Continuous Car Driving Intent Detection Using Structural Pattern Recognition

Sukhan Lee, *Life Fellow, IEEE*, Muhammad Qasim Khan, and Mohd Nizam Husen

***Abstract*— The early detection of a driver's intention prior to the initiation of actual maneuvering is to offer effective means of assisting the driver in times of safety. Conventional approaches in predicting a driver's intention include framing the motion parameters of the car and driver as well as the eye gaze sequence of the driver into the Hidden Markov Model (HMM) for analysis. However, their power of describing driver behavior is not only limited but also their performance, indicated by the recognition rate relative to the amount of preceding time in early detection, is not satisfactory and needs further improvement. This paper presents an approach for early detection of a driver's intention by modeling and analyzing driver behavior by using structural pattern recognition based on context-free and context-sensitive grammars. We specifically structured a sequence of the driver's eye fixations as well as of the vehicle speed, steering angle and signaling into a sequence of symbolic vectors to form sentences representing a specific driver behavior. It turns out that the proposed approach resulted in an average of 70.5% and 80% recognition rates at the respective 2 and 1 second preceding time to the actual initiation of maneuvering behavior. This performance can be compared to 56.8% and 69% by the state-of-the-art conventional results.**

***Index Terms*— Advanced driver assistance system, context-free grammar, context-sensitive grammar, driver intent, eye tracking, intelligent transportation system, pattern recognition, structural pattern recognition, syntactic pattern recognition.**

I. Introduction

ADVANCED driver assistance systems (ADAS) are becoming necessary features of modern vehicles to further improve their safety from road accidents. The World Health Organization (WHO) reported that there are around 50 million road traffic accidents worldwide each year, which result in serious injuries, with 1.25 million fatalities among them [1], [2]. The report also mentioned that the most contributing factor to road accidents is human error. More than 90% of road accidents are due to human mistakes as one of the contributing factors, and 57% are solely caused by an error from the human side [3], [4]. The most fundamental factor in human error is known to be fading of visual and/or auditory attention. In particular, it is the fallible mental capacity of humans [3] that induces limitations not only in information processing but also in prolonged attention. Therefore, an interactive system that can warn the driver of fading attention prior to the initiation of actual maneuvering should be of great value. More specifically, most of the road accidents are caused by misbehavior of a driver either when the maneuver is not appropriately announced to the environment so that the intention is misunderstood or the intended maneuver is not well attuned to the traffic situation. Hence, the main problem that is focused in this work is early detection and prediction of the driver's intention before the initiation of actual maneuvering. Successful detection of the driver's intention is key to the effective prevention of road accidents, leading to the development of new generation automotive safety systems. Here, every millisecond counts for saving from serious accidents in time-critical situations.

Conventional approaches to the prediction of driver's maneuvering behavior have been predominantly based on the Hidden Markov Model (HMM) and its variants [5]–[7], applied to the history of vehicle motion, the driver's head position [8] and/or the eye gaze sequence of the driver [9]. However, we found that their performance indicated by the recognition rate at the amount of preceding time for early detection is not satisfactory and needs further improvement.

This paper presents an early intent prediction of driving maneuver by using the syntactic or structural pattern recognition approach. Syntactic pattern recognition is a methodology that expresses each pattern in terms of a structural configuration of its components. The recognition of a pattern is realized by analyzing the pattern structure according to a given set of grammar or rules. The proposed work takes four types of concurrent data values as a vector for the generation of grammar to produce a comprehensive maneuver behavior by detecting lane change *and* turn as opposed to the previous studies, which mainly focused on a single maneuver behavior. The outcome of this study is intended to serve as an ADAS by predicting driver's maneuvering intention based on driver's eye fixations and car's motion data, particularly in predicting lane change

Manuscript received January 20, 2019; revised June 28, 2019, August 27, 2019, and November 18, 2019; accepted December 16, 2019. This work was supported in part by the 3D Recognition Project of Korea Evaluation Institute of Industrial Technology (KEIT) under Grant 10060160, in part by the Robocarechair: A Smart Transformable Robot for Multi-Functional Assistive Personal Care Project of KEIT, sponsored by the Korean Ministry of Trade, Industry and Energy (MOTIE), under Grant P0006886, and in part by the Institute of Information and Communication Technology Planning and Evaluation (IITP) Grant sponsored by the Korean Ministry of Science and Information Technology (MSIT), AI Graduate School Program, under Grant 2019-0-00421. The Associate Editor for this article was X. Ma. *(Corresponding author: Sukhan Lee.)*

Sukhan Lee and Muhammad Qasim Khan are with the Department of Electrical and Computer Engineering, Intelligent Systems Research Institute, Sungkyunkwan University (SKKU), Suwon 440-746, South Korea (e-mail: lsh1@skku.edu; qasim@skku.edu).

Mohd Nizam Husen is with the Universiti Kuala Lumpur Malaysian Institute of Information Technology (UniKL MIIT), Kuala Lumpur 50250, Malaysia (e-mail: mnizam@unikl.edu.my).

Digital Object Identifier 10.1109/TITS.2019.2961928

and turns intent, early enough to prevent road accidents. This would eventually lead to improving the safety of the driver, passengers and surrounding vehicles and pedestrians.

The organization of this paper is as follows: Section 2 reviews the related work in the relationship between driver behavior and traffic accidents, the recognition of driver behavior and intention as well as the fundamentals of structural pattern recognition. In Section 3, the problem definition and contribution of the proposed work is provided. Section 4 continues the explanation of our proposed approach with the process of experimental investigations. The experimental results are discussed in Section 5. Section 6 concludes the paper and points to potential future work.

II. Related Work

A. Relationship Between Driver Behavior and Traffic Accidents

Studies have shown that inept turning at road intersections and poor lane-change maneuvering are among the major causes of road crashes. According to statistics, over 30% of all crashes happen when a vehicle engages in a maneuver encroaching into the path of another vehicle due to turning or lane change [10]–[13]. The accidents caused by turning, in particular left turning and lane change, are often catastrophic in terms of injuries and fatalities as the other vehicle involved in the accident is caught off-guard, let alone posing a serious danger to pedestrians, motorcyclists and bicyclists [10], [11]. Understanding how these accidents occur could help design intelligent vehicles and transportation systems that can bring the number of crashes down. A number of factors contribute to drivers making an unsafe disposition in turning and lane change maneuvers, including driver fatigue, distraction from multitasking, operating under the influence of alcohol or drugs, as well as environmental factors, such as weather conditions [10]. More often than not, the drivers involved in accidents fail to take proper precautions to accurately assess whether or not it is safe to start a maneuver. Driver distraction from multitasking critically hinders taking proper precautions as drivers falsely believe that they can both drive and talk, use a cell phone, mess with electronics or handle pets and children [10], [12]. In this paper, we propose to analyze the structural patterns embedded in the behavior of drivers engaged in turning and lane change that help detect the failure of taking proper precautions even before the intended maneuvers.

B. Recognition of Driver Behavior and Intention

The most common implementation approach adopted by earlier researchers in the field of driving behavioral pattern recognition is based on HMM and its variants [14]–[17]. They adopted machine learning and feature-based pattern recognition techniques [18]–[20]. Moreover, the main focus of previous researchers was usually concentrated on a single maneuver behavior, such as lane change only or turn only [16], [18], [21], [22]. For example, the authors of [5] used HMM-based steering behavior models to distinguish between lane changing and keeping. They also differentiated between emergency and normal lane changes. The authors of [6] made use of the easily accessible environment and vehicle signals in their work based on HMM for continuous driver intention detection. Similarly, a real-time lane change intent prediction approach based on statistical recognition of previous offline classifiers was developed in [8]. The authors of [7] employed coupled HMMs in their approach. Other studies employed the use of the relevance vector machine to predict driver intentions to change lanes [23], brake [24] and turn [25]. Moreover, the authors of [9] made an effort to infer the intentions of the driver by gaze behavior data. They used artificial neural network models in learning data in order to predict the driving maneuvers.

