ASHESI UNIVERSITY COLLEGE

CLASSIFYING ROAD SURFACES USING SMARTPHONE ACCELEROMETERS FOR INFORMED ROAD TRANSPORTATION

UNDERGRADUATE THESIS

B.Sc. Computer Science

Nutifafa Amedior

2020

ASHESI UNIVERSITY COLLEGE

CLASSIFYING ROAD SURFACES USING SMARTPHONE ACCELEROMETERS FOR INFORMED ROAD TRANSPORTATION

UNDERGRADUATE THESIS

Undergraduate Thesis submitted to the Department of Computer Science, Ashesi University College in partial fulfilment of the requirements for the award of Bachelor of Science degree in Computer Science.

Nutifafa Amedior

2020

DECLARATION

I hereby declare that this thesis is the result of my original work and that no part of it has been presented for another degree in this university or elsewhere.

Candidate's Signature:

………………………………………………………………………………………...

Candidate's Name:

………………………………………………………………………………………...

Date:

………………………………………………………………………………………...

I hereby declare that the preparation and presentation of this thesis were supervised in accordance with the guidelines on the supervision of thesis laid down by Ashesi University.

Supervisor's Signature:

………………………………………………………………………………………

Supervisor's Name:

………………………………………………………………………………………

Date:

………………………………………………………………………………………

Acknowledgments

To my supervisor, whose encouragement and academic advice helped me undertake this project.

Abstract

Road navigation applications such as Google Maps and Apple Maps provide routing information to their continuously increasing number of users, enabling them to get from one destination to another. These applications provide information such as routes and traffic conditions which influence the time taken to travel by the user of the information. However, these navigation services are lacking in providing road surface quality information. Road surface quality information of a route not only influences the time taken to travel the route, but also provide salient information on the comfort of travel for the passenger and the effect the terrain will have on the vehicle.

This work builds on previous work by further developing and characterizing a Logistic Regression (LR) algorithm for classifying road surface quality using accelerometer data sourced from mobile devices in moving vehicles along four different types of roads: very good, good, bad, and very bad roads.

Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 Background

Road travel is the most popular mode of transportation, accounting for approximately 80% of passenger-kilometres, traveled on the vast network of roads all over the world [17]. Roads enable the secure transport of people, goods, and ideas, affecting quality and equity in products and service distribution and, ultimately, human outcomes [6].

In vehicular transport, the feeling of travelling over a smooth paved road is different from that of a bumpy dirt road. A road passenger's concern as they travel includes the comfort they feel in the vehicle, aside from arriving at their destination. The decision to travel on a road or not is limited if there is no alternative mode of transportation. The comfort of travelling on the road is influenced by the road surface quality, making up the travel experience of the passenger. Road surface quality is also a concern for government officials in road construction and maintenance for prolonging the service life of roads for commuters and their vehicles. Road surface quality monitoring is essential for passengers in informing their travel and for government officials in advancing road transport development.

Road surface quality relates to the roughness of a road. The American Society for Testing and Materials (ASTM) defines road roughness as "the deviation of a surface with characteristic dimensions that affect the vehicle dynamics and ride quality" [13]. Improving road surface quality must start with measurement. Recent technology has made possible inexpensive sensors embedded in smartphones that can enable the collection of data for road surface quality monitoring. These inexpensive sensors include 3-axis accelerometers, gyroscopes, and GPS that can aid in collecting, processing, and classifying data for road surface quality monitoring.

The use of passenger's smartphones to collect road surface data through participatory sensing [5] provides a cost-effective means of aggregating a dataset for characteristics of roads travelled. The dataset can then be used by classification algorithms to classify the dataset to identify different classes of road. Providing road conditions gives routing insight to passengers and drivers as an indication of their ride quality. Providing road condition information also aids government departments to improve the assessment for road construction and maintenance.

1.2 Prior work

This thesis builds on prior work done ar Ashesi University. Vorgbe [18], Boohene [5], Abeo [1], and Maxwell Aladago have explored the collection of accelerometer and GPS data from smartphones to classify road surface quality using Android devices. Vorgbe's initial work focused on collecting accelerometer and GPS data, and then implementing and training a logistic regression classifier to classify the data on the road surfaces recorded into one of three classes: bad, fair, and good. Boohene's work focused on building a mobile application to allow individuals to collect data about the quality of roads they travel on and view the data on a map interface. Abeo's work explored other machine learning algorithms to improve the classification results; the best resulting algorithm was of decision trees. . The difficulty in distinguishing between fair roads from other roads was an issue identified and this was key in driving subsequent work. Maxwell Aladago explored the classification algorithms: logistic regression, support vector machines and a multilayer perceptron using a newly curated dataset to classify road surfaces into one of four classes: very bad, bad, good, and very good Increasing the number of classes from three to four seeks to improve the precision of classification [2], allowing weaker significant features to strongly contribute to the classification in the increased classes. Prior work of linear and non-linear approaches has performed better on binary classification, but these approaches have been unable to classify well beyond two classes.

1.3 Purpose of research

The purpose of this research is to characterize the behaviour of the logistic regression classification algorithm to understand how the classification algorithm translates into a realistic situation. The classification algorithm is expected to give the intended users of road condition information a better sense of roads and not wrong information to frustrate them.

1.4 Problem statement

Despite the improvement in classification results from classification algorithms in prior work, there has not been a study into how the classification translates into real world situations for use. The gap in the precision of the road condition classification still lingers.

1.5 Research approach

The approach of this research will use the dataset by Maxwell Aladago and a Logistic Regression algorithm to describe the behaviour of the classification algorithm. Logistic regression is used in this study because its gives unbiased, low-varianced results, and its better performance on linearly seperable data as the different types of roads are linearly seperable. This is sought to improve the precision of classification by increasing the number of classes from three (bad, fair and good) to four (very bad, bad, good and very good).

1.6 Research questions

This research seeks to answer the following research questions:

1. What are the best set of features to extract for the algorithm?
2. What is the effect of device orientation on the algorithm's performance?
3. What is the effect of vehicle type on the algorithm's performance?

1.7 Structure of research paper

This thesis is structured as follows: Chapter 1 introduces the research project and provides background information on prior research done which this thesis furthers. Chapter 2 discusses related work to the research on road surface quality classification. Chapter 3 describes the methodology used in this research. Chapter 4 discloses and analyses the results of this research, and Chapter 5 summarizes the study conducted, concludes the thesis paper, and provides recommendations.

Chapter 2: Related Work

2.1 Prior work

Much research has gone into the problem of road quality classification and its subsequent monitoring. These research have followed varied approaches and explored different vehicles, leaving much to be learnt and strategies to build on. Some of these approaches have measured up to the international roughness index (IRI), and other studies have validated these approaches.

Vorgbe [18] using the logistic regression algorithm in his work with a held-out test was able to reliably distinguish between good and bad roads at a true positive rate of 92% as well as good and fair roads at a true positive rate of 83%. The logistic regression algorithm was, however, unable to reliably distinguish between bad and fair roads, fair roads from good and bad roads combined as well as bad roads from good and fair roads.

Abeo [1] explored a logistic regression, decision tree, random forest, K Nearest Neighbors, and Support Vector Machine algorithms. The performance of the algorithms was evaluated using 10-fold cross-validation. The decision tree had the best accuracy for classifying a road as good, fair, or bad producing true positives of 97% accuracy for bad roads, 81% accuracy for fair roads, and 93% accuracy for good roads. The other algorithms were reliably able to classify data belonging to good or bad however were unable to reliably classify roads as fair, being more likely to classify them as good. The decision tree was proposed in place of Vorgbe’s algorithm for its best overall accuracy of 92% with a precision of 92% and recall of 90% as more likely to accurately predict a new data point to its true class.

Maxwell Aladago presents a study seeking to develop a method to detect and classify road surfaces using a newly automatically was curated dataset. The classification considered four classes of roads: very good, good, bad, and very bad. A multinomial logistic regression performed for all the four road types obtained an F1-score of 0.59. A binary logistic regression between very good roads and very bad roads obtained an F1-score of 0.96 and an accuracy of 96%. Very good and good roads versus very bad and bad roads recorded an f1-score of 0.87. Good roads only versus very bad roads obtained an f1-score of 0.89.

Maxwell Aladago developed a method that reliably distinguished between very good and very bad roads and good roads from very bad roads. These results are like Vorgbe's initial results using a logistic regression algorithm.

This paper seeks understand the behaviour of the logistic regression algorithm on how it will translate into realistic situations of use. This can help to determine a model that can best be applied to the road surface quality classification problem.

2.2 Similar projects

Neural Networks.

Neural Networks architectures have improved approaches to object character recognition by varying the depth and breadth of the network. Convolutional Neural Networks (CNNs) make reliable and correct assumptions about the nature of images. The ImageNet classification with a deep convolutional neural network variant developed by Krizhevsky et al. [10] has been the most effective neural network at image classification. It comes as no surprise that deep CNNs are used in image classification tasks such as large-scale image recognition [16], and identifying patterns in urban environments [3]. CNNs have also been used in detecting roads in high-resolution aerial images [9] and deep learning on road satellite imagery for road quality classification [6].

Cadamuro, Muhebwa, and Taneja [6] in their work develop a model for monitoring road infrastructure quality using satellite imagery and relate the quality of intercity roads to economic activity. CNNs and an auto-encoder network structure were trained on the satellite images to compare which structure performed better. All the CNNs were trained over ten epochs of the data, augmented by random flips vertically and horizontally to prevent overfitting on the data. The auto-encoder was also trained overnight on the data for 20 epochs and subsequently regressed with an L2 penalty to enable the network to extract features for any road. The results show a better performance achieved by the auto-encoder on the key regression metric due to its superior generalization performance on unseen data. The best-case results of 0.79 R2 value for the regression of a standard train-test split instance and 0.35 R2 value for a harder held-out regression serve a good potential to generalize, the importance for real-world application.

The International Roughness Index (IRI) [14] established in 1986 has been the world metric for calculating the roughness of a road from a measured longitudinal road profile, yielding an index that classifies a road as one type or another.

An approach that has been used for road surface quality monitoring is signal processing. Forslöf and Jones [11] classified roads by collecting accelerometer amplitude levels and vehicle speed from smartphones in a moving vehicle. Regression analysis was performed on the data on two options for roughness. The first option was an eIRI (estimated IRI) based on a peak and root mean square (RMS) vibration analysis correlated to Swedish laser measurements on paved roads for classification of single points and stretches of road. The second option was a CIRI (calculated IRI) based on the quarter-car simulation (QCS) for sampling where the sensitivity of the device is calibrated by the owner to their reference to implement the IRI. The Roadroid [11] software bundle was passed the information from the smartphones as a cloud-based web geographical information systems (GIS). This use of a smartphone is demonstrated as an efficient, scalable, and cost-effective way of classifying and monitoring road surfaces.

The use of smartphones as a tool for measuring road surface quality has been tested for its efficiency and scalability. Yehaneh et al. [7], in their work, sought to validate the use of smartphones for measuring roughness along roads. By calculating an equilibrium between the (RMS) and the IRI from a sample of accelerometer outputs for a route with different pavement conditions, they discovered a correlation that allowed for the use of the inexpensive, easy to implement, and widely accessible tools of smartphones as a viable tool for measuring road quality. They conclude that smartphones can be deployed to estimate pavement roughness at an adequate level of precision and accuracy. This validation has encouraged this approach in other studies [8,13,15].

2.3 Research gap

Based on the literature reviewed, a knowledge gap is identified on the effect of orientation of the device at data collection and the vehicle used in data collection, on the performance of the classification algorithm. Also, there is a lack of clear objective distinction between data points to help distinguish between classes.

Chapter 3: Methodology

3.1 Machine learning

Machine learning is a branch of artificial intelligence where computer systems can learn from collected data, identify patterns in the data, and be able to make decisions with little or no human intervention. Machine learning uses data analysis methods to automate analytical model building. Machine learning algorithms are domain biased, given the available data to be effective at solving problems and evaluating their performance.

This paper explores a machine learning algorithm as an approach to solve the road surface quality classification problem. Vorgbe and Abeo evaluated Linear Regression, Logistic Regression, Support Vector Machines (SVM), K Nearest Neighbor (KNN), Decision Tree (DT) and Random Forests (RF) algorithms for classifying road surface quality.

The machine learning algorithm explored in this paper is the Logistic Regression classification algorithms to improve the four-type road classification, i.e., Very Bad, Bad, Good, and Very Good.

3.2 Data Collection

This study used the raw, cleaned, and labelled data collected by Maxwell Aladago. Three vehicles: a Toyota pickup truck (referred to as Pick up), a minivan (referred to as Minivan) and a green Hyundai Sports Utility Vehicle (SUV) (referred to as Green car) were each fitted with three Samsung Galaxy tablets and an Infinix Note 3 phone for collecting data. The tablets and a smartphone equipped with a tri-axial accelerometer and GPS sensors ran a custom-built Android application that recorded data from the accelerometer and GPS sensors of the devices. The tablets were fitted at fixed orientations, one vertically, a second horizontally (with the screen facing up), and the third diagonally (slanted) relative to the vehicle's bonnet. The phone was held by a data collector, roaming freely (free roam) during the data collection process. The data collector annotated the segments of road driven on with a semantic label representing the quality of the road as one of the four road types.

The data collection application was initialized to start recording when the vehicle was moving. The application recorded simultaneously on all four devices accelerometer and GPS readings at a frequency of 5Hz (i.e., five data point readings in a second). The features recorded by the application were:

1. Timestamp
2. GPS coordinates: Longitude and Latitude
3. Speed of the vehicle in m/s
4. The tri-axial accelerations along the X, Y and Z axes
5. A flag of presence or absence of an anomaly at that point, e.g., a speed ramp
6. A categorical label of the road quality by the data collector's annotation

The road circuit used for the data collection was within the Greater Accra and Eastern region of Ghana.

3.3 Road Types

The classification of roads for use in this study are defined below:

1. Very good road: Paved roads which cause little to no vibrations on the vehicle
2. Good road: Roads which cause minimal amounts of vibrations on the vehicle
3. Bad road: Roads that are rough and cause large vibrations on the vehicle
4. Very bad roads: Roads that are barely motorable and roads with large potholes.

The discrepancy of overlapping road types, which leads to noisy labels was addressed by considering data segments, data points that can be traversed in 10 seconds as windows for feature extraction.

3.4 Labelling and Cleaning

Every data point was associated with a road type label as recorded by the data collector. The tablets recorded accelerometer, and GPS data of the vehicle without labels in their fixed orientations. The labels for road segments, as annotated by the data collector was associated with all other data points via the closest timestamp. To enforce timestamp on all data points, time was sourced from the GSM network provider used by all the four devices, MTN.

Cleaning after labelling removed all data points with vehicle speed values of less than 0.001m/s before further processing. These points are characteristic of when the vehicle is either at a standstill or slow-moving. Additional cleaning removed data points that remained unlabeled. These points are indicative of the beginning of the data collection when a road type is not initially assigned.

3.5 Data Distribution

The distribution of the data points for all road types to each vehicle is shown in Table 3.1 below. The distribution of the data points for each vehicle on all orientations including the free roam device is shown in Table 3.2 below.

	Road Type (%)			
Vehicle	Very Bad	Bad	Good	Very Good
Green car	32	14	31	23
Pickup	36	13	15	36
Minivan	32	12	30	26

Table 3.1: Distribution of road types across each vehicle

Vehicle Type	Orientation	Road Type (%)			
		Very Bad	Bad	Good	Very Good
Green car	Free roam	32	14	31	23
	Horizontal	32	15	31	23
	Slanting	31	14	31	24
	Vertical	32	14	31	24
Pickup	Free roam	36	12	16	36

	Horizontal	36	13	15	37
	Slanting	36	13	15	36
	Vertical	36	13	15	36
Minivan	Free roam	34	13	27	26
	Horizontal	32	12	31	26
	Slanting	31	12	31	26
	Vertical	31	11	31	27

Table 3.2: Distribution of road types for all orientations across each vehicle

3.6 Feature Extraction

The features extracted in this study were data segments of data points collected in 10 second windows as the vehicle moved. Each data segment consisted of 50 data points. Three sets of features, each from Francis Vorgbe, Anthony Abeo and Maxwell Aladago, were extracted on the data and considered in this study. These sets of features were to be compared to see which set gives the best performance of the logistic regression algorithm. A data segment is also referred to as a window in this study.

