

RESEARCH

Open Access



Feature selection for driving fatigue characterization and detection using visual- and signal-based sensors

Khadidja Henni^{1*}, Neila Mezghani², Charles Gouin-Vallerand¹, Perrine Ruer¹, Youssef Ouakrim² and Évelyne Vallières¹

*Correspondence:

khadidja.henni@liceftelug.ca

¹ Centre de recherche LICEF, université TELUQ, 5800 Rue St-Denis, Montréal, QC, Canada

Full list of author information is available at the end of the article

Abstract

Driving fatigue detection has been the subject of several studies which relied on a set of features describing driver's facial expressions, driving behaviors, and bio-signals. The purpose of this study is to improve driving fatigue detection by identifying the adequate set of features which accurately characterize fatigued drivers. The considered features are derived from non-intrusive sensors; they are related to the changes in driving behavior and visual facial expressions. The relevance is first investigated by several feature selection methods. Second, a meta-analysis was performed to investigate method agreement about the relevance of each feature in driving fatigue recognition. Support vector machine and DBSCAN classifiers were used to detect fatigue by means of the identified features. Experimental analyses are performed on a real-world database, collected through the computer vision system "FaceLab" and car sensors, from 66 senior drivers when driving an instrumented car on a highway. Results provide a list of the most discriminative features, which enhances the classification average accuracy to 89.13%.

Keywords: Driving fatigue, Head and eye tracking, Facial expression signals, Driver's visual behavior, Feature selection methods

Introduction

Nowadays, about 1.25 million people die each year as a result of road traffic crashes (World Health Organization WHO). Without sustained action, road traffic crashes are predicted to become the leading cause of death by 2030, including among the growing population of senior drivers. 10–20% of all traffic accidents are made by non-vigilant drivers and 60% of these accidents are due to fatigue (Awake Consortium 2001). Fatigue is defined as a transitional state between awakenings and sleeping. It affects the skills required for a safe driving, by increasing the driving error frequency as well as their amplitudes and variability. It also reduces the driver perception and decision-making capability to control the vehicle (Sahayadhas et al. 2012). Fatigue depends on several endogenous and exogenous factors, for example, age, motivation and driving time. Otmani et al. (2005) demonstrated that drivers are fatigued after 20 min of driving, and countermeasures must be considered before the driver reaches

the final stage of fatigue 'somnia' However, the problem is due to the difficulty for the driver to perceive his real level of fatigue, as it is sensation based and subjective. Authors in Sáez-Francàs et al. (2014) confirm the risk of non-fatigue perception and how it affects decision making.

In intelligent transportation systems (ITS) field, several research projects have focused on the detection of the driver's state (fatigue, sleepiness, and inattention), which can reduce accidents and improve the public safety (Azim et al. 2014; Bergasa and Nuevo 2006; Ji et al. 2004; Jin et al. 2013). Nowadays, ITS community is giving wide attention to the development of autonomous cars, but despite the progress of these researches, the trust to adopt the driving of fully autonomous cars is not high among many potential users (older adults), whereas greater than 50% of these users prefer helper cars (Abraham et al. 2016), which can be considered as countermeasures when driver's state is unsafe. Hence, the automatic identification of the driver's state, especially fatigue, is still a challenging task.

The main objective of driver's fatigue identification researches was to propose prototypes which capture, track and store driver behaviors signals. These signals are then analyzed and various characteristics are extracted to describe driver behaviors and identify its state. Several features have been used to detect driving fatigue, especially those describing visual facial expressions signals. The relationships between driving fatigue and driver behavior characteristics have been investigated experimentally and performed on a limited set of features such as PERCLOS, head nodding (Of and Carriers 1998), steering wheel movements (Otmani et al. 2005) and the standard deviation of lane position (Thiffault and Bergeron 2003). The use of these features is not substantiated, then the results reported in several studies are not consistent, although they are based on the same datasets (Azim et al. 2014; Bergasa and Nuevo 2006). In addition, some conclusions have been considered with respect to simulation studies, although they are not necessarily applicable and nor meaningful in real-world driving.

