Driving Simulation for Driver Classification Algorithms

Lucas Bruck, *Student Member, IEEE,* Carlos Vidal, *Member, IEEE,* Gary Newton, Jr., *Member, IEEE,* and Ali Emadi, *Fellow, IEEE*

***Abstract*—Driving simulation allows the creation of virtual realities that are highly immersive, safe, controlled, and repeatable. Its use is seen throughout the entire design process of a vehicle, from virtual calibration to validation of systems and components. It reduces the cost and time consumed in testing and prototyping. Besides being used to develop vehicular systems, driving simulation is a great tool for studying the drivers. Driver behavior is a factor that greatly influences, energy consumption, vehicle dynamics, component wear and safety. By identifying the driver style and behavior, inteligent and connected systems can adapt and perform their tasks with customizable driver-centered calibrations. This work proposes the guidelines to create and conduct driving simulation experiments with the purpose of designing learning algorithms for driver behavior classification. The necessary hardware and software capabilities are outlined and a list of significant influencing factors and features are presented. The relevant driving data is also discussed and the tools to label it are presented in detail. A pilot study is conducted using long short-term memory (LSTM) artificial neural network (ANN) to classify the driver behavior between low, medium and high aggressive. The features used were the gradients of the driver controls and the vehicle longitudinal and lateral accelerations. Results show accuracy of 94.62% of the learning algorithm after a 8-minute training with 558 observations of 10 seconds each. The observations included driving data in residential and highway scenarios. The results show the method is promising and the prospect is that this type of algorithms could be implemented in real-time with systems such as autonomous driving, energy management, and vehicle stability.**

***Index Terms*—driver style, driving simulation, machine learning, virtual reality.**

I. Introduction

THE advent of computer-aided engineering (CAE) enabled the automotive industry to shorten the vehicle design process, remove cost, and to improve the performance of systems and components. In the simulation world, components are exchanged or edited easily, and the test cases are consistent and repeatable. This allows for reducing the number of prototyped components, rapid-effective calibration of systems, and coherent definition of performance parameters. It also makes it easier to compare different vehicles in a back-to-back mode. This has been a challenge previously. Nevertheless, a key factor to be considered is the human component. The way drivers execute driving tasks (e.g. drive away, stops, cornering, lane changes, etc.) impacts considerably the vehicle dynamics, component wear, and energy consumption.

To fill this void and keep up with the increasing speed of virtual vehicle design, driving simulators are introduced. Driving simulators come in a number of varieties made up of both hardware and software. The hardware enables the type of driving immersion that is commiserate with where the vehicle is in the engineering process. These range from simple desktop systems that consist of a steering wheel and pedals in a cubical to a full vehicle cockpit on a multi-degree of freedom (DOF) platform that requires its own site and infrastructure. All simulators capture real driver inputs, imitate the response of a road vehicle in real-time, accounting for that driver's perceptual system. From a software standpoint driving simulator usually consists in vehicle dynamics model, scenario builder, kinematics algorithm, visual, auditory, haptic, and motion cues [1].

In a driving scenario, the driver is the operator that provides commands. The vehicle is the mechanism that responds dynamically. The term driver behavior is used as to classify the style in which the driver provides the inputs when conducting a driving task.The following studies show how driver behavior is directly linked to energy consumption [2], [3], and [4], motion stability [5], and safety [6] and [7].

Since simulators provide a protected environment for otherwise safety-risk testing scenarios, their use for the testing of human-machine interfaces (HMI) and connected/autonomous (C/A) systems has been increased over the years. Experiments include the evaluation of driver interaction with secondary devices (such as touchscreen monitor and steering wheel controls) [8], driver-machine transitions for automated driving systems [9], and the assessment of interfaces that enhance trust between passengers and autonomous driver systems [10].

In addition, driver behavior is subject of study of several works. In [11] a compact simulator is used to evaluate the behavior of the drivers at traffic light intersections. In [12], a driver performance index is proposed based on reaction time, aggressiveness, skill, and vehicle jerk.

Several parameters that influence driver behavior are reviewed in [13] and divided between environmental factors and human factors. The work also provides a survey on driving style recognition methods. It reviews the sensors and inputs usually used, the classes in which the drivers are divided, and the different algorithms. It also highlights the importance of the environmental conditions and human factors. Methods used for classifying driver style include rule-based (RB) [14], fuzzy-logic (FL) [15], Gaussian mixture model (GMM) [16], and

Lucas Bruck, Carlos Vidal, and Ali Emadi are with the McMaster Automotive Resource Centre, McMaster University, Hamilton, ON L8P 0A6 Canada (e-mail: bruckl@mcmaster.ca, vidalc2@mcmaster.ca emadi@mcmaster.ca).

Gary Newton is with VI-Grade, Canton, MI. USA (email: gary.newton@vi-grade.com).

hidden Markov model (HMM) [17] and [18].

Given the difficulty of stablishing rules and complexity in developing models that generalize the driver style and behavior, rule-based methods (RB and FZ) and model-based methods (GMM and Markov chain) are increasingly being replaced by machine learning (ML) algorithms, especially in the form of artificial neural networks (ANN). The ANNs are adaptive, flexible, deal well with large amount of data and do not require previous knowledge of the system. A comprehensive comparative study of different types of ANNs for driver behavior classification is performed in [19].

In [20], a shallow learning ANN acchieves accuracy of 90% in classifying the driver style with three features: driver's throttle, vehicle acceleration, and velocity. In [21], another shallow learning algorithm leveraging support vectoring machine (SVM) and fuzzy c-means (FCM) is applied with more than 44 features extracted from the drive cycle to achieve accuracy of 92.86%. Bayes network is also applied as showed in [21] to estimate driver's lane change intent through driving style. In this work, the accuracy of the estimations is 78.2% using festures that include time to collision and time gap between ego and front vehicle. The work in [22] investigates the benefits of using deep learning over shallow learning algorithms for classic machine learning based structures. By adding hidden layers the accuracy of the ANN improves from 70.1% (one hidden layer) to 99.8% (25 hidden layers). The features used included the linear and the rotational acceleration of the vehicle in the three axis (longitudinal, lateral, and vertical). Finally, the authors of [23] develop a residual convolutional network (RCN), using steering wheel angle, vehicle velocity, engine load, and speed to perform the driver style classification. This method showed accuracy of 99.3%.

Artificial neural network algorithms require large amount of training data, though. And it is important to account for the driving influencing factors and conditions in the test cases so there is no bias in the final algorithm. Also relevant is the selection of input signals that will be used for the training and prediction.