Feature set 1: Vorgbe's Features

1. Z-Mean: the mean of all z-axial reading in a window
2. Z-Var: the variance of the z values in a window
3. Z-SD: the standard deviation of all z values in a window
4. Z-HPeak: the highest z-axial value recorded in a window
5. Z-LTrough: the lowest z-axial value recorded in each window
6. Z-DiffMean: the mean difference between successive peaks and troughs of z-axial readings in a window
7. Z-DiffVar: the variance of the difference between successive peaks and troughs of z-axial readings in a window
8. Z-DiffSD: the standard deviation of the difference between successive peaks and troughs of z-axial readings in a window

9. X-Var: the variance of the x-axial values in a window
10. X-SD: the standard deviation of all x-axial values in a window
11. X-HPeak: the highest x-axial value recorded in a window
12. X-LTrough: the lowest x-axial value recorded in a window
13. Y-Var: the variance of the y-axial values in a window
14. Y-SD: the standard deviation of all y-axial values in a window

Feature set 2: Abeo's Features

1. Z-Mean: the mean of all z-axial reading in a window
2. Z-Var: the variance of the z values in a window
3. Z-SD: the standard deviation of all z values in a window
4. Z-HPeak: the highest z-axial value recorded in a window
5. Z-LTrough: the lowest z-axial value recorded in each window

Feature set 3: Maxwell Aladago' Features

1. X-Peak: the highest accelerometer reading along the X-coordinate
2. Y-Peak: the highest accelerometer reading along the Y-coordinate
3. Z-Peak: the highest accelerometer reading along the Z-coordinate
4. X-Var: the variance of the accelerometer readings along the X-coordinate
5. Y-Var: the variance of the accelerometer readings along the Y-coordinate
6. Z-Var: the variance of the accelerometer readings along the Z-coordinate
7. X-Trough: the lowest accelerometer reading along the X-coordinate
8. Y-Trough: the lowest accelerometer reading along the Y-coordinate
9. Z-Trough: the lowest accelerometer reading along the Z-coordinate
10. The highest speed recorded in the segment
11. The lowest speed recorded in the segment
12. The variation of the speed values within the segment

3.7 Accelerometer reorientation

The tri-axial accelerometer sensor detects linear accelerations in a 3-dimensional frame of a_x, a_y, a_z by measuring the inertial forces in the longitudinal direction along the x-axis, the transverse direction along the y-axis and the z-axis which is perpendicular to the xy-plane directions [2]. In a real-world scenario, the orientation of a mobile device in a vehicle can vary. To resolve the orientation of the tri-axial directions of the data collecting device and the vehicle, the accelerometer data can be reoriented through Euler Angles as used in [2]. This reorientation transforms the accelerometer axial orientation to match the orientation of the vehicle [4].

A vehicle in a stationary position will have only the acceleration of gravity recorded along the z-axis.

$$a_x = 0\ m/s^2;$$

$$a_y = 0\ m/s^2;$$

$$a_z = 9.81\ m/s^2 = 1g \quad (1)$$

Given the XYZ sequence, a rotation around the x-axis by an angle α (roll angle), and one around the y-axis by β (pitch angle) is done. Where

$$\alpha = \tan^{-1}(a_{y'} / a_{z'}) \text{ in the range } [-\pi: \pi] ;$$

$$\beta = \tan^{-1}(-a_{x'} / (\sqrt{(a_{y'})^2 + (a_{z'})^2})) \text{ in the range } [-\pi/2; \pi/2] \quad (2)$$

The reoriented accelerometer values are estimated using equations 3, 4, and 5 where ***c*** and ***s*** represent cosine and sine respectively of the angles.

$$a_{xreor} = c_\beta\ a_x' + s_\beta\ s_\alpha\ a_y' + c_\alpha\ s_\beta\ a_z'; \quad (3)$$

$$a_{yreor} = c_\alpha\ a_y' - s_\alpha\ a_z'; \quad (4)$$

$$a_{zreor} = -\ s_\beta\ a_x' + c_\beta\ s_\alpha\ a_y' + c_\beta\ c_\alpha\ a_z' \quad (5)$$

The reorientation is done to compare performance of the classification algorithm between the data points in their primarily collected orientations and their reoriented orientations.

3.8 Classification Algorithm

The Logistic Regression algorithm used was implemented using Scikit-learn [3]. Logistic Regression is a linear classification algorithm that utilizes a sigmoid function to quell the value generated by the classification algorithm into a value in the range 0 to 1[1]. The output value represents the probability of the input data belonging to a class. Thus, given two classes A and B, the output for the of the Logistic Regression algorithm will be a probability of a test data point belonging to A, and the compliment of the probability is the probability of the test data belonging to class B.

The implementation of Scikit-learn Logistic Regression algorithm used in this study used the Limited-memory Broyden-Fletcher-Goldfarb-Shanno algorithm (LBFGS) solver, the one-vs.-rest approach, and L2 regularization to minimize the following optimization cost function:

$$min_{w,c} \; ||w||_1 \; C \sum_{i=1}^{n} log(exp(-y_i(X^T{}_i w + c)) + 1)$$

Through trials, best parameters for implementing the Logistic Regression algorithm in Scikit-Learn were as follows:

```
penalty = l2, C = 0.1, max_itera = 1e6,
class_weight = balanced, multi_class = ovr and
solver = 'lbfgs'
```

3.9 Classification Scenarios

The performance of the classification algorithm was tested using two binary scenarios and a multiclass classification scenario.

Binary classification:

I. Very Good (VG) vs. Very Bad (VB): Only data windows labelled as very good roads and very bad roads were considered for the binary classification.

II. Very Good & Good (VGG) vs. Very Bad & Bad (VBB): Data windows labelled as very good were combined with windows labelled as good for one class and data windows labelled as very bad were combined with windows labelled as bad for the second class in the binary classification.

Multiclass classification:

I. Very Bad (VG) vs. Bad (B) vs. Good (G) vs. Very Good (VG): Data windows labelled for each road type was considered as a separate class in multiclass classification.

3.10 Experiments

Experiments were grouped to answer each research question in this study, with each experiment done by the classification scenarios. The experiments to answer each question are outlined below.

I. To determine what set of features works best for the algorithm, each fixed orientation along each vehicle was used to train and test the classification algorithm for the classification scenarios varying each set of features.

II. To determine the effect of device orientation on algorithm performance, the best performing set of features from I above and the green car were considered for the classification scenarios varying data by each orientation. Data by each orientation was used to train the classification algorithm and tested on the same orientation and every other orientation.

    Also, two datasets, one of primary orientations and a second of reoriented data points from the green car were split to train and test on respectively for comparison to see the change in performance of the algorithm when data points are oriented to match the orientation of the car.

III. To determine the effect of vehicle type on the algorithm performance, the best performing set of features from I above and data for the slanting orientation by each vehicle were considered for the classification scenarios varying the vehicle type.

Data collected in the slanting orientation for each vehicle was used to train the classification algorithm and tested on the same slanting orientation of the same vehicle and every other vehicle's data for the slanting orientation.

IV. To determine the accuracy of the algorithm on all the dataset, the best set of features from I are used to train and test primary oriented and reoriented data points. To further gain an objective distinction between data points, the probability of test data being predicted to their ground truth label and their predicted label was visualized for analysis.

Chapter 4: Results

This section shows the results obtained from conducting the four groups of experiments described in the methodologyand the findings thereof. It describes each experiment conducted with the logistic regression algorithm, a table of the results of the experiment and narrative explanations of the results. It goes on to analyze and attempt to explain findings of the group of experiments.

4.1 Group I Experiment Results

Experiments in this group trained and tested for each of the three vehicles, for each of the four orientations, each of the three sets of features in the binary classification scenario of Very Good (VG) vs. Very Bad (VB) and Very Good & Good (VGG) vs. Very Bad & Bad (VBB) and multiclass classification of Very Bad (VG) vs. Bad (B) vs. Good (G) vs. Very Good (VG). The set of features that gave the most occurrences of the best performance is adjudicated as the best set of features.

Green car – Free roam orientation

Green car - Free roam orientation		
Vorgbe		
	Binary - VG vs. VB	
Accuracy	95.5%	
F1-score	95.4%	
	Confusion matrix	
	Predicted class	
True class	VB	VG
VB	52.2%	1.5%
VG	3.0%	43.3%
Abeo		
	Binary - VG vs. VB	
Accuracy	89.6%	
F1-score	89.4%	
	Confusion matrix	
	Predicted class	
True class	VB	VG

VB	50.0%	3.7%
VG	6.7%	39.6%
Maxwell Aladago		
	Binary - VG vs. VB	
Accuracy	94.0%	
F1-score	94.0%	
	Confusion matrix	
	Predicted class	
True class	VB	VG
VB	50.0%	3.7%
VG	2.2%	44.0%

Number of test data = 134

Table 4.1: Test results for all feature sets on Green car - Free roam orientation data in Very Good vs. Very Bad classification

1. In binary VG v VB, Francis' features were best forming by a slight margin with the LR algorithm, in terms of accuracy (95.5%) and f1-score (95.4%), followed by Maxwell Aladago features (accuracy = 94%, f1-score = 94%) then Anthony's features (accuracy = 89.6%, f1-score = 89.4%). Further analysis of the confusion matrices shows that Francis' features enabled the LR algorithm better to classify VB data points than VG data points. In contrast, Maxwell Aladago's features enabled the algorithm better classify VG data points than VB data points.

Green car - Free roam orientation		
Vorgbe		
	Binary - VGG vs. VBB	
Accuracy	84.6%	
F1-score	84.5%	
	Confusion matrix	
	Predicted class	
True class	VBB	VGG
VBB	39.6%	5.8%
VGG	9.6%	45.0%

Abeo		
	Binary - VGG vs. VBB	
Accuracy	81.3%	
F1-score	81.2%	
	Confusion matrix	
	Predicted class	
True class	VBB	VGG
VBB	37.9%	7.5%
VGG	11.3%	43.3%
Maxwell Aladago's		
	Binary - VGG vs. VBB	
Accuracy	84.6%	
F1-score	84.5%	
	Confusion matrix	
	Predicted class	
True class	VBB	VGG
VBB	37.9%	7.5%
VGG	7.9%	46.7%

Number of test data = 240

Table 4.2: Test results for all feature sets on Green car - Free roam orientation data in Very Good and Good vs. Very Bad and Bad classification

2. In binary VGG vs. VBB, both Francis' and Maxwell Aladago's features achieved the same level of accuracy (84.6%), and f1-score (84.5%) and Anthony's features were achieving the lower performance (accuracy = 81.3%, f1-score = 81.2%). An analysis of the confusion matrices shows Francis' features better at classifying VBB data points, and Maxwell Aladago's features better at classifying VGG data points.

Green car - Free roam orientation				
Vorgbe				
	Multiclass			
Accuracy	57.9%			
F1-score	53.6%			

	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	26.3%	2.9%	0.8%	0.0%
B	6.3%	4.2%	5.0%	0.0%
G	4.2%	5.0%	12.9%	8.8%
VG	0.0%	2.5%	6.7%	14.6%
Abeo				
	Multiclass			
Accuracy	52.1%			
F1-score	47.4%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	23.3%	3.8%	2.5%	0.4%
B	7.9%	2.5%	4.6%	0.4%
G	5.0%	4.2%	12.5%	9.2%
VG	0.4%	4.6%	5.0%	13.8%
Maxwell Aladago's				
	Multiclass			
Accuracy	55.0%			
F1-score	50.1%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	23.3%	5.0%	1.3%	0.4%
B	7.1%	2.9%	4.6%	0.8%
G	3.8%	3.8%	11.3%	12.1%
VG	0.0%	3.3%	2.9%	17.5%

Number of test data = 240

Table 4.3: Test results for all feature sets on Green car - Free roam orientation data in multiclass classification

3. In multiclass classification, Francis' features achieved the best performance (accuracy = 57.9%, f1-score = 53.6%) among the three sets of features with the LR algorithm. Maxwell Aladago's features had better performance (accuracy =

55%, f1-score = 50.1%) followed by Anthony's features performance (accuracy = 52.1%, f1-score = 47.4%). The confusion matrix gives additional insight on Francis' features being better suited at classifying VB and B roads, whereas Maxwell Aladago's features better suited at classifying G and VG roads. It is also worth adding that Francis's features did not misclassify any VG road as either VB or B.

Green car – Vertical orientation

Green car - Vertical orientation		
Vorgbe		
	Binary - VG vs. VB	
Accuracy	91.3%	
F1-score	91.2%	
	Confusion matrix	
	Predicted class	
True class	VB	VG
VB	49.3%	4.3%
VG	4.3%	42.0%
Abeo		
	Binary - VG vs. VB	
Accuracy	74.6%	
F1-score	74.4%	
	Confusion matrix	
	Predicted class	
True class	VB	VG
VB	42.0%	11.6%
VG	13.8%	32.6%
Maxwell Aladago's		
	Binary - VG vs. VB	
Accuracy	96.4%	
F1-score	96.4%	
	Confusion matrix	
	Predicted class	
True class	VB	VG
VB	51.4%	2.2%
VG	1.4%	44.9%

Number of test data = 138

Table 4.4: Test results for all feature sets on Green car - Vertical orientation data in Very Good vs. Very Bad classification

4. In binary VG v VB, Maxwell Aladago's features were best performing with accuracy and f1-score of 96.4%. Francis' features had the better performance with an accuracy of 91.3% and f1-score of 91.2%. Anthony's features followed with the lowest performance of accuracy 74.6% and f1-score of 74.4%.

Green car - Vertical orientation		
Vorgbe		
	Binary - VGG vs. VBB	
Accuracy	86.2%	
F1-score	86.1%	
	Confusion matrix	
	Predicted class	
True class	VBB	VGG
VBB	40.5%	6.1%
VGG	7.7%	45.7%
Abeo		
	Binary - VGG vs. VBB	
Accuracy	68.8%	
F1-score	68.8%	
	Confusion matrix	
	Predicted class	
True class	VBB	VGG
VBB	34.4%	12.1%
VGG	19.0%	34.4%
Maxwell Aladago's		
	Binary - VGG vs. VBB	
Accuracy	88.7%	
F1-score	88.7%	
	Confusion matrix	
	Predicted class	
True class	VBB	VGG
VBB	40.9%	5.7%

VGG	5.7%	47.8%

Number of test data = 247

Table 4.5: Test results for all feature sets on Green car – Vertical orientation data in Very Good and Good vs. Very Bad and Bad classification

5. In binary VGG v VBB, Maxwell Aladago's features were best performing with an accuracy and f1-score of 88.7%. Francis' features gave a slightly lower performance of accuracy 86.2% and f1-score of 86.1%. Anthony's features gave the lowest performance with accuracy and f1-score of 68.8%.

Green car - Vertical orientation				
Vorgbe				
	Multiclass			
Accuracy	63.6%			
F1-score	57.4%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	30.0%	2.0%	2.0%	0.0%
B	5.7%	4.0%	2.4%	0.4%
G	3.2%	4.5%	11.7%	9.7%
VG	0.8%	2.4%	3.2%	17.8%
Abeo				
	Multiclass			
Accuracy	43.3%			
F1-score	36.1%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	25.1%	4.5%	3.2%	1.2%
B	5.7%	1.6%	4.0%	1.2%
G	8.1%	5.3%	11.3%	4.5%
VG	4.9%	4.5%	9.7%	5.3%
Maxwell Aladago's				
	Multiclass			

Accuracy	59.5%			
F1-score	53.3%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	27.5%	4.0%	0.8%	1.6%
B	5.7%	3.2%	3.2%	0.4%
G	1.2%	3.6%	11.3%	13.0%
VG	0.0%	0.8%	6.1%	17.4%

Number of test data = 247

Table 4.6: Test results for all feature sets on Green car - Vertical orientation data in multiclass classification

6. In multiclass classification, Francis' features produced the best performance (accuracy = 63.6%, f1-score = 57.4%) followed by Maxwell Aladago's features (accuracy = 59.5%, f1-score = 53.3%) and then Anthony's features with the poorest performance (accuracy = 43.3%, f1-score = 36.1%).

Green car – Horizontal orientation

Green car - Horizontal orientation		
Vorgbe		
	Binary - VG vs. VB	
Accuracy	91.2%	
F1-score	90.9%	
	Confusion matrix	
	Predicted class	
True class	VB	VG
VB	53.7%	3.7%
VG	5.1%	37.5%
Abeo		
	Binary - VG vs. VB	
Accuracy	91.2%	
F1-score	91.0%	
	Confusion matrix	
	Predicted class	
True class	VB	VG

VB	52.2%	5.1%
VG	3.7%	39.0%
Maxwell Aladago’s		
	Binary - VG vs. VB	
Accuracy	97.8%	
F1-score	97.8%	
	Confusion matrix	
	Predicted class	
True class	VB	VG
VB	55.9%	1.5%
VG	0.7%	41.9%

Number of test data = 136

Table 4.7: Test results for all feature sets on Green car - Horizontal orientation data in Very Good vs. Very Bad classification

7. In binary VG v VB, Maxwell Aladago’s features were best performing with performance and f1-score of 97.8%, Francis’ and Anthony’s features both achieved 91.2% accuracy and f1-score of 91%.