In addition to HMM and gaze detection, driver's intention detection has been studied by using other methods, such as neural networks and features like safety margins and road maps (e.g. [26]–[30]). The authors of [28] developed a combined framework by using model-based and learning-based techniques for intention estimation and motion prediction, respectively. With the help of two types of cost functions (i.e., environment-based and interaction-based), this study has benefits in terms of native extensibility and a decrease in the computation burden. The work in [29] aimed at preventive safety by estimating the longitudinal jerk of the vehicle acted by a human driver using Kalman filter theory to infer the driver intention. In [30], the early detection of a driver's intention was realized by using deep learning approach, where RNNs and LSTM units were integrated to fuse the inside and outside vehicle features for predicting a maneuver mode from a partial temporal context. The inside and outside vehicle features used in the paper were, respectively, the motion of facial points captured by a face camera and driving context on the street map defined based on the street scene, GPS and vehicle speed. The results of [30] reported the respective 84.5% and 90.5% of the precision as well as 77.1% and 87.4% of recall performance when using the out of the box and customized optimal face tracker, together with 3.5 seconds of advanced prediction of maneuvers. Although the outside scene context on the street map used as an additional feature was expected to bring forth better performance at normal driving situations, such outside scene context features, apart from the features involved in direct vehicle control by the driver, may incur an adverse effect on predicting the abnormal behavior of the driver that may lead to a possible accident. Deep recurrent networks, such as LSTM, may offer a powerful tool for representing and identifying the temporal maneuvering patterns of a driver. However, they are limited in terms of interpretability or explainability, an important feature for identifying a driver's intent as it facilitates reviewing of a decision, especially, when abnormal behaviors are detected. Note that the time measured for advanced prediction of maneuvers may be correlated with the speed of a vehicle. For instance, one tends to make maneuvering decisions earlier as the vehicle speed becomes higher. Therefore, the prediction time is better to be evaluated relative to the vehicle speed used in experiments.

The above-mentioned studies related to driver intention recognition mainly employ statistical feature-based pattern recognition approaches. On the other hand, in this work, we develop a linguistic-based syntactic pattern recognition approach. Earlier applications of the syntactic or structural pattern recognition approach include identification of handwritten numerals [31], recognition of speech patterns [32], Chinese character recognition [33], chromosome classification and identification of bubble and spark chamber events [34], and ECG recognition [35]. It is notable that the use of the syntactic approach for driving behavior intent detection has never been proposed in literature. We considered the spatial and temporal context of the features associated with the driver, vehicle and driving environment important for car driving intent detection. The linguistic-based syntactic pattern recognition offers more descriptive power than conventional approaches with the capability of structural or contextual representation of features, such as a graph representation and of forming various complexity of patterns based on diverse grammatical structures, including context-free grammar (CFG) and context-sensitive grammar (CSG).

C. Fundamentals of Structural Pattern Recognition

In the structural pattern recognition, each pattern is represented by a structural composition of its components. Then, a pattern is recognized by analyzing and describing its structure based on a particular set of rules. In syntactic pattern recognition, this is accomplished by defining distinct and appropriate grammars that represent the structure of each pattern class syntactically [34]. A syntactic pattern recognition system generally consists of three main parts, namely, preprocessing, pattern representation and syntax analysis [34], [36]. Note that pattern recognition problems could not be solved solely by statistical decision-theoretic approaches since the relational or structural information embedded in certain patterns, which is critical for pattern recognition, cannot be quantified in the feature vector space for classification. Typical examples of pattern recognition problems where structural information play a significant role are scene analysis and picture recognition. In syntactic pattern recognition, the recognition process includes not only assigning a pattern to a particular class but also describing the structural information embedded in a pattern. Syntactic pattern recognition starts with defining words as basic building blocks of the patterns, based on which complex sentences representing patterns can be constructed and described by the governing grammatical rules. Classification is accomplished by parsing a sentence to identify the grammar of a particular class that can generate the sentence. Figure 1 illustrates the classification process in syntactic pattern recognition.

As detailed in [37], a grammar comprises of the four entities as follows;

$$G = (V_N, V_T, P, S) \tag{1}$$

1) Nonterminals: This set is denoted as V_N (or N) and contains *nonterminal symbols* or *variables,* which are used as intermediate quantities in the generation of an outcome. The outcome solely consists of terminal symbols.

2) Terminals: This is a set of *primitive symbols* or *terminal*, denoted V_T(alternatively $\sum$). In several applications,

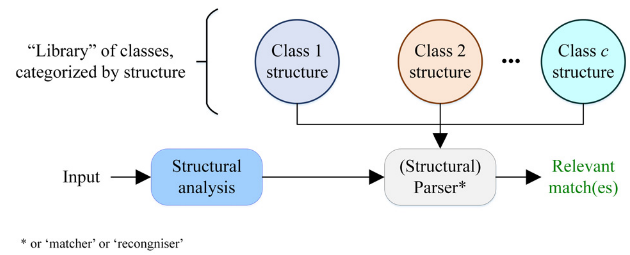


Fig. 1. Process of classification in syntactic pattern recognition.

the choice of the primitives or terminal set was difficult. In such cases, this process has a large component of 'art' rather than 'science'.

3) Production Rules: The set of productions is represented by P. This is a set of rewriting rules that allow previous substitutions. It is the set of productions, coupled with terminal symbols.

4) Starting Symbols: A root or starting symbol is denoted by S, where $S \epsilon V_N$. Note that V_N and V_T are disjoint sets, i.e., $V_N \cap V_T = \emptyset$.

Due to the availability of structural information in the driving pattern, we chose the syntactic pattern recognition approach for the car driving behavior pattern classification. The acts of driving behavior, such as *lane change right* can be expressed by a string of symbols or a sentence that represents a sequence of structural information associated with driver eye fixations, vehicle speed, steering angle and signaling. The sentences collected from various acts of driving behavior are represented by the context-free grammar (CFG) and context-sensitive grammar (CSG), based on which a parser is developed.

III. Problem Definition and Proposed Approach

As stated, the main problem defined in this paper is early detection and prediction of the driver's intention prior to the initiation of actual maneuvering. Unlike conventional approaches, we put our emphasis on identifying the true driver's intent regardless of whether the driver is engaged in normal or abnormal behaviors. Furthermore, we consider it important to have a prediction system that allows the interpretation or explanation of a decision, especially to be able to further analyze abnormal behaviors.

The framework of structural pattern recognition that we adopted has an advantage in interpreting or explaining the decision directly through parsing, which can be compared with the approach of deep recurrent networks that are limited in interpretability or explainability. We took into consideration the spatial and temporal context of the features associated with the driver, vehicle and driving environment important for car driving intent detection. However, we put our emphasis more on predicting the driver's true intention regardless whether or not it represents normal or abnormal behaviors, such that we consider only the features that are directly associated with the behavior of a driver engaged in vehicle control without resorting to environmental features that may incur an adverse effect on predicting the true driver's intent under abnormal behaviors. In general, our goal is to predict the true maneuvering intention of a driver in advance regardless of whether the driver's intent may lead to abnormal or accidental behaviors. If the intended maneuvering behavior is to lead to an accident, we would like to prevent such an accident from happening based on the proposed approach. The proposed structural pattern recognition approach offers a means of better interpreting whether the identified driver's intent and behavior lead to an accident or not through parsing induced pattern interpretation. Note that, in contrast to the conventional representation of grammars based on a non-vector form of symbols, our approach utilizes a vector form of symbols in order to accommodate four types of concurrent data entities present at a particular eye fixation for the generation of grammars. The four types of concurrent data entities are visual attention, vehicle speed, steering wheel angle and indicator signals. Specifically, a descriptor of driving context is constructed by concatenating the information from the sensors into the following 4-element feature vector:

$$v = [\alpha \ \beta \ \varphi \ \omega]^T \tag{2}$$

Eye fixation (α) is a fixation of the driver's visual attention. At each α, the eye fixation (α), vehicle speed (β), steering wheel angle (φ) and signal indication (ω) are combined to define a vector symbol. Then, a sentence is formed by a set of vectors collected in time. It is notable that the eye fixation, which indicates the visual attention of the driver, represents the key information among the four vector elements. The *eye tracker* cameras detect eye gaze and fixation. Whenever the eye fixation is kept as fixed at a particular location for a minimum time of 150 ms, the system considers it as a *fixation*. The selection of 150 ms as the threshold value is based on the result of our initial investigation. We observed that the fixation time less than 150 ms would be too short to ensure an actual attention, whereas the fixation time greater than 150 ms would result in a loss of useful eye fixation data.

IV. Driving Intent Detection Using Structural Pattern Recognition

The syntactic recognition of a driver's behavior is realized based on our proposed approach, as depicted in figure 2. The approach has two phases: *Learning* and *Recognition*.

The learning phase covers the steps when the system learns the training data of the driver's behaviors and comes out with the grammar for each behavior pattern. The grammars will then be utilized by the parser during the recognition phase.

