei and Jinwu Wu, "Road condition monitoring using on-board Three-axis Accelerometer and GPS Sensor," 2011 6th International ICST Conference on Communications and Networking in China (CHINACOM), 2011, pp. 1032-1037, doi: 10.1109/ChinaCom.2011.6158308.

[4] Tian Lei, Abduallah A. Mohamed, Christian Claudel, An IMU-based traffic and road condition monitoring system, HardwareX, Volume 4, 2018, e00045, ISSN 2468-0672, https://doi.org/10.1016/j.ohx.2018.e00045.

[5] Li, Y.; Liu, C.; Shen, Y.; Cao, J.; Yu, S.; Du, Y. RoadID: A Dedicated Deep Convolutional Neural Network for Multipavement Distress Detection. Journal of Transportation Engineering, Part B:Pavements 2021, 147, 04021057

[6] Tsai, Y.C.; Kaul, V.; Mersereau, R.M. Critical assessment of pavement distress segmentation methods. Journal of transportation engineering 2010, 136, 11–19

[7] Llopis-Castello, D.; Paredes, R.; Parreno-Lara, M.; Garcia-Segura, T.; Pellicer, E. Automatic Classification and Quantification of Basic Distresses on Urban Flexible Pavement through Convolutional Neural Networks. Journal of Transportation Engineering, Part B: Pavements 2021, 147, 04021063

[8] Alipour, M.; Harris, D.K.; Miller, G.R. Robust pixel-level crack detection using deep fully convolutional neural networks. Journal of Computing in Civil Engineering 2019, 33, 04019040.

[9] Sattar, S.; Li, S.; Chapman, M. Road surface monitoring using smartphone sensors: A review. Sensors 2018, 18, 3845

[10] Varona, B.; Monteserin, A.; Teyseyre, A. A deep learning approach to automatic road surface monitoring and pothole detection. Personal and Ubiquitous Computing 2020, 24, 519–534 392

[11] Chatterjee, A.; Tsai, Y.C. Training and testing of smartphone-based pavement condition estimation models using 3d pavement data. Journal of Computing in Civil Engineering 2020, 34, 04020043

[12] Jia, Shuo and Hui, Fei and Li, Shining and Zhao, Xiangmo and Khattak, Asad. (2020). LSTM-CNN for Abnormal Driving Behavior Recognition. IET Intelligent Transport Systems. 14. 397 10.1049/iet-its.2019.0200. 398

[13] Sepp Hochreiter, Jurgen Schmidhuber; Long Short-Term Memory. Neural Comput 1997; 9 (8): 1735–1780. doi: https://doi.org/10.1162/neco.1997.9.8.1735

[14] Pascale Maul* , Marc Mueller* , Fabian Enkler* , Eva Pigova* , Thomas Fischer* , Lefteris Stamatogiannakis; BeamNG.tech Technical Paper

[15] Ramesh, A.; Nikam, D.; Balachandran, V.N.; Guo, L.;Wang, R.; Hu, L.; Comert, G.; Jia, Y. Cloud-Based Collaborative Road-Damage Monitoring with Deep Learning and Smartphones. Sustainability 2022, 14, 8682. https://doi.org/10.3390/su14148682

[16] M. A. Ameddah, B. Das and J. Almhana, "Cloud-Assisted Real-Time Road Condition Monitoring 406 System for Vehicles," 2018 IEEE Global Communications Conference (GLOBECOM), 2018, pp. 1-6, doi: 10.1109/GLOCOM.2018.8647334.

[17] Y. Yuan, M. S. Islam, Y. Yuan, S.Wang, T. Baker and L. M. Kolbe, "EcRD: Edge-Cloud Computing Framework for Smart Road Damage Detection andWarning," in IEEE Internet of Things Journal, vol. 8, no. 16, pp. 12734-12747, 15 Aug.15, 2021, doi: 10.1109/JIOT.2020.3024885.

[18] Pham, V.; Pham, C.; Dang, T. Road damage detection and classification with detectron2 and faster r-cnn. In Proceedings of the 2020 IEEE International Conference on Big Data (Big Data), Atlanta, GA, USA, 10–13 December 2020; pp. 5592–5601.

[19] Alfian, Rio and Ma'arif, Alfian and Sunardi, Sunardi. (2020). Noise Reduction in the Accelerometer and Gyroscope Sensor with the Kalman Filter Algorithm. Journal of Robotics and Control (JRC). 2. 180-189. 10.18196/jrc.2375.

[20] Suwandi, Bondan et al. "Vehicle Vibration Error Compensation on IMU-accelerometer Sensor Using Adaptive Filter and Low-pass Filter Approaches." J. Inf. Process. 27 (2019): 33-40.