Moreover, the development of ITS applications is limited by certain constraints, such as (1) the response time: if the fatigue detection time is slow, the drivers can fall asleep and the risk of accident is greater, (2) the accuracy: it is important to detect rigorously fatigue, without distracting or stressing the driver by false alarms, and (3) the computational and storage costs: ITS are generally embedded systems, which limit their storage capacities. In addition to that, classification techniques are more accurate when they are provided with more relevant features, but are slower and suffer from the curse of dimensionality phenomenon (Friedman and Bias 1997). Therefore, it is important to reduce the features into a relevant, non-redundant subset and without loss of information, as well as providing a clear and precise definition of these features.

The aim of this paper was first, to list all quoted features and standardize their formulations. Second, to investigate a hypothesis about the relationships between driving fatigue and the vehicle speed in real driving environment, a statistical analysis was performed on driving behaviors, and significant changes were reported in the use of acceleration pedal. Therefore, new features are considered to characterize fatigue. Finally, a subset of relevant features that would best characterize driving fatigue is identified, and classification algorithms demonstrate their efficiency in discriminating fatigued and non-fatigued drivers.

To investigate the importance of fatigue driving features, different feature selection techniques (FS) have been combined to design an ensemble learning for ranking (Duin 2002; Guan et al. 2014). FS methods are effective to lessen the curse of dimensionality problem, they are used in several data engineering fields, such as medical image analysis (Rakoczy et al. 2013), bio-informatics and gene expression data analysis (Mitra and Majumder 2004), as well as text mining and natural language processing (Rahate and Emmanuel 2013).

This research is interested on senior drivers who are vulnerable to fatigue driving and potential customers of automated car or especially helper cars. The experiments are carried out on a fairly large and diversified database. Data are signals collected through the eye tracking system "FaceLab" and car sensors, from a real driving sample of 66 senior drivers on a highway (around 1 h driving for each driver with some wearing glasses), under different climatic and lighting conditions.

The remainder of this paper is organized as follows. The previous works in driver state detection are described in "Related works" section. The driving fatigue characterization tools, the dataset, the experimental protocol and the considered features, both existing and proposed are presented in "Method" section. "Experimental results" section presents various established experiments. Finally, "Conclusion" section concludes the paper and discusses the approaches for future studies.

Related works

Within the last years, several research works have investigated driving fatigue, they can be regrouped according to the used technology: (1) intrusive systems (Engström et al. 2005), which measure fatigue based on bio-signals such as heart rate collected by electrocardiogram (ECG) sensors, skin temperature, etc., all of which are efficient systems but less usable because of the body-mounted sensors. (2) Non-intrusive systems use cameras or eye-tracking systems based on infrared vision to extract face features describing fatigue, or/and, use car sensors mounted on the acceleration pedal and the steering wheel, which measure the driving behavior such as steering wheel movement (SWM), standard deviation of lane position (SDLP) (Engström et al. 2005; Joly et al. 2018).

Facial expressions are the most used characteristics in fatigue detection. The work of Ji et al. (2004) is one of the first studies that investigated driving fatigue through technological tools. Authors used the means-shift algorithm and Kalman filter to extract six visual expressions from videos: PERCLOS, average eye closure speed, frequency of head nodding, percentage of saccades, distribution of gaze in time and frequency of yawing. These features were modeled using Bayesian Networks and evaluated on a randomly selected image sequence that contains 13620 frames. Authors in Sigari et al. (2013) have extracted signals from 27 video sequences, captured in real conditions. Videos are preprocessed and several features are computed such as PERCLOS, eyelid distance changes (ELDC), eye closure rate and head rotation (ROT). Then, the fuzzy expert system was modeled to detect driver fatigue and distraction. 10 drivers were the subject of experiments conducted by Bergasa and Nuevo (2006) to detect the driver alertness by a fuzzy expert system. They have analyzed the facial expressions related to the eyes: PERCLOS, eye closure duration, blink frequency, and to the

head as: head nodding frequency, face position, fixed gaze. These features enabled the detection of alertness at 98% accuracy. But on the same drivers (the driving videos database is public), in Azim et al. (2014) have shown that based on only two features: (1) PERCLOS and (2) Mouth state, tracked and extracted by the Kalman filter and the Adaboost algorithm, the fuzzy expert system can detect the driver state with the accuracy of up to 100%. A recent study used on a set of face landmark and textures of the eye and mouth regions to detect driver drowsiness. Features are tracked and extracted by Viola–Jones and fused by the Deep belief network. The experiments performed on 30 facial videos of passengers seat on the right of drivers, subjects are subjected to 6–8h of straight sleep deprivation, demonstrated a high accuracy in drowsiness detection (Zhao et al. 2018).

