Identifying Lane Changes Automatically using the GPS Sensor of Portable Devices

By

Lokin Lakshmindra Bindu Prasad

In partial fulfilment of the requirements for the degree of

Master of Science
In Mechanical Engineering

At the Delft University of Technology
to be defended publicly on Tuesday November 30th, 2021, at 16:15 hrs.

Student number:	5092396
Supervisors:	Dr.ir. Joost De Winter, TU Delft Dr. Pavlo Bazilinskyy, TU Delft ir.Tom Driessen, TU Delft
External Committee:	Dr. Dimitra Dodou, TU Delft

Acknowledgement

This project was successful thanks to my supervisors, Professor Joost De Winter, Pavlo Bazilinskyy, and Tom Driessen. They provided guidance and suggested procedures to make it a successful project that stretched for many months. They were active and involved themselves in any question I had asked. I am happy and honoured to be part of such a team.

I want to thank Tom Driessen again as he drove on the Dutch highways, while I was recording the data. As I did not have a valid European driving license, I could not drive and conduct the experiment by myself. Tom was very quick in planning and helped greatly in completing my experiment.

I would also like to express my gratitude towards my family for supporting me at every step morally and financially.

My friends have been an essential part of my master's study journey, and it would have been a lot more hectic and challenging to complete the project without their constant support.

I want to mention my sister Sonika B Prasad and my best friend Suhas Donthi for encouraging me at every step, especially when I felt low and unproductive. It is not easy to express how much they have helped me, but the least I could do is thank them heartfully.

Abstract

Automatic lane change identification can be used in warning systems that alert the driver when drifting off-lane, or in lane-specific navigation systems. Furthermore, they may help road authorities plan better road network design and infrastructure. Though lane-departure warning systems are already commercially available, these systems rely on on-board vehicle sensors. The availability of such systems can be increased if it became available to owners of smartphones or navigation devices. In this thesis, I propose and test several methods that rely solely on GPS data originating from portable devices (two smartphones, a GPS-equipped GoPro Max, and a USB GPS receiver), recorded during test rides on a Dutch highway (42.8 km) with a total of 64 lane changes. The methods rely on observing changes in the lateral offset between the GPS trajectories and the road geometry. The resulting identification accuracy of the best performing algorithm was achieved using the GoPro Max with an overall F1 score (harmonic mean of precision and recall) of 0.9 on the validation data using the signal of filtered projected lateral distance. It is concluded that GPS-equipped portable devices could be a suitable choice for identifying lane changes, provided there is improvement in the quality of GPS receiver chips with higher data collection frequency such as the ones used in GoPro Max.

Introduction

Driver behaviour has been the subject of much research over the last decades. The advancement of data analysis algorithms has enabled the evolution of this field of study. The development of new analysis methods has boosted the quality of driver behavior analysis and opened new application areas [1]. Driver behaviour analysis is an evolving area of research and is an essential catalyst for improving safety, sustainability, and overall driving experience [2].

Driver behaviour can be assessed based on various variables collected while driving, such as vehicle speed and acceleration, driver steering and braking inputs, etc. Although there has been significant progress in algorithmic driver assessment, one area in which there appears to be a lack of research is automated lane change detection. Automated lane change detection may serve a vast number of applications. For example, route navigation devices may provide more appropriate navigation instructions if it can be ascertained on which lane the driver is driving [3] or whether the driver has already made a lane change (e.g., highway entry or exit). In the same vein, in-vehicle advisory systems may advise drivers to change lanes to enhance traffic flow [4]. Furthermore, studies investigate the possibilities of in-car advisory systems that recommend where and when lane changes should be made, based on traffic flow circumstances and data available through vehicle sensors (e.g., [5]). The information provided to vehicle insurance companies can help decide premiums where instead of a flat charge, drivers must pay a premium based on their driving behaviour and degree of exposure [6].

The commercially available lane departure warning systems currently rely on sensors based on road infrastructure and vehicle-based sensors such as cameras [7]. However, the use of devices with lane change identification systems will increase if they are available as portable devices such as smartphones or other nomadic devices. Smartphones are becoming more common and may be well suited for driver assessment as sensing capabilities are improving for such devices, leading to increased research in smartphone-based telematics systems. The fact that smartphones are relatively low cost, widely available, and show increasing market penetration rates have made smartphones good candidates as sensing devices [8].

However, data quality from current smartphones is not comparable to the quality of data collection by mapping-grade or survey-grade GPS receivers [9]. This research project aims to discover whether smartphones can still provide or act as a valid substitute compared to other portable devices for applications from which driver behaviour can be analysed.

Background

This section will explain several methods of determining lane changes from the literature.

Faizan et al. [10] described an algorithm for identifying lane changes. Position accuracy of GPS sensors is generally in the range of 3 to 5 meters as explained by the authors, i.e., equal to or greater than the lane's width. As such accuracy is not sufficient to calculate lane changes, the authors have relied on the concept of relative accuracy. In Figure 1 below, red dots show adjacent GPS coordinates of a car on a straight and curving road, recorded by a GPS receiver, whereas the green dots depict the exact positions of the vehicle. GPS error could be present anywhere in the large, dashed circle representing the aforementioned measurement inaccuracy, which may arise due to due to atmospheric conditions, amongst others. However, any GPS device in the vicinity will encounter similar atmospheric conditions, and hence the errors are subject to only device-related errors represented by the smaller circle. The differences in heading angles of the known road and recorded GPS points are calculated to find the lateral or projected offset of the vehicle at every recorded instance to check if it exceeds a threshold value for it to be classified as a lane change scenario. Note that Figure 1 serves to provide an explanation of the concept of relative accuracy and does not truly represent actual recordings, which might have higher or lower errors. Although this algorithm identified all lane changes correctly but with nearly 1/10th of the recordings as false positives on a two-lane road. It had problems identifying lane changes on sharp curves due to a combined effect of the road direction and the driver behaviour.

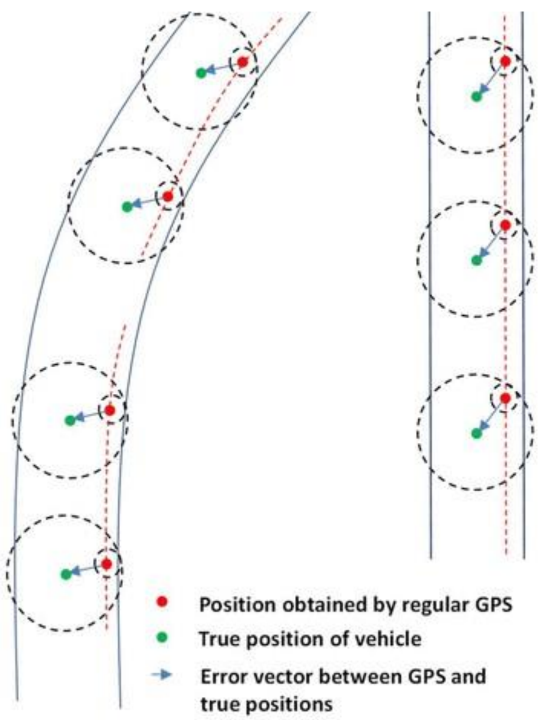

Figure 1: Explanation of Relative accuracy vs Absolute accuracy [10]

Yoshinde et al. [11] described a simple procedure of finding the distance from the centre to the GPS point recorded on the road. The absolute distance from the centre of the road varies as lane changes occur; changes in the absolute distance for a group of points would indicate a lane change.

Ramah et al. [12] described how smartphone sensors such as accelerometers and gyroscopes can be used to study the different types of manoeuvres during driving. They described the performance of deep learning classifiers, each achieving to recognize different sets of driving events such as turning, obstacle avoidances, lane change manoeuvres, and overtaking. Their research made clear that it is difficult to identify lane changes when the maneuver is very gentle.

Aim of the study

The above literature is instrumental in determining how GPS data can be used to identify lane changes on a highway. However, the above research was conducted using high-end GPS devices and low accuracy GPS smartphones, whereas the algorithms were described only cursorily. From the literature, it is still unclear whether everyday GPS devices can be used to accurately identify lane changes, and what algorithmic approach should be used for this purpose, hence serving as a research gap.

The present research study aims to fill this research void, by collecting lane change data on Dutch highways, and applying an algorithm for identifying lane changes. The study adopts the basic principles of the algorithms of the previously mentioned authors [10] [11] and verifies how lane change identification performance differs among the different portable devices when using the GPS sensor.

Method

An experiment was performed with the goal of obtaining a dataset that allows for identifying lane changes made when a vehicle is driven on a highway. This data was recorded by the student working on the project thesis while the car was driven by the project supervisor. While driving, lane changes on different sections of the highway were performed.

Hardware

The experiment was conducted on a road that started on the campus of Delft University of Technology, went around Rotterdam, and then went back to the campus, as shown in Figure 2. In total, this section was driven three times, leading to 179 km of data (highway and non-highway data which will be used later for analysis). The details of the car and the hardware recording systems used for data collection are mentioned in Table 1 and Table 2 respectively.

Table 1: Data collection car details

Car	Date	Experiment	Length (mm)	Width (mm)	Height (mm)	Curb weight (kg)
Peugeot 108 [13]	4th of June, 2021	Pilot	3475	1615	1460	815
KIA Picanto [14]	21st of October, 2021	Validation Dataset 1	3595	1595	1485	874
KIA Picanto [14]	21st of October, 2021	Validation Dataset 2	3595	1595	1485	874

Table 2: GPS receiver device details

GPS Devices	Frequency of GPS signal (Hz)	Chipset
GoPro Max	18	U-Blox M8 GNSS (chip) [15]
Huawei P20 Lite	1	HiSilicon Kirin 659 (A-GPS (US)) [16]
Samsung Galaxy S9	1	Exynos 9810 (A-GPS and Galileo) [17]

GlobalSat BU343-s4	1	SiRF Star IV GPS chipset [18]

The GoPro Max camera recorded the videos in Quad HD or 2560×1440 pixel resolution at 30 frames per second (FPS). The GoPro Max records GPS signals at a frequency of 18Hz. As the other devices recorded at 1 Hz, the GPS values of the GoPro Max are averaged to 1 Hz to maintain uniformity. The choice of the above devices was to compare two smartphones of different brands and their accuracy compared to that of a professional-grade GPS device. The choice of a GoPro Max camera was for recording the drive by having a stable mount on the car dashboard and recording the video of high quality as well as to serve as ground truth.

Figure 2: Map of the path travelled for Data collection

The recording was made using all four devices, which collected GPS traces of the travel. The journey included following the speed limit of 100 km/h and conducting lane changes on the highway at different intervals. Figure 2 shows the map of the route traveled to collect the data for the experiment.

Data Collection

First, the devices were acquainted with on a road in the campus of Delft University of Technology. This trial test helped to mount properly the camera on the dashboard, the GlobalSat GPS device mounted on top of the car, and changes to the smartphones settings to continue recordings even when the screen is not switched on.

The smartphones (Huawei P20 lite and Samsung Galaxy S9) were attached to the car dashboard using tape.