In the present work, we stablish the guidelines for creating driving simulation studies in which driver behavior classification algorithms can be developed. From the literature review, we conclude that besides being immersive, the virtual reality created in the simulation should support the testing of diverse demographics, realistic environments, various driving conditions, and tasks. Furthermore, the learning algorithm should be preferably a deep learning algorithm with the ability to leverage present and past information.

II. Driver Behavior Influening Factors

In this section, we analyze in detail the influencing factors on driver behavior as to define the appropriate test cases for the development of classification algorithms.

A. Enviroment

The term environment is used to describe the surroundings in which the subject driver is inserted. Here, we subdivide the environment in four categories: highway, commercial area, residential area, and proving ground.

1) Highway: Highway is the term used for express roads that connects towns or cities. Commuting using highways is part of the day of many drivers, hence the importance of testing under this environment. In addition, given its usual high-speed characteristic and presence of traffic, accidents on highways tend to be more impactful.

2) Comercial Area: This group includes areas where commercial activity takes place so there is a high traffic of pedestrians and vehicles. Examples are city centers and open markets. Although less fatal, the number of accidents tends to be higher in this type of environment.

3) Residential Area: These are neighborhood areas such as suburbs. The speed limits, traffic of vehicles and pedestrians are usually lower when compared to commercial areas. Nevertheless, level of distraction can be higher. The driving tasks in residential areas are similar to the commercial areas.

4) Proving ground: This environment includes race tracks and test facilities. They can be closed circuits, tracks that imitate real roads, large, paved areas for free testing, or areas marked with specific geometry such as circles, ovals, and straight lanes. Proving grounds are broadly used in industry since they provide a relatively repeatable and controlled environment for testing. For that reason, they are adopted specially for the testing of safety-related systems such as electronic stability control (ESC) and advanced driver assistance systems (ADAS).

B. Conditions

The conditions of each environment also play an important role in driver behavior. Here we subdivide the conditions into traffic, weather, lighting, and road quality.

1) Traffic: The number and flow of vehicles in an environment. Higher traffic conditions are more prone to accidents while lower traffic conditions are less.

2) Weather: This condition is associated directly with tire grip and visibility. In sunny-dry conditions, the vehicle is more responsive to driver actions (due to higher grip), requiring less skill in emergency situations. In rainy-wet conditions, vehicle response becomes less intuitive and driver skill is key in emergency scenarios. The impact is even higher for snowy-ice situations. Wind conditions might be also accounted for when in higher levels.

3) Lighting: Decent lighting level is key for performing driving tasks [24]. Lighting can be associated with weather conditions (sunny-bright, rainy-cloudy and fogy). In addition, time of the day also influences lighting (day, night), and even traffic (glare from opposite vehicle's headlights).

4) Road: This group of conditions is with respect to the type, and quality of the pavement. It includes rough roads, concrete, tarmac, asphalt, stone, among others. Also, the road condition and slope influences driver behavior.

C. Task

The driving tasks are the missions, or the instructions assigned to the testing subject (driver). Table I depicts a diagram of driving tasks associated with the previous listed environments that can be used for the design of experiments.

TABLE I
DRIVING TASKS ASSOCIATED WITH ENVIROMENTS

Environment	Driving Task
Highway	Merging Exiting Lane Change Overtaking Collision Avoidance Cornering Cruising Speed Reduction Speed Increase Pull over Sudden Stop
Commertial Area	Stop at Traffic Light Stop at Crosswalk Drive Away Yield Right/Left Turn Parking
Residential Area	Stop at Traffic Light Stop at Crosswalk All-way Stops Drive Away Yield Right/Left Turn Parking U-turn
Proving Ground	Constant Speed Cornering Constant Speed Cornering Acceleration Maneuver Deceleration Maneuver Right/Left Turn Swerve Circuit lap

Note that the driving task is usually environment and sometimes condition specific. For instance, a U-turn is not allowed in a highway environment, therefore it is not listed as a highway driving task.

III. APPARATUS

Driving simulators consist in a combination of hardware and software that, given real driver inputs, imitate the response of a road vehicle in real-time, accounting for the driver's perceptual system. A driving simulator usually consists in vehicle dynamics model, scenario builder, kinematics algorithm, visual, auditory, haptic, and motion cues. All of which explained in detail in [1].

A. Hardware

The hardware side includes the driver controls, display, sound, motion, and force-feedback systems. Driving is in its essence a visual task. Therefore, visual and motion cues must be set-up correctly and tuned to the simulator so the driver can get the best possible immersion and provide the best possible subjective evaluation.

Regarding the visual cues, there are three prime technologies. The first is providing the image through screens. Although usually the cheapest solution, screen edges and low field-of-view contribute to lesser immersion. Rounded screens, powered by high frequency projectors are the most adopted solution for driving simulators. Their large field-of-view is provided by the screen conicity and digital warping and blending of the different projections. Although they require dimmed environments (which is not always the case in the real world), they provide a highly immersive visual environment since they are proved to give a better perception of velocity and surroundings [25]. Last, head-mounted virtual reality devices have gained momentum in recent years. Their easiness of transport, store, and use make them the most versatile. In addition, the possibility of a complete field of view enhances the immersion. Their main disadvantage is the invasiveness of adding an apparatus to the driver's head and loss of velocity perception [26].

Besides the graphic display, the sound system and distribution is also important to create immersiveness, velocity perception, and driver awareness [27]. The state-of-the-art driving simulations opt for three-dimensional (3D) sound systems with speakers distributed inside the vehicle cabin [28].

In motion simulators (also called dynamic simulators) a motion system is part of the hardware. Motion simulators include mid- and high-fidelity apparatuses, depending on the number of DOF their systems provide [29], the most common being the hexapod with 6-DOF [28]. Redundancies in some degrees of freedom may also be applied to enhance the motion evelope, especially in the longitudinal and lateral direction. Examples are found in [30] and [31] (8-DOF), and in the 9-DOF Driver in Motion (DiM) structure detailed in [32], [33], and [34]. Some simulators also count on haptic actuators, such as force-feedback steering, active belts and seats.

The different apparatuses combined, contribute to the level of immersion provided by the simulator. That immersion is key to build trust in the reality presented.

