2017 Fall

ECE537/CIS568 Data Mining

Project Report

Project Title

Topic: On Road Dynamic scenario classification

Team Member

Kexun Chen

Takchoi Yu

2017-12-2

Topic: On Road Dynamic scenario Classification

1. Abstract: The growing volume of traffic requires higher level of traffic safety. The safety of driving can be significantly increased by using a system automatically interpret the traffic situation using the image sequence recorded with cameras mounted in a moving vehicle. A robust and reliable classification system can recognize the real traffic situation and remind the driver the potential danger and reduce the number of accidents. Our goal is to perform the vehicle on-road dynamic scenario classification based on the On-Board-Diagnostic (OBD) data, driver physiology data and video. Based on the data we get, in this project, we classify our on-road driving status into 5 basic categories, includes: lane keep, turn right, turn left, lane change left, lane change right. All these status should be clearly defined. Video data, OBD, physiology data are used to train our model thus we can get a model which can classify the dynamic scenario.

2. Problem Definition

In real-world situation, it can have many variations on the vehicle on-road scenario, such as lane change, lane keeping events. Stated in [i], shown as below Figure1, it describes several variations of lane changes and lane keepings.

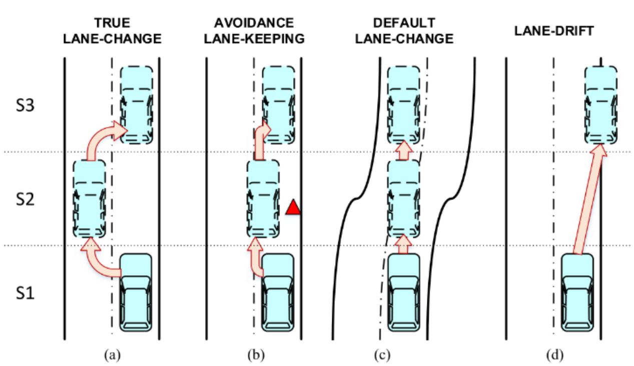


Figure1. Variations of Lane change and lane keeping.

In this project, the definition of true lane change scenario is adopted from the paper. That means, vehicle going from S1 to S2 is defined as left lane change, from S2 to S3 is defined right lane change. Turning shown is strictly defined as the scenario shown in the Figure2. Other scenario, examples, (b), (c), (d) and other scenarios are defined as GoStraight category.

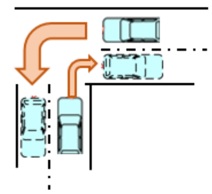


Figure2. Illustration of turning event

This problem can be categorized to time series event classification problem. By utilizing time sequence data within the event occurrence, analysis and extract feature patterns to classify the scenario.

3. Survey

Researchers had studied the problem and developed several approaches. In brief summary, they are physiological signal analysis, OBD data analysis, extended sensor signal analysis and computer vision approaches.

In [ii], authors proposed robust lane departure prediction based on physiological signal. The authors utilized different physiological signal – Electrocardiography (ECG), galvanic skin response (GSR) and respiration signal to derive other physiological signal features. 12 common statistical features are then computed on each derived signals and create 144 high dimension feature vectors. To reduce dimension, the paper has a study on the dimension reduction algorithm, with a detailed comparison on both joint mutual information (JMI) and principal components analysis (PCA) on the derived physiological features. Followed by training a support vector machine classifier.

In [iii], the paper suggests using Inertia motion Unit (IMU) in smartphone as an extended vehicle sensor, to convert the 3-dimension readings from gyroscope to vehicle steering direction. The data is then classified to different vehicle on-road event through analyzing the time series pattern of the gyroscope data.

In [iv], the authors combined different input signals (Yaw rate, Longitudinal acceleration, Lateral acceleration, Steering wheel angle, Wheel speed,…) for which the data and signals are generated from simulation device. Feature selection is performed on the augmented vehicle states and road surface situation data and are fed to SVM classifier to predict the driver's lane change intention.

In [v], authors purposed an image processing pipeline to detect lanes. By choosing specified color (yellow, white, blue) in YUV color channel, essential lane marks are segmented, followed by edge detection, pixels which present lane line are extracted. Hough transform holds the key process to transform the discontinuous pixels to continuous linear/quadratic line depending on the algorithm.