Figure 3: Experiment setup: (a) car used for data collection; (b)placement of smartphone and GoPro on the car dashboard; (c) laptop placement connected to the GlobalSat GPS receiver; (d) placement of GlobalSat GPS receiver on the car roof

Figure 3 shows the experimental setup used for collecting data in the car. The setup includes the four hardware devices mounted in the same position throughout the journey. There were two days of data collection. The first was conducted between 14:00 and 16:15 with an outside temperature of 25° C and sunny conditions. The first day of data collection was used as a pilot experiment to know and understand the data to identify lane changes. The second day of data collection was conducted between 10:30 and 13:30, with an outside temperature of 14° C with a mix of weather conditions such as sunny, overcast and rainfall. This data was collected to validate the algorithm used for identifying lane changes based on the data collected for the pilot experiment. The lane width on the highway where data was recorded is 3.25 m and multiple lane changes were made on the highway which consisted of 3 or more lanes.

Figure 4: Camera recording on a highway on the first day of data recording

Figure 4 is an example of how the highway recording on the camera looks like. The highways had lane markings and speed limits. The route consisted of driving in tunnels, under flyovers, and overpasses.

The driver was instructed to make naturalistic lane changes at speeds as desired and seemed fit for that scenario of traffic on the road. While driving on the highways, the lane changes were made on all types of road trajectories and at varying intervals to ensure the experiment consisted of different scenarios and lane changes.

Data processing

The data gathered through the different GPS devices were collected and stored as CSV files. These files were added onto Google Maps [19] on the QGIS software [20] platform to view the path travelled with more detail.

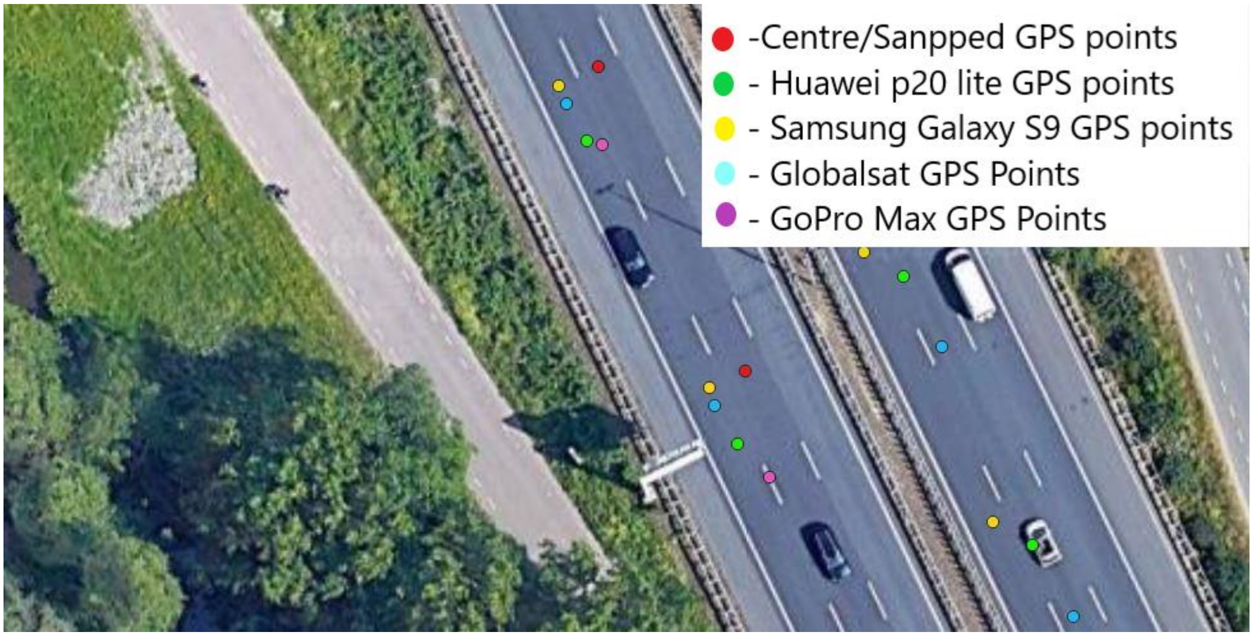

Figure 5: The different GPS device data projected on a map

The GPS points captured from the portable devices were snapped or moved to the centre of the road on which it was captured. Hence, no matter the variations in the path travelled, the actual shape of the road is obtained. This task was performed by using Google's Roads API [19]. This module has a function called 'snap to road'. The function takes the latitude and longitude values of the smartphones and snaps them to the values of the point at the centre of the road. The GPS points representing centre of the road are used for further calculations that involve heading angles and lateral distances between recorded GPS points and the centre of the road. Figure 5 shows the GPS points visualised on a QGIS software map [20].

After snapping the recorded GPS points to the nearest road, we have two layers with coordinates. The first layer contains the originally recorded GPS points, and the second layer contains the coordinates of the road centre points that lie closest each original GPS coordinate (see Figure 7).

Ground truth for identifying lane changes

The ground truth for identifying lane changes is the camera recording from the GoPro. The time at which the car's centre coincides with the lane marking is recorded manually as a lane change.

Figure 6: Camera recording of lane change instant on the A13 highway

For example, as shown in Figure 6, there was a lane change from lane number 2 to lane number 3 at the camera time of 6:00 min (red box) recorded from the journey's beginning. The red vertical lines in Figure 6 are edited on to the picture to explain the how the centre of the car is considered to move across lanes for the task of manual annotation. Similarly, all the lane changes are identified following the same procedure, serving as the complete dataset's ground truth (i.e., lane changes manually captured by the camera which captures videos at 30 FPS [21], hence each frame base of 30 FPS is synchronised with 1 Hz of GPS data collection).

In the following section I explain the two methods that were used to identify lane changes. The first method estimates the lateral distance that would be built up if the car proceeded to drive in its current direction relative to the road's direction. This method is named the 'Projected lateral distance to the centre of the road'. The second method is named 'Lateral distance to the centre of the road', which relies on changes in the distance between the GPS points and their road-snapped counterparts.

Projected lateral distance to the centre of the road

Smartphone GPS values are not entirely accurate as they obtain GPS data directly from satellites which may be prone to multipath effects and variations in recording due to tunnels and buildings during driving. The captured GPS values show the same path of travel but may have an offset from the true position, as observed in Figure 7. Hence, the recorded data, appears to follow the same path of travel but is not accurate as it is offset by a particular value.

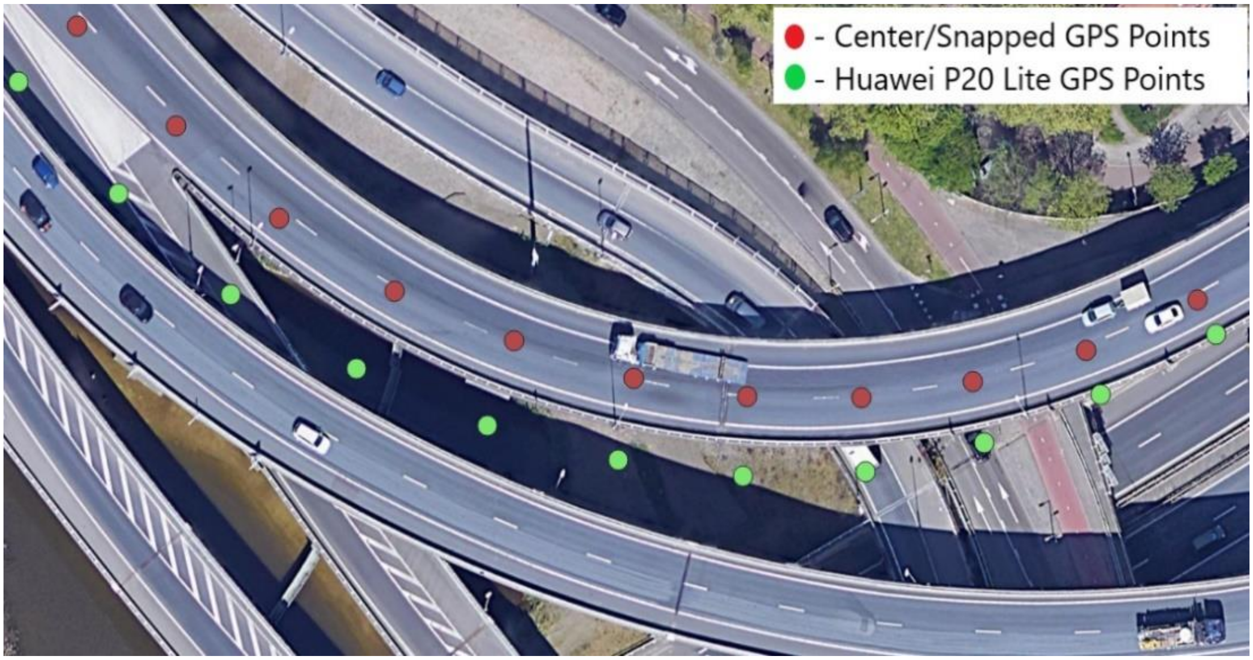

Figure 7: The recorded points (green) compared to the road centre points (red), are precise but not accurate

To solve the above-mentioned problem of the GPS point lying outside the road boundary, the concept of lateral deviation for every consecutive point of travel is calculated. An essential part of the algorithm is finding the difference between the road's heading angle and the vehicle to determine the lane change. After snapping the GPS points to the centre, the heading angle of the road is calculated using the azimuth angle formula for two subsequent centre of the road points. The same procedure is adopted to calculate the recorded GPS points heading angle. The difference between both angles is multiplied with the distance between two subsequent GPS recorded points to find the lateral deviation as explained in Equation 1 and shown in Figure 8, adopted from the research of Faizan et al [10]. Once this lateral or projected deviation crosses a threshold limit, the algorithm identifies it as a lane change event.

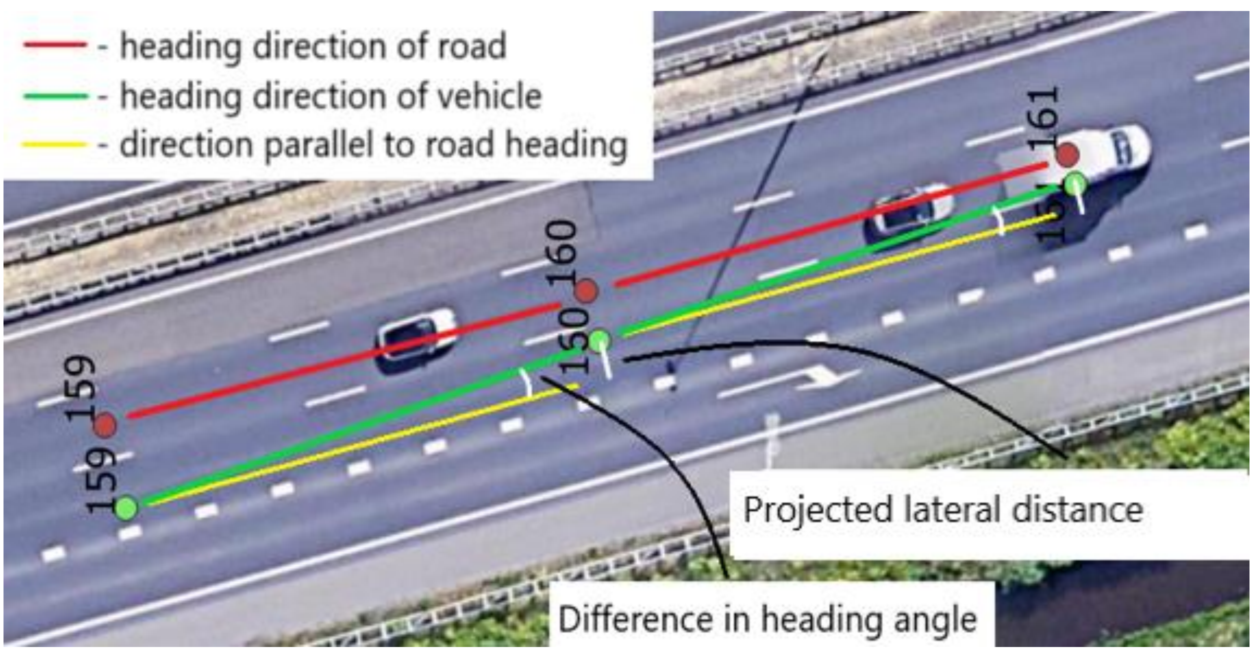

Figure 8: Explanation of heading angle and projected lateral distance

$$\textit{Relative lateral distance} = \textit{sine} * (\textit{Difference in heading angle of road and car}) * \textit{distance between two consecutive recorded GPS points}$$

Equation 1: Relative lateral distance calculation

Lateral distance from the centre of the road

The distance between the points of the smartphone/Globalsat/GoPro GPS and the centre of the road is calculated at every instance of recording (1 Hz).