B. Software

In essence, the software creates the inputs that will interact with driver through the hardware. Given the importance of the visual and sound cues to create driver imertion, the virual reality software is of great relevance. The virtual reality software must allow not only the creation of realistic enviroments but also the edit of the conditions. The VI-WorldSim is a graphic environment that allows for easy control of start/stop operations, run time control of lighting, time of day and creation of scenarios the driver needs to react to [35]. Including features like traffic, pedestrians, lighting, weather, and sensor enables the user to create anything from simple to complex scenarios that test the functionality of controls and algorithms. Using a high-quality graphics environment, built on an unreal graphics engine, it allows a significantly improved immersive subjective feel and drive. The closer to reality the driving experience in the simulator the better the feedback on the vehicle performance and the driving data measured. It also allows for building enviroments from scanned surfaces and actual road profiles that may be used for physical testing, it further connects the simulator to what will eventually be an actual test drive for enventual system validation.

Fig. 1. Residential area and highway scenarios created with VI-WorldSim graphic enviroment

Fig. 2. MARCdrive lab from outside (top) and inside the vehicle cabin (bottom)

The creation of scenarios consists in selecting the environment, defining conditions, and placing the traffic agents (pedestrians and vehciles). The software counts on a vast library of agents, including vehicles of all purposes, makes, and models and pedestrians from a wide range of demographics. The computer-controled agents can be programmed to act and behave with different levels of aggressivity, their trajectories can be defined by routes or they can be set to wander. Other functions allow to manage the traffic and trigger events. Figure 1 shows the quality of the graphics achieved with VI-WorldSim.

Many high-quality driving simulators include playback of real noise and vibration data synced in phase with the vehicle velocity and engine rotational speed being driven in the simulator. This source data can come from physical test in the case of production vehicles or from simulation sources in the case of protype or pre-production vehicles. Road, tire, and powertrain noise, coupled with aeroacoustics wind noise provides even more context to the driving experience and vehicle operation.

To add additional immersion and context an accurate noise and vibration (NVH) model can be coupled to the vehicle performance model inside the simulator. This allows for a more immersive drive and is representative of how drivers perceive the actual vehicle. Providing the right amount of NVH context puts the driver in correct cognitive mode to perceive the vehicle. These models can be complex and broken down into many different sources and cascade the various transfer paths and sources in the vehicle and this would make sense for a pure subjective ride and feel drive, but for our current type of work a more simplistic model that plays wind/road and powertrain noise signals to the driver.

A vehicle model that performs realistically with the driver inputs is also needed to generate realistic vehicle responses to driver commands. That is key for achieving driver immersion and accurate classification algorithms. VI-CarRealTime is a software that enables the creation of vehicle models by including information about the body, suspensions, wheels, brake, steering, and powertrain [36]. The 14-DOF of the model are achieved by the linear and rotational motion of the body in all three dimensions (6-DOF) and the vertical and rotational motion of each wheel (2-DOF per wheel). The tire model is the Pacejka [37]. It computes forces and momentum in all three directions in addition to relaxation.

The vehicle model is also the provider of the input signals for the ANN training and testing.

C. MARCdrive

An exemple of driving simulator is is located at the McMaster Automotive Resource Centre, Hamilton-Canada. The

MARCdrive lab depicted in Fig. 2 houses a static simulator. The complete vehicle cabin faces a cylindrical screen with 210 degrees of field-of-view that is powered by three projectors at 120 Hz. The three-dimensional surround system is provided by speakers distributed by the vehicle cabin. Soft handling motion cues are enabled by active seat and seatbelts. Haptic feedback is given through the active steering.

The virtual scenarios are created using the previously presented VI-WorldSim and the vehicle models are built using VI-CarRealTime.

IV. Methodology

The methodology consists in: defining the test cases that better suit the application, executing the test, collecting, and post processing the data that will be used by the driver behavior classification algorithm.

A. Scenarios

The term scenario refers to the combination of environment, conditions and driving task presented to the driver on the driving simulator. Depending on the application, some environments and conditions might be suppressed. For instance, in a take-over maneuver (autonomous driver to human driver and vice versa) that is only enabled in highway roads, the developer should focus on highway scenarios. The objective of running scenarios is to obtain driving data for training and testing of the algorithms.

B. Driving Data

To perform the driver classification it is important to carefully define the inputs to the algorithm. Driver controls are given through the pedals, steering wheel, shifter, and clutch (last two for manual vehicles). The vehicle responds dynamically with longitudinal, lateral, and vertical accelerations. Therefore, driver controls and vehicle response are key to analyze driver behavior.

Previous works have focused on using the vehicle response to classify driver behavior [38] and [39]. Although driver controls are signals available in the vehicle network, using them for the same task is not yet explored. The relevant inputs signals for driver behavior classification algorithms are depicted in Table II.

The inputs in Table II must be provided by the vehicle model embedded in the real-time simulation apparatus, preferably in a similar fashion as in the real vehicle network, so the algorithm designed in the simulation environment is easily transferable to a real vehicle.

In addition, the size of the logged data might be different depending on the length of each scenario devised. To avoid biases towards longer scenarios, it is recommended to partition the driving data in even time-series segments. That is called segmentation of the data and each segment is called an observation.

TABLE II
Input Signals for Driver Behavior Classification

Type	Signal
Driver Controls	Throttle Pedal Position
	Brake Pedal Position
	Steering Wheel Angle
	Gear Engaged (manual vehicles)
	Clutch State (manual vehicles)
Vehicle Response	Longitudinal Velocity
	Longitudinal Acceleration
	Longitudinal Deceleration
	Lateral Acceleration
	Vertical Acceleration
	Brake System Pressure
	System-specific

C. Defining Driver Behavior

The most common distribution of classes for driver behavior includes styles from aggressive to non-aggressive [13] and [19]. The aggressiveness of the driver is often perceived not in the magnitude of the command, but in the rate of change in its value. For instance, during a parking event the magnitude of the steering wheel angle assumes high values, even if performed by a non-aggressive driver. In contrast, in an aggressive lane change maneuver, the magnitude of the same signal is much lower. Hence, using the magnitude of the driver controls could mislead the algorithm. By using the rate of change in the controls (throttle gradient, brake pedal gradient, and steering wheel angle gradient) that issue is mitigated.

Furthermore, an observation might show different levels of aggressiveness for each control, e.g., a emergency braking maneuver with the vehicle going straight (no steering wheel input). Therefore, the labeled aggressiveness should vary accounting for the different levels of aggressiveness at each control.

D. Clustering

For classification algorithms, it is necessary to define the different classes and label the data to enable the training the ANN algorithm. To avoid biases, manual labelling of driving tasks is not recommended when the class needs interpretation (e.g. aggressivity). The k-means clustering is an unsupervised learning method that uses vector quantization to partition a number of observations into a predefined number of clusters (or classes). The algorithm defines each clusters position by creating centroids and an observation belongs to the cluster with the nearest mean to that centroid [40].