The recognition stage is executed when a driver is in driving mode, and his/her eye fixation together with the car's behaviors is collected as testing data in sentence form. Then, the sentence is used as an input to a parser. The parser performs syntax analysis based on the grammar formulated during the training phase. The outcome of this phase is the classification result, where the driving behavior pattern is recognized as a specific maneuver behavior. The detection part is executed continuously before a maneuver is initiated. In this way, early detection can be realized by detecting the intent of maneuver behavior.

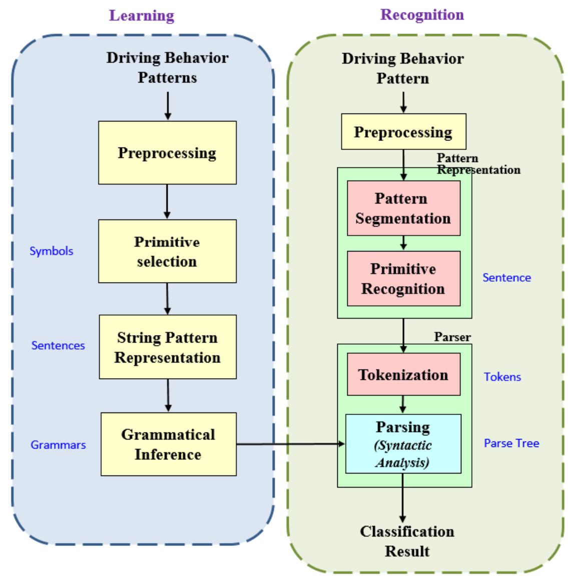


Fig. 2. The proposed car driving behavior pattern prediction using syntactic recognition approach.

The implementation of the proposed approach was conducted in the experimental study, as explained later.

A. Learning Phase

This section explains the processes implemented in the *learning* phase. The learning process of the training dataset is based on four driving behaviors:

- Lane Change Right
- Lane Change Left
- Turn Right
- Turn Left

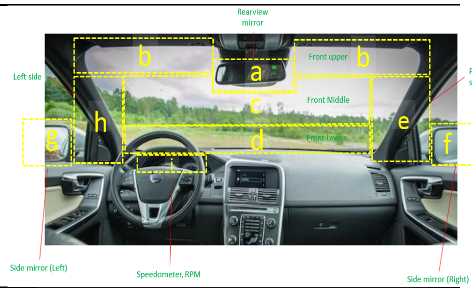
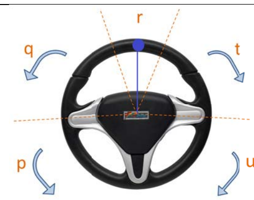
This phase involves three processes, which are primitive selection, representation of string pattern and grammatical inference.

1) Primitive Selection: The simplest sub-patterns, also known as *symbols* or *pattern primitives*, were selected for each specific driving behavior. Since symbols are the output of this process, their selection requires a thorough analysis of the training dataset.

For the *visual attention* data, the eye fixation regions on the scene are the selected areas and represented by symbols, and it is called Region of Interests (ROIs). For the *steering wheel angle*, *speed* and *turn signal indicator* data, their symbols are selected by analyzing the data patterns. Table I provides a description of the symbolic forms from the syntactic data together with the ROIs discovered from the driving scene. Based on (1), we develop V_T={a, b, c, d, e, f, g, h, i, k, l, m, n, p, q, r, t, u, x, y, z }.

As illustrated in table I, the symbols a through i represent the eye fixation of the driver towards the windscreen and surrounding eyesight, e.g., rear-view, right side- and left side-mirrors, etc. Meanwhile, the symbols k, l, m, and n represent the speed of the car with respect to the previous fixation, whether it is respectively the initial fixation, decreasing, unchanged or increasing in speed. The steering

TABLE I
DESCRIPTION OF DATA REPRESENTATION IN SYMBOLIC FORM

Data	Symbolic Form
Visual attention / Eye fixation (Region of Interest)	
Speed	k = initial/start symbol l = decrease (compare to previous) m = unchanged/static (compare to previous) n = increase (compare to previous)
Steering wheel angle	
Signal indicator	

wheel angle data is represented by symbol p, q, r, t, and u. When the steering wheel angle is 90° or greater to the left (−90° to −180°), the symbolic form is p. If the steering wheel angle is in between 30° and 90° to the left (−30° to −90°), the q symbol was used. If it is in between 30° to the left and 30° to the right (−30° to 30°), the r symbol is used. If it is in between 30° and 90° to the right (30° to 90°), the t symbol is used. When the angle is 90° or greater to the right (90° to 180°), the symbolic form is u. Intuitively, the steering position is r most of the time, while other angles represent higher movement, such as at sharp turns. Finally, for the signal indicator data, symbols x, y and z are set respectively when the left signal indicator is switched on, no signal indicator is switched on, and the right signal indicator is switched on as illustrated in table I. The steering wheel angle movement during normal driving may be small, perhaps, with less than ±30 degrees in "r" state. While, a large angle movement of the steering wheel, beyond "r", may be observed at a sharp turn, during parking and passing cars under traffic congestions. Note that the number of states defined for steering wheel angles can be extended beyond five such that the state "r" can be further elaborated for representing driving on straight and curved lanes as well as changing lanes. Notice, however, that the steering angle alone cannot immediately provide a sufficient power of discriminating different modes of driving. What is important for the discrimination of various modes of driving is the combination of "r" with other states, such as "a" …"i" representing the eye fixations of the driver on the

```
ccfcee  ←— Visual attention
kmnmll  ←— Speed
rrrttu  ←— Steering angle
yyzzzz  ←— Signal indicator
```

Fig. 3. A *sentence* in symbolic form consisting of a sequence of vectors.

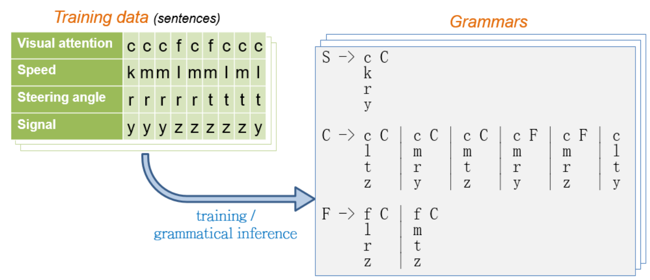


Fig. 4. Inference of grammar from sentences. Training datasets of the known behaviors are learned and grammars are formulated.

particular locations of the windscreen and surrounding mirrors, respectively, as well as "k", "l", "m" and "n" representing the patterns of speed variation of a vehicle with respect to the previous state of speed, as defined above.

2) String Pattern Representation: The output of this process is a *sentence*. Sentence is a structural formation of the symbols for each driving behavior. It is also called as a *string of symbols* because it is a concatenation of symbols. Sentence represents a pattern. So, this process involves analyzing the training dataset and finding the pattern of each driving behavior in a string or sentence format. An example of a sentence is shown in Figure 3.

3) Grammatical Inference: An automated grammar generation algorithm was developed to infer grammars based on *sentences* or *strings of symbols* collected from the training dataset. The training data are learned from their known class patterns to formulate grammars. Figure 4 depicts an example of the grammar inference process. After the grammar of each class/behavior is formulated, the learning phase is completed. The grammar is ready to be tested in the recognition phase, as explained in the next subsection, which explains the processes implemented in the *recognition* phase.

B. Recognition Phase

In the recognition phase, an unknown driving behavior pattern is given as input to the system. The input is randomly selected from the test dataset with known ground truth. The system classifies it as one of the driving behaviors based on the comparison to the learned grammars.

1) Pattern Representation: An unknown driving behavior pattern is segmented into primitives or symbols according to the Region of Interests (ROIs). Then, primitive recognition is carried out to recognize the symbols used in the structural pattern. The goal of this process is producing a sentence from the structural information.

2) Parsing: The sentence extracted from the pattern representation process is taken as input in the tokenization process.

Fig. 5. Experimental setup for eye tracking of the car driver.

Tokenization is implemented by breaking the sentence into chunks of symbols, which are meaningful elements called as tokens. Each token is in vector form because it contains four different types of data. Then, parsing is performed by the automated parser that we developed. This process is referred to as syntax analysis. The tokens become the input of the parsing process and the output is the classification decision. Parsing is the process of comparing the structural pattern of the tokens to the production rules of the grammars. If a sentence complies with all rules of a particular grammar, it means the sentence belongs to the grammar, which represents a particular class [38], [39]. For example, if an unknown driving behavior pattern (sentence) comply with all production rules of turn right grammar, it means that the behavior is classified as turn right.