Figure 9 and Figure 10 (captured from processing the data recorded in the experiments on QGIS software) show how lateral distance is calculated for the car travelling on a highway. The yellow points indicate the position of the centre of the road. The green points indicate the position of the car. The distance between the yellow and green points is the lateral distance of the car from the centre of the road for every instance (1 Hz) of data recorded. The value of change of lateral distance greater than the width of the lane or a multiple of the width of the lane, indicates a lane change event.

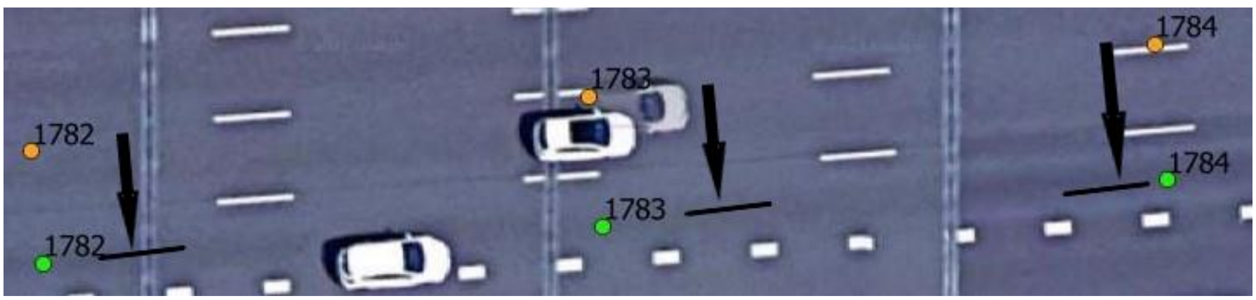

Figure 9: Lateral distance to the right of the centre of the road

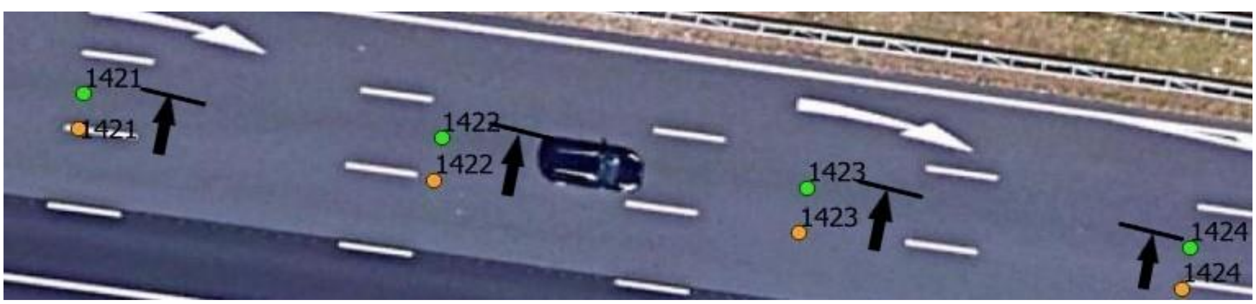

Figure 10: Lateral distance to the left of the centre of the road

To visualise the entire lateral position change while driving on a highway, Figure 11 represents the values varying from positive to negative. Negative distances indicate lateral positions to the left of the centre of the road whereas the positive values indicate lateral positions to the right of the centre of the road.

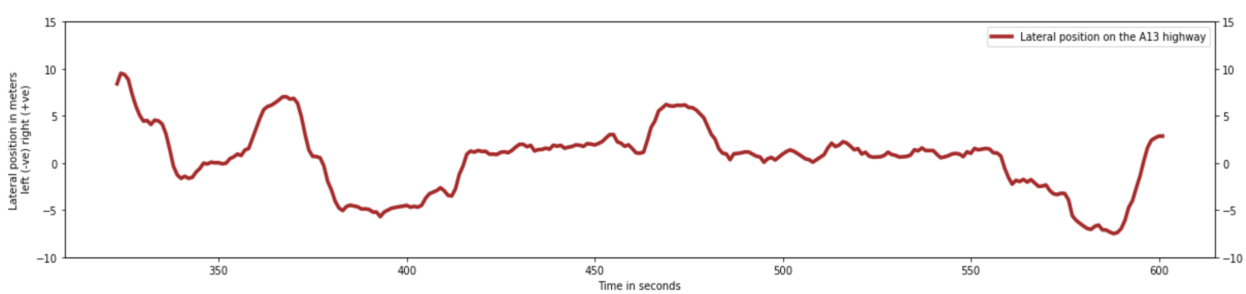

Figure 11: Lateral position of the vehicle on the A13 highway

Data for the experiment

This project has three datasets:

- Pilot experiment Data (Data processed on the A13 highway during the first trip)
- Validation Dataset 1 (Data processed on A13 and A15 highway during second and third trip)
- Validation Dataset 2 (Data processed on A13 and A15 highway during second and third trip)

The information on the three datasets is explained in Figure 12. The green arrows processed and used for the lane change identification task, whereas the red arrows represent the omitted data.

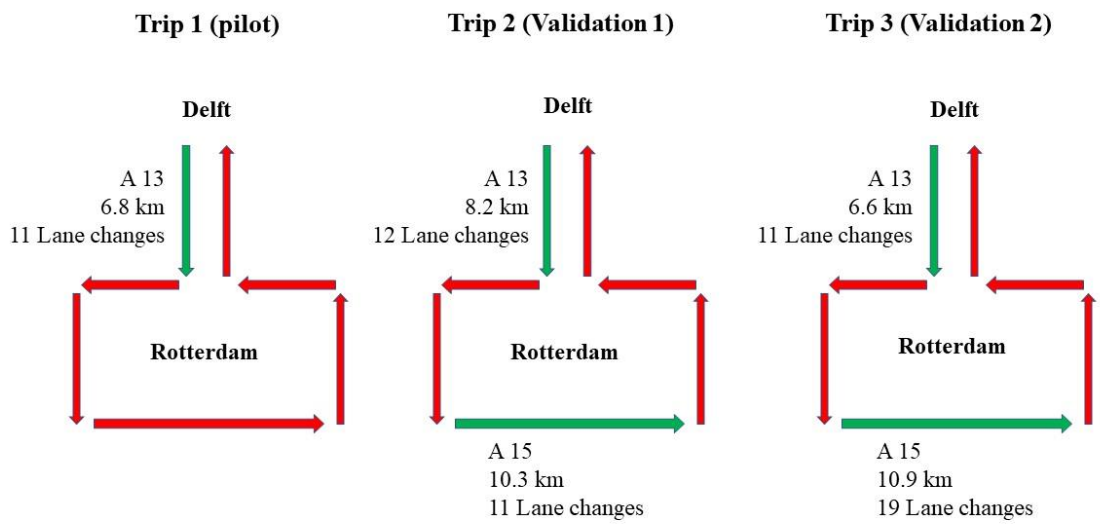

Figure 12: Choice of highway data considered for pilot experiment and validation datasets

Data analysed on the following two highways:

- A13 – The highway is relatively straight and has less curves
- A15 – The highway has more curves and not straight

The two motorways mentioned above provide a variation in the road shape and thus can help in analysing how the GPS data affects the lane identification algorithm which will be described later in the report.

Steps followed in identifying lane changes

Visualisation of the different data signals

First, the pilot experiment data is used to visualise the different signals, which will be used to identify lane changes. In this step of visualisation, the signals, such as lateral distance, heading angles, projected lateral distance and relative heading angle, will be plotted on a graph along with the ground truth to understand if the signals vary or provide a distinct variation in the presence of or near the ground truth values. Without much effort, a person can discern distinctions in line length, shape, orientation, distances, and colour (hue); these are referred to as "pre-attentive characteristics." The human visual system is a suitable pattern seeker when data is represented in a way that essential and informative patterns stand out [22]. Hence, spikes/peaks or sudden deviations in the data signal (e.g., lateral distance) in the vicinity of the ground truth of lane changes can serve as a good indication to know if the signal is a good choice for the task.

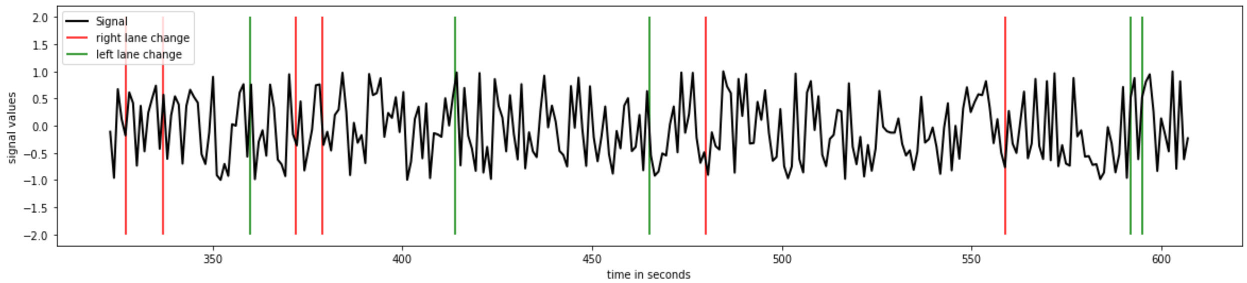

Figure 13: Data signal not showing a suitable pattern when plotted along with the ground truth representing actual lane changes

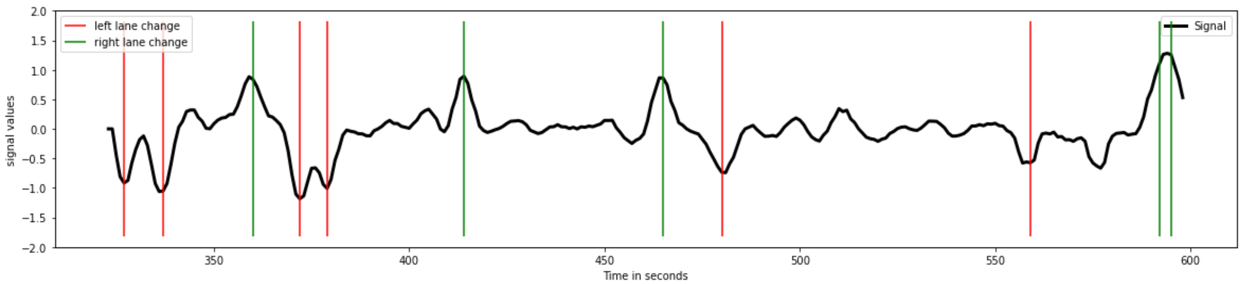

Figure 14: Data Signal such as lateral velocity showing a suitable pattern when plotted along with the ground truth representing actual lane changes

In Figure 13, it is observed that the signal does not show a pattern or variation in values when it approaches a ground truth value. However, in Figure 14, the signal shows a pattern or variation in values in the vicinity of an actual lane change. The aim is to identify such characteristic signals which will be analysed by an algorithm to identify lane changes.