In this project, OBD data and physiological signal are being studied as features to classify vehicle on-road dynamic scenario. Here also study the effect of data normalization, problem of data imbalanced, performance from MLP neural network.

4. Data Study

Dataset being used is dedicated by Intelligent System Lab in University of Michigan Dearborn (ISL-UMD) for Driver Work Load project. Only a subset of the data from the project is being studied. The subset consists of 6 trips with same driver driving around UMD road way with a passenger as an expert to label concurrent certain on-road events, which include vehicle on-road dynamic scenario. Each trip is with around 40 minutes long, and they all consist of recorded OBD data, physiological signals and vehicle-front-view video recorded from a shimmer device, a bio-harness device and a Go-Pro camera respectively.

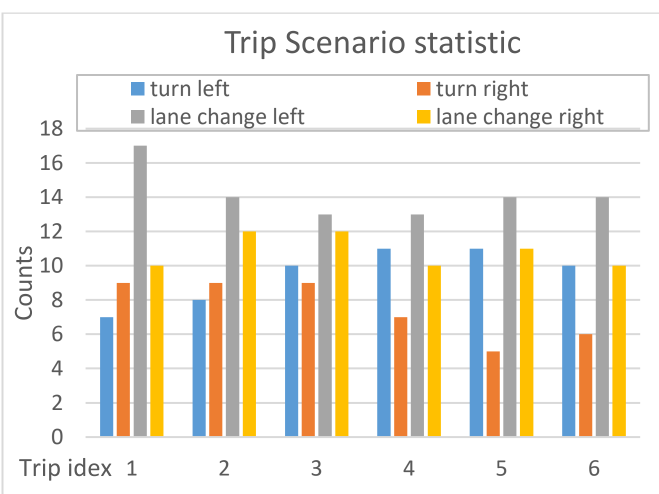


Figure3. Trip scenario statistic

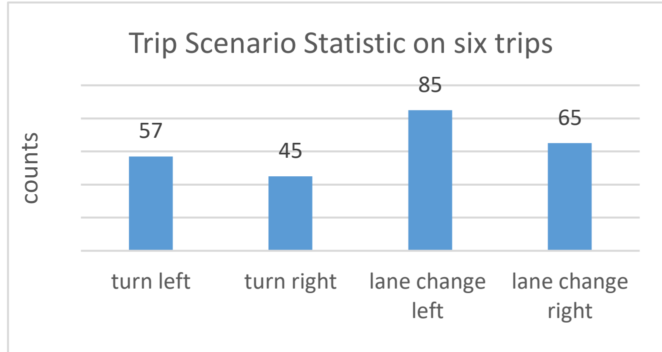


Figure4. Trip scenario statistic on six trips

Figure3 and 4 show the statistic of the data being studied. Noted that as the data is recorded for Work Load project, the timing of dynamic scenario event is actually not precise but only accurate to work load status. By watching the videos, it is observed that the event from original label often delay around 2 to 10 seconds relative to actual event occurrence.

Hence, it is necessary to relabel the event through watching the video to allow data-event matching.

Labeling software is a modified version origin from ISL-UMD for event labeling on video. Figure5 shows the graphical user interface of the modified labeling software developed in MATLAB.

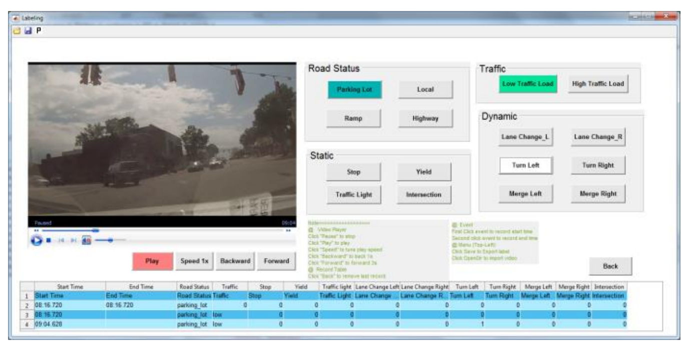


Figure5. Event labeling software GUI

This version allows fast-forward, backward moving on video time stamp. The added features allow fast and precise event-data matching with precision within 0 to 0.5 second.

Event definition is adopted from [i], event is started while the vehicle starts changing the state (direction) which also mentioned in Figure1 from Problem Definition, previous section in this report.