Green car - Horizontal orientation		
Vorgbe		
	Binary - VGG vs. VBB	
Accuracy	82.6%	
F1-score	82.4%	
	Confusion matrix	
	Predicted class	
True class	VBB	VGG
VBB	36.4%	9.3%
VGG	8.1%	46.2%
Abeo		
	Binary - VGG vs. VBB	
Accuracy	81.4%	
F1-score	81.2%	
	Confusion matrix	
	Predicted class	
True class	VBB	VGG

VBB	35.2%	10.5%
VGG	8.1%	46.2%
Maxwell Aladago's		
	Binary - VGG vs. VBB	
Accuracy	91.9%	
F1-score	91.8%	
	Confusion matrix	
	Predicted class	
True class	VBB	VGG
VBB	41.3%	4.5%
VGG	3.6%	50.6%

Number of test data = 247

Table 4.8: Test results for all feature sets on Green car – Horizontal orientation data in Very Good and Good vs. Very Bad and Bad classification

8. In binary VGG vs. VBB, Maxwell Aladago's features produced the best performance with accuracy 91.9% and f1-score of 91.8%. Francis' features were better performing with accuracy of 82.6% and f1-score 82.4%. Anthony's features were least performing with accuracy of 81.4% and f1-score of 81.2%.

Green car - Horizontal orientation				
Vorgbe				
	Multiclass			
Accuracy	54.7%			
F1-score	48.2%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	25.1%	3.2%	2.0%	2.4%
B	6.1%	2.0%	4.5%	0.4%
G	3.6%	3.6%	12.1%	10.9%
VG	0.4%	0.8%	7.3%	15.4%
Abeo				
	Multiclass			
Accuracy	54.3%			
F1-score	49.1%			

	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	25.5%	3.2%	3.2%	0.8%
B	5.7%	3.2%	1.6%	2.4%
G	4.0%	3.6%	11.3%	11.3%
VG	0.8%	0.8%	8.1%	14.2%
Maxwell Aladago’s				
	Multiclass			
Accuracy	62.3%			
F1-score	55.9%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	29.1%	2.4%	0.8%	0.4%
B	5.7%	3.2%	4.0%	0.0%
G	0.0%	5.3%	12.6%	12.6%
VG	0.0%	0.4%	6.1%	17.4%

Number of test data = 247

Table 4.9: Test results for all feature sets on Green car - Horizontal orientation data in multiclass classification

9. In multiclass classification, Maxwell Aladago’s features produced the best performance of 62.3% accuracy and 55.9% f1-score. Francis’ features produced better results with the performance of 54.7% accuracy and 48.2% f1-score. Anthony’s features produced the lowest performance of 54.3% accuracy and 49.1% f1-score.

Green car – Slanting orientation

Green car - Slanting orientation		
Vorgbe		
	Binary - VG vs. VB	
Accuracy	87.0%	
F1-score	87.0%	
	Confusion matrix	
	Predicted class	

True class	VB	VG
VB	46.3%	7.4%
VG	5.9%	41.9%
Abeo		
	Binary - VG vs. VB	
Accuracy	83.3%	
F1-score	83.2%	
	Confusion matrix	
	Predicted class	
True class	VB	VG
VB	46.3%	7.4%
VG	9.6%	38.2%
Maxwell Aladago’s		
	Binary - VG vs. VB	
Accuracy	95.0%	
F1-score	95.0%	
	Confusion matrix	
	Predicted class	
True class	VB	VG
VB	49.3%	4.4%
VG	0.7%	47.1%

Number of test data = 138

Table 4.10: Test results for all feature sets on Green car - Slanting orientation data in Very Good vs. Very Bad classification

10. In binary VG vs. VB, Maxwell Aladago’s features produced the best performance with accuracy and f1-score of 95%. Francis’ features produced the next best performance with accuracy and f1-score of 87%. Anthony’s features produced the lowest performance with accuracy and f1-score of 83%.

Green car - Slanting orientation		
Vorgbe		
	Binary - VGG vs. VBB	
Accuracy	81.5%	
F1-score	81.1%	
	Confusion matrix	
	Predicted class	

True class	VBB	VGG
VBB	33.6%	8.5%
VGG	10.1%	48.2%
Abeo		
	Binary - VGG vs. VBB	
Accuracy	75.0%	
F1-score	75.0%	
	Confusion matrix	
	Predicted class	
True class	VBB	VGG
VBB	35.6%	6.5%
VGG	18.6%	39.7%
Maxwell Aladago's		
	Binary - VGG vs. VBB	
Accuracy	88.7%	
F1-score	88.3%	
	Confusion matrix	
	Predicted class	
True class	VBB	VGG
VBB	35.6%	6.5%
VGG	4.9%	53.4%

Number of test data = 248

Table 4.11: Test results for all feature sets on Green car – Slanting orientation data in Very Good and Good vs. Very Bad and Bad classification

11. In binary VGG vs. VBB, Maxwell Aladago's features produced the best performance of accuracy and f1-score of 88%. Francis's features produced the next best performance with an accuracy of 82% and f1-score of 81.1%. The lowest performance was from Anthony's features, which had an accuracy and f1-score of 75%.

Green car - Slanting orientation				
Vorgbe				
	Multiclass			
Accuracy	56.0%			

F1-score	51.2%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	23.9%	2.0%	1.6%	2.4%
B	4.9%	3.2%	3.2%	0.8%
G	3.6%	4.0%	15.0%	10.5%
VG	1.2%	4.9%	4.9%	14.2%
Abeo				
	Multiclass			
Accuracy	48.8%			
F1-score	41.2%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	25.1%	3.6%	1.2%	0.0%
B	5.7%	2.8%	2.0%	1.6%
G	6.9%	4.5%	17.0%	4.9%
VG	6.1%	4.9%	10.1%	4.0%
Maxwell Aladago’s				
	Multiclass			
Accuracy	62.1%			
F1-score	55.9%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	25.5%	1.6%	0.4%	2.4%
B	6.1%	2.8%	2.0%	1.2%
G	2.8%	3.2%	14.6%	12.6%
VG	0.4%	0.4%	4.9%	19.4%

Number of test data = 248

Table 4.12: Test results for all feature sets on Green car - Slanting orientation data in multiclass classification

12. In multiclass classification, Maxwell Aladago’s features produced the best performance with an accuracy of 62.1% and f1-score of 55.9%. Francis’ features were the next best performing with an accuracy of 56% and f1-score of 51.2%.

Anthony’s features had the lowest performance with an accuracy of 48.8% and f1-score of 41.2%.

Pickup – Free roam orientation

Pickup – Free roam orientation		
Vorgbe		
	Binary - VG vs. VB	
Accuracy	90.7%	
F1-score	90.7%	
	Confusion matrix	
	Predicted class	
True class	VB	VG
VB	44.8%	2.6%
VG	6.7%	45.9%
Abeo		
	Binary - VG vs. VB	
Accuracy	90.2%	
F1-score	90.2%	
	Confusion matrix	
	Predicted class	
True class	VB	VG
VB	44.8%	2.6%
VG	7.2%	45.4%
Maxwell Aladago’s		
	Binary - VG vs. VB	
Accuracy	89.7%	
F1-score	89.7%	
	Confusion matrix	
	Predicted class	
True class	VB	VG
VB	44.3%	3.1%
VG	7.2%	45.4%

Number of test data = 194

Table 4.13: Test results for all feature sets on Pickup – Free roam orientation data in Very Good vs. Very Bad classification

13. In binary VG vs. VB, Francis’ features were best performing with accuracy and f1-score of 90.7%. Anthony’s features were the next best performing with

accuracy and f1-score of 90.2%. Maxwell Aladago's features produced the lowest performance with accuracy and f1-score of 89.7%.

Pickup – Free roam orientation		
Vorgbe		
	Binary - VGG vs. VBB	
Accuracy	83.7%	
F1-score	83.7%	
	Confusion matrix	
	Predicted class	
True class	VBB	VGG
VBB	42.8%	8.3%
VGG	8.0%	40.9%
Abeo		
	Binary - VGG vs. VBB	
Accuracy	83.0%	
F1-score	83.0%	
	Confusion matrix	
	Predicted class	
True class	VBB	VGG
VBB	42.4%	8.7%
VGG	8.3%	40.5%
Maxwell Aladago's		
	Binary - VGG vs. VBB	
Accuracy	86.7%	
F1-score	86.7%	
	Confusion matrix	
	Predicted class	
True class	VBB	VGG
VBB	45.1%	6.1%
VGG	7.2%	41.7%

Number of test data = 264

Table 4.14: Test results for all feature sets on Pickup – Free roam orientation data in Very Good and Good vs. Very Bad and Bad classification

14. In binary VGG vs. VBB, Maxwell Aladago's features produced the best performance with accuracy and f1-score of 86.7%. Francis' features were the next best performing features with accuracy and f1-score of 83.7%. Anthony's features produced the lowest performance with accuracy and f1-score of 83%.

Pickup – Free roam orientation				
Vorgbe				
	Multiclass			
Accuracy	67.4%			
F1-score	56.2%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	31.8%	4.2%	0.4%	1.9%
B	3.0%	6.1%	1.9%	1.9%
G	1.9%	4.5%	3.0%	4.5%
VG	1.9%	3.4%	3.0%	26.5%
Abeo				
	Multiclass			
Accuracy	64.4%			
F1-score	52.3%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	33.0%	2.7%	1.5%	1.1%
B	4.5%	3.4%	2.3%	2.7%
G	1.9%	3.4%	3.8%	4.9%
VG	2.7%	2.3%	5.7%	24.2%
Maxwell Aladago's				
	Multiclass			
Accuracy	66.7%			
F1-score	53.7%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG

VB	31.8%	3.4%	0.8%	2.3%
B	6.8%	3.8%	1.1%	1.1%
G	1.5%	1.9%	3.0%	7.6%
VG	2.7%	1.1%	3.0%	28.0%

Number of test data = 264

Table 4.15: Test results for all feature sets on Pickup – Free roam orientation data in multiclass classification

15. In multiclass classification, Francis's features were best performing with an accuracy of 67.4% and f1-score of 56.2%. Maxwell Aladago's features were next best performing with accuracy of 66.7% and f1-score 53.7%. Anthony's features were the least performing with accuracy of 64.4% and f1-score of 52.3%.

Pickup – Vertical orientation

Pickup – Vertical orientation		
Vorgbe		
	Binary - VG vs. VB	
Accuracy	92.0%	
F1-score	92.0%	
	Confusion matrix	
	Predicted class	
True class	VB	VG
VB	49.8%	2.5%
VG	5.5%	42.3%
Abeo		
	Binary - VG vs. VB	
Accuracy	87.1%	
F1-score	87.0%	
	Confusion matrix	
	Predicted class	
True class	VB	VG
VB	47.8%	4.5%
VG	8.5%	39.3%
Maxwell Aladago's		
	Binary - VG vs. VB	
Accuracy	92.0%	
F1-score	92.0%	

	Confusion matrix	
	Predicted class	
True class	VB	VG
VB	48.3%	4.0%
VG	4.0%	43.8%

Number of test data = 201

Table 4.16: Test results for all feature sets on Pickup – Vertical orientation data in Very Good vs. Very Bad classification

16. In binary VG vs. VB, Maxwell and Francis' features produced equal performance of accuracy and f1-score of 92%. Francis' features had a better classification for VB road segments whereas Maxwell Aladago's features had a better classification for VG road segments. Anthony's features gave the lowest performance of 87% for accuracy and f1-score.

Pickup – Vertical orientation		
Vorgbe		
	Binary - VGG vs. VBB	
Accuracy	86.8%	
F1-score	86.8%	
	Confusion matrix	
	Predicted class	
True class	VBB	VGG
VBB	46.5%	5.5%
VGG	7.7%	40.3%
Abeo		
	Binary - VGG vs. VBB	
Accuracy	76.2%	
F1-score	76.1%	
	Confusion matrix	
	Predicted class	
True class	VBB	VGG
VBB	40.7%	11.4%
VGG	12.5%	35.5%

Maxwell Aladago's		
	Binary - VGG vs. VBB	
Accuracy	89.7%	
F1-score	89.7%	
	Confusion matrix	
	Predicted class	
True class	VBB	VGG
VBB	48.4%	3.7%
VGG	6.6%	41.4%

Number of test data = 273

Table 4.17: Test results for all feature sets on Pickup – Vertical orientation data in Very Good and Good vs. Very Bad and Bad classification

17. In binary VGG vs. VBB Maxwell Aladago's features gave the best performance with accuracy and f1-score of 89.7%, followed by Francis' features with accuracy and f1-score of 86.8% and trailed by Anthony's features with accuracy and f1-score of 76%.

Pickup – Vertical orientation				
Vorgbe				
	Multiclass			
Accuracy	62.3%			
F1-score	48.8%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	31.9%	6.2%	1.1%	0.7%
B	5.5%	3.3%	2.2%	1.1%
G	2.2%	4.4%	2.2%	4.8%
VG	2.2%	4.0%	3.3%	24.9%
Abeo				
	Multiclass			
Accuracy	56.4%			
F1-score	43.9%			
	Confusion matrix			
	Predicted class			

True class	VB	B	G	VG
VB	32.2%	4.0%	2.2%	1.5%
B	5.5%	2.6%	1.8%	2.2%
G	4.4%	1.5%	2.6%	5.1%
VG	5.1%	5.1%	5.1%	19.0%
Maxwell Aladago's				
	Multiclass			
Accuracy	67.4%			
F1-score	52.5%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	34.8%	2.6%	0.7%	1.8%
B	6.6%	3.7%	1.1%	0.7%
G	2.6%	2.2%	2.2%	6.6%
VG	1.5%	2.6%	3.7%	26.7%

Number of test data = 273

Table 4.18: Test results for all feature sets on Pickup – Vertical orientation data in multiclass classification

18. In multiclass classification, Maxwell Aladago's features were best performing with accuracy of 67.4% and f1-score 52.5%. Followed by Francis' features which gave an accuracy of 62.3% and 48.8% f1-score. Anthony's features gave the lowest performance of 56.4% accuracy and f1-score of 43.9%.

Pickup – Horizontal orientation

Pickup – Horizontal orientation		
Vorgbe		
	Binary - VG vs. VB	
Accuracy	91.6%	
F1-score	91.6%	
	Confusion matrix	
	Predicted class	
True class	VB	VG
VB	45.3%	1.0%

VG	7.4%	46.3%
Abeo		
	Binary - VG vs. VB	
Accuracy	86.7%	
F1-score	86.7%	
	Confusion matrix	
	Predicted class	
True class	VB	VG
VB	43.8%	2.5%
VG	10.8%	42.9%
Maxwell Aladago’s		
	Binary - VG vs. VB	
Accuracy	94.6%	
F1-score	94.6%	
	Confusion matrix	
	Predicted class	
True class	VB	VG
VB	44.3%	2.0%
VG	3.4%	50.2%

Number of test data = 203

Table 4.19: Test results for all feature sets on Pickup – Horizontal orientation data in Very Good vs. Very Bad classification

19. In binary VG vs. VB Maxwell Aladago’s features produced the best performance with an accuracy and f1-xore of 94.6%. This was followed by Francis’ features which produced performance accuracy and f1-score of 91.6% and then by Anthony's features which produced an accuracy and f1-score of 86.7%.

Pickup – Horizontal orientation		
Vorgbe		
	Binary - VGG vs. VBB	
Accuracy	83.3%	
F1-score	83.3%	
	Confusion matrix	
	Predicted class	
True class	VBB	VGG

VBB	41.1%	6.9%
VGG	9.8%	42.2%
Abeo		
	Binary - VGG vs. VBB	
Accuracy	80.0%	
F1-score	80.0%	
	Confusion matrix	
	Predicted class	
True class	VBB	VGG
VBB	41.1%	6.9%
VGG	13.1%	38.9%
Maxwell Aladago’s		
	Binary - VGG vs. VBB	
Accuracy	89.1%	
F1-score	89.1%	
	Confusion matrix	
	Predicted class	
True class	VBB	VGG
VBB	42.2%	5.8%
VGG	5.1%	46.9%

Number of test data = 275

Table 4.20: Test results for all feature sets on Pickup – Horizontal orientation data in Very Good and Good vs. Very Bad and Bad classification

20. In binary VGG vs. VBB Maxwell Aladago’s features produced the best performance of accuracy and f1-score of 89.1%. This was followed by Francis’ features which produced accuracy and f1-score of 83.3% and then with Anthony's features which produce accuracy and f1-score of 80%.

Pickup – Horizontal orientation				
Vorgbe				
	Multiclass			
Accuracy	59.3%			
F1-score	48.1%			
	Confusion matrix			

	Predicted class			
True class	VB	B	G	VG
VB	29.1%	2.5%	0.4%	2.2%
B	6.5%	3.6%	0.7%	2.9%
G	2.9%	2.9%	2.9%	8.0%
VG	4.0%	3.6%	4.0%	23.6%
Abeo				
	Multiclass			
Accuracy	55.3%			
F1-score	42.1%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	28.7%	4.4%	0.0%	1.1%
B	7.6%	2.2%	0.4%	3.6%
G	3.6%	2.5%	1.8%	8.7%
VG	5.8%	5.1%	1.8%	22.5%
Maxwell Aladago’s				
	Multiclass			
Accuracy	68.7%			
F1-score	57.6%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	29.8%	1.8%	0.7%	1.8%
B	6.9%	4.0%	1.8%	1.1%
G	1.5%	1.1%	4.7%	9.5%
VG	1.1%	1.5%	2.5%	30.2%

Number of test data = 275

Table 4.21: Test results for all feature sets on Pickup – Horizontal orientation data in multiclass classification

21. In multiclass classification, Maxwell Aladago’s features were best performing with accuracy of 68.7% and f1-score of 57.6%. This was followed by Francis’ features which gave an accuracy of 59.3% and f1-score of 48.1%. Anthony's features gave the lowest performance of 55.3% accuracy and f-score of 42.1%.