Improvement of data signals

The above graphs from the pilot experiment data contain many short-term variations, making it challenging to identify the actual lane change events. Filtered data signals reveal distinct characteristics such as lower peaks when the car travels in the same lane and sharper or steep peaks when a lane-change manoeuvre occurs. The suitable data signals selected are filtered after visualisation by removing noise present in it. 1-dimensional filters (e.g., mean, and median filters) are employed to smoothen the signals.

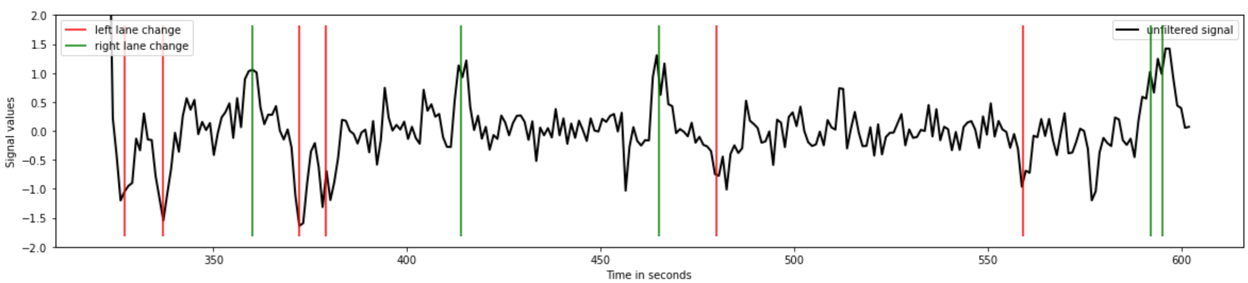

Figure 15: Unfiltered data signal containing noise

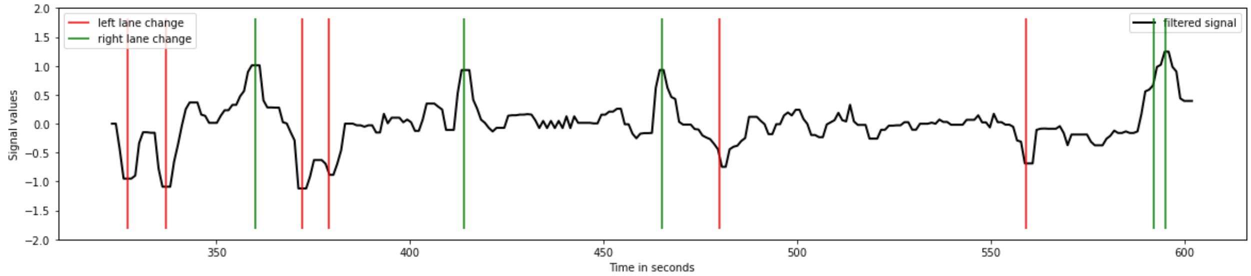

Figure 16: Filtered smooth data signal with noise removed

Figure 15 shows how a signal has a pattern to depict variations of lane changes in the vicinity of the ground truth but contains noise that might affect identifying false signal values as lane changes. To avoid errors, the signal is filtered, as shown in Figure 16, to provide a smoother curve so that it will be easier for an algorithm to distinguish between noise and actual variations or spikes in the signal values when it approaches the ground truth. The combination of filters and the parameters is explained in the results section.

Algorithm for automatic identification of lane changes

To measure the performance of the data signal in the algorithm the F1 score metric is used. The F1 score can be understood as a harmonic mean of precision and recall, with the best value being 1 and the poorest being 0 [23]. To understand F1 score it is important to understand what precision and recall means. The meaning can be explained better in terms of equations.

$$Precision = TP/(TP + FP)$$

Equation 2: Precision [24]

$$Recall = TP/(TP + FN)$$

Equation 3: Recall [24]

Precision is the fraction of true positive examples among the examples that the model classified as positive
Recall is the fraction of examples classified as positive, among the total number of positive examples

The F1 score is calculated using the ground truth (i.e., lane changes manually captured by the camera which captures videos at 30 FPS; hence each frame base of 30 FPS [21] is synchronized with 1 Hz of GPS data collection) and peak identification for the different signals. Different scenarios in the algorithm:

1. If the peaks and ground truth are within a window, the scenario is identified as a true positive.
2. If the window includes a ground truth without any peak, the scenario is identified as a false negative.
3. If the window includes a peak without ground truth, the scenario is identified as a false positive.
4. If the window does not include a peak nor a ground truth, the scenario is identified as a true negative.

For F1 score calculation, true negatives are not required and hence such values are not identified.

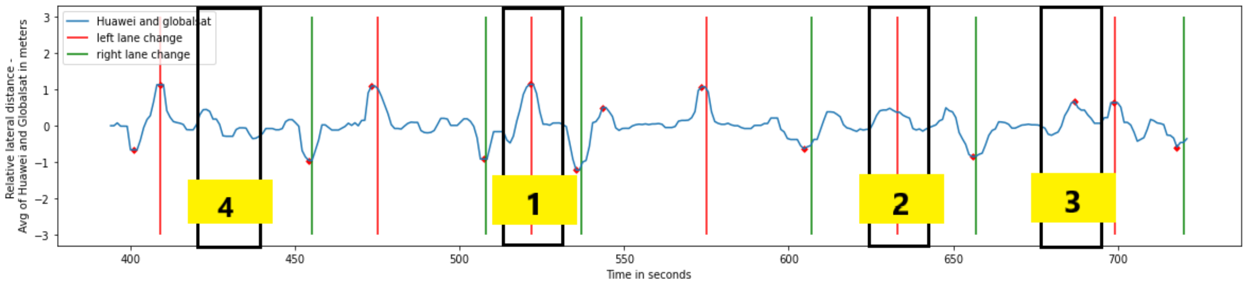

Figure 17: Identifying peaks and showing the ground truth

Figure 17 illustrates peaks of data signals along with the ground truth of the car during the drive on the highway (Note the actual window size is not shown by the black box in Figure 17, it is considerably wider to provide easy explanation of window, ground truth and peaks in the graph).

The procedure of obtaining peaks in a graph plot is by setting a threshold value. Once the signal crosses the threshold value, the peak is identified with a red point marker. The threshold is identified by analysing the graph plot for the signal. The procedure involves first identifying the threshold at which the car/vehicle stays within the lane limit during a regular or naturalistic drive on a highway. To obtain this threshold, a section of the data was taken during which there were no lane changes, and the drive was continuous without any traffic interruptions. The lateral distance was calculated for every pair of consecutive GPS point to know how far the car would travel laterally with every recording.

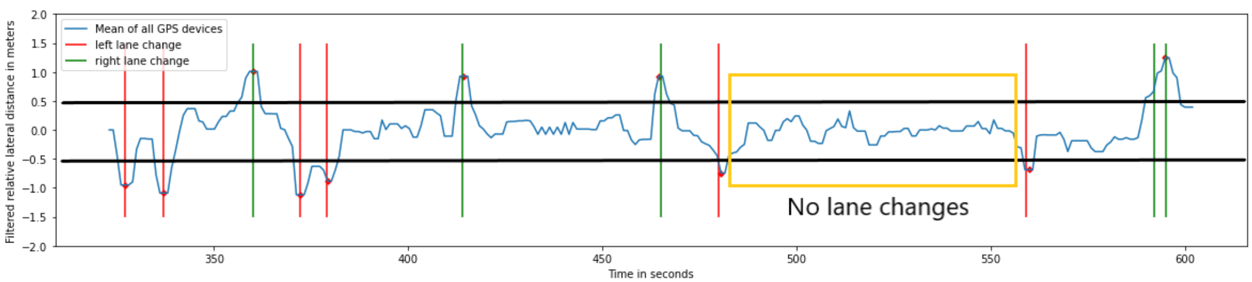

Figure 18: Explanation of the threshold explanation on A13 highway

For example, during the drive on the A13 highway, a section of the travel which lasted for 79 seconds (yellow box in Figure 18) had no lane change events. In this window, the graph plotted for filtered (median filter) relative lateral distance signal values do not cross the limit range of -0.5 to 0.5. Hence, a value above this limit will be identified as a lane change event. A similar procedure is followed to identify the thresholds for other signals used to identify lane changes.

For the algorithm to identify the true positives, false positives and false negatives, a window of size 3 is used as an average lane change lasts for 3 seconds or less[10]. Using these values of true, false positives and false negatives, the F1 score can be computed as shown in Equation 4. It represents how well the signal can be used for the purpose of lane change identification.

$$F1\ score = TP/(TP + 1/2\,(FP + FN))$$

Equation 4: Calculation of the F1 score

Finally, choosing the best signals for the task of lane change identification concludes the pilot experiment data analysis.

Further, the algorithm with the best choice of signals will be assessed on the Validation datasets 1 and 2 which includes GPS data from all four devices (2 smartphones, GlobalSat and GoPro Max).

Results

Visualising data from the pilot experiment

Data gathered from the GPS devices during the pilot experiment are visualised to understand which signal will most efficiently identify lane changes. Figure 19 shows the lateral position of the various GPS data obtained from the smartphones and the GlobalSat GPS device. In the above figure, note that 0 m on the Y-axis represents the center of the road. Figure 20 depicts the heading angle of the car using data from the different GPS recording devices.

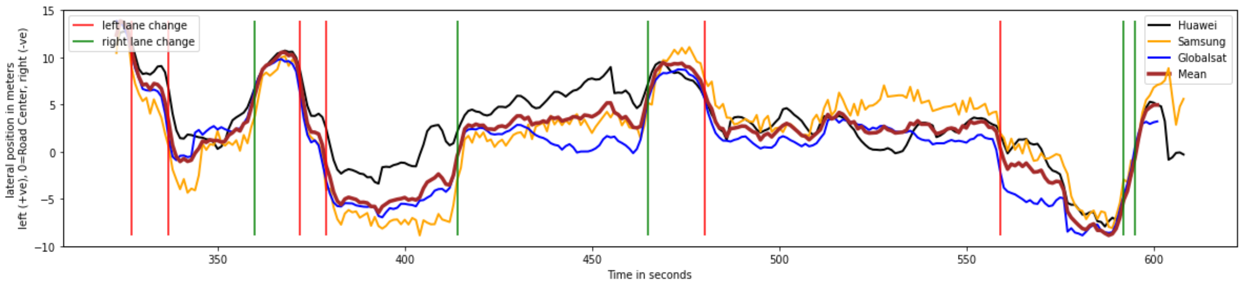

Figure 19: Visualisation of lateral position of the car on the road of pilot experiment. The vertical lines represent the actual moments of lane changes.