It is observed from the dataset that the period of events are in range 3 to 10 seconds and 5 seconds in average, thus in the experiment, event window length is defined as 5 seconds.

Mentioned in Problem Definition, all scenarios other than lane change and turn are defined as GoStraight category, it means there is much GoStraight class than the lane change and turning events. To deal with the data imbalanced, similar amount of GoStraight events to lane change/turning are randomly extracted without time overlapping.

Attributes	description	unit
Speed	Instant speed of vehicle	speed (mph)
GPS_longitude	GPS longitude	degree
GPS_latitude	GPS latitude	degree
GPS_heading	Heading	degree(0-360)
Long_accel	x-axis acceleration	Gram(g)
Lat_accel	y-axis acceleration	Gram(g)
Vector_accel	Combined xyz-vector acceleration	Gram(g)
Vert_accel	z-axis acceleration	Gram(g)
hr	Heart rate derived from ECG	Heart bit per min
scl	Skin conductance level derived from GSR	n/a

Table1. Attributes description

Data normalization is performed based on the properties of signals, properties of attributes are described in Table.1. Table.2 shows the summary for data normalization method.

Signal	Properties	Method
Speed(mph)	Bounded by road speed limit	Max-min normalization
GPS lat/long(degree)	Spatial relation	Scale, orientation normalization
Accelerometer data(g)	Small value range around 0	N/A
Heading(degree)	Bounded by 360 degree	Max-min normalization
Heart rate per min	Range around 60 to 120	Z-normalization
SCL (low freq components of GSR)	Range around 30 to 150	Z-normalization

Table2. Method of normalization

As there are spatial relationship between GPS Latitude and Longitude, to allow better data normalization, they are first converted to UTM coordinate with position and scale normalization. Afterwards, in order to have all segment invariant to its orientation, rotation is operated on the segment based on the angle of mean position to zero origin. Figure.6 illustrate the transformation of one left turn segment.

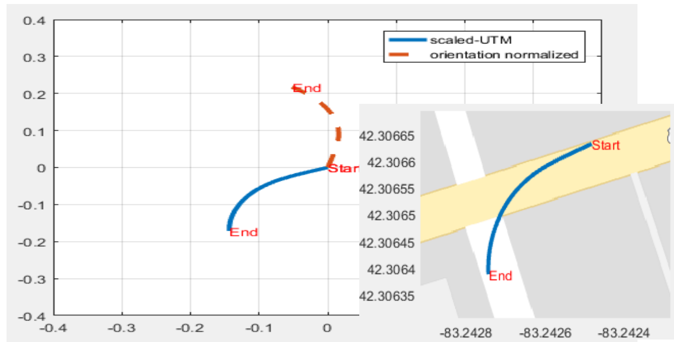


Figure6. Left: Scaled after UTM conversion (blue). Orientation normalized (red). Right: A left turn segment selected from one event and being plotted in Google map)

Feature Extraction is performed on the time series data in each event. Below shows the meaning of 12 statistic features.

* *Maximum value of a signal*
* *Minimum value of a signal*
* *Median value of a signal*
* *Standard derivation of a signal*
* *The interquartile range (IQR)*
* *Energy of a signal*
* *Zero-Crossing Rate (ZCR)*
* *Skewness of a signal*
* *Kurtosis of a signal*
* *Standard deviation of signal first derivative*
* *Root mean square of signal first derivative*
* *Difference between signal(t), signal(t+Δt)*

And it constructs 12 x 10 = 120 dimension of one feature vector. The statistic features are suggested from paper [ii], and being used to analysis lane change event.

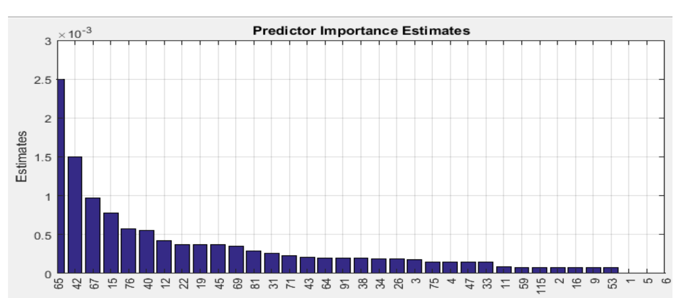


Figure7. Top 35 Features importance

Top 35 features are selected based on feature importance derived from decision tree. Figure7 shows the weighting of the features. And Table3 describes top 10 features after feature selection in descending order.