Pickup – Slanting orientation

Pickup – Slanting orientation		
Vorgbe		
	Binary - VG vs. VB	
Accuracy	90.1%	
F1-score	90.1%	
	Confusion matrix	
	Predicted class	
True class	VB	VG
VB	45.8%	1.5%
VG	8.4%	44.3%
Abeo		
	Binary - VG vs. VB	
Accuracy	89.2%	
F1-score	89.2%	
	Confusion matrix	
	Predicted class	
True class	VB	VG
VB	45.3%	2.0%
VG	8.9%	43.8%
Maxwell Aladago's		
	Binary - VG vs. VB	
Accuracy	95.1%	
F1-score	95.1%	
	Confusion matrix	
	Predicted class	
True class	VB	VG
VB	46.3%	1.0%
VG	3.9%	48.8%

Number of test data = 203

Table 4.22: Test results for all feature sets on Pickup – Slanting orientation data in Very Good vs. Very Bad classification

22. In binary VG vs. VB, Maxwell Aladago's features produced the best results with accuracy and f1-score of 95.1%. Followed by Francis' features with the next best results of accuracy and f1-score of 90.1%. Anthony's features gave the lowest performance with accuracy and f1-score of 89.2%.

Pickup – Slanting orientation		
Vorgbe		
	Binary - VGG vs. VBB	
Accuracy	82.5%	
F1-score	82.5%	
	Confusion matrix	
	Predicted class	
True class	VBB	VGG
VBB	38.9%	10.5%
VGG	6.9%	43.6%
Abeo		
	Binary - VGG vs. VBB	
Accuracy	82.9%	
F1-score	82.9%	
	Confusion matrix	
	Predicted class	
True class	VBB	VGG
VBB	40.0%	9.5%
VGG	7.6%	42.9%
Maxwell Aladago's		
	Binary - VGG vs. VBB	
Accuracy	89.8%	
F1-score	89.8%	
	Confusion matrix	
	Predicted class	
True class	VBB	VGG
VBB	44.7%	4.7%
VGG	5.5%	45.1%

Number of test data = 275

Table 4.23: Test results for all feature sets on Pickup – Slanting orientation data in Very Good and Good vs. Very Bad and Bad classification

23. In binary VGG vs. VBB, Maxwell Aladago's features produced the best results with accuracy and f1-score of 89.8%. Anthony's features produced the nest best

results with accuracy and f1-score of 82.9% whereas Francis' features produced the lowest results with accuracy and f1-score of 82.5%.

Pickup – Slanting orientation				
Vorgbe				
	Multiclass			
Accuracy	58.2%			
F1-score	47.0%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	26.9%	4.0%	1.5%	2.2%
B	8.0%	2.5%	0.7%	3.6%
G	1.5%	2.5%	3.6%	8.0%
VG	4.4%	1.5%	4.0%	25.1%
Abeo				
	Multiclass			
Accuracy	60.0%			
F1-score	42.9%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	29.8%	1.8%	0.4%	2.5%
B	8.0%	1.8%	0.7%	4.4%
G	1.8%	1.1%	1.1%	11.6%
VG	5.1%	1.1%	1.5%	27.3%
Maxwell Aladago's				
	Multiclass			
Accuracy	67.6%			
F1-score	56.4%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	29.8%	3.3%	0.7%	0.7%
B	7.3%	5.5%	1.5%	0.7%
G	0.4%	2.2%	3.3%	9.8%
VG	2.2%	2.2%	1.5%	29.1%

Number of test data = 275

Table 4.24: Test results for all feature sets on Pickup – Slanting orientation data in multiclass classification

24. In multiclass classification, Maxwell Aladago's features produced the best results with accuracy of 67.6% and f1-score of 56.4%. Anthony's features produced the next best results with accuracy of 60% and f1-score of 42.9%. Francis' featured produced the lowest performance with accuracy of 58.2% and f1-score of 47%.

White van – Free roam orientation

White van – Free roam orientation		
Vorgbe		
	Binary - VG vs. VB	
Accuracy	95.8%	
F1-score	95.8%	
	Confusion matrix	
	Predicted class	
True class	VB	VG
VB	51.9%	1.4%
VG	2.8%	44.0%
Abeo		
	Binary - VG vs. VB	
Accuracy	90.3%	
F1-score	90.2%	
	Confusion matrix	
	Predicted class	
True class	VB	VG
VB	49.5%	3.7%
VG	6.0%	40.7%
Maxwell Aladago's		
	Binary - VG vs. VB	
Accuracy	94.4%	
F1-score	94.4%	
	Confusion matrix	
	Predicted class	
True class	VB	VG
VB	50.0%	3.2%

VG	2.3%	44.4%

Number of test data = 216

Table 4.25: Test results for all feature sets on White van – Free roam orientation data in Very Good vs. Very Bad classification

25. In binary VG vs. VB, Maxwell Aladago's features produced the best results with accuracy and f1-score of 94.4%. Followed by Francis' features with the next best results of accuracy and f1-score of 95.8%. Anthony's features gave the lowest performance with accuracy of 90.3% and f1-score of 90.2%.

White van – Free roam orientation		
Vorgbe		
	Binary - VGG vs. VBB	
Accuracy	84.3%	
F1-score	84.2%	
	Confusion matrix	
	Predicted class	
True class	VBB	VGG
VBB	37.9%	5.4%
VGG	10.3%	46.3%
Abeo		
	Binary - VGG vs. VBB	
Accuracy	79.9%	
F1-score	79.8%	
	Confusion matrix	
	Predicted class	
True class	VBB	VGG
VBB	36.0%	7.3%
VGG	12.7%	43.9%
Maxwell Aladago's		
	Binary - VGG vs. VBB	
Accuracy	83.2%	
F1-score	83.1%	
	Confusion matrix	

	Predicted class	
True class	VBB	VGG
VBB	37.1%	6.2%
VGG	10.6%	46.1%

Number of test data = 369

Table 4.26: Test results for all feature sets on White van – Free roam orientation data in Very Good and Good vs. Very Bad and Bad classification

26. In binary VGG vs. VBB, Francis' features produced the best results with accuracy of 84.3% and f1-score of 84.2%. Followed by Maxwell Aladago's features with the next best results of accuracy of 83.2% and f1-score of 83.1%. Anthony's features gave the lowest performance with accuracy of 79.9% and f1-score of 79.8%.

White van – Free roam orientation				
Vorgbe				
	Multiclass			
Accuracy	60.2%			
F1-score	55.8%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	25.2%	5.7%	2.2%	0.5%
B	2.4%	4.1%	2.7%	0.5%
G	3.3%	6.0%	15.4%	7.9%
VG	0.8%	1.9%	6.0%	15.4%
Abeo				
	Multiclass			
Accuracy	55.0%			
F1-score	49.9%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	26.3%	5.1%	1.1%	1.1%
B	2.7%	3.5%	2.7%	0.8%

G	4.1%	6.5%	11.1%	10.8%
VG	1.9%	2.4%	5.7%	14.1%
Maxwell Aladago's				
	Multiclass			
Accuracy	59.9%			
F1-score	55.5%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	24.4%	7.0%	1.6%	0.5%
B	3.0%	4.1%	2.2%	0.5%
G	3.3%	6.2%	13.6%	9.5%
VG	0.3%	2.7%	3.3%	17.9%

Number of test data = 369

Table 4.27: Test results for all feature sets on White van – Free roam orientation data in multiclass classification

27. In multiclass classification, Francis' features produced the best results with accuracy of 60.2% and f1-score of 55.8%. Followed by Maxwell Aladago's features with the next best results of accuracy of 59.9% and f1-score of 55.5%. Anthony's features gave the lowest performance with accuracy of 55% and f1-score of 49.9%.

White van – Vertical orientation

White van – Vertical orientation		
Vorgbe		
	Binary - VG vs. VB	
Accuracy	87.4%	
F1-score	87.4%	
	Confusion matrix	
	Predicted class	
True class	VB	VG
VB	43.4%	5.6%
VG	7.0%	44.1%
Abeo		
	Binary - VG vs. VB	

Accuracy	70.6%	
F1-score	70.6%	
	Confusion matrix	
	Predicted class	
True class	VB	VG
VB	34.3%	14.7%
VG	14.7%	36.4%
Maxwell Aladago's		
	Binary - VG vs. VB	
Accuracy	94.4%	
F1-score	94.4%	
	Confusion matrix	
	Predicted class	
True class	VB	VG
VB	44.8%	4.2%
VG	1.4%	49.7%

Number of test data = 143

Table 4.28: Test results for all feature sets on White van – Vertical orientation data in Very Good vs. Very Bad classification

28. In binary VG vs. VB, Maxwell Aladago's features produced the best results with accuracy and f1-score of 94.4%. Followed by Francis' features with the next best results of accuracy of and f1-score of 87.4%. Anthony's features gave the lowest performance with accuracy and f1-score of 70.6%.

White van – Vertical orientation		
Vorgbe		
	Binary - VGG vs. VBB	
Accuracy	83.3%	
F1-score	83.1%	
	Confusion matrix	
	Predicted class	
True class	VBB	VGG
VBB	35.4%	6.9%
VGG	9.8%	48.0%
Abeo		

	Binary - VGG vs. VBB	
Accuracy	70.3%	
F1-score	70.2%	
	Confusion matrix	
	Predicted class	
True class	VBB	VGG
VBB	32.1%	10.2%
VGG	19.5%	38.2%
Maxwell Aladago’s		
	Binary - VGG vs. VBB	
Accuracy	90.2%	
F1-score	89.9%	
	Confusion matrix	
	Predicted class	
True class	VBB	VGG
VBB	35.4%	6.9%
VGG	2.8%	54.9%

Number of test data = 246

Table 4.29: Test results for all feature sets on White van – Vertical orientation data in Very Good and Good vs. Very Bad and Bad classification

29. In binary VGG vs. VBB, Maxwell Aladago’s features produced the best results with accuracy of 90.2% and f1-score of 89.9%. Followed by Francis’ features with the next best results of accuracy of 83.3% and f1-score of 83.1%. Anthony’s features gave the lowest performance with accuracy of 70.3% and f1-score of 70.2%.

White van – Vertical orientation				
Vorgbe				
	Multiclass			
Accuracy	62.2%			
F1-score	57.8%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG

VB	24.8%	2.4%	1.6%	0.4%
B	4.9%	4.1%	2.4%	1.6%
G	4.5%	2.0%	15.9%	8.1%
VG	3.3%	0.4%	6.1%	17.5%
Abeo				
	Multiclass			
Accuracy	52.8%			
F1-score	48.0%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	22.4%	2.8%	0.8%	3.3%
B	4.9%	2.8%	2.0%	3.3%
G	7.7%	3.3%	13.0%	6.5%
VG	6.9%	2.0%	3.7%	14.6%
Maxwell Aladago's				
	Multiclass			
Accuracy	62.2%			
F1-score	58.4%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	21.5%	5.3%	0.4%	2.0%
B	4.5%	4.5%	3.3%	0.8%
G	1.2%	2.8%	16.3%	10.2%
VG	0.0%	0.8%	6.5%	19.9%

Number of test data = 246

Table 4.30: Test results for all feature sets on White van – Vertical orientation data in multiclass classification

30. In multiclass classification, Maxwell Aladago's features produced the best results with accuracy of 62.2% and f1-score of 58.4%. Followed by Maxwell Aladago's features with the next best results of accuracy of 62.2% and f1-score of 57.8%. Anthony's features gave the lowest performance with accuracy of 52.8% and f1-score of 48%.

White van – Horizontal orientation

White van – Horizontal orientation		
Vorgbe		
	Binary - VG vs. VB	
Accuracy	86.7%	
F1-score	86.3%	
	Confusion matrix	
	Predicted class	
True class	VB	VG
VB	52.4%	7.7%
VG	5.6%	34.3%
Abeo		
	Binary - VG vs. VB	
Accuracy	85.7%	
F1-score	85.2%	
	Confusion matrix	
	Predicted class	
True class	VB	VG
VB	52.1%	8.0%
VG	6.3%	33.6%
Maxwell Aladago's		
	Binary - VG vs. VB	
Accuracy	92.7%	
F1-score	92.4%	
	Confusion matrix	
	Predicted class	
True class	VB	VG
VB	54.9%	5.2%
VG	2.1%	37.8%

Number of test data = 286

Table 4.31: Test results for all feature sets on White van – Horizontal orientation data in Very Good vs. Very Bad classification

31. In binary VG vs. VB, Maxwell Aladago's features produced the best results with accuracy of 92.7% and f1-score of 92.4%. Followed by Francis' features with the next best results of accuracy of 86.7% and f1-score of 86.3%. Anthony's features gave the lowest performance with accuracy of 85.7% and f1-score of 85.2%.

White van – Horizontal orientation		
Vorgbe		
	Binary - VGG vs. VBB	
Accuracy	80.3%	
F1-score	80.3%	
	Confusion matrix	
	Predicted class	
True class	VBB	VGG
VBB	39.4%	7.8%
VGG	11.9%	40.9%
Abeo		
	Binary - VGG vs. VBB	
Accuracy	78.2%	
F1-score	78.2%	
	Confusion matrix	
	Predicted class	
True class	VBB	VGG
VBB	40.7%	6.5%
VGG	15.3%	37.5%
Maxwell Aladago's		
	Binary - VGG vs. VBB	
Accuracy	90.1%	
F1-score	90.1%	
	Confusion matrix	
	Predicted class	
True class	VBB	VGG
VBB	42.8%	4.4%
VGG	5.5%	47.4%

Number of test data = 477

Table 4.32: Test results for all feature sets on White van – Horizontal orientation data in Very Good and Good vs. Very Bad and Bad classification

32. In binary VGG vs. VBB, Maxwell Aladago's features produced the best results with accuracy and f1-score of 90.1%. Followed by Francis' features with the next best results of accuracy and f1-score of 80.3% Anthony's features gave the lowest performance with accuracy and f1-score of 78.2%.

White van – Horizontal orientation				
Vorgbe				
	Multiclass			
Accuracy	53.2%			
F1-score	45.7%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	28.1%	3.6%	0.6%	2.9%
B	6.1%	1.9%	1.0%	2.9%
G	4.4%	3.8%	10.9%	9.4%
VG	2.3%	2.9%	6.7%	12.4%
Abeo				
	Multiclass			
Accuracy	51.2%			
F1-score	43.7%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	28.9%	4.0%	1.3%	1.0%
B	6.9%	2.1%	1.5%	1.5%
G	4.4%	5.9%	9.6%	8.6%
VG	2.9%	3.8%	7.1%	10.5%
Maxwell Aladago's				
	Multiclass			
Accuracy	65.8%			
F1-score	57.7%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	31.2%	1.9%	0.6%	1.5%
B	6.7%	2.7%	1.9%	0.6%
G	1.3%	5.0%	12.8%	9.4%
VG	0.6%	1.0%	3.6%	19.1%

Number of test data = 477

Table 4.33: Test results for all feature sets on White van – Horizontal orientation data in

multiclass classification

33. In multiclass classification, Maxwell Aladago's features produced the best results with accuracy of 65.8% and f1-score of 57.7%. Followed by Francis' features with the next best results of accuracy of 53.2% and f1-score of 45.7%. Anthony's features gave the lowest performance with accuracy of 51.2% and f1-score of 43.7%.

White van – Slanting orientation

White van – Slanting orientation		
Vorgbe		
	Binary - VG vs. VB	
Accuracy	93.0%	
F1-score	93.0%	
	Confusion matrix	
	Predicted class	
True class	VB	VG
VB	45.1%	3.5%
VG	3.5%	47.9%
Abeo		
	Binary - VG vs. VB	
Accuracy	90.8%	
F1-score	90.8%	
	Confusion matrix	
	Predicted class	
True class	VB	VG
VB	43.7%	4.9%
VG	4.2%	47.2%
Maxwell Aladago's		
	Binary - VG vs. VB	
Accuracy	94.4%	
F1-score	94.3%	
	Confusion matrix	
	Predicted class	
True class	VB	VG
VB	43.7%	4.9%
VG	0.7%	50.7%

Number of test data = 142

Table 4.34: Test results for all feature sets on White van – Slanting orientation data in Very Good vs. Very Bad classification

34. In binary VG vs. VB, Maxwell Aladago's features produced the best results with accuracy of 94.4% and f1-score of 94.3%. Followed by Francis' features with the next best results of accuracy and f1-score of 93%. Anthony's features gave the lowest performance with accuracy and f1-score of 90.8%.