An example of the application of a k-means algorithm is given in Fig. 3. Maximum longitudinal velocity and acceleration are extracted from each observation, during driving missions on highway and residential area scenarios. The measured data is scattered in the plot and the k-means algorithm labels each sample without knowing its origin. The figure shows how the k-means algorithm is effective in assigning each

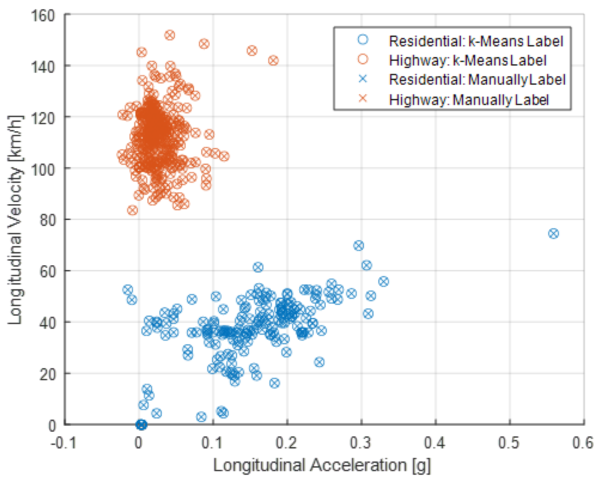

Fig. 3. Comparison of k-means and manual labelling of two different enviroments using longitudinal velocity and acceleration as features

observation to the correct cluster. Overall, high-speed and low-acceleration measurements belong to the highway scenarios and low-speed high-acceleration measurements belog to the residential area scenarios. In this particular example manual labelling is not complex, but for less intuitive labelling such as level of aggressiviness, the k-means is a powerful tool.

E. Training and Validation

With the k-means clustering algorithm, different discretizations of relative aggressiveness can be obtained. Once each observation is automatically labeled by the clustering algorithm, it is ready to be used for training the neural network. Once all observations are labeled, the data set is ready for training.

The observations are randomly selected as training or validation data. The amount of training data is 75% and the validation data, the remaining 25%. The test data is equal to the validation data.

F. Machine Learning Algorithm

Lastly, it is important to define which type of algorithm and structure will be used to perform the driver behavior classification. As the literature reviewed in the introduction, supervised deep learning algorithms have shown better performance. It is also imporant to account for some chacteristics of the problem when choosing the ANN topology. In this case, the inputs are time-series signals, and the outputs are categories (levels of aggressiveness).

Furhtermore, it is reasonoable to assume that driver behavior is affected by current and past situations presented to the drivers in the scenario, i.e., aggressivity might build up or smooths down over time.

V. Pilot Study

This section applies the previously defined guidelines to exemplify the design of a driver behavior classification. For the driving data acquisition, the MARCdrive driving simulator was used. The vehicle model is representative of a small-sized electric city car which is created using VI-CarRealTime and Matlab/Simulink [41]. The driving data measured from the model is the driver controls and the output vehicle responses as defined in Table II. The objective is to create an algorithm that accurately classifies driver behavior between three levels of relative agrressiveness.

A. Experiments

For this pilot study, the scenarios built on the driving simulator account for two influencing factors, environment (highway and residential) and traffic (with and without). Table III summarizes the scenarios created for the pilot studies.

Five drivers are subjected to the scenarios. For the scenarios with traffic, the behavior of the computer-controlled pedestrians and vehicles is set to "wander" and "swarm", respectively. That means that the pedestrians will walk around the map and their interaction with the human-driven vehicle is incidental. As for the vehicles, the "swarm" function means that they will be constantly placed around the human-driven vehicle in order to create the perception of traffic.

Most of the driving tasks are "incidental" since they depend on driver's decision and interaction with the computer-controlled agents. Some driving tasks though, are "imposed" since they are part of the instructions given to the driver. Specific instructions are given for scenarios HW 02, RS 01, and RS 02. In HW 01, the driver is simply told to drive for the duration of 10 minutes. The instructions are necessary in scenario HW02 to make sure the driver performs all the key driving tasks for the highway scenario without traffic. As for the scenarios RS01 and RS02, instructions are necessary to guide the driver throughout the mock-up residential area. Table IV details the instructions given for all four scenarios. In all scenarios, the drivers are told to drive and comply to the driving regulations as they would in a real-world experiment.

Driving data is logged at a sampling frequency of 100 Hz. All the data for all drivers and scenarios are combined and then segmented in observations that have equal lenght of 10 seconds.

B. Labelling Mehtod

Driver style is assessed in this pilot study from the perspective of aggressiveness in a scale from 1 to 3, 1 being the lowest level of relative aggressiveness and 3 being the highest. As

TABLE III
Pilot Studies Scenarios and Influencing Factors

Environmnet	Condition	Scenario ID
Highway	with traffic	HW 01
	without traffic	HW 02
Residential Area	with traffic	RS 01
	without traffic	RS 01

TABLE IV
INSTRUCTIONS GIVEN FOR EACH SCENARIO

Scenario	Task
HW 01	Free driving Duration: 10 min
HW 02	Merge to highway Lane change: left Lane change right Cruising (1 min) Speed Increase: +20 km/h Speed Decrease: -20 km/h Pull over Duration: 6 min
RS 01 and RS 02	Full stop At intersection: turn left At intersection: turn right At traffic light: turn left At intersection: turn right At intersection: turn right At intersection: turn right At traffic light: turn left At intersection: keep straight At intersection: turn right At the end of the road: park Duration: 2 min

TABLE V
LSTM PARAMETERS

Overall Parameters	
Features	Steering Wheel Gradient Brake Pedal Gradient Throttle Gradient Longitudinal Acceleration Longitudinal Deceleration Lateral Acceleration (total: 6)
Responses	Low Aggressiveness Medium Aggressiveness High Aggressiveness (total: 3)
Number of Hidden Units	10
Training Parameters	
Number of Training Observations	558
Number of Test Observations	186
Max Epochs	200
Gradient Threshold	0.01
Initial Learning Rate	0.1
Learn Rate Drop Factor	0.5
Learn Rate Drop Period	100

mentioned before, the aggressiveness of the driver is peceived through the gradient of the action on the controls (steering wheel, brake pedal, and throttle) as well as in the respective vehicle dynamic response (lateral acceleration, longitudinal acceleration, and longitudinal deceleration, respectively).