Feature	Feature description
65	Std of **Lat-accel**
42	IGR of **heading**
67	Median of **Lat-accel**
15	Min of **Y-coordinate**
76	Difference of start,end value of **vector_accel**
40	Difference of start,end value of **heading**
12	RMS of first derivative of **Speed**
22	Kurtosis of **Y-coordinate**
19	Median of **Y-coordinate**

Table3. Description of top 10 features

Data Visualization allows understanding of data distribution. We compare the distribution of the times series data between the data with and without feature extraction method. Using PCA dimension reduction method to project the data to 2 dimension presentation, it is observed that feature extraction method able to divides the data to its class more effectively.

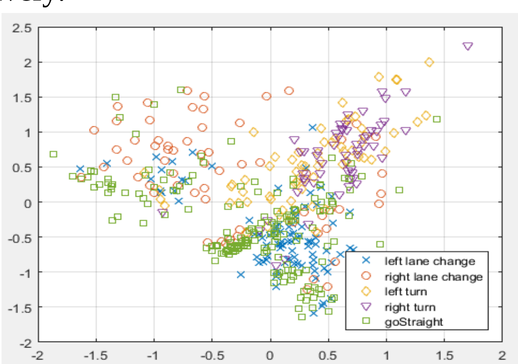


Figure8. Data visualization of data with feature extraction

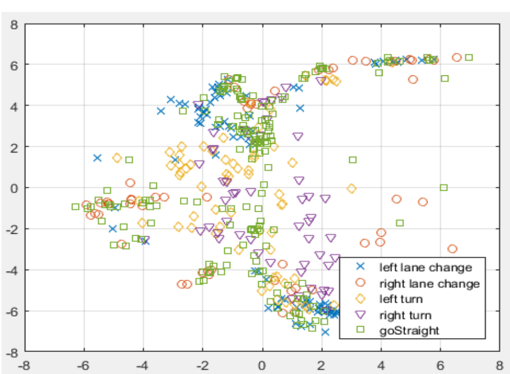


Figure9. Data visualization of data without feature extraction

6. Experiment

Crossing Validation is used to validate the neural network classification model. In this project, 10-Fold cross validation is performed to have a better credibility of experiment result for small dataset.

For each fold of cross validation, samples are also stratified to make sure even distribution of events from each class.

Heavily Overfitting without feature extraction is observed using neural network classifier. There are total 10 (signals) x 10 (Hz) x 5 (seconds), 500 dimension of feature in the dataset, eventually greater than the number of samplings in the dataset. While in the following experiment, only *X-coordinate*, *Y-coordinate*, *Lat-acceleration*, *Heading* attributes are selected as input (200 dimension). The selection is based on the contribution of feature importance of top 35 features described in previous section.

Neural network is used for classifier. While training the model, all weights are randomly initialized, to prevent overfitting, regularization term with rate 1.0 is used to error function to penalize the scale of weights in the network.

$$E_R(\theta) = E(\theta) + \quad \lambda\Omega(\omega)$$

$$\Omega(\omega) = \frac{1}{2}\omega^T\omega$$

Where $E(\theta)$ denotes error function, ω denotes the weight vector, λ denotes the regularization term. Levenberg-Marquardt algorithm is adopted for training in backpropagation. The criteria to stop training is I. Validation fails at 7 times, II. Number of epoch meets 100.

Comparison with/without segment orientation normalization. The experiment is being tested with/without feature extraction method. Shown in Figure10, with/without feature extraction method, generally using orientation normalization give better model accuracy.

Comparison with/without feature extraction is experimented and illustrated in Figure11, it illustrates the comparison of 10-fold validation results between with/without feature extraction method among different neural network architecture.