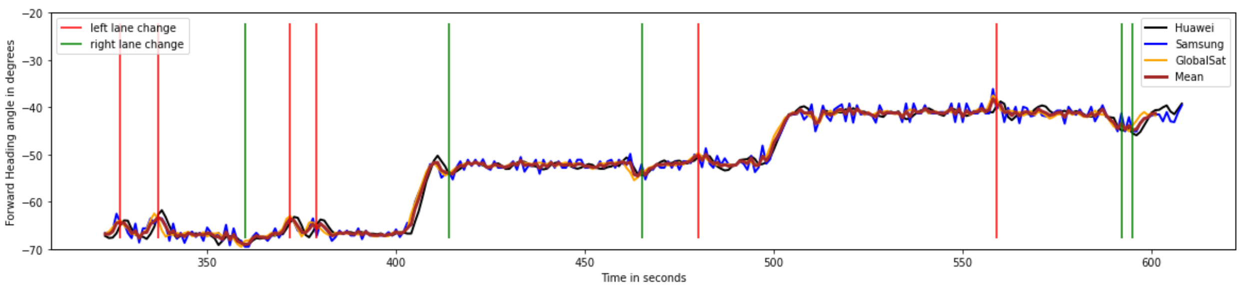

Figure 20: Visualisation of forward heading angle of moving vehicle of pilot experiment. The vertical lines represent the actual moments of lane changes.

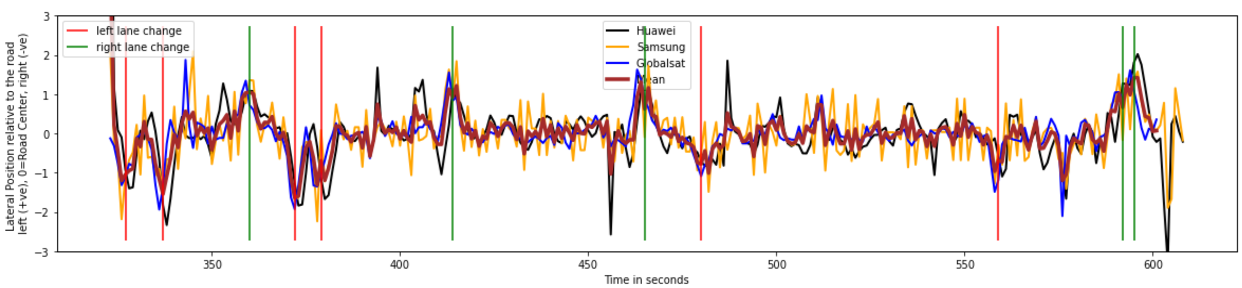

Figure 21: Visualisation of lateral position of vehicle relative to the road of pilot experiment. The vertical lines represent the actual moments of lane changes.

Figure 21 shows the vehicle's lateral position on the road relative to the center of the road. The peaks correspond with the lane changes, i.e., as the relative lateral distance increases during the event of a lane change. Figure 22 shows the heading angle in degrees of the vehicle relative to the heading angle of the road/highway.

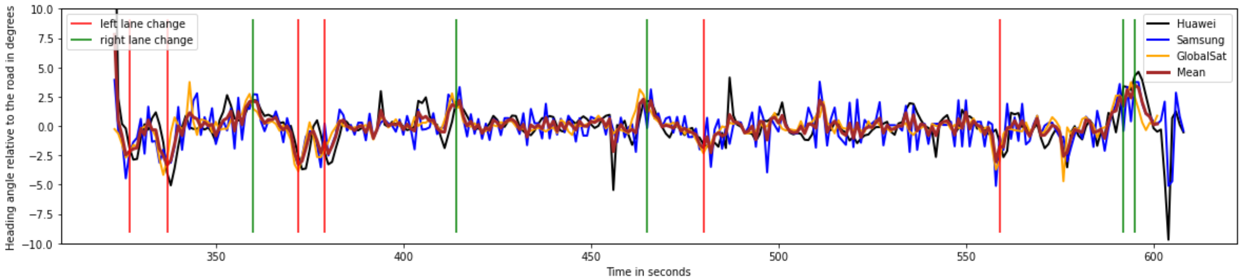

Figure 22: Visualisation of relative heading angle of road and the GPS recorded points by the devices of the pilot experiment. The vertical lines represent the actual moments of lane changes.

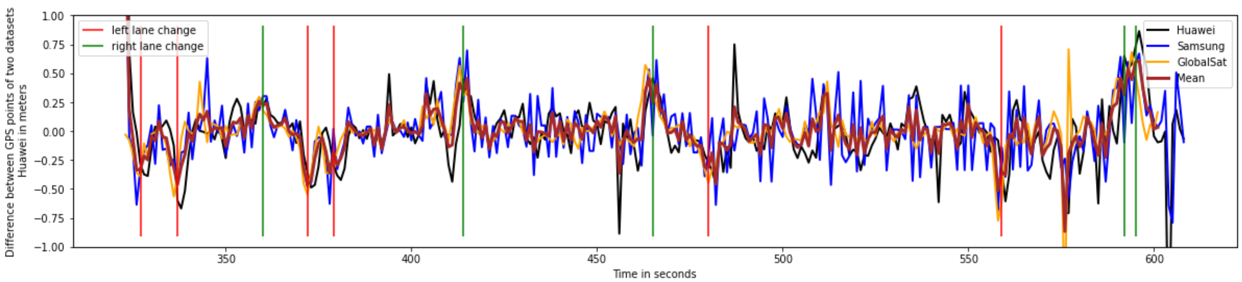

Figure 23: Difference between distance of subsequent GPS points of the centre of the road and the GPS points recorded by the devices. The vertical lines represent the actual moments of lane changes.

The difference between the distance of two subsequent GPS points of the centre of the road and that of the GPS points recorded by the device is represented in Figure 23 (i.e., the derivative of the signals shown in Figure 21). The difference in the distance is expected to be small when the car travels in the same direction of the road. Figure 17 shows that large values, which correspond to lateral movement, tend to occur with lane changes.

From the above explorative visualisation of the different signals to capture information on lane changes, it was decided to use lateral distance, relative lateral distance, and relative heading angle for the automatic identification of lane changes of the vehicle on a highway.

Improving data signals from pilot experiment

For the selected signals (lateral distance, relative lateral distance, and relative heading angle) filters are applied to remove noise. The choice of filters and their parameters are explained below. An example of signal improvement is explained with the lateral distance signal. Figure 24 shows an unfiltered mean lateral distance signal of the pilot experiment GPS receiving devices and Figure 25, Figure 26 are filtered mean lateral distance data signals. Figure 27 depicts a new signal called estimated of lateral velocity, derived from lateral distance.

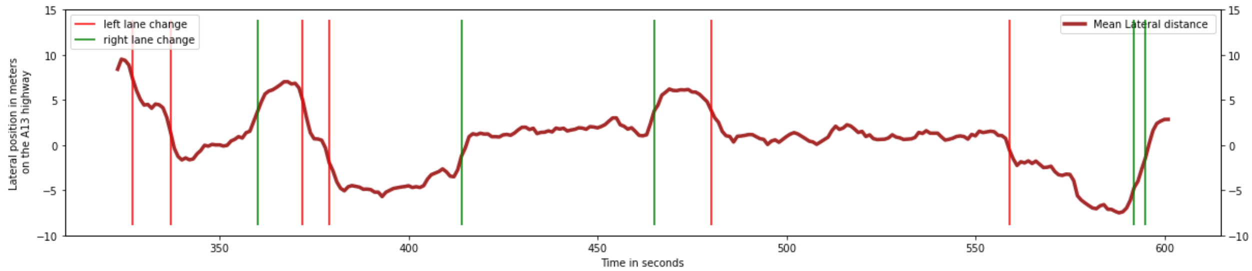

Figure 24: Visualisation of Unfiltered Mean Lateral Distance of the pilot experiment.

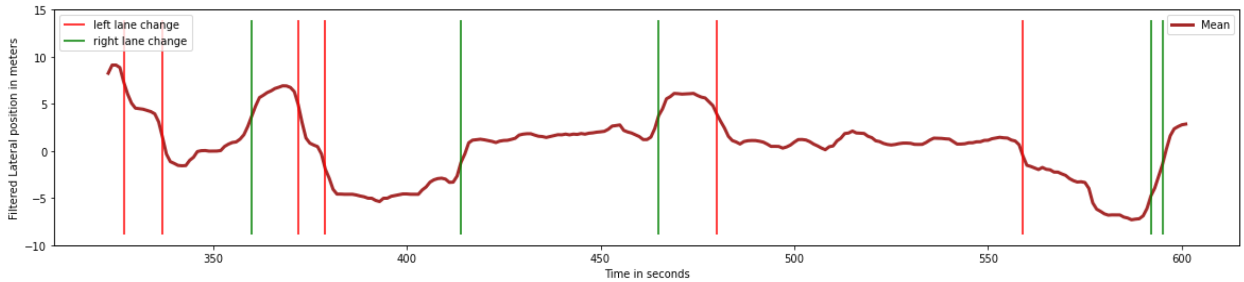

Figure 25: Visualisation of Mean Lateral Distance of the pilot experiment with median filter only

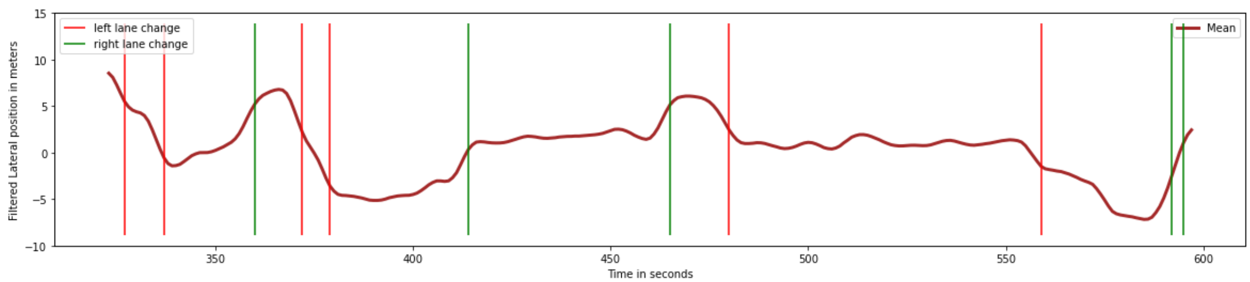

Figure 26: Visualisation of Mean Lateral Distance of the pilot experiment with median and moving average filter.

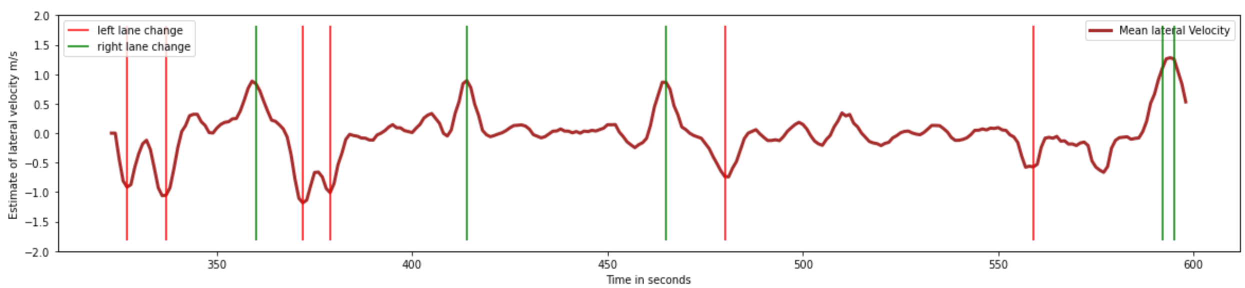

Figure 27: Visualisation of estimated of Lateral Velocity of vehicle of pilot experiment.