A first-level k-means clustering algorithm is used here for labelling the level of aggressiveness of the drivers at each control, individually. For each control, the same three levels of relative aggressiveness are considered: low, medium, and high. The driver inputs considered are the gradients of the controls. The gradients indicate the rate of change of the driver inputs and are more representative of the agressiveness of the driver. Therefore, for each observation, the maximum value of the steering wheel angle gradient, brake pedal gradient, and throttle pedal gradient is saved and the respective vehicle response with it.

As for the the overall aggressivenes of the driver, that should be the combination of the indexes generated for each control. In the end, there will be 27 possible combinations ($3^3 = 27$). Those combinatios are used as inputs to the second-level k-means algorithm that will then define the boundaries of the three classes of overall level of relative aggressiveness of the driver.

C. Long Short-Term Memory Recurrent Neural Network

Here, the long short-term memory (LSTM) recurrent neural network (RNN) algorithm is investigated. The LSTM RNN is indicated when it is necessary to model long sequences with long term dependencies [42]. Applications in the automotive field can be found in [43] where LSTM is used for battery state of charge prediction. Despite being a RNN, the usual long-memory limitation (caused by gradient shrink in the back propagation learning) does not apply to the LSTM [44]. This ability is especially relevant for driver behavior classification.

Table V shows the parameters used for the LSTM learning algorithm. The layers of the classification algorithm consist of input layer, LSTM layer (with 10 hidden units) a fully connected layer, a SoftMax function layer, and a classification layer. The input layer is responsible for receiving the time-series data from the observations. The LSTM layer creates the structure for the long-term learning algorithm. The fully connected layer reduces the output vector to the number of clusters previously defined. The SoftMax function computes the output belongings and the classification layer assign it to the cluster of equals. The data from the second-level k-means algorithm is used as the label for the driver behavior classification algortihm.

D. Results

This section assesses the results of the labelling method, driver behavior classification algorithm, and the classification of aggressiveness at individual controls.

1) Labelling Algorithm: The labelling method used two levels of the k-means algorithm. The first defines the level of relative aggressiveness for each driver control and the second defines the clusters for the overall driver behavior. Figures 4, 5, and 6 show the clusters for the agressiveness at the steering, brake pedal, and throttle, respectively.

Then, the second-level k-means clustering takes the outputs of the first level as features. Figure 7 shows the clustering results for the 27 possible combinations of the aggressiveness of each control and the results of the logged data after first and second level of clustering. It is possible to see that the data measured from the experiments fills almost every spot

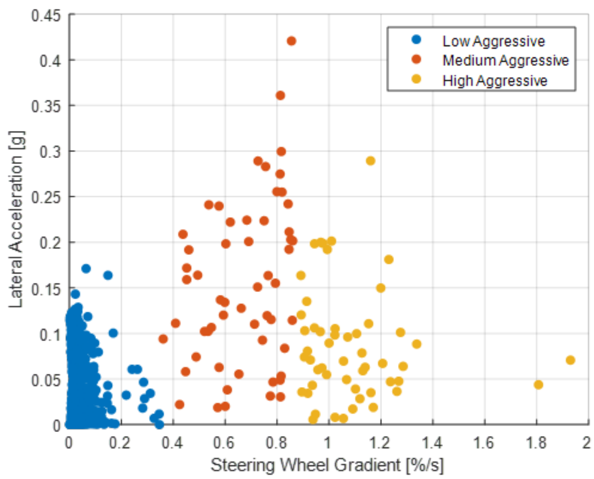

Fig. 4. Levels of aggressiveness at the steering wheel. Features are the gradient of the steering wheel angle (x-axis) and lateral acceleration (y-axis)

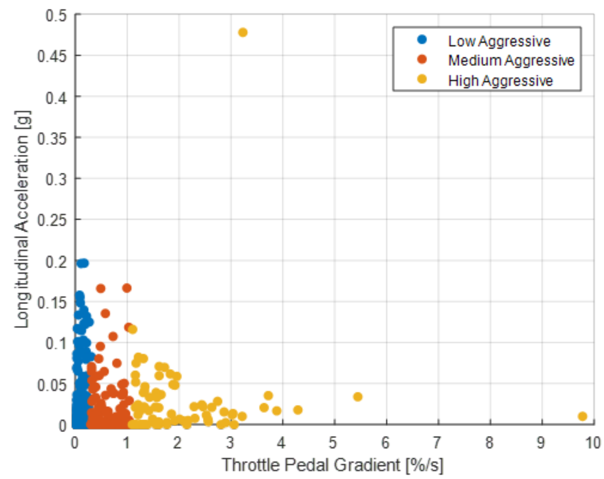

Fig. 6. Levels of aggressiveness at the throttle. Features are the gradient of the throttle position (x-axis) and longitudinal acceleration (y-axis)

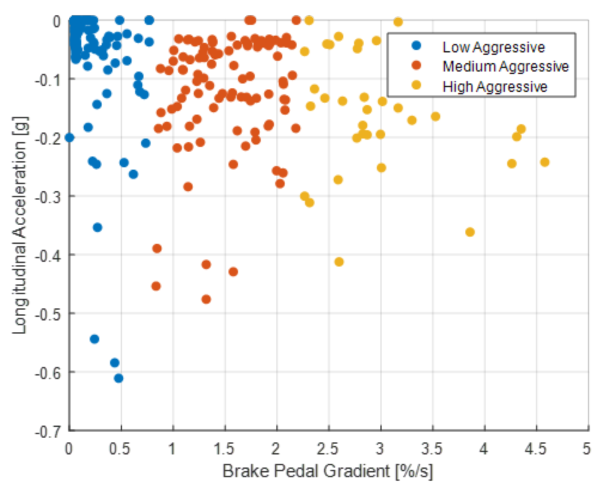

Fig. 5. Levels of aggressiveness at the brake pedal. Features are the gradient of the brake pedal position (x-axis) and longitudinal deceleration (y-axis

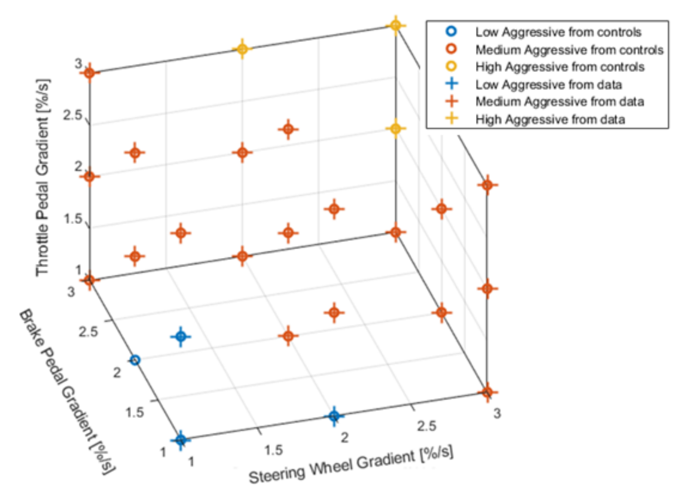

Fig. 7. Levels of overall driver aggressiveness. The 'o' terms are defined from the combinations of the aggressiveness of the controls and the '+' terms are clustered observations from the driving data measured. Features are the previously defined levels of aggressiveness at steering wheel (x-axis), brake pedal (y-axis), and throttle (z-axis)

(96.3%) of the predefined possible combinations of the levels of aggressiveness from the controls.