C. Experimental Study

1) Experimental Setup: To begin the experimentation, we set up the equipment of the *Eye Tracker* and *Scene*. The Eye Tracker equipment consists of three cameras and two infra-red (IR) flashes attached on top of a vehicle dashboard. The eye tracker apparatus is used to track the motion of participants' pupils of both eyes at the rate of 60Hz, while the movements of both eyes are also sampled at the rate of 25Hz based on the eye movement equipment. The tracked pupils, together with the eye movements, determined the eye gaze directions. The scene equipment consists of a scene camera and screen displaying the driving simulator. All the cameras and IRs are connected to a laptop that runs the eye tracker. Figure 5 illustrates the experimental simulation setup. The tool used for eye tracking is Smart Eye Pro, while MAPPS is the tool used to perform analysis and preprocessing of driving data.

2) Eye Tracking: Eye detection and tracking are required before data collection takes place. Smart Eye Pro is used as a tool for eye tracking [40], [41]. A calibration process has to be carried out to make the tracking of the eye accurate before any data collection on the driving session is implemented. Specifically, camera, world coordinate system (WCS), gaze and scene calibrations (using 9p calibration) have been carried out. Infrared light is used to create a reflection in the eye's cornea that can be captured by cameras. To get a more accurate and robust tracking performance, our solutions track each eye separately. To this end, the iris and pupil of each eye were identified [41] for two separate monocular feeds, from which a consensus gaze is extracted. Figure 6 shows the successful eye tracking based on camera, WCS and gaze calibration. It is notable that the driver's visual range is narrower in such setups as compared to real-life applications. The proposed approach, however, is also suitable for real-life applications with minor adjustments as the pattern of eye movements for a specific action (e.g. turn right) essentially remains the same whether the driver is driving in a real environment or using a driving simulator.

Fig. 6. Eye tracking module tracks the movement of the eye pupils. Red line shows the eye fixation direction, blue ovals shows the detected eyes, and green line shows the angle of the eyes.

3) Data Collection: In order to train and test the maneuver intent linguistic-based grammar as a classifier, naturalistic driving data was collected by using City Car Driving simulator. A total of 49 subjects with varying nationalities, ages ranging from their 20s to 40s (mean = 31.7, standard deviation = 3.9), and had various amounts of driving experience, ranging from 3 years to 18 years, were involved in the data collection stage. For each participant, the experiments are conducted in a repetition of three to five times, which are enough to remove out outliers.

a) Training data: A training dataset, D, containing almost 500 minutes of driving data has been collected. A total of 500 car driving behavior data with an equal number from each of the four different behaviors (i.e., right turn, left turn, lane change right, and lane change left) has been collected for learning purposes. The drivers drove a specifically given route of the journey where the journey includes changing lanes rightwards and leftwards, right turn and left turn. The subjects were briefed to drive in their completely natural driving behaviors. This dataset will then be used to be learned to infer a grammar that is capable of generating a set of patterns.

Only training samples with "good" lane change and turn behaviors were chosen in training dataset. This training dataset was split into two separate subsets for cross-validation training. The set D_{train} with 400 samples (100 samples for each behavior) were used for grammar inferring. The rest 100 samples (25 samples for each behavior), D_{test}, were used for assessing the performance during learning.

b) Testing data: An independent testing dataset, T, were collected to test the grammars. A total of 595 data extracted from approximately 540 minutes of completely uncontrolled driving sessions has been collected to test the grammars in order to check the efficiency of the recognition system. Table II shows the breakdown of the testing data.

TABLE II
DESCRIPTION OF TESTING DATASET

Driving Behavior	Number of Data
Lane Change Right	252
Lane Change Left	168
Turn Right	98
Turn Left	77
Total	595

Fig. 7. Data collection with eye tracking where eye tracking denoted by a blue circle (left) and data analysis and preprocessing of eye tracking (right).

Each test drivers were given 10 minutes to drive anywhere they wished in the city. From the numbers of testing data breakdown, it was observed that the drivers made more attempt to make a rightward lane change and very less on making a left turn during uncontrolled driving sessions.

c) Data analysis: MAPPS, an eye tracker analysis tool was used to set the region of interest (ROI) on the scene camera and transforming the output of tracked eye movement as two-dimensional data (estimated gaze direction vector position based on fixed tracking camera) into two-dimensional intersections with gaze direction and scene. From this process, coordinate data are changed into symbolic ROI data. Figure 7 shows the snapshot of data collection and data analysis process.

The data frame of the Smart Eye Pro, which is used to record the driving session, was 60Hz, which means 60 data per second. Considering the average spending time of lane change is over 7 seconds [5], the data quantity is very large. However, MAPPS can filter out inaccurate data and only select the data that have estimated intersection with the scene. This process leads to brief and clear datasets for the construction of grammars.

D. An Illustrative Example of Recognition Procedure

This subsection elaborates in detail the procedure in the recognition phase of the proposed car driving syntactic recognition.

1) Driving Behavior Pattern Acquisition: This step is performed by collecting raw data from a driving session where the driver performed a particular behavior pattern. Figure 8 shows the acquisition of driving behavior pattern where the ground truth is a right turn.

2) Pattern Representation: The processes involved in this step are pattern segmentation and primitive recognition. A raw

Fig. 8. Driving behavior pattern acquisition from a driving session. The unknown driving behavior is taken as an input to the system.

```
ccfcee
kmnmll
rrrttu
yyzzzz
```

Fig. 9. Example of a sentence that contains four types of information: visual attention of the driver, speed of the car, steering wheel angle of the car, and the signal indicator of the car.

driving data is taken as input, and the output is a sentence in symbolic form. Pattern segmentation is the process of defining the region of interests/symbols. The primitive recognition recognized the symbols used in the structural pattern and produced a sentence based on the structural information. The sample of a sentence is shown in figure 9.

3) Parsing: This process was performed by our automated parser. The parser utilized the sentence formed by the previous step and took it as input. Tokenization process is implemented, where the sentence is breaking into chunks of symbols that are meaningful elements called tokens. The list of tokens becomes input for the parsing process. The parsing process is the process of comparing the structural pattern of the tokens to the production rules of the grammars. Figure 10 elaborates the parsing processes that include tokenization, syntax analysis, and decision making; the snapshot of the parser output is also shown in the figure.

4) Pattern Classification: The structural pattern of a sequence of symbols (sentence) that comply with all production rules of a particular pattern grammar means the sentence is part of the language of that grammar. Figure 11 shows the summary of the recognition process where notification is presented as the output at the end of the syntactic recognition system. Apart from a notification, the output of this system may also be sent to another system in the car, e.g. brake system, in order for that system to make any necessary precautions for the safety of the driver and passengers.

The length of a character sequence for the input to individual driving modes may not be fixed but varied for successful parsing. Here, the fundamental unit of window size for a character sequence is chosen to be 6. This is because we found that for most of the cases, parsing is successful within the window. However, when more units are necessary for successful parsing, additional units are to be recruited. Note that the sampling interval for generating a character is 150ms due to eye fixation.

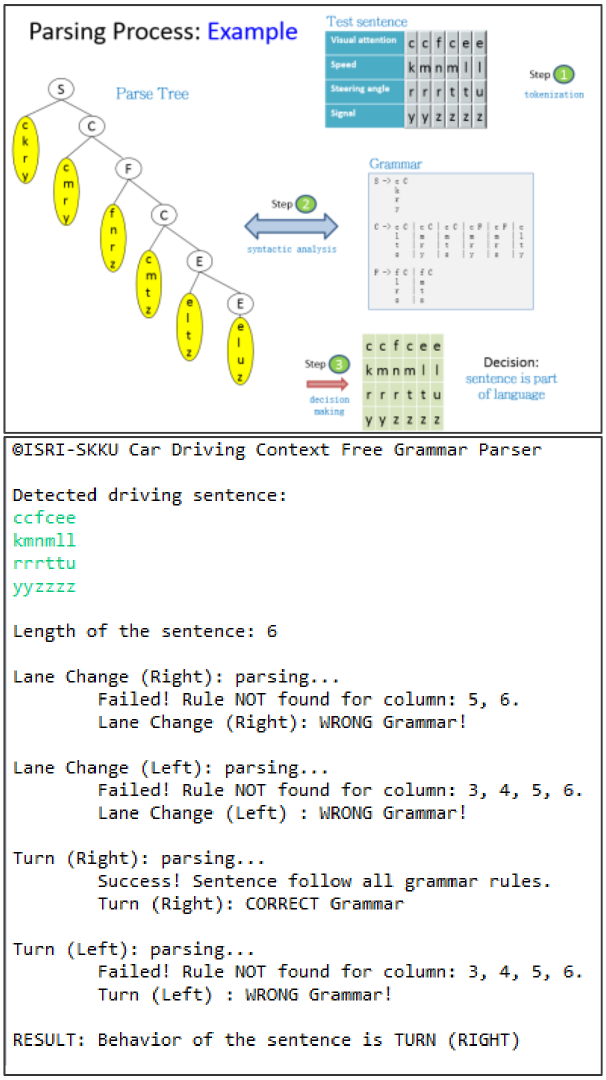


Fig. 10. Parsing process explanation with parsing tree (top) and example of output of parser program (bottom).