The estimate of lateral velocity as seen in Figure 27 shows distinct peaks or variations in the signal. The corresponding equation to derive this signal is shown in Equation 5.

$$Estimated\ Lateral\ Velocity =$$
$$numpy.diff(moving\ average(sp.signal.medfilt(lateral\ distance, 5), 5))$$

Equation 5: Estimate of lateral velocity derived from lateral distance using filters

Another example (Equation 6) is to use only the median filter on the signal of relative lateral distance (Figure 15). This approach yielded prominent peaks at lane change events, again based on visual inspection.

$$Filtered\ relative\ lateral\ distance = sp.signal.medfilt(relative\ lateral\ distance, 5)$$

Equation 6: Filtered relative distance derived from relative lateral distance

Median filters, moving average filter, and the difference between two subsequent values were used in combination or individually for a signal.

- 1D-Median filter: filter used for smoothing signal and suppression of noise.
- Moving average filter: A moving average is a method of analysing data points by calculating the averages of various subsets of the entire data set.
- Difference: When calculating the n-th order discrete difference along a given axis, we utilize the function. The difference in the filtered value of the lateral distance is an estimate of lateral velocity.

The parameter values for median filter can be adjusted. After verifying which value suits best for the recorded signals, it was observed that 5 was the value to observe distinct peaks visually at lane change moments in the graph. Hence, median and moving average of values for 5 seconds are calculated.

Result of running the algorithm on the collected data of the pilot experiment

The pilot experiment was performed to find the best signal for identifying lane changes. Hence, a mean of the GPS signals from different devices used in the pilot experiment is calculated for each signal and measuring its performance with the results shown in Table 3.

Signal	F1 score
Mean estimated lateral velocity	0.846
Mean relative lateral distance	0.606
Mean filtered projected lateral distance	0.952
Mean filtered heading angle	0.740

Table 3: Results of the performance of pilot experiment signals on the algorithm

Hence for further assessing the performance of different GPS devices, the signals of estimate of lateral velocity and filtered projected lateral distance is used. In the following evaluation of the validation sets, the individual performance of the different GPS devices and their mean is assessed using the earlier explained algorithm.

Results of validation dataset 1

F1 score of the different GPS devices on the highway of A13 (shown in Table 4) having 12 lane changes along with 1 aborted lane change, to rate the performance of lane change identification

Device	F1 score (estimated lateral velocity)	F1 score (filtered projected lateral position)
Mean of all GPS devices	0.583	0.869
Huawei P20 lite smartphone	0.457	0.647
Samsung Galaxy S9 smartphone	0.461	0.615
GlobalSat Bu-343s4 GPS device	0.666	0.615
Go-Pro	0.692	0.916
Combination of smartphone and Globalsat device	0.592 (Huawei p20 lite smartphone + GlobalSat Bu-343s4 GPS device)	0.846 (Huawei p20 lite smartphone + GlobalSat Bu-343s4 GPS device)

Table 4: Results of A13 highway data of Validation Dataset 1

F1 score of the different GPS devices on the highway of A15 (shown in Table 5) having 11 lane changes and no aborted lane change, to rate the performance of lane change identification

Device	F1 score (estimated lateral velocity)	F1 score (filtered projected lateral position)
Mean of all GPS devices	0.727	0.916
Huawei P20 lite smartphone	0.540	0.600
Samsung Galaxy S9 smartphone	0.516	0.800
GlobalSat Bu-343s4 GPS device	0.692	0.818
Go-Pro	0.740	0.842
Combination of smartphone and GlobalSat device	0.720 (Huawei p20 lite smartphone + GlobalSat Bu-343s4 GPS device)	0.869 (Samsung Galaxy S9 smartphone + GlobalSat Bu-343s4 GPS device)

Table 5: Results of A15 highway data of Validation Dataset 1

Results of validation dataset 2

F1 score of the different GPS devices on the highway of A13 (shown in Table 6) having 11 lane changes and no aborted lane change to rate the performance of lane change identification

Device	F1 score (estimated lateral velocity)	F1 score (filtered projected lateral position)
Mean of all GPS devices	1.000	1.000
Huawei P20 lite smartphone	0.740	0.760
Samsung Galaxy S9 smartphone	0.916	0.960
GlobalSat Bu-343s4 GPS device	0.833	1.000
Go-Pro	0.859	1.000
Combination of smartphone and GlobalSat device	0.916 (Samsung Galaxy S9 smartphone + GlobalSat Bu-343s4 GPS device)	1.000 (Samsung Galaxy S9 smartphone + GlobalSat Bu-343s4 GPS device)

Table 6: Results of A13 highway data of Validation Dataset 2

F1 score of the different GPS devices on the highway of A15 (shown in Table 7) having 19 lane changes and no aborted lane change to rate the performance of lane change identification

Device	F1 score (estimated lateral velocity)	F1 score (filtered projected lateral position)
Mean of all GPS devices	0.588	0.894
Huawei P20 lite smartphone	0.744	0.864
Samsung Galaxy S9 smartphone	0.615	0.833
GlobalSat Bu-343s4 GPS device	0.769	0.833
Go-Pro	0.800	0.842
Combination of smartphone and GlobalSat device	0.705 (Huawei p20 lite smartphone + GlobalSat Bu-343s4 GPS device)	0.923 (Huawei p20 lite smartphone + GlobalSat Bu-343s4 GPS device)

Table 7: Results of A15 highway data of Validation Dataset 2

Discussion

Analysis of results of validation dataset 1:

The performance of the detection of lane changes is low with the validation dataset 1. A few possibilities for the low performance are:

- The overcast was high during the data collection and hence can be a direct result on the performance.
- Traffic was high and few instances of lane changes were made but not completed, hence these movements also contributed to false positive results.
- The low performance is mainly observed for the smartphones. A possible reason for such low performance can be attributed to not detecting the GPS points at few instances as shown below in Figure 28.

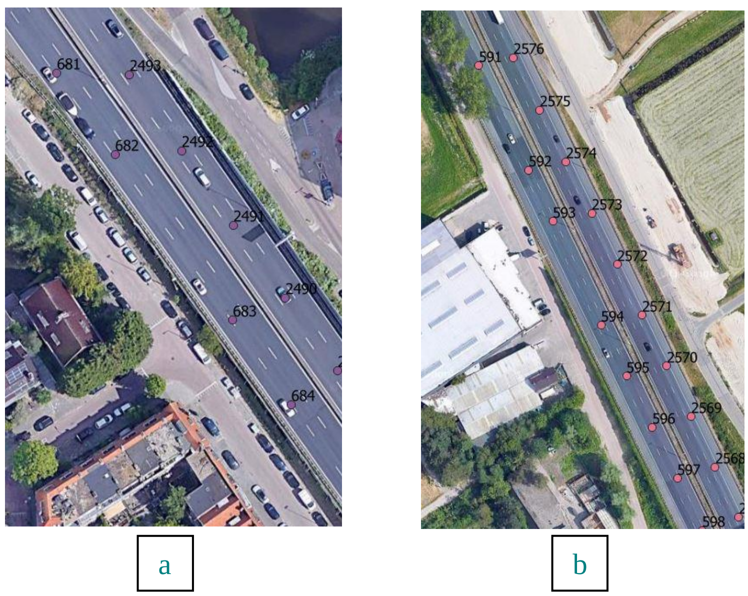

Figure 28: GPS points not detected at equal frequencies by smartphones; a) Huawei p20 lite; b) Samsung Galaxy S9

Analysis of results of validation dataset 2:

Compared to the A13 highway, the performance on the A15 highway is lower in the second validation data set. Hence an analysis was made to understand why the results varied greatly. After observing the A15 highway closely on QGIS software:

- The centre road points were snapped to another linking road due to the low accuracy of smartphone GPS points as shown in Figure 29.

- De-merging lanes at a few sections of the road increased the false lane change values due to varying centre positions of the road points as shown in Figure 30.