Once the data is labeled, we can train and test the learning algorithm.

2) Driver Behavior Classification Algorithm: After 8 minutes and 31 seconds of training the LSTM ANN is able to perform the driver behavior classifications with accuracy of 94.62%. This performance is remarkable considering the short training period and the size of the ANN. The test data is classified within 0.45 seconds. Considering the the test data contains 186 observations, the average prediction time is 0.002 second for predicting the driver behavior of an observation that is 10 seconds long.

3) Classification of Aggressiveness at Individual Controls: In addition to the previous algorithm, the same LSTM ANN structure was trainned with the same observations to identify the level of aggressiveness at each control individually, i.e., steering wheel, brake pedal, and throttle. For that test, the labels used are straight from the first-level k-means algorithm and the input features to the LSTM ANN are the same used for the clustering method. The trained algorithms achieved 99.46%, 100%, and 98.92% of accuracy for the identification of the aggressiveness at the steering wheel, brake pedal, and throttle, respectively. That shows how easier it is to identify the aggressiveness at an individual control compared to overall driver aggressiveness.

Such individualized approach might be particularly usefull for energy management systems and chassis control systems as discussed in the following section.

VI. Applications

The spectrum of applications for driver behavior classification algorithms is very broad, from conventional systems to advanced driver assistance systems (ADAS). Any system that is calibrated based on driver actions and/or vehicle response can make use of this information to switch between modes of operation. The following sections list some possible applications.

A. Energy Management Systems

Energy management systems (EMS) translate driver intention into power demand, and therefore energy consumption [45]. They are key in electrified powertrains but also in conventional ones with internal combustion engines (ICE). By being able to predict driver's longitudinal behavior in the form of acceleration, the EMS can better estipulate the power demand from the energy source, whether coming from fuel tank or battery. The benefit of adding such information is highlighted in systems such as the one presented in [46]. In this work, the knowledge of the driver's deceleration behavior is used to maximize the usage of battery power during vehicle operation, thus saving fuel.

B. Chassis Control Systems

Chassis control systems include electronic stability control (ESC), anti-lock brake system (ABS), active suspensions, electronic power steering (EPS), drive-by-wire systems, among others. Such systems either assist the driver in safety-risk situations (ESC and ABS) or create comfortable driving experience (EPS and drive-by-wire). In any case, they need to be calibrated. The calibration is often different between model versions, market, and target consumer. For instance, a sportive vehicle will have an ESC calibration with more flexible thresholds of safety when compared to a commercial version of the same model.

Therefore, the chassis control systems could also be calibrated differently for different driver styles and perform accordingly given the style identified in real-time.

C. Advanced Driver Assistance Systems

Advanced driver assistance systems (ADAS) are control algorithms that leverage the sensors and the modules mounted in the vehicle to assist drivers in several different driving tasks. They include collision alert features such as forward collision warning, lane departure warning, and pedestrian detection. They also include active dynamic controls such as automatic emergency braking, adaptive cruise control, lane keeping assist, and park assist. Being able to classify drivers style and predict they behavior is definitely an asset in predicting collisions and giving warnings.

D. Autonomous Systems

The levels of autonomous driving vary from 0 to 5 as defined in [47] The most common systems share driving responsibilities with the driver (levels 1 to 4). That means transition systems must be designed to blend human-machine driving in all sorts of situations. In addition, the way a driver performs the driving task is directly related to the way he/she expects the task to be conducted by the machine. Therefore, identifying driver style might be pivotal to create high-performance blending systems, autonomous acceptance and build trust.

VII. Conclusion and Prospects

This study has outlined guidelines for driving simulation experiments for data aquisition in order to designing driver behavior classification algorithms. In addition, a pilot study was conducted to show an effective data labeling process and ANN training. The results showed better performance for individualized approaches (where the level of aggressiveness at each control is identified individually). Next projects should incorporate such classification algorithms to actual vehicular control systems listed in the applications sections for evaluating the benefits of such information. For instance, energy management systems might leverage the classification of aggressiveness at the throttle and chassis control systems might be designed for using the classification at the steering wheel and brake pedal.

References

[1] L. Bruck, B. Haycock, and A. Emadi, "A review of driving simulation technology and applications," *IEEE Open Journal of Vehicular Technology*, vol. 2, pp. 1–16, 2021.

[2] S. M. Pampel, S. L. Jamson, D. L. Hibberd, and Y. Barnard, "How i reduce fuel consumption: An experimental study on mental models of eco-driving," *Transportation Research Part C: Emerging Technologies*, vol. 58, pp. 669–680, 2015.

[3] Y. Wu, X. Zhao, J. Rong, and Y. Zhang, "How eco-driving training course influences driver behavior and comprehensibility: a driving simulator study," *Cognition, Technology & Work*, vol. 19, no. 4, pp. 731–742, 2017.

[4] A. Jamson, D. L. Hibberd, and N. Merat, "The design of haptic gas pedal feedback to support eco-driving," in *Proceedings of the Seventh International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design*. University of Iowa, 2013, pp. 264–270.

[5] J. Yoon, W. Cho, J. Kang, B. Koo, and K. Yi, "Design and evaluation of a unified chassis control system for rollover prevention and vehicle stability improvement on a virtual test track," *Control Engineering Practice*, vol. 18, no. 6, pp. 585–597, 2010.

[6] A. Ali, A. Elnaggarz, D. Reichardtz, and S. Abdennadher, "Gamified virtual reality driving simulator for asserting driving behaviors," in *2016 1st International Conference on Game, Game Art, and Gamification (ICGGAG)*. IEEE, 2016, pp. 1–6.

[7] T. Imamura, T. Ogi, E. T. C. Lun, Z. Zhang, and T. Miyake, "Trial study of traffic safety education for high school students using driving simulator," in *2013 IEEE International Conference on Systems, Man, and Cybernetics*. IEEE, 2013, pp. 4606–4611.