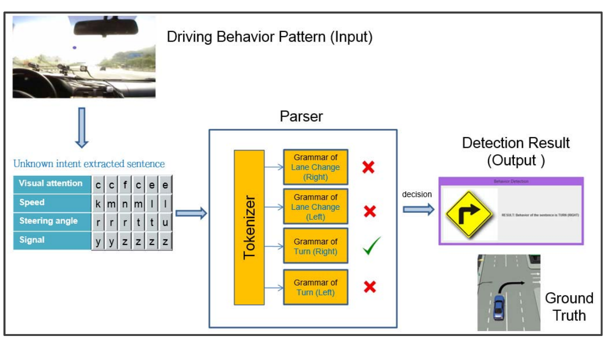


Fig. 11. Summary of the recognition phase. It shows how an unknown driving behavior is taken as an input and it is being detected as "turn right" behavior as the classification results.

V. Experimental Results

This section elaborates the results from our experimental study, which include the sentence detection accuracy, continuously narrowed down intent detection, comparison of performance with different eye conditions, performance of driving behavior intent detection, and performance comparison of driving behavior intent detection.

TABLE III

Driving Behavior Detection Accuracy (Confusion Matrix)

System Prediction	Actual Maneuver				Accuracy
	Lane Change Right	Lane Change Left	Turn Right	Turn Left	
Lane Change Right	252	0	0	0	100%
Lane Change Left	0	168	0	0	100%
Turn Right	0	0	98	0	100%
Turn Left	0	0	0	77	100%

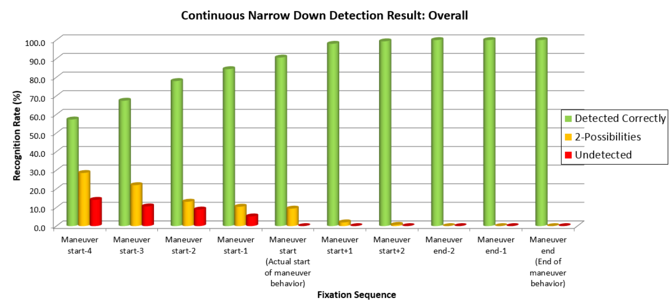


Fig. 12. The test result of continuously narrowed down detection of driver's intent. Note that driver's intent can be well recognized even before the start of the actual maneuver.

A. The Sentence Detection Accuracy

Table III shows the confusion matrix of the system testing to evaluate the performance of our proposed classification model. A total of 595 testing data of driving runs (in sentences) with known true values has been used in the testing so as to test the reliability of the parser and grammars. The result indicates that our parser and each of the four grammars that have been formulated to perform 100% accuracy. We believe that the reason for the high performance rate is because of the comprehensive grammar inference process with numerous numbers of training samples.

B. The Continuously Narrowed Down Early Detection

Further investigation was conducted to see the possibility of early detection of driver's intent even before the starting of the actual maneuver. The investigation was carried out by continuously narrowing down the class possibilities, as shown in figure 12, following the sequence of eye fixation for all data, regardless of fixation patterns. The same 595 testing data, *T*, from four different driving behavior patterns as used in the preceding investigation were used.

Figure 12 depicts the test results of continuously narrowed down early detection. The x axis of the figure represents the sequence of eye fixations, while the y axis represents the intent recognition rate. From the graph, it could be observed that at

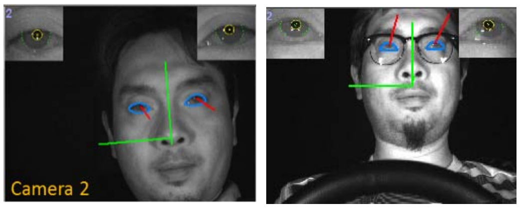


Fig. 13. Example of eye and pupil detection and tracking from drivers with naked eyes (left) and eyeglasses (right).

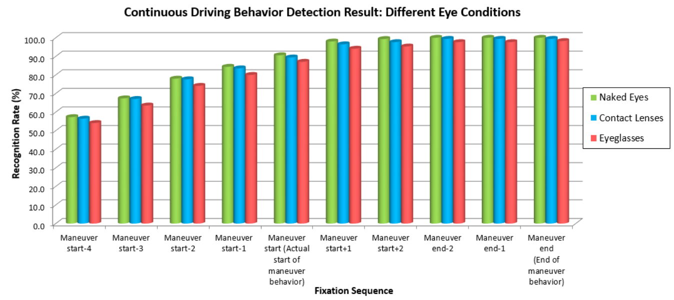


Fig. 14. Result of continuous driving behavior intent detection with different eye conditions. All eye conditions i.e. naked eyes, with contact lenses, and with eyeglasses perform good results using the proposed method.

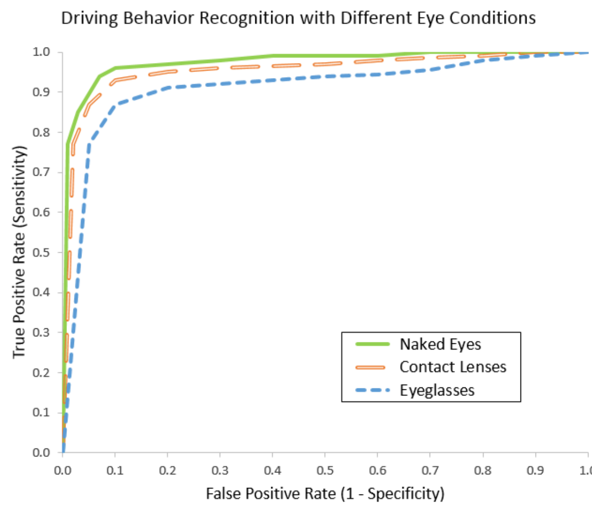


Fig. 15. Performance comparison of lane change right (LCR) driving behavior intent detection with different eye conditions: 1) naked eyes, 2) contact lenses, and 3) eyeglasses.

fixation *maneuver start-2* (i.e. 2 fixations before the start of the actual maneuver), 78% correct intent detection is achieved. Since the average time of an eye fixation is around 600 ms, it means that with an average of 1.2 seconds before the start of the actual maneuver, the system detects 78% correct intent and have that amount of time to notify the driver or another subsystem in the car of any warning if the intended maneuver is not safe to be executed. Furthermore, a correct driver intent recognition of 90.6% recognition rate is achieved at as early as fixation *maneuver start*. This means that an average of 90.6% correct intent detection is successfully achieved in less than 600 ms after the start of the actual maneuver. This result shows that the proposed method is able to steadily detect the driving intents before the start of the actual maneuver (i.e. *maneuver start*).

C. The Comparison of Performance With Different Eye Conditions

Since the proposed approach uses eye fixation data as the key information for classification, it is sensible to test the proposed approach against different eye conditions of the drivers. This investigation is intended to compare the recognition performance of our proposed method when the drivers are in three different eye conditions:

- Naked eyes
- With contact lenses
- With eyeglasses

For clarification, clear contact lenses and clear eyeglasses were used by the drivers during the testing period. Figure 13 shows an image of how the different eye conditions did not affect the proposed system's eye detection and tracking.

For testing purposes, sentences in symbolic form for all driving behaviors (i.e., lane change right, lane change left, turn right, turn left) from drivers with all three different eye conditions were tested against grammars of all learned behaviors. The sentences consist of eye fixations of the driver and car properties data before a driving maneuver starts, continuously followed with during the maneuver starts until completion. A binary recognition is expected for each sentence, whether the correct driving behavior was detected or not at any time between the start and finish of the sentences.