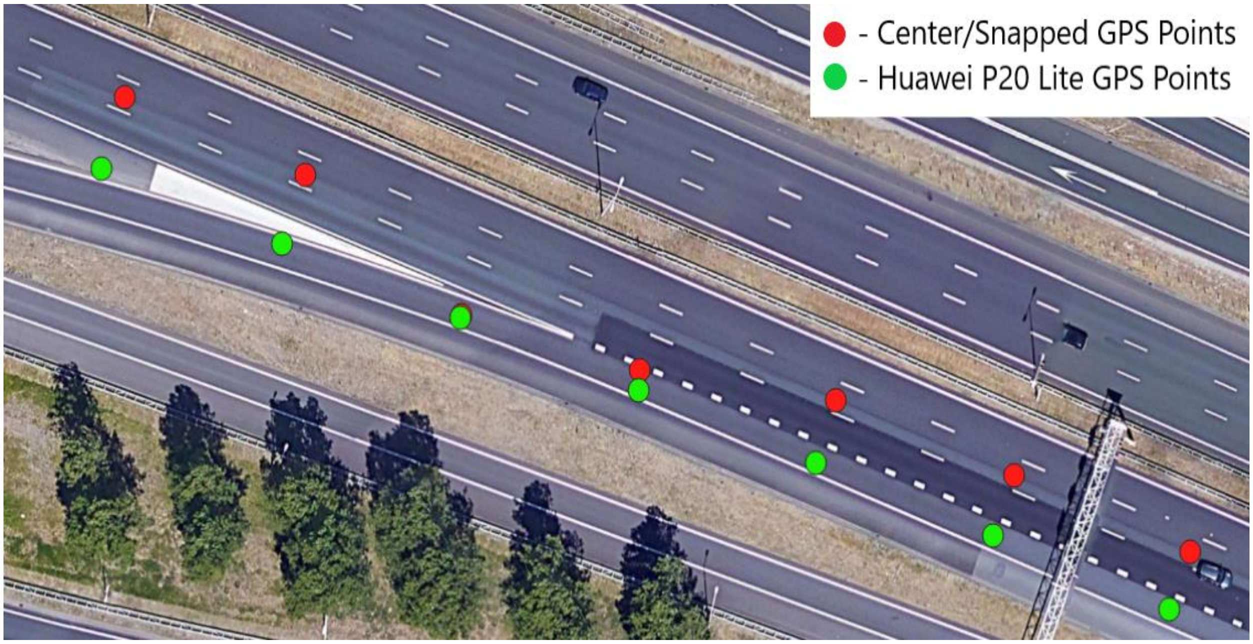

Figure 29: Points not snapped to the centre of the road accurately

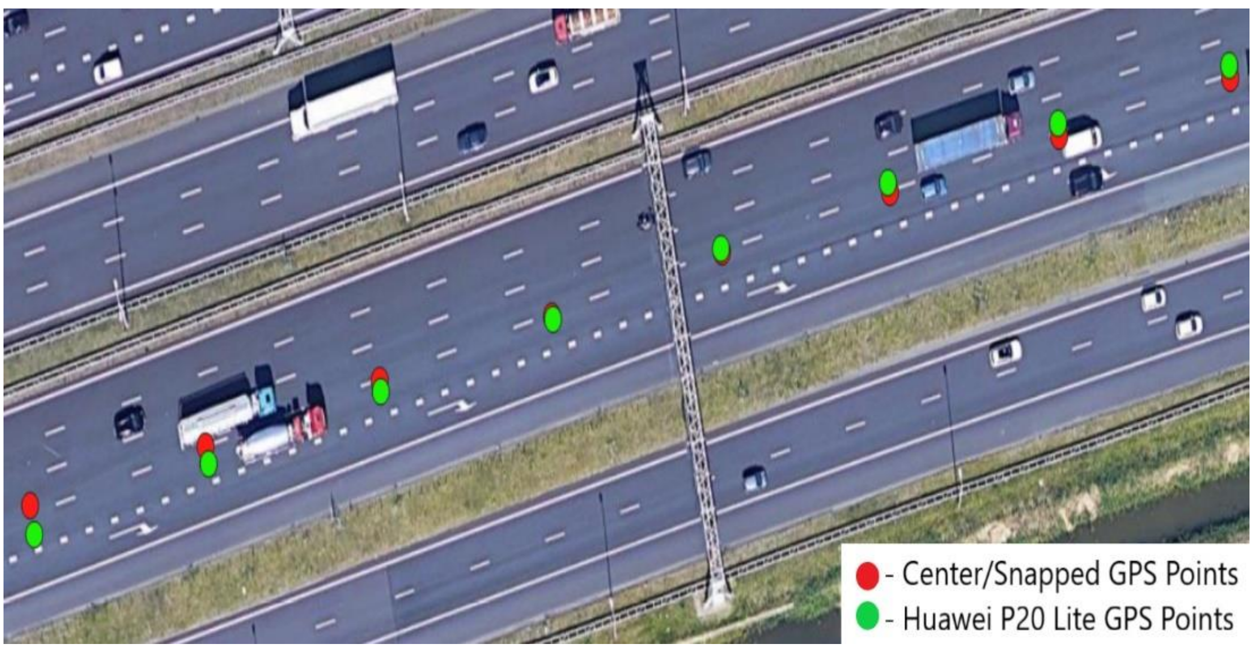

Figure 30: Centre of the road points are not at the actual centre due to merging of lanes

Analysis of signals from the two validation datasets:

- From the above results, in most validation dataset cases, the filtered projected lateral distance signal provides a higher F1 score than the estimated lateral velocity signal in identifying lane changes. Apart from being an estimation of lateral velocity, the signal has more variations within the threshold boundary when observed. These may be due to irregular movements influenced by traffic.

- Among the GPS capturing devices, the GoPro Max provides the best results of an overall F1 score of 0.9 using the filtered projected lateral distance signal in the validation datasets 1 and 2. GlobalSat provided the second highest F1 score in most test cases, which is followed by the two smartphones.

- The Huawei P20 lite outperforms the Samsung Galaxy S9 on a few datasets and highways, while the Samsung Galaxy S9 outperforms the Huawei P20 lite in the rest of the data. In the end the two smartphones provide results which do not exceed each other and thus show similar overall performance.

- The mean of all GPS devices ranks higher in performance for most cases of the validation datasets than all other devices and combinations of devices but gets outperformed mainly by GoPro Max and the combination of smartphone and GlobalSat in few cases the validation datasets.

- The F1 scores are low for a few signals in different datasets, this can be attributed to the fact that the values of thresholds are based and are tailored to the pilot experiment data. This may provide a form of underfitting or overfitting on human optimization methods.

- The threshold for the signals may vary from the pilot experiment to the validation data as it a method of human optimization used which may have overfitted or underfitted the validation datasets leading to results varying from the expected performance.

- Along with the five scores, an extra F1 score is calculated for the best performing smartphone with the GlobalSat GPS receiver. Such data is desirable if the combination provides results better than their individual performance respectively. As a GlobalSat device is relatively cheap compared to a GoPro Max, the combination of data with a smartphone can provide better accuracy than smartphones individually in identifying lane changes.

Results from the validation datasets can conclude that the lateral velocity signal proved to be less efficient in identifying lane changes than the filtered projected lateral distance, a modification of the method described by Faizan et al [10].

Future developments may focus on improving the quality of GPS sensors in smartphones. The GoPro GPS sensor works better than the smartphone's GPS sensors. Hence the GPS chip u-Blox M8 GNSS can be used in a smartphone if it can be made compatible with smartphones to improve the accuracy of the received GPS points.

In this project, the median and moving average filters help smoothen out the 1-dimensional signal by removing noise. The results show that the use of such filters helps visualise the signals better and eventually helps researchers decide which combination of filters provides good results by the algorithm in identifying lane changes. Exploring the different combinations of filters and their parameter values may further improve the results. Determining the optimal filtering techniques is seen as a worthwhile future research topic.

The above data analysis was conducted on straight and mildly curved roads. The characteristics of flat terrain and relatively minor sections of the motorway covered by canopies of trees may not provide the same result when tested on motorways on hilly areas and covered by trees. Hence the performance of the lane identification algorithm can differ in other countries with variations in road characteristics.

Limitations

The snap-to-road function was used for directing or correcting the recorded GPS points to the centre of the road. When this function is executed, the points obtained are correctly snapped to the centre of the road when there is no multipath effect. However, the points get misaligned when there the car goes under tunnels and is obstructed by flyovers (Figure 31), which leads to mistakes in calculations of the heading angle as the points will depict different sections of the road.

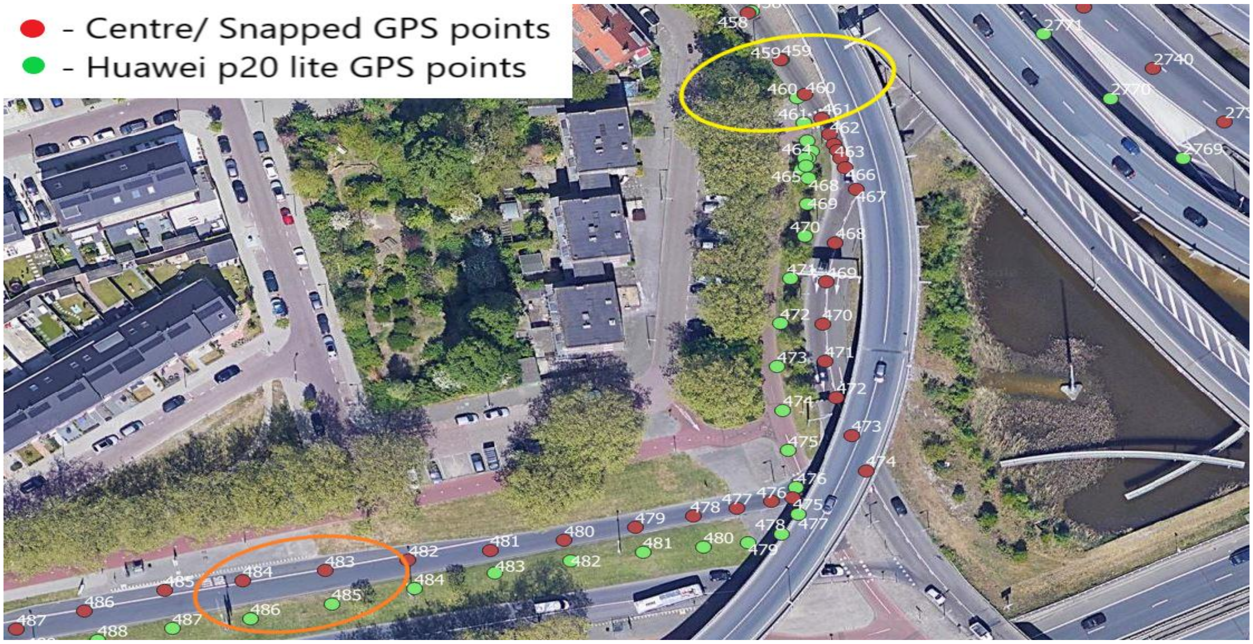

Figure 31: Misalignment of original and center of the road GPS points due to multipath effect

Prospects

- Camera sensors use high amount of smartphone battery capacity by recording videos to capture driving data [25]. Using GPS sensors leads to longer data capturing when power sources are limited within the vehicle.
- Although camera-based sensors are growing more common, they are also attracting criticism due to privacy issues. This is mirrored in the legislation of various countries restricting its usage. For example, restricted usage in the countries of Germany, Austria [26] [27] and strongly discouraged in Switzerland. Hence, using only GPS sensors is easier to gather data in all regions/countries.
- The recorded data can serve as information to road map data that could be used for road developments [28] and traffic management [29] considering sections of roads where there may be frequent lane changes.
- The data can show the significance of driver education and awareness in reducing risky lane changing behaviour and its positive impact on road safety [30]
- Cameras capturing road frames for data, at times are obstructed by changes in light or blockage by vehicles. There are scenarios of road signs obstructed due to graffiti on them. The above problem can be avoided by resorting to GPS or a combination of GPS and IMU to gather road data and driving manoeuvers [31].

Conclusions

From the experiment results, it can be concluded that smartphones can be used to identify lane changes to ultimately identify driver behaviour. However, it would be better if smartphones had improved quality of GPS receiver chips, which have higher accuracy than the current ones. Such improvement in quality would make smartphones highly beneficial to be considered a device for data collection to identify lane changes.

The signals for identifying lane changes can be further improved in filtering noise and providing cleaner data to identify lane changes. Well-processed data is much easier for the algorithm to detect actual lane changes. Different combinations of filters and parameters are yet to be explored, which is good future research in lane change identification.

An important concept noticed is that the GoPro GPS was recorded at a higher frequency of 18 Hz, whereas the smartphones and the standard GPS receiver GlobalSat records GPS values at 1 Hz. This higher frequency may be a possible reason for its higher accuracy, as was seen in the results.

Further research of the algorithm on different terrain such as mountains will help understand how the algorithm can be improved to identify the change of lanes by the car more accurately if it shows low performance.

The accuracy of GPS is low or close to nil in tunnels. As this cannot be avoided, future research on the combination of GPS and IMU may help solve the problem of losing information inside tunnels.

References

[1] K. Zinebi and N. Souissi, "Driver Behavior Analysis Methods: Applications oriented study," *The 3rd International Conference on Big Data, Cloud and Applications,* no. 1, pp. 1-15, 2018.

[2] E. Mantouka, E. Barmpounakis, E. Vlahogianni and J. Golias, "Smartphone sensing for understanding driving behavior: Current practice and challenges," *International Journal of Transportation Science and Technology,* 2020.

[3] T. Song, N. Capurso, X. Cheng, J. Yu, B. Chen and W. Zhao, "Enhancing GPS With Lane-Level Navigation to Facilitate Highway Driving," *IEEE Transactions on Vehicular Technology,* vol. 66, pp. 4579-4591, 2017.

[4] S. Yao, V. Knoop and B. Arem, "Optimizing Traffic Flow Efficiency by Controlling Lane Changes: Collective, Group, and User Optima," *Transportation Research Record: Journal of the Transportation Research Board,* vol. 2622, pp. 96-104, 2017.

[5] P. del Rio Barral, "In-car advisory system for lane-changing in a connected vehicle environment," *Aalto University. School of Engineering,* p. 53+7, 2018.

[6] D. I. Tselentis, G. Yannis and E. I. Vlahogianni, "Innovative Insurance Schemes: Pay as/how You Drive," *Transportation Research Procedia,* vol. 14, pp. 362-371, 2016.