White van – Slanting orientation		
Vorgbe		
	Binary - VGG vs. VBB	
Accuracy	82.0%	
F1-score	81.5%	
	Confusion matrix	
	Predicted class	
True class	VBB	VGG
VBB	32.7%	6.9%
VGG	11.0%	49.4%
Abeo		
	Binary - VGG vs. VBB	
Accuracy	80.8%	
F1-score	80.3%	
	Confusion matrix	
	Predicted class	
True class	VBB	VGG
VBB	32.7%	6.9%
VGG	12.2%	48.2%
Maxwell Aladago's		
	Binary - VGG vs. VBB	
Accuracy	91.0%	
F1-score	90.5%	
	Confusion matrix	
	Predicted class	
True class	VBB	VGG
VBB	33.9%	5.7%

VGG	3.3%	57.1%

Number of test data = 245

Table 4.35: Test results for all feature sets on White van – Slanting orientation data in Very Good and Good vs. Very Bad and Bad classification

35. In binary VGG vs. VBB, Maxwell Aladago's features produced the best results with accuracy of 91% and f1-score of 90.5%. Followed by Francis' features with the next best results of accuracy of 82% and f1-score of 81.5%. Anthony's features gave the lowest performance with accuracy of 80.8% and f1-score of 80.3%.

White van – Slanting orientation				
Vorgbe				
	Multiclass			
Accuracy	60.4%			
F1-score	55.9%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	22.9%	2.9%	0.8%	2.4%
B	3.7%	3.7%	2.4%	0.8%
G	4.1%	5.7%	15.1%	8.6%
VG	0.4%	1.6%	6.1%	18.8%
Abeo				
	Multiclass			
Accuracy	55.5%			
F1-score	49.6%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	25.3%	2.0%	1.2%	0.4%
B	4.5%	2.4%	2.0%	1.6%
G	4.1%	5.7%	16.3%	7.3%
VG	1.2%	5.3%	9.0%	11.4%
Maxwell Aladago's				
	Multiclass			
Accuracy	61.6%			
F1-score	55.5%			

	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	24.1%	2.9%	0.8%	1.2%
B	2.9%	2.4%	3.3%	2.0%
G	1.2%	4.1%	13.9%	14.3%
VG	0.0%	0.4%	5.3%	21.2%

Number of test data = 245

Table 4.36: Test results for all feature sets on White van – Slanting orientation data in multiclass classification

36. In multiclass classification, Maxwell Aladago's features produced the best results with accuracy of 61.6% and f1-score of 55.5%. Followed by Francis' features with the next best results of accuracy of 60.4% and f1-score of 55.9%. Anthony's features gave the lowest performance with accuracy of 55.5% and f1-score of 49.6%.

Findings

1. Maxwell Aladago's features maintain the highest level of performance among most of the orientations across the different vehicles in the exception of the free roam orientation where Francis' features perform better. However, the better performance of Francis' features is by a small margin of no more than 2%
2. The free roam orientation is the closest to reality orientation a user will have their device in for collecting data. This raises the question on whether a prescribed orientation must be given for the collection of data.
3. From the results observed, given a relative fixed point, Maxwell Aladago's features outperform the other set of features.

4. From the results obtained along primary oriented data, Maxwell Aladago's features transfer best across all the orientations along all three vehicles. Making Maxwell Aladago's features the best set of features to use.
5. This enables the conclusion of Maxwell Aladago's features as the best set of features to use.

4.2 Group II Experiment Results

In experiments for this group, for the green car, for each of the four orientations, data from each each of the orientations was trained and tested on a hold out dataset from the same orientation and dataset of everyother orientation. the binary classification scenario of Very Good (VG) vs. Very Bad (VB) and Very Good & Good (VGG) vs. Very Bad & Bad (VBB) and multiclass classification of Very Bad (VG) vs. Bad (B) vs. Good (G) vs. Very Good (VG). Also, the dataset of all primary orientations combined and the dataset of reoriented data points all from the green car were each trained and tested on a hold-out dataset. The changes in the performance of the algorithm was assessed to understand the effect of orientation on the algorithm's performance.

Free roam

Free roam vs. Free roam				
Binary classification - VG vs. VB				
Accuracy	94%			
F1-score	94%			
	Confusion matrix			
	Predicted class			
True class	VB	VG		
VB	50.0%	3.7%		
VG	2.2%	44.0%		
Binary classification - VGG vs. VBB				
Accuracy	85%			
F1-score	84%			
	Confusion matrix			
	Predicted class			

True class	VBB	VGG		
VBB	37.9%	7.5%		
VGG	7.9%	46.7%		
Multiclass classification				
Accuracy	55%			
F1-score	50%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	23.3%	5.0%	1.3%	0.4%
B	7.1%	2.9%	4.6%	0.8%
G	3.8%	3.8%	11.3%	12.1%
VG	0.0%	3.3%	2.9%	17.5%

Number of test data: VG vs. VB = 134; VBB vs. VGG and Multiclass = 240

Table 3.37: Test results for training on Free roam orientation data and testing on Free roam orientation data

1. In the Free roam vs. Free roam for the LR algorithm, binary VG vs. VB achieved an accuracy and f1-score of 94%. The binary VGG vs. VBB achieved accuracy of 85% and f1-score of 84% and the multiclass classification achieved accuracy of 55% and f1-score of 50%.

Free roam vs. Vertical				
Binary classification - VG vs. VB				
Accuracy	88%			
F1-score	87%			
	Confusion matrix			
	Predicted class			
True class	VB	VG		
VB	58.2%	0.7%		
VG	11.1%	29.9%		
Binary classification - VGG vs. VBB				
Accuracy	73%			
F1-score	72%			
	Confusion			

	matrix			
	Predicted class			
True class	VBB	VGG		
VBB	46.2%	1.2%		
VGG	26.1%	26.5%		
Multiclass classification				
Accuracy	45%			
F1-score	37%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	30.8%	1.4%	0.0%	0.2%
B	12.5%	2.0%	0.2%	0.3%
G	8.2%	14.9%	2.1%	4.8%
VG	1.3%	9.0%	1.9%	10.3%

Number of test data: VG vs. VB = 541; VBB vs. VGG and Multiclass = 985

Table 4.38: Test results for training on all Free roam orientation data and testing on Vertical orientation data

2. In the Free roam vs. Vertical for the LR algorithm, the binary VG vs. VB achieved an accuracy of 88% and f1-score of 87%. The binary VGG vs. VBB achieved accuracy of 73% and f1-score of 72% and the multiclass classification achieved accuracy of 45% and f1-score of 37%.

Free roam vs. Horizontal				
Binary classification - VG vs. VB				
Accuracy	94%			
F1-score	94%			
	Confusion matrix			
	Predicted class			
True class	VB	VG		
VB	55.6%	2.2%		
VG	3.6%	38.6%		
Binary classification - VGG vs. VBB				
Accuracy	86%			

F1-score	86%			
	Confusion matrix			
	Predicted class			
True class	VBB	VGG		
VBB	42.0%	4.6%		
VGG	9.8%	43.7%		
Multiclass classification				
Accuracy	59%			
F1-score	54%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	29.0%	1.7%	1.1%	0.4%
B	7.2%	4.0%	2.5%	0.5%
G	2.0%	8.7%	11.1%	8.1%
VG	0.5%	3.0%	5.3%	14.8%

Number of test data: VG vs. VB = 552; VBB vs. VGG and Multiclass = 985

Table 4.39: Test results for training on all Free roam orientation data and testing on Horizontal orientation data

3. In the Free roam vs. Horizontal for the LR algorithm, the binary VG vs. VB achieved an accuracy of 94% and f1-score of 94%. The binary VGG vs. VBB achieved accuracy of 86% and f1-score of 86% and the multiclass classification achieved accuracy of 59% and f1-score of 54%.

Free roam vs. Slanting				
Binary classification - VG vs. VB				
Accuracy	93%			
F1-score	93%			
	Confusion matrix			
	Predicted class			
True class	VB	VG		
VB	57.5%	0.7%		
VG	6.4%	35.5%		
Binary classification - VGG vs. VBB				

Accuracy	81%			
F1-score	81%			
	Confusion matrix			
	Predicted class			
True class	VBB	VGG		
VBB	44.9%	1.7%		
VGG	16.9%	36.4%		
Multiclass classification				
Accuracy	51%			
F1-score	44%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	30.3%	1.5%	0.0%	0.6%
B	10.8%	2.5%	0.6%	0.3%
G	5.4%	11.8%	6.2%	6.7%
VG	1.1%	5.8%	4.1%	12.2%

Number of test data: VG vs. VB = 550; VBB vs. VGG and Multiclass = 988

Table 4.40: Test results for training on all Free roam orientation data and testing on Slanting orientation data

4. In the Free roam vs. Slanting for the LR algorithm, the binary VG vs. VB achieved an accuracy of 93% and f1-score of 93%. The binary VGG vs. VBB achieved accuracy of 81% and f1-score of 81% and the multiclass classification achieved accuracy of 51% and f1-score of 44%.

Vertical

Vertical vs. Vertical				
Binary classification - VG vs. VB				
Accuracy	98%			
F1-score	98%			
	Confusion matrix			
	Predicted class			
True class	VB	VG		
VB	55.9%	1.5%		

VG	0.7%	41.9%		
Binary classification - VGG vs. VBB				
Accuracy	92%			
F1-score	92%			
	Confusion matrix			
	Predicted class			
True class	VBB	VGG		
VBB	41.3%	4.5%		
VGG	3.6%	50.6%		
Multiclass classification				
Accuracy	62%			
F1-score	56%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	29.1%	2.4%	0.8%	0.4%
B	5.7%	3.2%	4.0%	0.0%
G	0.0%	5.3%	12.6%	12.6%
VG	0.0%	0.4%	6.1%	17.4%

Number of test data: VG vs. VB = 136; VBB vs. VGG and Multiclass = 247

Table 4.41: Test results for training on Vertical orientation data and testing on Vertical orientation data

5. In the Vertical vs. Vertical for the LR algorithm, binary VG vs. VB achieved an accuracy and f1-score of 98%. The binary VGG vs. VBB achieved accuracy of 92% and f1-score of 92% and the multiclass classification achieved accuracy of 63% and f1-score of 56%.

Vertical vs. Free roam				
Binary classification - VG vs. VB				
Accuracy	84%			
F1-score	84%			
	Confusion matrix			
	Predicted class			
True class	VB	VG		
VB	42.6%	16.1%		
VG	0.2%	41.1%		
Binary classification - VGG vs. VBB				

Accuracy	80%			
F1-score	79%			
	Confusion matrix			
	Predicted class			
True class	VBB	VGG		
VBB	27.9%	19.0%		
VGG	1.0%	52.0%		
Multiclass classification				
Accuracy	48%			
F1-score	41%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	19.2%	4.0%	5.5%	4.0%
B	2.6%	1.7%	5.9%	4.1%
G	0.4%	0.6%	4.8%	24.2%
VG	0.1%	0.0%	0.7%	22.2%

Number of test data: VG vs. VB = 533; VBB vs. VGG and Multiclass = 957

Table 4.42: Test results for training on all Vertical orientation data and testing on Free roam orientation data

6. In the Vertical vs. Free roam for the LR algorithm, the binary VG vs. VB achieved an accuracy of 84% and f1-score of 84%. The binary VGG vs. VBB achieved accuracy of 80% and f1-score of 79% and the multiclass classification achieved accuracy of 48% and f1-score of 41%.

Vertical vs. Horizontal				
Binary classification - VG vs. VB				
Accuracy	93%			
F1-score	93%			
	Confusion matrix			
	Predicted class			
True class	VB	VG		
VB	82.6%	9.9%		
VG	0.9%	6.7%		

Binary classification - VGG vs. VBB				
Accuracy	86%			
F1-score	85%			
	Confusion matrix			
	Predicted class			
True class	VBB	VGG		
VBB	33.3%	13.2%		
VGG	0.9%	52.6%		
Multiclass classification				
Accuracy	53%			
F1-score	46%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	23.1%	4.1%	3.7%	1.3%
B	2.8%	2.0%	6.8%	2.6%
G	0.0%	0.4%	6.5%	23.1%
VG	0.2%	0.3%	1.3%	21.7%

Number of test data: VG vs. VB = 345; VBB vs. VGG and Multiclass = 989

Table 4.43: Test results for training on all Vertical orientation data and testing on Horizontal orientation data

7. In the Vertical vs. Horizontal for the LR algorithm, the binary VG vs. VB achieved an accuracy of 93% and f1-score of 93%. The binary VGG vs. VBB achieved accuracy of 86% and f1-score of 85% and the multiclass classification achieved accuracy of 53% and f1-score of 46%.

Vertical vs. Slanting				
Binary classification - VG vs. VB				
Accuracy	95%			
F1-score	95%			
	Confusion matrix			
	Predicted class			
True class	VB	VG		

VB	53.5%	4.7%		
VG	0.2%	41.6%		
Binary classification - VGG vs. VBB				
Accuracy	89%			
F1-score	89%			
	Confusion matrix			
	Predicted class			
True class	VBB	VGG		
VBB	37.8%	8.9%		
VGG	2.3%	51.0%		
Multiclass classification				
Accuracy	59%			
F1-score	53%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	25.4%	3.8%	1.4%	1.7%
B	4.9%	3.4%	4.5%	1.5%
G	0.6%	2.1%	10.6%	16.7%
VG	0.1%	0.1%	3.7%	19.3%

Number of test data: VG vs. VB = 550; VBB vs. VGG and Multiclass = 988

Table 4.44: Test results for training on all Vertical orientation data and testing on Slanting orientation data

8. In the Vertical vs. Slanting for the LR algorithm, the binary VG vs. VB achieved an accuracy of 95% and f1-score of 95%. The binary VGG vs. VBB achieved accuracy of 89% and f1-score of 89% and the multiclass classification achieved accuracy of 59% and f1-score of 53%.

Horizontal

Horizontal vs. Horizontal				
Binary classification - VG vs. VB				
Accuracy	95%			
F1-score	95%			
	Confusion matrix			

	Predicted class			
True class	VB	VG		
VB	48.6%	4.3%		
VG	0.7%	46.4%		
Binary classification - VGG vs. VBB				
Accuracy	89%			
F1-score	88%			
	Confusion matrix			
	Predicted class			
True class	VBB	VGG		
VBB	35.5%	6.5%		
VGG	4.8%	53.2%		
Multiclass classification				
Accuracy	62%			
F1-score	56%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	25.4%	1.6%	0.4%	2.4%
B	6.0%	2.8%	2.0%	1.2%
G	2.8%	3.2%	14.5%	12.5%
VG	0.4%	0.4%	4.8%	19.4%

Number of test data: VG vs. VB = 138; VBB vs. VGG and Multiclass = 248

Table 4.45: Test results for training on Horizontal orientation data and testing on Horizontal orientation data

9. In the Horizontal vs. Horizontal for the LR algorithm, binary VG vs. VB achieved an accuracy and f1-score of 95%. The binary VGG vs. VBB achieved accuracy of 89% and f1-score of 88% and the multiclass classification achieved accuracy of 62% and f1-score of 56%.

Horizontal vs. Vertical				
Binary classification - VG vs. VB				
Accuracy	92%			
F1-score	92%			
	Confusion matrix			
	Predicted class			

True class	VB	VG		
VB	56.0%	3.0%		
VG	5.2%	35.9%		
Binary classification - VGG vs. VBB				
Accuracy	79%			
F1-score	79%			
	Confusion matrix			
	Predicted class			
True class	VBB	VGG		
VBB	44.8%	2.6%		
VGG	18.5%	34.1%		
Multiclass classification				
Accuracy	51%			
F1-score	42%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	30.1%	0.4%	0.1%	1.8%
B	11.6%	2.6%	0.2%	0.6%
G	6.4%	11.0%	2.7%	9.9%
VG	0.9%	3.5%	2.2%	15.9%

Number of test data: VG vs. VB = 541; VBB vs. VGG and Multiclass = 985

Table 4.46: Test results for training on all Horizontal orientation data and testing on Vertical orientation data

10. In the Horizontal vs. Vertical for the LR algorithm, the binary VG vs. VB achieved an accuracy of 92% and f1-score of 92%. The binary VGG vs. VBB achieved accuracy of 79% and f1-score of 79% and the multiclass classification achieved accuracy of 51% and f1-score of 42%.

Horizontal vs. Free roam				
Binary classification - VG vs. VB				
Accuracy	84%			
F1-score	84%			
	Confusion			

	matrix			
	Predicted class			
True class	VB	VG		
VB	43.2%	15.6%		
VG	0.8%	40.5%		
Binary classification - VGG vs. VBB				
Accuracy	82%			
F1-score	81%			
	Confusion matrix			
	Predicted class			
True class	VBB	VGG		
VBB	32.1%	14.8%		
VGG	3.4%	49.6%		
Multiclass classification				
Accuracy	57%			
F1-score	53%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	23.3%	0.7%	6.2%	2.5%
B	5.1%	3.1%	5.3%	0.6%
G	1.5%	2.3%	14.7%	11.6%
VG	0.2%	0.5%	6.4%	15.9%

Number of test data: VG vs. VB = 533; VBB vs. VGG and Multiclass = 957

Table 4.47: Test results for training on all Horizontal orientation data and testing on Free roam orientation data

11. In the Horizontal vs. Free roam for the LR algorithm, the binary VG vs. VB achieved an accuracy of 84% and f1-score of 84%. The binary VGG vs. VBB achieved accuracy of 82% and f1-score of 81% and the multiclass classification achieved accuracy of 57% and f1-score of 53%.