Figure 14 illustrates the result of continuous driving behavior intent detection with different eye conditions (i.e., naked eyes, contact lenses and eyeglasses). It is a detailed result of the fixation sequence from fixation *maneuver start-4* until fixation *end*, where *end* is the last fixation of the maneuver behavior. The result shows that the intent detection of drivers with all different eye conditions performed good results. The result also shows that at all fixations, intent detection of the drivers with naked eyes performed the best, while intent detection of the drivers with contact lenses performed marginally less with not more than 2% from the drivers with naked eyes. Meanwhile, intent detection of the drivers with eyeglasses performed the least, with not more than 5% lesser from the drivers with naked eyes.

Sensitivity measures the proportion of positive examples that are correctly identified (also known as the 'true positive' rate or 'hit' rate). Meanwhile, *specificity* measures the proportion of negative examples that are correctly identified (also known as the 'true negative' rate). The false positive rate (FPR) indicates the number of examples incorrectly identified as positive examples, and it can be shown that Specificity = (1-FPR). The threshold, T, was varied from 0.0 to 1.0 in order to create plots of the true positive rate (sensitivity) and false positive rate (1-Specificity), known as Receiver Operating Characteristics (ROC) curves, which give a graphical representation of the tradeoff between false alarms and higher detection rate of the phenomenon of interest with changing T.

Figure 15 is the performance comparison in the form of ROC curves of driving behavior recognition with different eye conditions (naked eye, contact lenses and eyeglasses) for lane change right (LCR) driving behavior. Based on the results, we found that the performance of the proposed method for drivers with naked eyes and contact lenses were almost the same. The system recognition from drivers with naked eyes was higher, as predicted. However, for the drivers with eyeglasses, the system performance was slightly lower, while it is still in a high acceptance rate. With good recognition rate from all three different eye conditions, where the recognition rate for all the three different eye conditions was above 87%, the proposed method is suitable to be implemented in the car regardless of whether the drivers have naked eyes, wearing contact lenses or wearing eyeglasses.

The main reason for the good performance regardless of the eye conditions is the use of infrared (IR) to detect the pupil of the eyes where the IR lights are reflected by the pupil and detected by the cameras. With the IR light, the pupil is very black, making it is easy to locate the pupil and iris of the eyes. This result is supported by the identical results produced by [42] and [43]. Additionally, a statement by the Medical Research Council (MRC) of Cognition and Brain Sciences Unit in the United Kingdom [44] states that contact lenses are not a problem for eye tracking, and eyeglasses are not a problem in eye tracking except in the case of reflections, i.e. when the orientation of the eyeglasses in such that a clear reflection of the IR light source. The report also mentioned that eyeglasses with scratches or in dirty condition might create a problem in eye tracking.

D. The Performance of Driving Behavior Intent Detection

This investigation was carried out to observe how quick our proposed approach could detect and predict the driver's intent in time prior to the actual maneuver, based on the driving behavior information. Each sentence from the testing data was tested *based on time* as opposed to base on fixation from previous evaluations. Each sentence was tested at 2.0, 1.5, 1.0, 0.5 and 0.0 seconds prior to the actual maneuver of their respective driving behavior pattern.

Figure 16 depicts the results of four different driving behavior in ROC curves between 0.0 and 2.0 seconds in 0.5 seconds intervals before the actual maneuver of the respective driving behavior. The results show that the prediction model of lane change right and left are better than turn right and turn left. The performance of the predictive models decreased as the time before the maneuver was increased as indicated by the ROC curves that are moving away from the (0.0, 1.0) point of the (FPR, TPR). The results also indicate that the lane change right had an accuracy of 87% and 75% at 1.0 and 2.0 seconds, respectively, before the actual maneuver of the behavior. Meanwhile, lane change left had an accuracy of 90% and 82% at 1.0 and 2.0 seconds, respectively, before the maneuver of the behavior. Turn right and left behaviors have both performed almost a similar accuracy of 82% at 0.0 seconds before a maneuver and 73% at 1.0 second before the maneuver.

Based on the good performance described above, where the proposed method manages to detect driving intent between

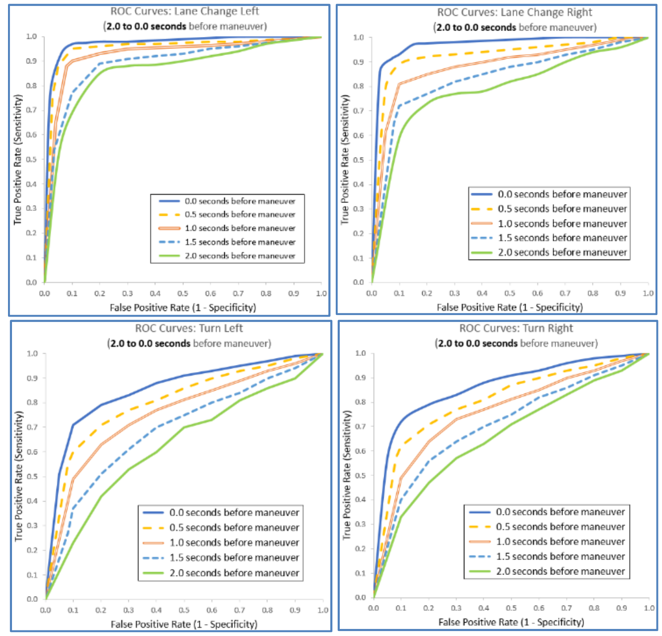


Fig. 16. Performance of the proposed method at different times before actual maneuver (2.0, 1.5, 1.0, 0.5, 0.0 seconds) for all 4 behaviors: *lane change right*, *lane change left*, *turn right*, and *turn left*.

0.0 and 2.0 seconds before the start of a driving maneuver, this will give enough time for an alert to be given to the driver if there are dangers upon executing the maneuver. Also, if the detected intent and actual maneuver is not consistent, a warning should be given.

E. The Performance Comparison of Driving Behavior Intent Detection

We compared the performance of our proposed driving behavior intent detection by using the syntactic approach to the existing non-syntactic methods proposed by Morris *et al.* [8] and Lethaus *et al.* [9]. Morris *et al.* employed multiple cameras within the car to detect the car and driver's movement to predict the intent. They did not take eye gaze into consideration. Meanwhile, Lethaus *et al.* only took into account the eye gaze of the driver in predicting the diving behavior intent. As already being thoroughly explained earlier, our proposed method takes into account the eye fixation of the driver as the main property together with three car's properties data, i.e. speed, steering wheel angle and signal activation.

Figure 17 shows the test result that summarizes the performance comparison of driving behavior intent detection by the time between 2 seconds and 0 seconds, with 0.5 seconds interval before the actual driving maneuvers took place. From the results, it is observed that at 2.0 seconds before the actual maneuver, the performance of our proposed method is better by 13.3% and 13.7% respectively in comparison to [8] and [9]. The better performance accomplished by our proposed method continued for 1.5 seconds, 1.0 seconds, 0.5 seconds, and finally at 0.0 seconds before the maneuver started, where at 0.0 seconds, our proposed method performs 5.3% and 5.6% better than [8] and [9], respectively.

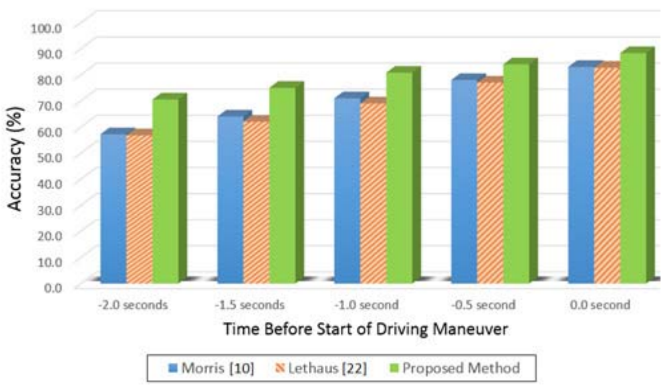


Fig. 17. Comparison of accuracy rate for driving behavior intent detection of the proposed method using syntactic approach and existing non-syntactic approach: Morris *et al.* and Lethaus *et al.*, at different time before maneuver (2.0, 1.5, 1.0, 0.5, 0.0 seconds) for all driving behaviors.