[8] S. Yun, T. Teshima, and H. Nishimura, "Human–machine interface design and verification for an automated driving system using system model and driving simulator," *IEEE Consumer Electronics Magazine*, vol. 8, no. 5, pp. 92–98, 2019.

[9] L. Morra, F. Lamberti, F. G. Pratticó, S. La Rosa, and P. Montuschi, "Building trust in autonomous vehicles: Role of virtual reality driving simulators in hmi design," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 10, pp. 9438–9450, 2019.

[10] E. Benedito and A. Dòria-Cerezo, "Influence of cooperative-controlled driving in the traffic flow," in *2018 IEEE International Conference on Industrial Technology (ICIT)*. IEEE, 2018, pp. 1795–1800.

[11] M. Almallah, R. Alfahel, Q. Hussain, W. K. Alhajyaseen, and C. Dias, "Empirical evaluation of drivers' start-up behavior at signalized intersection using driving simulator," *Procedia Computer Science*, vol. 170, pp. 227–234, 2020.

[12] D. Michalík, M. Jirgl, J. Arm, and P. Fiedler, "Developing an unreal engine 4-based vehicle driving simulator applicable in driver behavior analysis—a technical perspective," *Safety*, vol. 7, no. 2, p. 25, 2021.

[13] C. M. Martinez, M. Heucke, F.-Y. Wang, B. Gao, and D. Cao, "Driving style recognition for intelligent vehicle control and advanced driver assistance: A survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 3, pp. 666–676, 2017.

[14] S. Sarker and M. M. Haque, "A data change rule based empirical framework for labeling unlabeled time series driving data," in *2020 2nd International Conference on Advanced Information and Communication Technology (ICAICT)*, 2020, pp. 476–479.

[15] S. Fernandez and T. Ito, "Driver classification for intelligent transportation systems using fuzzy logic," in *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, 2016, pp. 1212–1216.

[16] C. Miyajima, Y. Nishiwaki, K. Ozawa, T. Wakita, K. Itou, K. Takeda, and F. Itakura, "Driver modeling based on driving behavior and its evaluation in driver identification," *Proceedings of the IEEE*, vol. 95, no. 2, pp. 427–437, 2007.

[17] C. Tran, A. Doshi, and M. M. Trivedi, "Modeling and prediction of driver behavior by foot gesture analysis," *Computer Vision and Image Understanding*, vol. 116, no. 3, pp. 435–445, 2012.

[18] C. Deng, C. Wu, N. Lyu, and Z. Huang, "Driving style recognition method using braking characteristics based on hidden markov model," *PloS one*, vol. 12, no. 8, p. e0182419, 2017.

[19] M. H. Alkinani, W. Z. Khan, and Q. Arshad, "Detecting human driver inattentive and aggressive driving behavior using deep learning: Recent advances, requirements and open challenges," *IEEE Access*, vol. 8, pp. 105 008–105 030, 2020.

[20] Z.-J. Cheng, L.-W. Jeng, and K. Li, "Behavioral classification of drivers for driving efficiency related adas using artificial neural network," in *2018 IEEE International Conference on Advanced Manufacturing (ICAM)*, 2018, pp. 173–176.

[21] X. Li, W. Wang, and M. Roetting, "Estimating driver's lane-change intent considering driving style and contextual traffic," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 9, pp. 3258–3271, 2019.

[22] F. Martinelli, F. Marulli, F. Mercaldo, and A. Santone, "Neural networks for driver behavior analysis," *Electronics*, vol. 10, no. 3, p. 342, 2021.

[23] N. Abdennour, T. Ouni, and N. B. Amor, "Driver identification using only the can-bus vehicle data through an rcn deep learning approach," *Robotics and Autonomous Systems*, vol. 136, p. 103707, 2021.

[24] B. Haycock, J. Campos, N. Koenraad, M. Potter, and S. Advani, "Creating headlight glare in a driving simulator," *Transportation research part F: traffic psychology and behaviour*, vol. 61, pp. 93–106, 2019.

[25] F. Colombet, D. Paillot, F. Mérienne, and A. Kemeny, "Visual scale factor for speed perception," *Journal of Computing and Information Science in Engineering*, vol. 11, no. 4, 2011.

[26] C. Zöller, A. Müller, L. Eggert, H. Winner, and B. Abendroth, "Applicability of head-mounted displays in driving simulation," in *Proceedings of the Driving Simulation Conference 2019 Europe VR*, A. Kemeny, F. Colombet, F. Merienne, and S. Espié, Eds. Strasbourg, France: Driving Simulation Association, 2019, pp. 9–15.

[27] R. Ramkhalawansingh, B. Keshavarz, B. Haycock, S. Shahab, and J. L. Campos, "Age differences in visual-auditory self-motion perception during a simulated driving task," *Frontiers in Psychology*, vol. 7, p. 595, 2016. [Online]. Available: https://app.dimensions.ai/details/publication/pub.1014944210andhttps://www.frontiersin.org/articles/10.3389/fpsyg.2016.00595/pdf

[28] E. Blana. (1996, Dec) A survey of driving research simulators around the world. Institute for Transport Studies, University of Leeds. Leeds, UK. [Online]. Available: http://eprints.whiterose.ac.uk/2110/

[29] N. Mohajer, H. Abdi, K. Nelson, and S. Nahavandi, "Vehicle motion simulators, a key step towards road vehicle dynamics improvement," *Vehicle System Dynamics*, vol. 53, no. 8, pp. 1204–1226, 2015.

[30] J. Pitz, M.-T. Nguyen, G. Baumann, and H.-C. Reuss, "Combined motion of a hexapod with a xy table system for lateral movements," in *presented at Driving Simulation Conference*, 2014.

[31] S. Kharrazi, B. Augusto, and N. Fröjd, "Vehicle dynamics testing in motion based driving simulators," *Vehicle system dynamics*, vol. 58, no. 1, pp. 92–107, 2020.

[32] M. Bruschetta, C. Cenedese, and A. Beghi, "A real-time, mpc-based motion cueing algorithm with look-ahead and driver characterization," *Transportation research part F: traffic psychology and behaviour*, vol. 61, pp. 38–52, 2019.

[33] Y. Chiew, M. A. Jalil, and M. Hussein, "Kinematic modeling of driving simulator motion platform," in *2008 IEEE Conference on Innovative Technologies in Intelligent Systems and Industrial Applications*. IEEE, 2008, pp. 30–34.

[34] L. Bruck, S. Veldhuis, and A. Emadi, "Selection method of a driving simulator motion system," in *2019 IEEE Transportation Electrification Conference and Expo (ITEC)*. IEEE, 2019, pp. 1–6.