Horizontal vs. Slanting				
Binary classification - VG vs. VB				
Accuracy	93%			

F1-score	93%			
	Confusion matrix			
	Predicted class			
True class	VB	VG		
VB	53.1%	5.1%		
VG	1.5%	40.4%		
Binary classification - VGG vs. VBB				
Accuracy	86%			
F1-score	86%			
	Confusion matrix			
	Predicted class			
True class	VBB	VGG		
VBB	40.7%	6.0%		
VGG	8.5%	44.8%		
Multiclass classification				
Accuracy	57%			
F1-score	50%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	28.4%	0.6%	1.4%	1.9%
B	9.1%	2.4%	2.1%	0.6%
G	3.1%	6.9%	9.6%	10.4%
VG	0.2%	1.9%	4.6%	16.6%

Number of test data: VG vs. VB = 550; VBB vs. VGG and Multiclass = 988

Table 4.48: Test results for training on all Horizontal orientation data and testing on Slanting orientation data

12. In the Horizontal vs. Slanting for the LR algorithm, the binary VG vs. VB achieved an accuracy of 93% and f1-score of 93%. The binary VGG vs. VBB achieved accuracy of 86% and f1-score of 86% and the multiclass classification achieved accuracy of 57% and f1-score of 50%.

Slanting

Slanting vs. Slanting
Binary classification - VG vs. VB

Accuracy	96%			
F1-score	96%			
	Confusion matrix			
	Predicted class			
True class	VB	VG		
VB	51.4%	2.2%		
VG	1.4%	44.9%		
Binary classification - VGG vs. VBB				
Accuracy	89%			
F1-score	89%			
	Confusion matrix			
	Predicted class			
True class	VBB	VGG		
VBB	40.9%	5.7%		
VGG	5.7%	47.8%		
Multiclass classification				
Accuracy	60%			
F1-score	53%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	27.5%	4.0%	0.8%	1.6%
B	5.7%	3.2%	3.2%	0.4%
G	1.2%	3.6%	11.3%	13.0%
VG	0.0%	0.8%	6.1%	17.4%

Number of test data: VG vs. VB = 138; VBB vs. VGG and Multiclass = 247

Table 4.49: Test results for training on Slanting orientation data and testing on Slanting orientation data

13. In the Slanting vs. Slanting for the LR algorithm, binary VG vs. VB achieved an accuracy and f1-score of 96%. The binary VGG vs. VBB achieved accuracy of 89% and f1-score of 89% and the multiclass classification achieved accuracy of 60% and f1-score of 53%.

Slanting vs. Horizontal				
Binary classification - VG vs. VB				
Accuracy	93%			

F1-score	93%			
	Confusion matrix			
	Predicted class			
True class	VB	VG		
VB	52.4%	5.4%		
VG	1.1%	41.1%		
Binary classification - VGG vs. VBB				
Accuracy	87%			
F1-score	87%			
	Confusion matrix			
	Predicted class			
True class	VBB	VGG		
VBB	36.7%	9.8%		
VGG	3.1%	50.4%		
Multiclass classification				
Accuracy	63%			
F1-score	57%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	26.0%	2.0%	2.2%	2.0%
B	4.5%	3.2%	5.7%	0.9%
G	0.5%	1.8%	16.6%	11.0%
VG	0.1%	0.4%	5.6%	17.5%

Number of test data: VG vs. VB = 552; VBB vs. VGG and Multiclass = 989

Table 4.50: Test results for training on all Slanting orientation data and testing on Horizontal orientation data

14. In the Slanting vs. Horizontal for the LR algorithm, the binary VG vs. VB achieved an accuracy of 93% and f1-score of 93%. The binary VGG vs. VBB achieved accuracy of 87% and f1-score of 87% and the multiclass classification achieved accuracy of 63% and f1-score of 57%.

Slanting vs. Vertical	
Binary classification - VG vs. VB	

Accuracy	94%			
F1-score	94%			
	Confusion matrix			
	Predicted class			
True class	VB	VG		
VB	56.2%	2.8%		
VG	3.1%	37.9%		
Binary classification - VGG vs. VBB				
Accuracy	85%			
F1-score	85%			
	Confusion matrix			
	Predicted class			
True class	VBB	VGG		
VBB	43.2%	4.2%		
VGG	10.6%	42.0%		
Multiclass classification				
Accuracy	56%			
F1-score	49%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	28.2%	1.8%	0.5%	1.8%
B	9.8%	2.7%	1.8%	0.6%
G	3.6%	6.6%	10.3%	9.6%
VG	0.3%	1.9%	5.7%	14.6%

Number of test data: VG vs. VB = 541; VBB vs. VGG and Multiclass = 985

Table 4.51: Test results for training on all Slanting orientation data and testing on Vertical orientation data

15. In the Slanting vs. Vertical for the LR algorithm, the binary VG vs. VB achieved an accuracy of 94% and f1-score of 94%. The binary VGG vs. VBB achieved accuracy of 85% and f1-score of 85% and the multiclass classification achieved accuracy of 56% and f1-score of 49%.

Slanting vs. Free roam				
Binary classification - VG vs. VB				
Accuracy	87%			

F1-score	87%			
	Confusion matrix			
	Predicted class			
True class	VB	VG		
VB	73.6%	19.3%		
VG	0.9%	6.2%		
Binary classification - VGG vs. VBB				
Accuracy	85%			
F1-score	84%			
	Confusion matrix			
	Predicted class			
True class	VBB	VGG		
VBB	34.0%	13.0%		
VGG	2.5%	50.6%		
Multiclass classification				
Accuracy	58%			
F1-score	54%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	21.8%	5.3%	2.8%	2.7%
B	3.3%	3.9%	6.2%	0.8%
G	0.9%	1.8%	14.3%	13.1%
VG	0.1%	0.1%	4.4%	18.4%

Number of test data: VG vs. VB = 337; VBB vs. VGG and Multiclass = 957

Table 4.52: Test results for training on all Slanting orientation data and testing on Free roam orientation data

16. In the Slanting vs. Free roam for the LR algorithm, the binary VG vs. VB achieved an accuracy of 87% and f1-score of 87%. The binary VGG vs. VBB achieved accuracy of 85% and f1-score of 84% and the multiclass classification achieved accuracy of 58% and f1-score of 54%.

Reoriented and Primary oriented Green car data

Reoriented data – Green car vs. Green car	
Binary classification - VG vs. VB	

Accuracy	92%			
F1-score	92%			
	Confusion matrix			
	Predicted class			
True class	VB	VG		
VB	50.2%	5.9%		
VG	2.4%	41.5%		
Binary classification - VGG vs. VBB				
Accuracy	85%			
F1-score	85%			
	Confusion matrix			
	Predicted class			
True class	VBB	VGG		
VBB	39.2%	9.6%		
VGG	5.7%	45.5%		
Multiclass classification				
Accuracy	60%			
F1-score	53%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	27.7%	2.0%	1.3%	2.9%
B	6.7%	3.3%	3.4%	1.5%
G	1.8%	3.9%	10.4%	12.0%
VG	0.2%	1.1%	3.6%	18.2%

Number of test data: VG vs. VB = 544; VBB vs. VGG and Multiclass = 980

Table 4.53: Test results for training on reoriented data and testing on reoriented data from the Green car

17. Using reoriented data points, the logistic regression algorithm achieved a performance of 92% accuracy and 91% f-score for the VG vs. VB binary classification. The performance of the algorithm was 85% for accuracy and 85% for f1-score in the VGG vs. VBB binary classification. In the multiclass classification, the logistic regression algorithm achieved a performance of 59% accuracy and 53% f1-score.

Primary oriented data – Green car vs. Green car				
Binary classification - VG vs. VB				
Accuracy	96%			
F1-score	96%			
	Confusion matrix			
	Predicted class			
True class	VB	VG		
VB	53.7%	2.4%		
VG	1.1%	42.8%		
Binary classification - VGG vs. VBB				
Accuracy	89%			
F1-score	89%			
	Confusion matrix			
	Predicted class			
True class	VBB	VGG		
VBB	42.1%	6.6%		
VGG	4.8%	46.4%		
Multiclass classification				
Accuracy	65%			
F1-score	59%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	29.7%	2.2%	1.2%	0.7%
B	6.4%	4.2%	3.5%	0.8%
G	1.4%	2.9%	13.4%	10.5%
VG	0.4%	0.5%	4.6%	17.6%

Number of test data: VG vs. VB = 544; VBB vs. VGG and Multiclass = 980

Table 4.54: Test results for training on primary oriented data and testing on primary oriented data from the Green car

18. Using primary oriented data points, the logistic regression algorithm achieved a performance of 97% accuracy and 96% f1-score for the VG vs. VB binary classification. The performance of the algorithm was 86% for accuracy and 86% for f1-score in the VGG vs. VBB binary classification. In the multiclass classification, the logistic regression algorithm achieved a performance of 65% accuracy and 59% f1-score.

Findings

1. In the Free roam orientation, the LR algorithm performs best against the Free roam orientation itself and the Horizontal orientation. Performance reduces slightly against the slanting orientation and then reduces much further against the Vertical orientation. Holding the device in free roam orientation during data collection is a possible explanation to the better performance among all orientations except the Vertical orientation. Holding the mobile phone in the palm during data collection will have most likely been in a horizontal or slanted position at the time of annotation. This positioning in hand will have allowed for data points collected in the Free roam orientation to liken that of data points collected in the Free roam orientation.

2. In the Vertical orientation, the LR algorithm performs best against the Vertical orientation itself, the Horizontal orientation and the Slanting orientation. Performance however reduces against the Free roam orientation.

3. In the Horizontal orientation, the LR algorithm performs best against the Horizontal orientation itself, the Vertical orientation and the Slanting orientation. The performance of the algorithm however reduces against the Free roam orientation.

4. In the Slanting orientation, the LR algorithm performs best against the Slanting orientation itself, the Horizontal orientation and the Vertical orientation. Performance of the algorithm however reduces against the Free roam orientation.

5. In the Binary classifications (i.e. VG vs. VB and VGG vs. VBB) and Multiclass classifications, even though the performance of the algorithm reduces when a given

orientation is tested along some other orientation, the performance of the algorithm does not reduce beyond an appreciable amount on the level of performance. Each orientation able to generalize well along other orientations.

6. The confusion matrix of the Multiclass classifications shows continually a difficulty of in distinguishing between Very Bad and Bad labelled data as well as distinguishing between Very Good and Good labelled data.

7. Using the reoriented and primary oriented data points, the performance of the algorithm does not reduce beyond an appreciable amount when compared to the pairwise orientation test results. Given that the reorientation algorithm resolves the orientation of a data point to a fixed point, and the performance of the logistic regression algorithm does not reduce by a significant amount, the reorientation of data points is a good decision to be done on collected data.

8. This shows that the effect of orientation on the performance of the algorithm is not very significant.

4.3 Group III Experiment Results

Experiments in this group trained and tested for each of the three vehicles, a hold-out dataset from the vehicle itself and dataset from each other vehicle in the binary classification scenario of Very Good (VG) vs. Very Bad (VB) and Very Good & Good (VGG) vs. Very Bad & Bad (VBB) and multiclass classification of Very Bad (VG) vs. Bad (B) vs. Good (G) vs. Very Good (VG). The changes in the performance of the algorithm was assessed to understand the effect of vehicle type on the algorithm's performance.

Green car

Green car vs. Green car				
Binary classification - VG vs. VB				
Accuracy	95%			
F1-score	95%			

	Confusion matrix			
	Predicted class			
True class	VB	VG		
VB	48.6%	4.3%		
VG	0.7%	46.4%		
Binary classification - VGG vs. VBB				
Accuracy	89%			
F1-score	89%			
	Confusion matrix			
	Predicted class			
True class	VBB	VGG		
VBB	35.5%	6.5%		
VGG	4.8%	53.2%		
Multiclass classification				
Accuracy	62%			
F1-score	56%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	25.4%	1.6%	0.4%	2.4%
B	6.0%	2.8%	2.0%	1.2%
G	2.8%	3.2%	14.5%	12.5%
VG	0.4%	0.4%	4.8%	19.4%

Number of test data: VG vs. VB = 138; VBB vs. VGG and Multiclass = 248

Table 4.55: Test results for training on Green car data and testing on Green car data

1. In the Green car vs. Green car for the LR algorithm, binary VG vs. VB achieved an accuracy and f1-score of 95%. The binary VGG vs. VBB achieved accuracy of 89% and f1-score of 89% and the multiclass classification achieved accuracy of 62% and f1-score of 56%.

Green car vs. Pickup				
Binary classification - VG vs. VB				
Accuracy	90%			
F1-score	90%			

	Confusion matrix			
	Predicted class			
True class	VB	VG		
VB	42.8%	7.4%		
VG	2.6%	47.2%		
Binary classification - VGG vs. VBB				
Accuracy	86%			
F1-score	86%			
	Confusion matrix			
	Predicted class			
True class	VBB	VGG		
VBB	41.0%	8.7%		
VGG	4.9%	45.3%		
Multiclass classification				
Accuracy	61%			
F1-score	54%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	28.6%	2.4%	0.3%	5.8%
B	4.5%	4.7%	2.1%	1.4%
G	0.5%	2.7%	6.6%	3.6%
VG	0.3%	1.6%	14.0%	20.8%

Number of test data: VG vs. VB = 811; VBB vs. VGG and Multiclass = 1099

Table 4.56: Test results for training on Green car data and testing on Pick up data

2. In the Green car vs. Pickup for the LR algorithm, binary VG vs. VB achieved an accuracy and f1-score of 90%. The binary VGG vs. VBB achieved accuracy of 86% and f1-score of 86% and the multiclass classification achieved accuracy of 61% and f1-score of 54%.

Green car vs. White van				
Binary classification - VG vs. VB				
Accuracy	93%			
F1-score	93%			

	Confusion matrix			
	Predicted class			
True class	VB	VG		
VB	52.8%	3.2%		
VG	3.7%	40.3%		
Binary classification - VGG vs. VBB				
Accuracy	87%			
F1-score	87%			
	Confusion matrix			
	Predicted class			
True class	VBB	VGG		
VBB	39.6%	4.9%		
VGG	7.9%	47.7%		
Multiclass classification				
Accuracy	63%			
F1-score	56%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	29.0%	1.4%	0.4%	1.6%
B	6.6%	2.7%	1.4%	1.3%
G	1.8%	4.2%	15.8%	8.2%
VG	0.4%	1.8%	8.0%	15.3%

Number of test data: VG vs. VB = 568; VBB vs. VGG and Multiclass = 980

Table 4.57: Test results for training on Green car data and testing on White van data

3. In the Green car vs. White van for the LR algorithm, binary VG vs. VB achieved an accuracy and f1-score of 93%. The binary VGG vs. VBB achieved accuracy of 87% and f1-score of 87% and the multiclass classification achieved accuracy of 63% and f1-score of 56%.

Pickup

Pickup vs. Green car				
Binary classification - VG vs. VB				
Accuracy	95%			

F1-score	95%			
	Confusion matrix			
	Predicted class			
True class	VB	VG		
VB	55.6%	2.2%		
VG	2.7%	39.5%		
Binary classification - VGG vs. VBB				
Accuracy	88%			
F1-score	88%			
	Confusion matrix			
	Predicted class			
True class	VBB	VGG		
VBB	41.1%	5.5%		
VGG	6.1%	47.4%		
Multiclass classification				
Accuracy	55%			
F1-score	46%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	28.3%	2.7%	0.9%	0.3%
B	7.0%	3.0%	2.1%	2.1%
G	2.1%	4.7%	3.4%	19.7%
VG	1.1%	1.2%	0.7%	20.5%

Number of test data: VG vs. VB = 552; VBB vs. VGG and Multiclass = 989

Table 4.58: Test results for training on Pickup data and testing on Green car data

4. In the Pickup vs. Green car for the LR algorithm, binary VG vs. VB achieved an accuracy and f1-score of 95%. The binary VGG vs. VBB achieved accuracy of 88% and f1-score of 88% and the multiclass classification achieved accuracy of 55% and f1-score of 46%.

Pickup vs. Pickup				
Binary classification - VG vs. VB				
Accuracy	95%			

F1-score	95%			
	Confusion matrix			
	Predicted class			
True class	VB	VG		
VB	46.3%	1.0%		
VG	3.9%	48.8%		
Binary classification - VGG vs. VBB				
Accuracy	90%			
F1-score	90%			
	Confusion matrix			
	Predicted class			
True class	VBB	VGG		
VBB	44.7%	4.7%		
VGG	5.5%	45.1%		
Multiclass classification				
Accuracy	68%			
F1-score	56%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	29.8%	3.3%	0.7%	0.7%
B	7.3%	5.5%	1.5%	0.7%
G	0.4%	2.2%	3.3%	9.8%
VG	2.2%	2.2%	1.5%	29.1%

Number of test data: VG vs. VB = 203; VBB vs. VGG and Multiclass = 275

Table 4.59: Test results for training on Pickup data and testing on Pickup data

5. In the Pickup vs. Pickup for the LR algorithm, binary VG vs. VB achieved an accuracy and f1-score of 95%. The binary VGG vs. VBB achieved accuracy of 90% and f1-score of 90% and the multiclass classification achieved accuracy of 68% and f1-score of 56%.