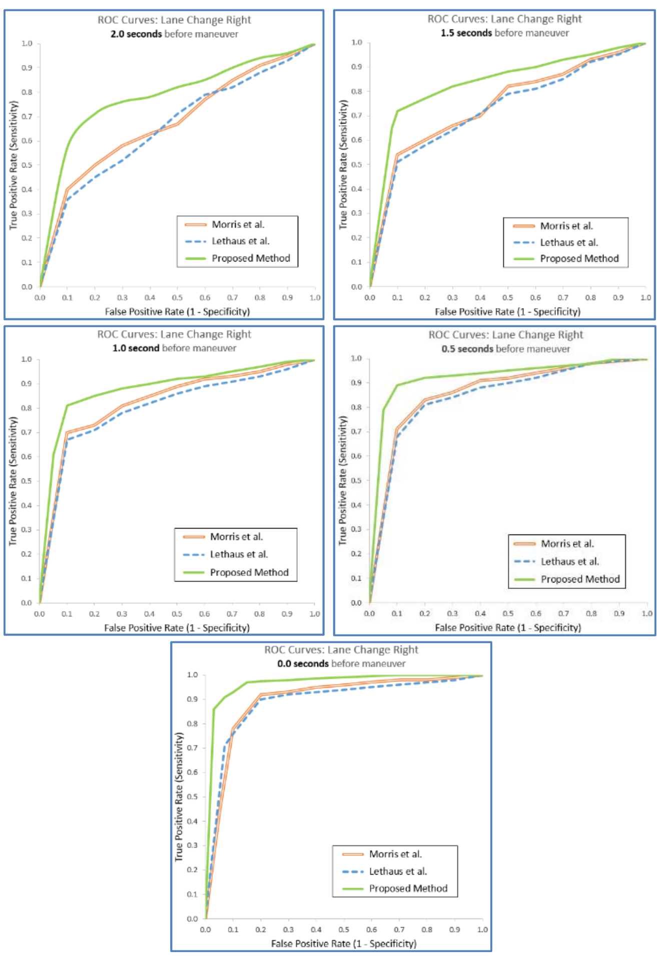


Fig. 18. Performance comparison for the proposed method using syntactic approach and the existing non-syntactic approach, at different time before actual maneuver (2.0, 1.5, 1.0, 0.5, 0.0 seconds) of *lane change right*.

We also present figure 18 that depicts the ROC curves comparison of *lane change right* driving behavior intent between the proposed method and non-syntactic methods [8] and [9]. The results are presented in a different time before the start of the actual maneuver, i.e., 2.0, 1.5, 1.0, 0.5 and 0.0 seconds.

It is concluded that the classifier of our proposed method has better predictive power than [8] and [9]. This means that by taking the eye fixation and car properties data using linguistics-based syntactic approach as being employed by our proposed method is better than taking only car properties data using statistical approach as being employed by [8]. Moreover it is better than taking only eye gaze data using statistical approach as being employed by [9]. The result reveals that by early detection of a driver's intent, the system can notify the driver or another subsystem in the car of any warning if the intended maneuver is not safe. In this way, a safe driving environment for life and property can be realized.

VI. Conclusion

In this work, a linguistic-based syntactic recognition approach with a context-free and context-sensitive grammars is presented for early detection of a driver's intention to improve driving safety. The results from the experimental study show that the proposed approach is able to reliably detect the driving behaviors intent before the start of the actual maneuver with an average of 70.5%, 75.0% and 80.8% accuracy, at 2.0, 1.5, and 1.0 seconds, respectively. The outcome of our proposed approach can be utilized to alert the driver about the possibility of safety risk during driving. Alternatively, the information can be communicated to another system in the car for further action for safety reasons. For example, sending it to the brake and stability control system or engine control system. For future work, further analysis and implementation shall be carried out to realize faster detection and higher recognition rate. It is possible to include in the data structure the additional information from surrounding vehicles and infrastructures [45], such as outside street scene, vehicle location on a map, blind spots and information on nearby vehicles, as well as more information on a driver's condition, such as sleepiness and eye fixations irrelevant to driving [46]. Additional information from the surrounding context is expected to result in better performance for predicting the driver's intent under the driver's normal maneuvering behaviors. However, it is possible that such additional information apart from the direct vehicle control by the driver may result in an adverse effect on predicting the intent of the driver engaged in abnormal or accidental behaviors. It would be interesting to obtain the two predictions, one from the features representing driver's direct vehicle control and another from the features representing outside scene context, separately, such that we can have a comparative analysis and fusion of these two results. The effect of this extension in terms of the prediction time and reliability is subject to further studies.

References

[1] M. Q. Khan and S. Lee, "A comprehensive survey of driving monitoring and assistance systems," *Sensors*, vol. 19, no. 11, p. 2574, Jun. 2019. [Online]. Available: https://www.mdpi.com/1424-8220/19/11/2574

[2] *Global Status Report On Road Safety 2015*, World Health Org., Geneva, Switzerland, 2015.