[7] W. Chen, W. Wang, K. Wang, Z. Li, H. Li and S. Liu, "Lane departure warning systems and lane line detection methods based on image processing and semantic segmentation: A review," *Journal of Traffic and Transportation Engineering,* vol. 7, no. 6, pp. 748-774, 2020.

[8] G. Castignani, T. Derrmann, R. Frank and T. Engel, "Driver behavior profiling using smartphones: A low-cost platform for driver monitoring," *IEEE Intelligent Transportation Systems Magazine,* no. 1, pp. 91-102, 2015.

[9] K. Merry and P. Bettinger, "Smartphone GPS accuracy study in an urban environment," *PLoS ONE,* vol. 14, no. 7, pp. 1-19, 2019.

[10] M. Faizan, S. Hussain and M. I. Hayee, "Design and Development of In-Vehicle Lane Departure Warning System using Standard Global Positioning System Receiver," *Transportation Research Record,* vol. 2673, no. 8, pp. 648-656, 2019.

[11] Y. Sekimoto, Y. Matsubayashi, H. Yamada, R. Imai, T. Usui and H. Kanasugi, "Lightweight lane positioning of vehicles using a smartphone GPS by monitoring the distance from the center line," *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC,* pp. 1561-1565, 2012.

[12] S. E. Ramah, A. Bouhoute, K. Boubouh and I. Berrada, "One Step Further Towards Real-Time Driving Maneuver Recognition Using Phone Sensors," *IEEE Transactions on Intelligent Transportation Systems,* pp. 1-13, 2021.

[13] A. E. team, "Auto Express," Peugeot, 5 April 2019. [Online]. Available: https://www.automobiledimension.com/model/peugeot/108.

[14] K. UK, "KIA Picanto Specification," KIA, 2021. [Online]. Available: https://www.kia.com/uk/new-cars/picanto/specification/.

[15] M. Kirschenbaum, "Hypoxic Extreme Electronics," 29 October 2019. [Online]. Available: https://gethypoxic.com/blogs/technical/gopro-max-teardown.

[16] G. Arena, "GSM Arena," Kirin 710, July 2019. [Online]. Available: https://www.gsmarena.com/huawei_p20_lite_(2019)-9703.php.

[17] S. E. 9810, "Samsung," Exynos 9810, [Online]. Available: https://www.samsung.com/semiconductor/minisite/exynos/products/mobileprocessor/exynos-9-series-9810/.

[18] G. W. Group, "GlobalSat Worldcom Group," GlobalSat Worldcom Group, 18 November 2011. [Online]. Available: https://www.globalsat.com.tw/ftp/download/GMouse_Win_UsersGuide-V1.0.pdf.

[19] g. developers, "Google Maps Platform Documentation," Google, [Online]. Available: https://developers.google.com/maps/documentation. [Accessed 23 November 2021].

[20] Q. D. Team, "QGIS Geographic Information System," QGIS Association, 2021. [Online]. Available: https://www.qgis.org/. [Accessed 22 November 2021].

[21] G. Team, "GoPro," GoPro, [Online]. Available: https://gopro.com/en/us/shop/cameras/max/CHDHZ-202-master.html.

[22] S. Few, "Tapping the Power of Visual Perception," Perceptual edge, 4 September 2004. [Online]. Available: http://www.perceptualedge.com/articles/ie/visual_perception.pdf.

[23] J. Korstanje, "Towrads data science," 31 August 2021. [Online]. Available: https://towardsdatascience.com/the-f1-score-bec2bbc38aa6.

[24] C. Goutte and E. Gaussier, "A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation," *Lecture Notes in Computer Science,* vol. 3408, pp. 345-359, 2005.

[25] S. V. Rajaraman, M. Siekkinen and M. A. Hoque, "Energy consumption anatomy of live video streaming from a smartphone," *IEEE 25th Annual International Symposium on Personal, Indoor, and Mobile Radio Communication (PIMRC),* pp. 2013-2017, 2014.

[26] W. Paul, B. Pascal, K. Erik, B. Sebastian and B. Jürgen, "Privacy Dashcam – Towards Lawful Use of Dashcams Through Enforcement of External Anonymization," *Springer International Publishing,* vol. 10436, pp. 183--201, 2017.

[27] Veronika, "Dashcams – safety versus privacy? The Austrian perspective," Schonherr, 27 November 2020. [Online]. Available: https://www.schoenherr.eu/content/dashcams-safety-versus-privacy-the-austrian-perspective/.

[28] L. Zhang, C. Chen, J. Zhang, S. Fang, J. You and J. Guo, "Modeling Lane-Changing Behavior in Freeway Off-Ramp Areas from the Shanghai Naturalistic Driving Study," *Journal of Advanced Transportation,* vol. 2018, pp. 0968-090X, 2018.

[29] Rijkswaterstaat, "Roads and waterways," Rijkswaterstaat: Ministry of Infrastructure and water management, 2020 December 18. [Online]. Available: https://www.rijkswaterstaat.nl/en/mobility/roads-and-waterways.

[30] S. Mohamed, "Factors affecting lane change crashes," *IATSS Research,* vol. 44, no. 2, pp. 155-161, 2020.

[31] S. Gilroy, "Overcoming Occlusion in the Automotive Environment," *IEEE Transactions on Intelligent Transportation Systems,* 2019.

Appendices

Appendix A: Absolute lane position

An attempt to find the absolute lane position of the vehicle was made as an addition to the research on identifying lane changes to this project. As Open Street Maps (OSM) provides information on the number of lanes, the Overpass API allows extracting the number of lanes closest to the GPS points recorded during the drive. Each 'way' or 'edge' represents a road, and this will contain values such as lane numbers once accessed through the overpass API. When the lane numbers are obtained corresponding to the GPS recordings, the absolute lane position is calculated by the distance the point lies away from the centre of the road (lateral distance).

Results of absolute lane positioning

In the attempt to identify the absolute lane position of the vehicle using GPS data, requests were made through the Open Street Map API to obtain the lane numbers which were linked to the GPS points. The lane numbers were not present for all points and even after assuming 3 lanes (as it was manually checked for the highway on which it was driven) for A13 highway, the accuracy of the GPS points was low, which was not enough to identify the lane position on which it was traveling.

	osmid	highway	maxspeed	oneway	length	geometry	lanes	name	ref	bridge	tunnel
275	[308977074, 560370763, 7532526]	motorway	80	True	1183.441	LINESTRING (598471.7571407361 5754954.48173535...	3	Rijksweg A13	A13	yes	NaN
276	[308977074, 560370763, 7532526]	motorway	80	True	1183.441	LINESTRING (598471.7571407361 5754954.48173535...	3	Rijksweg A13	A13	yes	NaN

Figure 32: OSM API providing information on the number of lanes for A13 highway

	osmid	highway	maxspeed	oneway	length	geometry	lanes	name	ref	bridge	tunnel
30	[121834300, 131656652, 585710740, 509251734]	unclassified	50	False	555.779	LINESTRING (595071.3432733205 5760717.11048093...	NaN	Thijsseweg	NaN	NaN	NaN
31	[121834300, 131656652, 585710740, 509251734]	unclassified	50	False	555.779	LINESTRING (595071.3432733205 5760717.11048093...	NaN	Thijsseweg	NaN	NaN	NaN

Figure 33: OSM API not providing information on the number of lanes for different section of A13 highway

Figure 32 and Figure 33 show the results of OSM API on obtaining information on the number of lanes on the same highway of A13 at different sections. As the number of lanes is not available at every section of the highway, the possibility of estimating the absolute lane position using lateral distance of the vehicle is very low.

Appendix B : Snapping of GPS points of devices to the centre of the road

The GPS points from the devices need to be snapped or moved to the centre of the road to calculate heading angle or lateral distance. The procedure would first gather the list of GPS points representing the road on which data must be processed (for example, a stretch of A13 highway from delft towards Rotterdam). The points will then be loaded as a list of latitude and longitude values. These values will be moved or snapped to the centre of the nearest road using Google Roads API, which uses the function of ‘snap_to_road’, as shown in Figure 34, along with an API key. The function snap to the road can process 100 points at once. Due to this, we need to split our data into lists of 100 values each. Once the points are snapped, the new list of points (each list containing 100 GPS points) will then be merged into one JSON file for further calculations.

```
n=99

output_validation_2_huawei_a13=[map_client.snap_to_roads(validation_2_huawei_list1[i:i + n])
        for i in range(0, len(validation_2_huawei_list1), n)]
```

Figure 34: Using the function of snap_to_road from Google Roads API

Appendix C: Peak Identification code of the algorithm using python Jupyter notebook

Every signal after crossing a threshold value is identified as a peak marked with a red marker. The threshold values are chosen as explained in the methods section of the report.

```
In [1150]:
from scipy.signal import chirp, find_peaks, peak_widths
#from oct2py import octave
peaks2_globalsat_method1 = find_peaks(validation_2_a13_data_method1_globalsat_list_1, height = (0.5, None), distance=3)
peaks3_globalsat_method1 = find_peaks(validation_2_a13_data_method1_globalsat_list_2, height = (0.5, None), distance=3)

#plt.plot(check)
heights_2 = peaks2_globalsat_method1[1]['peak_heights']
heights_3 = peaks3_globalsat_method1[1]['peak_heights']
#plt.plot(peaks,check[peaks], "x")
peak_pos2 = x5[peaks2_globalsat_method1[0]]
peak_pos3 = x5[peaks3_globalsat_method1[0]]
#plt.plot(peak_pos,heights_1, "x")

fig = plt.figure(figsize=(20, 4))
ax = fig.subplots()
ax.plot(x5[0:270],validation_2_a13_data_method1_globalsat_list_1)
#ax.set_xlim([146:429])
ax.scatter(peak_pos2, heights_2, color = 'r', s = 15, marker = 'D')
ax.scatter(peak_pos3, -1*heights_3, color = 'r', s = 15, marker = 'D')

ax5 = ax1.twinx()
ax5.axes.yaxis.set_ticks([])
xs = [-3,3]
plt.vlines(x = [207,248,287,331,359,430], ymin = min(xs), ymax = max(xs),
           colors = 'red',
           label = 'left lane change')
plt.legend(loc = 'upper left')
#plt.vlines.legend

plt.vlines(x = [234,274,307,353,412,450], ymin = min(xs), ymax = max(xs),
           colors = 'green',
           label = 'right lane change')
plt.legend(loc = 'upper left')

Out[1150]: <matplotlib.legend.Legend at 0x2e7e49d72e0>
```

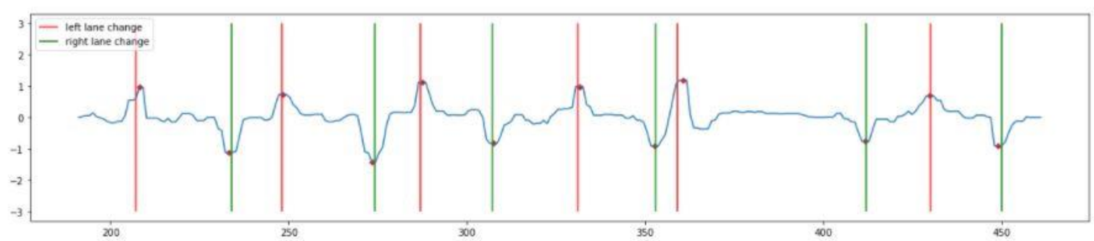

Figure 35: Peak identification algorithm