Pickup vs. White van				
Binary classification - VG vs. VB				
Accuracy	94%			

F1-score	94%			
	Confusion matrix			
	Predicted class			
True class	VB	VG		
VB	53.9%	2.1%		
VG	4.2%	39.8%		
Binary classification - VGG vs. VBB				
Accuracy	88%			
F1-score	88%			
	Confusion matrix			
	Predicted class			
True class	VBB	VGG		
VBB	40.1%	3.9%		
VGG	8.5%	47.5%		
Multiclass classification				
Accuracy	58%			
F1-score	49%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	29.1%	2.2%	0.5%	0.6%
B	6.7%	3.1%	1.5%	0.7%
G	3.0%	6.3%	4.9%	15.8%
VG	0.7%	3.1%	1.0%	20.7%

Number of test data: VG vs. VB = 568; VBB vs. VGG and Multiclass = 971

Table 4.60: Test results for training on Pickup data and testing on White van data

6. In the Pickup vs. White van for the LR algorithm, binary VG vs. VB achieved an accuracy and f1-score of 94%. The binary VGG vs. VBB achieved accuracy of 88% and f1-score of 88% and the multiclass classification achieved accuracy of 58% and f1-score of 49%.

White van

White van vs. Green car	
Binary classification - VG vs. VB	

Accuracy	95%			
F1-score	95%			
	Confusion matrix			
	Predicted class			
True class	VB	VG		
VB	54.2%	3.6%		
VG	1.1%	41.1%		
Binary classification - VGG vs. VBB				
Accuracy	89%			
F1-score	89%			
	Confusion matrix			
	Predicted class			
True class	VBB	VGG		
VBB	39.1%	7.4%		
VGG	3.8%	49.6%		
Multiclass classification				
Accuracy	61%			
F1-score	58%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	25.1%	3.9%	1.2%	2.0%
B	4.6%	5.1%	3.4%	1.2%
G	0.7%	2.6%	12.8%	13.8%
VG	0.1%	0.8%	4.1%	18.5%

Number of test data: VG vs. VB = 552; VBB vs. VGG and Multiclass = 989

Table 4.61: Test results for training on White van data and testing on Green car data

7. In the White van vs. Green car for the LR algorithm, binary VG vs. VB achieved an accuracy and f1-score of 95%. The binary VGG vs. VBB achieved accuracy of 89% and f1-score of 88% and the multiclass classification achieved accuracy of 61% and f1-score of 58%.

White van vs. Pickup
Binary classification - VG vs. VB

Accuracy	93%			
F1-score	93%			
	Confusion matrix			
	Predicted class			
True class	VB	VG		
VB	43.9%	6.3%		
VG	0.9%	49.0%		
Binary classification - VGG vs. VBB				
Accuracy	88%			
F1-score	88%			
	Confusion matrix			
	Predicted class			
True class	VBB	VGG		
VBB	39.7%	10.1%		
VGG	1.7%	48.5%		
Multiclass classification				
Accuracy	57%			
F1-score	51%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	24.3%	5.6%	0.8%	6.4%
B	3.0%	4.0%	4.3%	1.5%
G	0.3%	1.5%	6.9%	4.7%
VG	0.1%	0.9%	14.5%	21.3%

Number of test data: VG vs. VB = 811; VBB vs. VGG and Multiclass = 1099

Table 4.62: Test results for training on White van data and testing on Pickup data

8. In the White van vs. Pickup for the LR algorithm, binary VG vs. VB achieved an accuracy and f1-score of 93%. The binary VGG vs. VBB achieved accuracy of 88% and f1-score of 88% and the multiclass classification achieved accuracy of 57% and f1-score of 51%.

White van vs. White van	
Binary classification - VG vs. VB	

Accuracy	94%			
F1-score	94%			
	Confusion matrix			
	Predicted class			
True class	VB	VG		
VB	43.7%	4.9%		
VG	0.7%	50.7%		
Binary classification - VGG vs. VBB				
Accuracy	91%			
F1-score	91%			
	Confusion matrix			
	Predicted class			
True class	VBB	VGG		
VBB	33.9%	5.7%		
VGG	3.3%	57.1%		
Multiclass classification				
Accuracy	62%			
F1-score	55%			
	Confusion matrix			
	Predicted class			
True class	VB	B	G	VG
VB	24.1%	2.9%	0.8%	1.2%
B	2.9%	2.4%	3.3%	2.0%
G	1.2%	4.1%	13.9%	14.3%
VG	0.0%	0.4%	5.3%	21.2%

Number of test data: VG vs. VB = 142; VBB vs. VGG and Multiclass = 245

Table 4.63: Test results for training on White van data and testing on White van data

9. In the White van vs. White van for the LR algorithm, binary VG vs. VB achieved an accuracy and f1-score of 94%. The binary VGG vs. VBB achieved accuracy of 91% and f1-score of 91% and the multiclass classification achieved accuracy of 62% and f1-score of 55%.

Findings

1. Training the LR algorithm on data from the Green car, the LR algorithm performs best against data from the Green car itself in the binary VG vs. VB classification. Performance of the algorithm reduced slightly against data from the White van and reduced further, testing against data from the Pickup.
2. In the binary VGG vs. VBB, the algorithm performs best against data from the Green car, reducing in performance slightly against data from the White van and further against the Pickup.
3. In the multiclass classification, the algorithm achieves the best performance testing against data from the White van, reducing in performance slightly against data from the Green car and the Pickup.
4. Training the LR algorithm on data from the Pickup, the LR algorithm performs best against data from the Green car and the Pickup itself in the binary VG vs. VB classification. Performance of the algorithm however reduced slightly against data from the White van.
5. In the binary VGG vs. VBB, the algorithm performs best against data from the Pickup, reducing in performance slightly to be same against data from the White van and further against the Green car.
6. In the multiclass classification, the algorithm achieves the best performance testing against data from the Pickup itself, reducing in performance greatly against data from the White van. Performance reduced further testing against data from the Green car.
7. Training the LR algorithm on data from the White van, the LR algorithm performs best against data from the Green car in the binary VG vs. VB classification.

Performance of the algorithm reduced slightly against data from the White van itself and reduced slightly, testing against data from the Pickup.

8. In the multiclass classification, the algorithm achieves the best performance testing against data from the Green car, reducing in performance slightly against data from the White van itself and further against data from the Pickup.

9. The findings above allow for the conclusion that the type of vehicle for collecting data has a minimal effect on the algorithm's performance.

4.4 Group IV Experiment Results

Experiments in this group trained and tested on in their primary and reoriented forms. A hold-out dataset from each of the primary and reoriented data was used to test the classification algorithm in the multiclass classification of Very Bad (VG) vs. Bad (B) vs. Good (G) vs. Very Good (VG). The changes in the performance of the algorithm was assessed to understand the behaviour of the classification algorithm when datapoints are of mixed orientations or a relative fixed point.

Primary oriented data

Primary oriented data				
Number of test data			1361	
Accuracy	65%			
f1-score	56%			
Confusion matrix				
	Predicted class			
True class	VB	B	G	VG
VB	29.2%	2.1%	0.6%	1.2%
B	5.4%	2.5%	2.4%	1.5%
G	2.2%	4.4%	9.4%	9.0%
VG	0.5%	2.4%	3.5%	23.7%

Table 4.64: Test results for training and testing on all primary oriented data

Using the primary oriented data for a multiclass classification, the algorithm achieved an accuracy of 65% and f1-score of 56% testing on 1361 road segments after being trained on 12,244 primary oriented data points.

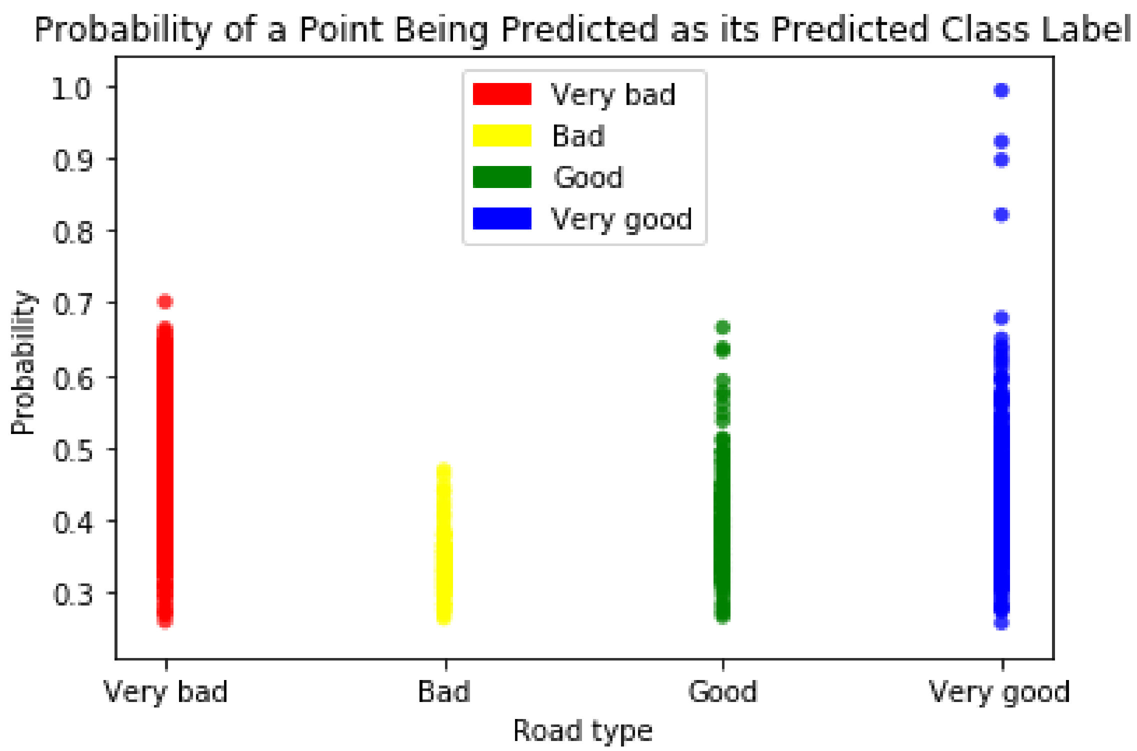

Figure 4.1: Distribution of primary oriented test data probability of being predicted as predicted class label

Visualizing the probability of the test data points being predicted as their predicted class label, all predictions had at least a probability of 0.25 chance of belonging to their predicted class. Majority of the predictions remain below a 0.5 probability of belonging to their predicted class. A very small proportion of test data had chances above 0.8 to belong to their predicted class.

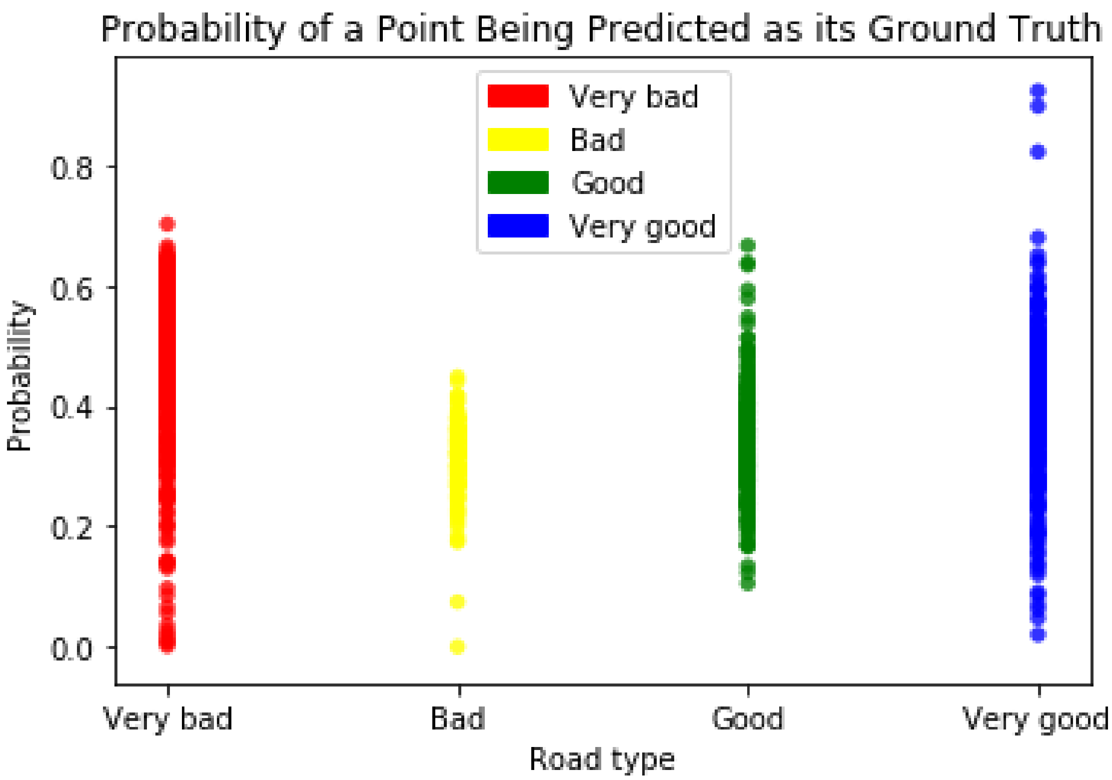

Figure 4.2: Distribution of primary oriented test data probability of being predicted as ground truth class label

Visualizing the probability of the test data being predicted as its ground truth, majority of test data had at least 0.25 chance at being predicted as their ground truth. Some data points however have had zero or near zero chance at being predicted as its ground truth. A very small proportion of test data had chances above 0.8 to belong to their ground truth class.

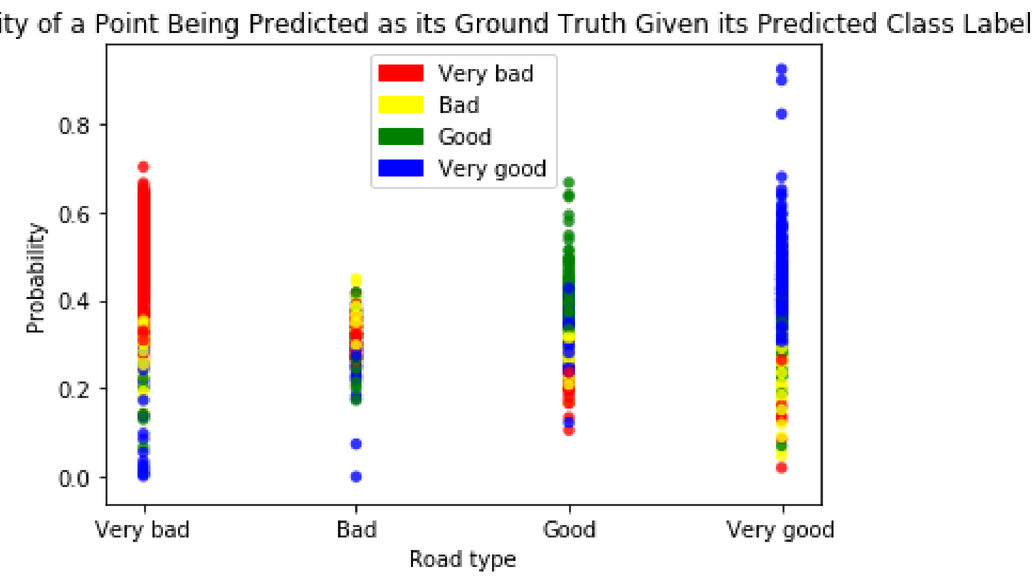

Figure 4.3: Distribution of primary oriented test data probability of being predicted as ground truth class label given predicted class label

Visualizing the probability of a point being predicted as its ground truth given its predicted class label, each data point is graphed at the probability of its ground truth colored with the label of its predicted class. Majority of data points with less than 0.25 chance of belonging to their ground truth labels of Very bad or Bad were predicted as belonging to the Very good class. Majority of data points with less than 0.25 belonging to their ground truth label of Good were predicted as belonging to the Very bad road class. Majority of data points with less than 0.25 chance of belonging their ground truth class of Very good road were classified as Bad roads. Lastly, when the confidence level is above 0.45, the classification algorithm correctly predicts the classification of a data point.

Reoriented data

Reoriented data				
Number of test data			1361	
Accuracy	60%			
f1-score	51%			

Confusion matrix				
	Predicted class			
True class	VB	B	G	VG
VB	26.3%	3.5%	0.7%	2.6%
B	5.6%	2.6%	1.8%	1.8%
G	3.1%	4.2%	6.9%	10.8%
VG	0.6%	1.8%	3.1%	24.6%

Table 4.65: Test results for training and testing on all reoriented data

Using the reoriented data for a multiclass classification, the algorithm achieved an accuracy of 60% and f1-score of 51% testing on 1361 road segments after being trained on 12,244 primary oriented data points.