[3] M. Green and J. Senders. (2013). *Human Errors in Road Accidents*. Accessed: Jan. 12, 2019. [Online]. Available: http://www.visualexpert.com/Resources/roadaccidents.html
[4] S. Singh, "Critical reasons for crashes investigated in the National Motor Vehicle Crash Causation Survey," Nat. Highway Traffic Saf. Admin., Washington, DC, USA, Traffic Saf. Facts Crash Stats. Tech. DOT HS 812 115, Feb. 2015.
[5] N. Kuge, T. Yamamura, O. Shimoyama, and A. Liu, "A driver behavior recognition method based on a driver model framework," SAE Tech. Paper 0148–7191, 2000.
[6] H. Berndt, J. Emmert, and K. Dietmayer, "Continuous driver intention recognition with hidden Markov models," in *Proc. 11th Int. IEEE Conf. Intell. Transp. Syst.*, Oct. 2008, pp. 1189–1194.
[7] N. Oliver and A. P. Pentland, "Graphical models for driver behavior recognition in a smartcar," in *Proc. IEEE Intell. Vehicles Symp.*, Oct. 2000, pp. 7–12.
[8] B. Morris, A. Doshi, and M. Trivedi, "Lane change intent prediction for driver assistance: On-road design and evaluation," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2011, pp. 895–901.
[9] F. Lethaus, M. R. Baumann, F. Köster, and K. Lemmer, "Using pattern recognition to predict driver intent," in *Proc. Int. Conf. Adapt. Natural Comput. Algorithms*. Berlin, Germany: Springer, 2011, pp. 140–149.
[10] B. Sen, J. D. Smith, and W. G. Najm, "Analysis of lane change crashes," Nat. Highway Traffic Saf. Admin., Washington, DC, USA, Tech. Rep. HS-809 571, 2003.
[11] E.-H. Choi, "Crash factors in intersection-related crashes: An on-scene perspective," Nat. Highway Traffic Saf. Admin., Washington, DC, USA, Tech. Rep. DOT HS-811 366, 2010.
[12] D. W. Harwood *et al.*, "Safety effectiveness of intersection left-and right-turn lanes," Federal Highway Admin., Office Saf. Res. Develop., Washington, DC, USA, Tech. Rep. FHWA-RD-02-089, 2002.
[13] W. Najm, J. D. Smith, and D. L. Smith, *Analysis of Crossing Path Crashes*. Cambridge, MA, USA: John A. Volpe National Transportation Systems Center, 2001.
[14] C. Song, X. Yan, N. Stephen, and A. A. Khan, "Hidden Markov model and driver path preference for floating car trajectory map matching," *IET Intell. Transp. Syst.*, vol. 12, no. 10, pp. 1433–1441, Dec. 2018.
[15] M. Muñoz, B. Reimer, J. Lee, B. Mehler, and L. Fridman, "Distinguishing patterns in drivers' visual attention allocation using hidden Markov models," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 43, pp. 90–103, Nov. 2016, doi: 10.1016/j.trf.2016.09.015.
[16] H. Hou, L. Jin, Q. Niu, Y. Sun, and M. Lu, "Driver intention recognition method using continuous hidden Markov model," *Int. J. Comput. Intell. Syst.*, vol. 4, no. 3, p. 386, 2011, doi: 10.1080/18756891.2011.9727797.
[17] R. Fu, H. Wang, and W. Zhao, "Dynamic driver fatigue detection using hidden Markov model in real driving condition," *Expert Syst. Appl.*, vol. 63, pp. 397–411, Nov. 2016, doi: 10.1016/j.eswa.2016.06.042.
[18] J. Tang, F. Liu, W. Zhang, R. Ke, and Y. Zou, "Lane-changes prediction based on adaptive fuzzy neural network," *Expert Syst. Appl.*, vol. 91, pp. 452–463, Jan. 2018, doi: 10.1016/j.eswa.2016.06.042.
[19] W. Zhu, J. Miao, J. Hu, and L. Qing, "Vehicle detection in driving simulation using extreme learning machine," *Neurocomputing*, vol. 128, pp. 160–165, Mar. 2014.
[20] P. Kumar, M. Perrollaz, S. Lefevre, and C. Laugier, "Learning-based approach for online lane change intention prediction," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2013, pp. 797–802.
[21] M. Beggiato *et al.*, "Lane change prediction: From driver characteristics, manoeuvre types and glance behaviour to a real-time prediction algorithm," in *UR: BAN Human Factors in Traffic*. Wiesbaden, Germany: Springer Vieweg, 2018, pp. 205–221.
[22] J. Krumm, "A Markov model for driver turn prediction," in *Proc. World Congr. Soc. Automot. Eng. (SAE)*, Detroit, MI USA, Apr. 2008, Paper 2008-01-0195. [Online]. Available: https://www.microsoft.com/en-us/research/wp-content/uploads/2016/12/sae-markov-distribute.pdf
[23] A. Doshi and M. Trivedi, "On the roles of eye gaze and head dynamics in predicting driver's intent to change lanes," *IEEE Trans. Intell. Transport. Syst.*, vol. 10, no. 3, pp. 453–462, Sep. 2009.
[24] J. C. Mccall and M. M. Trivedi, "Driver behavior and situation aware brake assistance for intelligent vehicles," *Proc. IEEE*, vol. 95, no. 2, pp. 374–387, Feb. 2007.
[25] S. Yuanhsien Cheng and M. Trivedi, "Turn-intent analysis using body pose for intelligent driver assistance," *IEEE Pervasive Comput.*, vol. 5, no. 4, pp. 28–37, Oct. 2006.
[26] G. Lu, B. Cheng, Q. Lin, and Y. Wang, "Quantitative indicator of homeostatic risk perception in car following," *Safety Sci.*, vol. 50, no. 9, pp. 1898–1905, Nov. 2012, doi: 10.1016/j.ssci.2012.05.007.
[27] M. Liu, G. Lu, Y. Wang, and Z. Zhang, "Analyzing drivers' crossing decisions at unsignalized intersections in China," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 24, pp. 244–255, May 2014, doi: 10.1016/j.trf.2014.04.017.
[28] M. Bahram, C. Hubmann, A. Lawitzky, M. Aeberhard, and D. Wollherr, "A combined model-and learning-based framework for interaction-aware maneuver prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 6, pp. 1538–1550, Jun. 2016, doi: 10.1109/tits.2015.2506642.
[29] A. Bisoffi, F. Biral, M. Da Lio, and L. Zaccarian, "Longitudinal jerk estimation of driver intentions for advanced driver assistance systems," *IEEE/ASME Trans. Mechatronics*, vol. 22, no. 4, pp. 1531–1541, Aug. 2017, doi: 10.1109/tmech.2017.2716838.
[30] A. Jain, H. S. Koppula, S. Soh, B. Raghavan, A. Singh, and A. Saxena, "Brain4cars: Car that knows before you do via sensory-fusion deep learning architecture," 2016, *arXiv:1601.00740*. [Online]. Available: https://arxiv.org/abs/1601.00740
[31] F. Ali and T. Pavlidis, "Syntactic recognition of handwritten numerals," *IEEE Trans. Syst., Man, Cybern.*, vol. 7, no. 7, pp. 537–541, Jul. 1977.
[32] R. DeMori, "Syntactic recognition of speech patterns," in *Syntactic Pattern Recognition and Applications*. Berlin, Germany: Springer, 1977, pp. 65–94.
[33] W. W. Stallings, "Chinese character recognition," in *Syntactic Pattern Recognition, Applications*, K. S. Fu Ed. Berlin, Germany: Springer, 1977, pp. 95–123.
[34] K. S. Fu, "Introduction to syntactic pattern recognition," in *Syntactic Pattern Recognition, Applications*. Berlin, Germany: Springer, 1977, pp. 1–30.
[35] P. Trahanias and E. Skordalakis, "Syntactic pattern recognition of the ECG," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 7, pp. 648–657, Jul. 1990.
[36] M. N. Husen, S. Lee, and M. Q. Khan, "Syntactic pattern recognition of car driving behavior detection," Presented at the 11th Int. Conf. Ubiquitous Inf. Manage. Commun., Beppu, Japan, 2017.
[37] R. J. Schalkoff, "Pattern recognition," *Wiley Encyclopedia of Computer Science and Engineering*. Hoboken, NJ, USA: Wiley, 2007.
[38] J. Earley, "An efficient context-free parsing algorithm," *Commun. ACM*, vol. 13, no. 2, pp. 94–102, 1970.
[39] S. P. Abney, "Parsing by chunks," in *Principle-Based Parsing*. Dordrecht, The Netherlands: Springer, 1991, pp. 257–278.
[40] D. Jo and D. Shin, "Driver's behavioral pattern in driver assistance system," *Math. Problems Eng.*, vol. 15, no. 5, pp. 579–586, 2014.
[41] *SE PRO-Smart Eye*. Accessed: Jan. 10, 2019. [Online]. Available: https://smarteye.se/research-instruments/se-pro/
[42] S. Y. Gwon, C. W. Cho, H. C. Lee, W. O. Lee, and K. R. Park, "Robust eye and pupil detection method for gaze tracking," *Int. J. Adv. Robotic Syst.*, vol. 10, no. 2, p. 98, Feb. 2013.
[43] R. Youmaran and A. Adler, "Using red-eye to improve face detection in low quality video images," in *Proc. Can. Conf. Electr. Comput. Eng.*, 2006, pp. 1940–1943.
[44] *Common Eye Tracking Problems*. Accessed: Jan. 16, 2019. [Online]. Available: http://imaging.mrc-cbu.cam.ac.uk/meg/EyeTrackingProblems
[45] M. M. Trivedi, T. Gandhi, and J. Mccall, "Looking-in and looking-out of a vehicle: Computer-vision-based enhanced vehicle safety," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 1, pp. 108–120, Mar. 2007, doi: 10.1109/tits.2006.889442.
[46] G. Li, F. Zhu, T. Zhang, Y. Wang, S. He, and X. Qu, "Evaluation of three in-vehicle interactions from drivers' driving performance and eye movement behavior," in *Proc. 21st Int. Conf. Intell. Transp. Syst. (ITSC)*, Nov. 2018, pp. 2086–2091.

Sukhan Lee (Life Fellow, IEEE) received the B.S. and M.S. degrees in electrical engineering from Seoul National University, South Korea, in 1972 and 1974, respectively, and the Ph.D. degree in electrical engineering from Purdue University, West Lafayette, IN, USA, in 1982. From 1983 to 1997, he was with the Department of Electrical Engineering and Computer Science, University of Southern California. From 1990 to 1997, he was with the Jet Propulsion Laboratory, California Institute of Technology. From 1998 to 2003, he was the Executive Vice President and the Chief Research Officer with the Samsung Advanced Institute of Technology. Since 2003, he has been a Professor of information and communication engineering with Sungkyunkwan University. His research interest in the areas of cognitive robotics, intelligent systems, deep learning, and micro/nano electro-mechanical systems. He is currently a Life Fellow of the Korea National Academy of Science and Technology.

Muhammad Qasim Khan received the B.S. and M.S. degrees in electrical engineering in 2008 and 2012, respectively. He is currently pursuing the Ph.D. degree with the Intelligent Systems Research Institute (ISRI), Sungkyunkwan University, Suwon, South Korea. His research interests include intelligent transportation systems and pattern recognition.

Mohd Nizam Husen received the B.S. degree in computing from the University of Portsmouth, U.K., in 1997, the M.S. degree in information technology from Universiti Utara Malaysia, Malaysia, in 2007, and the Ph.D. degree in computer engineering from Sungkyunkwan University, Suwon, South Korea, in 2017. From 1997 to 2008, he was a Lecturer with the Universiti Kuala Lumpur (UniKL), Malaysia, where he was promoted as a Senior Lecturer, in 2008. He is currently the Deputy Dean (Academic & Technology) at UniKL. He is the author of more than 20 research articles. His research interests include intelligent systems, visual attention, pattern recognition, and indoor localization. He is also a Reviewer of the IEEE COMMUNICATIONS LETTERS and *Springer Mobile Networks and Applications Journal*.