[35] Vi-worldsim: A new, user-friendly, fully integrated graphic environment that accelerates vehicle development using driving simulators. [Accessed: 2021-09-30]. [Online]. Available: https://www.vi-grade.com/en/products/vi-worldsim/

[36] Vi-carrealtime: One vehicle model from concept to sign-off. [Accessed: 2021-09-27]. [Online]. Available: https://www.vi-grade.com/en/products/vi-carrealtime/

[37] H. Pacejka, *Tire and vehicle dynamics*. Elsevier, 2005.

[38] D. Filev, J. Lu, K. Prakah-Asante, and F. Tseng, "Real-time driving behavior identification based on driver-in-the-loop vehicle dynamics and control," in *2009 IEEE International Conference on Systems, Man and Cybernetics*. IEEE, 2009, pp. 2020–2025.

[39] D. Dörr, D. Grabengiesser, and F. Gauterin, "Online driving style recognition using fuzzy logic," in *17th international IEEE conference on intelligent transportation systems (ITSC)*. IEEE, 2014, pp. 1021–1026.

[40] D. Arthur and S. Vassilvitskii, "k-means++: The advantages of careful seeding," Stanford, Tech. Rep., 2006.

[41] Math. graphics. programming. matlab is a programming and numeric computing platform used by millions of engineers and scientists to analyze data, develop algorithms, and create models. [Accessed: 2021-09-27]. [Online]. Available: https://www.mathworks.com/products/matlab.html

[42] M. Haußmann, D. Barroso, C. Vidal, L. Bruck, and A. Emadi, "A novel multi-mode adaptive energy consumption minimization strategy for p1-p2 hybrid electric vehicle architectures," in *2019 IEEE Transportation Electrification Conference and Expo (ITEC)*, 2019, pp. 1–6.

[43] C. Vidal, P. Malysz, P. Kollmeyer, and A. Emadi, "Machine learning applied to electrified vehicle battery state of charge and state of health estimation: State-of-the-art," *IEEE Access*, vol. 8, pp. 52 796–52 814, 2020.

[44] Y. L. Murphey, J. Park, Z. Chen, M. L. Kuang, M. A. Masrur, and A. M. Phillips, "Intelligent hybrid vehicle power control—part i: Machine learning of optimal vehicle power," *IEEE Transactions on Vehicular Technology*, vol. 61, no. 8, pp. 3519–3530, 2012.

[45] A. Emadi, *Advanced electric drive vehicles*. Boca Raton, FL: CRC Press, Oct 2014.

[46] L. Bruck and A. Emadi, "A new regen-based energy management strategy for online control of hybrid powertrains," in *2021 IEEE Transportation Electrification Conference Expo (ITEC)*, 2021, pp. 762–766.

[47] S. Taxonomy, "Definitions for terms related to on-road motor vehicle automated driving systems," *Society of Automotive Engineers (SAE): Troy, MI, USA*, 2014.

Lucas Bruck (Student Member, IEEE) received his B.S. in Mechanical Engineering in 2014, from the Federal University of Sao Joao del-Rei, Brazil (UFSJ). Following his graduation, he worked in industry with virtual vehicle dynamics, driving simulation, and calibration of chassis control systems. In 2018 he has joined the McMaster Automotive Resource Centre (MARC) at Mcmaster University as a M.S. student and is currently persuing his Ph.D. degree at the same university. His main field of expertise is energy management, vehicle dynamics, chassis active controls and safety systems.

PLACE PHOTO HERE

Carlos Vidal (Member, IEEE) Insert Bio

Gary Newton, Jr (Member, IEEE) is the Director for Simulator Sales in the Americas where he engages with customers across the automotive industry to advance virtual prototyping and validation. Gary strongly believes that VI-grades simulator technology can help bring vehicles to market faster in a way that is less expensive and safer. He seeks opportunities to assist engineering teams to embrace the digital transformation of the automotive development process. Gary has more than 28 years of automotive experience and multiple technical credits in both conferences and peer reviewed publications and journals. He also supports student engineering efforts to help develop the next generation of automotive engineers. Currently Gary is a member in good standing with SAE supporting many conferences and events as an organizer and technical contributor. This year he was added a contributing member to SAE's Vehicle Dynamics Standards Committee and was nominated for the SAE McFarland Award for his technical contributions. Recently Gary become a member of IEEE and The Tire Society. Gary holds degrees in business and marketing from Schoolcraft College and has attended multiple automotive industry training events.

Ali Emadi (Fellow, IEEE) received the B.S. and M.S. degrees in electrical engineering with highest distinction from the Sharif University of Technology, Tehran, Iran, in 1995 and 1997, respectively, and the Ph.D. degree in electrical engineering from Texas A&M University, College Station, TX, USA, in 2000. He is the Canada Excellence Research Chair Laureate at McMaster University in Hamilton, Ontario, Canada. He is also the holder of the NSERC/FCA Industrial Research Chair in electrified powertrains and Tier I Canada Research Chair in transportation electrification and smart mobility. Before joining McMaster University, Dr. Emadi was the Harris Perlstein Endowed Chair Professor of Engineering and Director of the Electric Power and Power Electronics Center and Grainger Laboratories at the Illinois Institute of Technology in Chicago, where he established research and teaching facilities as well as courses in power electronics, motor drives, and vehicular power systems. He was the Founder, Chairman, and President of Hybrid Electric Vehicle Technologies, Inc. (HEVT) - a university spin-off company of Illinois Tech. Currently, he is the President and Chief Executive Officer of Enedym Inc. and Menlolab Inc.–two McMaster University spin-off companies. He is the principal author/coauthor of over 500 journal and conference papers as well as several books including Vehicular Electric Power Systems (2003), Energy Efficient Electric Motors (2004), Uninterruptible Power Supplies and Active Filters (2004), Modern Electric, Hybrid Electric, and Fuel Cell Vehicles (2nd ed, 2009), and Integrated Power Electronic Converters and Digital Control (2009). He is also the Editor of the Handbook of Automotive Power Electronics and Motor Drives (2005) and Advanced Electric Drive Vehicles (2014). He is the co-editor of the Switched Reluctance Motor Drives (2018). Dr. Emadi was the Inaugural General Chair of the 2012 IEEE Transportation Electrification Conference and Expo (ITEC) and has chaired several IEEE and SAE conferences in the areas of vehicle power and propulsion. He was the founding Editor-in-Chief of the IEEE TRANSACTIONS ON TRANSPORTATION ELECTRIFICATION from 2014 to 2020.