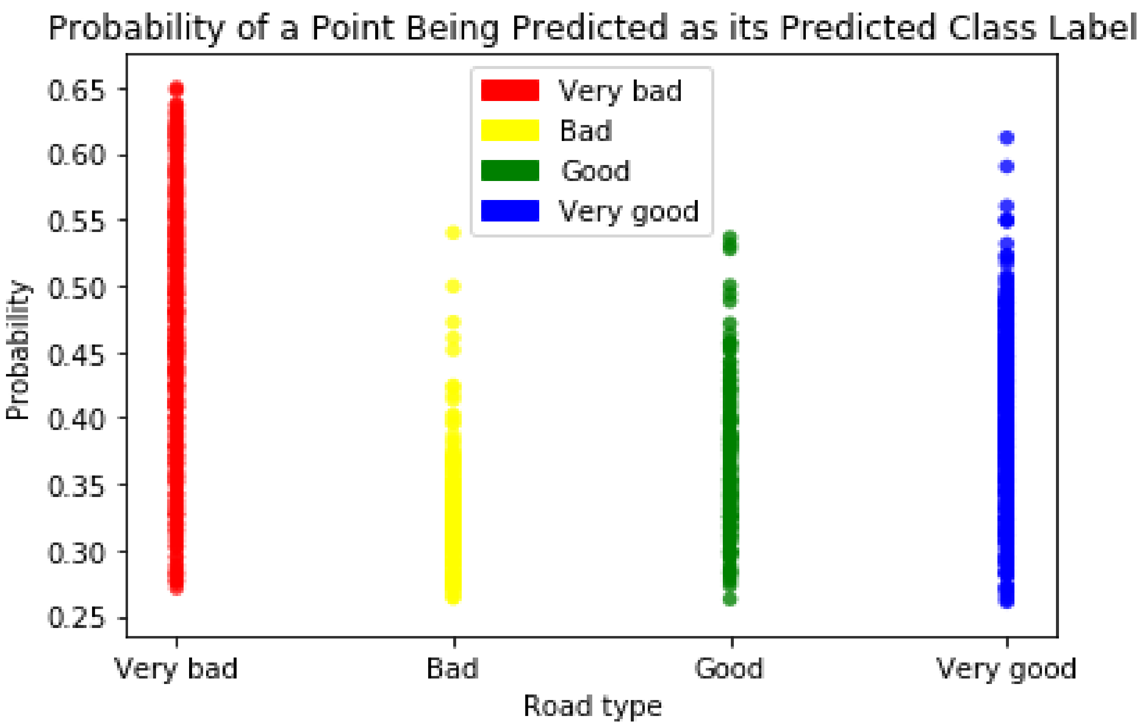

Figure 4.4: Distribution of reoriented test data probability of being predicted as predicted class label

Visualizing the probability of the test data points being predicted as their predicted class label, all predictions had at more than a probability of 0.25 chance of belonging to their predicted class. Majority of the predictions remain below a 0.5 probability of belonging to their predicted class.

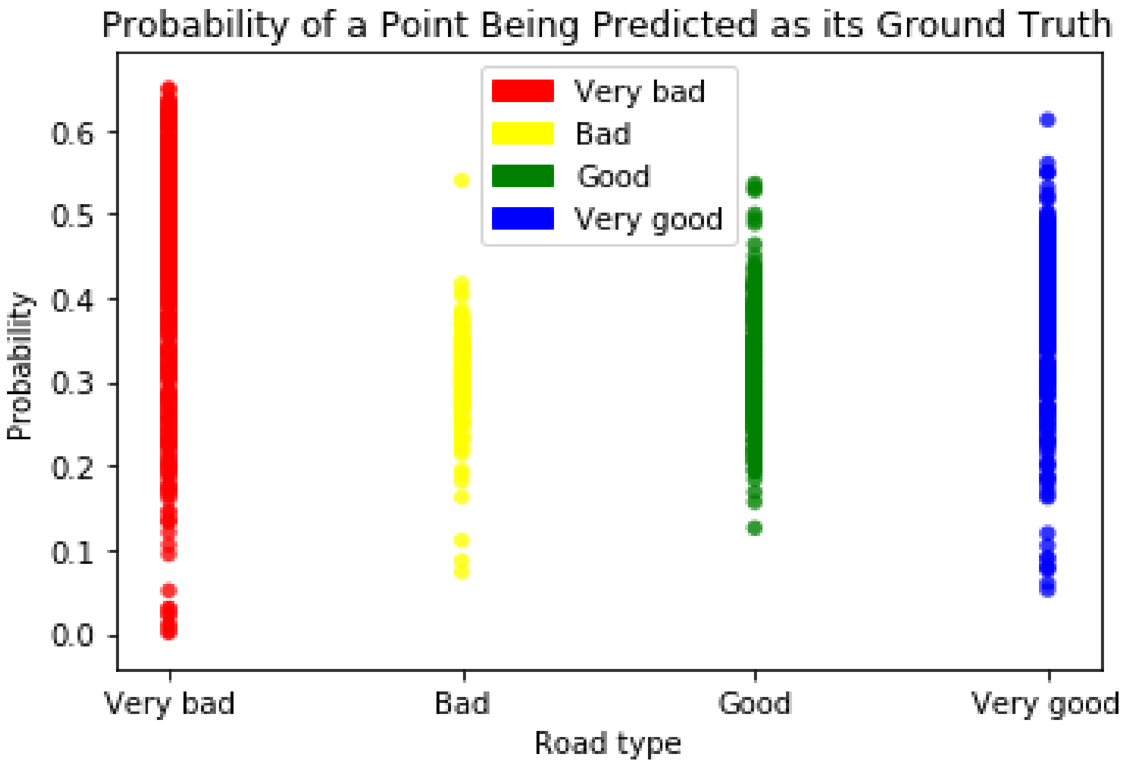

Figure 4.5: Distribution of reoriented test data probability of being predicted as ground truth class label

Visualizing the probability of the test data being predicted as its ground truth, majority of test data had at least 0.25 chance at being predicted as their ground truth. Some data points however have had zero or near zero chance at being predicted as its ground truth.

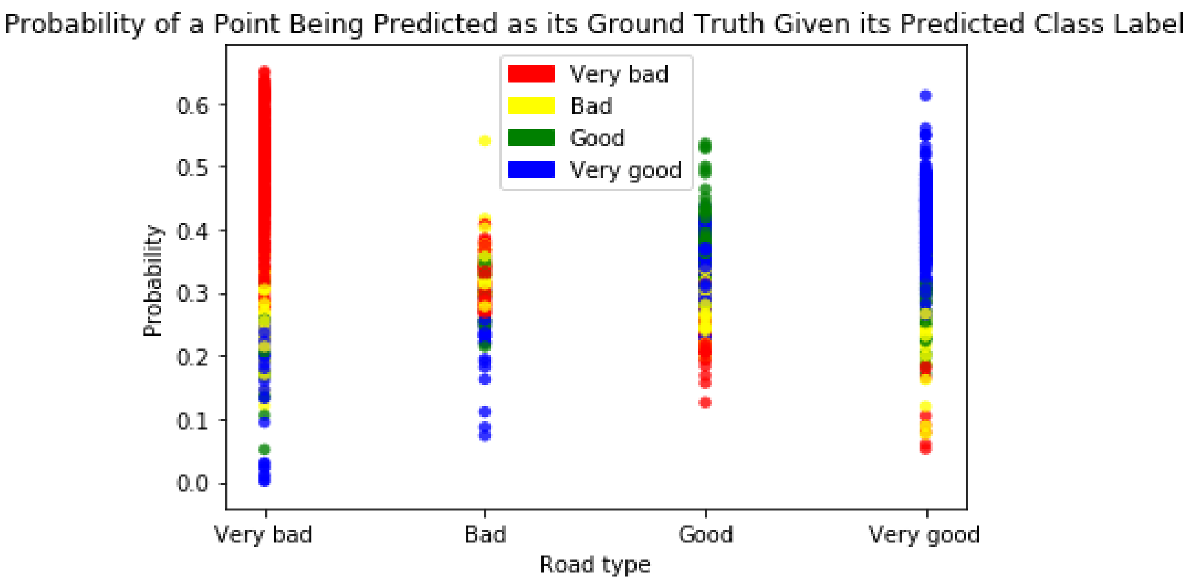

Figure 4.6: Distribution of reoriented test data probability of being predicted as ground truth class label given predicted class label

Visualizing the probability of a point being predicted as its ground truth given its predicted class label, each data point is graphed at the probability of its ground truth colored with the label of its predicted class. Majority of data points with less than 0.25 chance of belonging to their ground truth labels of Very bad or Bad were predicted as belonging to the Very good class. Majority of data points with less than 0.25 belonging to their ground truth label of Good or Very good were predicted as belonging to the Very bad road class. Majority of data points with chances above 0.25 belonging to their ground truth label of Bad were predicted as Very bad roads. When a confidence level is above 0.41, the classification algorithm correctly predicts the class of the data point.

Findings

1. Training and testing the algorithm with the primary oriented data compared to the reoriented data points reduced the performance of the classification algorithm.

2. From the visualization of the probability of the test data points being predicted as their predicted class label, the reoriented data points increased the probability of test data predicted as their predicted class to all be above a 0.25 chance.
3. From the visualization of the probability of the test data being predicted as its ground truth, the reoriented data points keeps the probability of test data to be predicted as its ground truth to not go above a 0.8 chance.

Chapter 5: Conclusion and Recommendations

5.1 Summary

This study aimed at developing a Logistic Regression algorithm to classify stretches of road travelled along in a moving vehicle into different qualities using mobile phone accelerometer sensors data. The stretches of roads were labelled by segments of stretches travelled in 10 second time frames, referred to as road segments/windows.

The study used Scikit-Learn Logistic Regression, three sets of features, and data collected along four different orientations from three different vehicles to attempt a four-class classification: very bad, bad, good and very good. These classifications represented different qualities of road surfaces travelled.

An evaluation for the best set of features was done, the third of which by Maxwell Aladago was learnt as the best set of features among the three considered. An examination into the effect of device orientation during data collection on the algorithm's performance showed an insignificant effect. Findings from experiments also show that the type of vehicle used in data collections has a minimal effect on the classification algorithm's performance.

The Logistic Regression algorithm trained and tested on the combined dataset of four orientations from each vehicle in a four-class multiclass classification achieved an accuracy of 65% and f1-score of 56% using the primary oriented data. Using Euler's Angles to reorient data points, training and testing the combined dataset on the Logistic Regression algorithm achieved an accuracy of 60% and f1-score of 51%.

The algorithm is unreliable at distinguishing between Bad and Good road segments. Also, the algorithm has a sometimes classifies very good roads as very bad roads and vice versa.

The findings of this study align with the findings from prior work that the multiclass classification problem is difficult. The data sourcing has significant amount or errors with mislabeled data and poorly distinguishable characteristics for bad from very bad roads and very good from good roads.

5.2 Limitations

This study was limited by the similarity of data points among very bad and bad as well as very good and good. This could have resulted from the methodology of data collection where the stretch of road was continuously changing. A way to eliminate this phenomenon will be to select long distances of road for each road surface quality and collect data solely along each stretch. This will give proper definitions to each class of road, improving the taxonomy of road quality for classification.

5.4 Further Work

Logistic Regression algorithm is domain biased to linearly separable data. The mislabeling of the dataset has reduced the ability to linearly separate the dataset into the four distinct classes, hence the classifier in this study not being very robust.

The robustness of the classifier can be improved by compiling a new dataset and transforming the axial orientation of the accelerometer data to match the vehicle.

A new dataset should be compiled by using long-distance stretches for each type of road surface quality. This will give each type of road surface quality linearly distinguishable characteristics from the others. With proper definitions to each road type, a better taxonomy for road quality will be available for the classifier.

Transforming the axial orientation of the accelerometer data using the formula mentioned in this study will make the data match the axial orientation of the vehicle. With corresponding orientations, feature sets which have greater focus on the z-axis such as those used by Vorgbe will improve the performance of the classification algorithm.

A new data set should be compiled by using long-distance stretches for each type of road surface quality. Each road surface quality will yield linearly distinguishable characteristics from the others with appropriate definitions to each road type; for a better taxonomy to a road quality will be available for the classifier.

Transforming the axial orientation of the accelerometer data (using the formula mentioned in this study) will be equivalent to the axial orientation of the vehicle. With corresponding orientations, feature sets with greater focus on the z-axis including those used by Vorgbe could improve the performance of the classification algorithm.

References

[1] Anthony Anabila Abeo. 2018. Evaluating and choosing a machine learning algorithm for classifying road surface quality data. Applied Project. Ashesi University, Accra, Ghana. Retrieved October 14, 2019 from http://hdl.handle.net/20.500.11988/400

[2] Felix Abramovich and Marianna Pensky. 2019. Classification with many classes: challenges and pluses. arXiv:1506.01567 [math, stat] (July 2019). Retrieved December 9, 2019 from http://arxiv.org/abs/1506.01567

[3] Adrian Albert, Jasleen Kaur, and Marta C. Gonzalez. 2017. Using Convolutional Networks and Satellite Imagery to Identify Patterns in Urban Environments at a Large Scale. In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '17, ACM Press, Halifax, NS, Canada, 1357–1366. DOI:https://doi.org/10.1145/3097983.3098070

[4] V. Astarita, M. V. Caruso, G. A. Danieli, D. C. Festa, V. P. Giofre, T. luele, and R. Vaiana. 2012. A mobile application for road surface quality control: UNIquALroad. Procedia - Social and Behavioral Sciences 54, (2012), 1135–1144. DOI:https://doi.org/10.1016/j.sbspro.2012.09.828

[5] Kwabena Gyang Boohene. 2017. Automated collection and visualization of road quality data to aid driver navigation. Applied Project. Ashesi University, Accra, Ghana. Retrieved October 14, 2019 from http://hdl.handle.net/20.500.11988/399

[6] Gabriel Cadamuro, Aggrey Muhebwa, and Jay Taneja. 2019. Street smarts: measuring intercity road quality using deep learning on satellite imagery. In

Proceedings of the Conference on Computing & Sustainable Societies - COMPASS 19, ACM Press, Accra, Ghana, 145–154. DOI:https://doi.org/10.1145/3314344.3332493

[7] Sayna Firoozi Yeganeh, Ahmadreza Mahmoudzadeh, MohammadAmin Azizpour, and Amir Golroo. 2017. Validation of Smartphone_Based Pavement Roughness Measures. AJCE 1, 2 (December 2017). DOI:https://doi.org/10.22060/ajce.2017.13105.5328

[8] Marius Hoffmann, Michael Mock, and Michael May. Road-quality classification and bump detection with bicycle-mounted smartphones. 5.

[9] David Hutchison, Takeo Kanade, Josef Kittler, Jon M. Kleinberg, Friedemann Mattern, John C. Mitchell, Moni Naor, Oscar Nierstrasz, C. Pandu Rangan, Bernhard Steffen, Madhu Sudan, Demetri Terzopoulos, Doug Tygar, Moshe Y. Vardi, Gerhard Weikum, Volodymyr Mnih, and Geoffrey E. Hinton. 2010. Learning to Detect Roads in High-Resolution Aerial Images. In Computer Vision – ECCV 2010, Kostas Daniilidis, Petros Maragos and Nikos Paragios (eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 210–223. DOI:https://doi.org/10.1007/978-3-642-15567-3_16

[10] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. 2017. ImageNet classification with deep convolutional neural networks. Commun. ACM 60, 6 (May 2017), 84–90. DOI:https://doi.org/10.1145/3065386

[11] Lars Forslöf and Hans Jones. 2015. Roadroid: Continuous Road Condition Monitoring with Smart Phones. JCEA 9, 4 (April 2015). DOI:https://doi.org/10.17265/1934-7359/2015.04.012

[12] Fabian Pedregosa, Gael Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, and

Vincent Dubourg. 2011. Scikit-learn: Machine learning in Python. Journal of machine learning research 12, Oct (2011), 2825–2830.

[13] Seyed Yousef Sadjadi. 2017. Investigating for Road Roughness using Smartphone Sensors. 11, (2017), 8.

[14] Mrcrmsr W Savnns. On the Calculation of International Roughness Index from Longitudinal Road Profile. 12.

[15] Eko Budi Setiawan and Hadi Nurdin. 2019. Road Quality Assessment Using International Roughness Index Method and Accelerometer on Android. LKJITI (August 2019), 62. DOI:https://doi.org/10.24843/LKJITI.2019.v10.i02.p01

[16] Karen Simonyan and Andrew Zisserman. 2014. Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv:1409.1556 [cs] (September 2014). Retrieved September 25, 2019 from http://arxiv.org/abs/1409.1556

[17] Sustainable Mobility for All. 2017. Global Mobility Report 2017. Sustainable Mobility for All, Washington DC. Retrieved from https://openknowledge.worldbank.org/bitstream/handle/10986/28542 License: CC BY 3.0 IGO

[18] Francis Delali Vorgbe. 2014. Classification of Road Surface Quality Using Android Smartphone Devices. Thesis. Ashesi University, Accra, Ghana. Retrieved October 14, 2019 from http://hdl.handle.net/20.500.11988/226