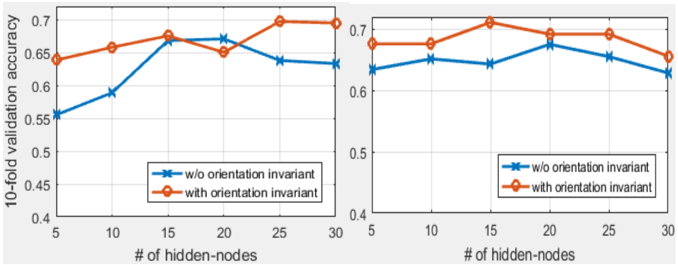


Figure10. Left: comparison without using feature extraction. Right: comparison using feature extraction

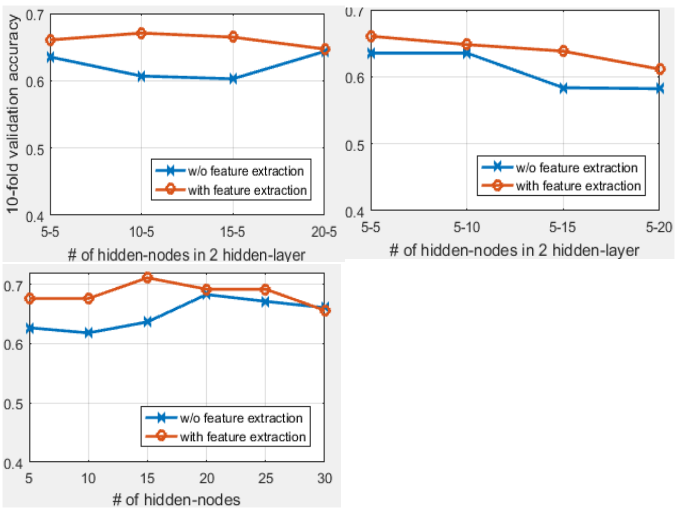


Figure11. Cross validation results of with/without feature extraction.

It is observed that the optimal accuracy 0.7112 locates at one hidden layer of 15 hidden nodes.

7. Conclusion and discussion

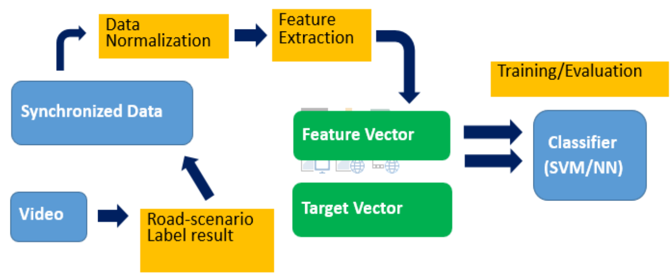
In this project, we developed a modified labeling software with added features in GUI to ease event labeling on video in UMD-ISL, and manually label the events with high time precision mapping between label from video and data from sensor. We tried different way to normalize the data dependent to signal properties, and develop orientation normalization method to independent trip segment. And experiment result supports segment orientation normalization able to increase model accuracy. We also implemented stratified cross validation method to model neural network classification model. Compare with paper[ii], we use decision tree to select top 35 important features instead of using PCA method to extract features. It is also observed that top features do not cover any physiological signal in feature extraction section.

During experiment, it is found that in PCA dimension reduction for data visualization, top one principal component consists of 90 percent explained variant among all components, which is weird. And it is found that one statistic feature (Energy of Signal) being extracted is a large scale value feature as its definition. So in the experiment section, mean normalization is done to all data after feature extraction.

It is still lots of work to have more data to increase credibility of the model accuracy. Increase data size can also prevent overfitting of larger size of neural network or using greater number of features without feature extraction. And it is worth to try other signal existed in the original dataset (i.e. other derivation of physiological signal, respiration rate so on).

Appendix.

Project Modules



8. Reference

[i] Zheng, Yang, and John HL Hansen. "Lane-Change Detection from Steering Signal using Spectral Segmentation and Learning-based Classification." IEEE Transactions on Intelligent Vehicles (2017).

[ii] Hong Zhao, Dev Kocchar, Yi Lu Murphey, and Paul Watta, "Robust prediction of lane departure based on driver physiological signals," SAE World Congress & Exhibition

[iii] Chen, Dongyao, Kyong-Tak Cho, Sihui Han, Zhizhuo Jin, and Kang G. Shin. "Invisible sensing of vehicle steering with smartphones." In *Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services*, pp. 1-13. ACM, 2015.

[iv] Kim, Il-Hwan, Jae-Hwan Bong, Jooyoung Park, and Shinsuk Park. "Prediction of Driver's Intention of Lane Change by Augmenting Sensor Information Using Machine Learning Techniques." *Sensors* 17, no. 6 (2017): 1350.

[v] Somasundaram, Gayathiri. "Lane Change Detection and Tracking for a Safe-Lane Approach in Real Time Vision Based Navigation Systems." (2011).