After several attempts and contributions employing machine-learning techniques such as mean-shift, Adaboost and Viola–Jones, as well as other image analysis tools. Recent works and automobile manufacturers are addressing more eye-tracking systems, like: (1) ‘FaceLAB’, which is able to maintain tracking integrity even lighting and movement perturbations, this technology is used in recent prototypes of vehicle systems as Volkswagen. (2) ‘Smart eyes’, which is commercialized in Europe by Volvo and BMW. It is also used in Ahlstrom et al. (2013) to extract information about the gaze direction and its quality. Ahlstrom et al. (2013) used AttenD to study the driver distraction, on a database of seven drivers. Smart eyes has also been used by Jin et al. (2013), in a simulator to extract blink frequency, gaze direction, fix time and the PERCLOS, then these features are used by Support Vector Machine (SVM) to detect sleepiness among twelve drivers. Miyaji and Kawanaka (2010) have used Adaboost and SVM to classify bio-signals and facial expression data coming from ECG and the eye tracking system FaceLAB. Data are simulation experiment results, where cognitive and arithmetic charges are applied on eight drivers. Engström et al. (2005) used data from 48 drivers in a simulator and 24 in a real driving environment, to illustrate the effects of visual and cognitive loads on driver performances, when driving simulators and real vehicles on highways, data were collected by FaceLab. The proposed prototype is based on facial expressions related to the gaze tracking, as well as to bio-signals coming from intrusive systems (EEG) and to vehicle behavior (maintaining voice and steering wheel).

Vehicle-based measures (driving behavior) have been also involved in fatigue detection. Micro-corrections in the steering wheel are reduced for drowsy drivers, following this hypothesis the SWM and SDLP have been the most used features, SWM is derived from steering wheel angle signal and it has been adopted by Nissan and Renault manufacturers. The efficiency of these features has been demonstrated in simulations studies (Sahayadhas et al. 2012). However, driving and road conditions related to steering wheel micro-corrections are not the same in simulators and real-world environment, for instance, driver’s risk perception is reduced in simulators, straight line does not always exist in road, and the width of roads and vehicles is variable and controls the extent of micro-corrections, the possibility to make steering wheel micro-corrections and their extent when there is an overtaking manoeuvre. In addition to that, these features are not specific to fatigue and can be caused by other impaired driving types such the influence of alcohol and drugs (Sahayadhas et al. 2012). Therefore, SWM and SDLP are less-effective features for predicting fatigue in real-word driving.

In short, a large amount of research works have studied the human behavior while driving, using different tools to capture and extract the data (eye-tracking systems, camera with videos processing and machine learning tools, etc.); these data were processed to extract various features such as PERCLOS (Bergasa and Nuevo 2006; Ji et al. 2004; Jin et al. 2013; Kong et al. 2015; Friedrichs and Yang 2010; Darshana et al. 2015; Sun et al. 2017) and blink frequency (Bergasa and Nuevo 2006; Jin et al. 2013; Friedrichs and Yang 2010; Sun et al. 2017, 2015). (More details are giving in "Driver's behavior features extraction" section). Finally, regarding the driver-state identification, three main approaches have been adopted: threshold and probabilistic (Ji et al. 2004), expert system (Azim et al. 2014; Bergasa and Nuevo 2006), and classification (Ji et al. 2004; Jin et al. 2013; Kong et al. 2015; Friedrichs and Yang 2010; Miyaji and Kawanaka 2010; Zhao et al. 2018). In this work, a literature study revealed that works use different features, chosen randomly or according to an unjustified cause. In addition, these works give different definitions to the same feature, for example, Kong et al. (2015) defines the PERCLOS as being percentage of times in 30 s, when the eye is closed at 30%; while others define it as the proportion of time in 3 min, when the eyes are at least 80% closed. Thus, we gather a list of all features related to fatigue, this list included the most used features in the literature and which are sorted by FS methods, according to their order of importance (relevance and non-redundant). It is validated on a set of 66 senior drivers and the database was collected by the eye-tracking system 'seeing machines' FaceLAB, while above research works used small databases composed of at least 30 real drivers.

Method

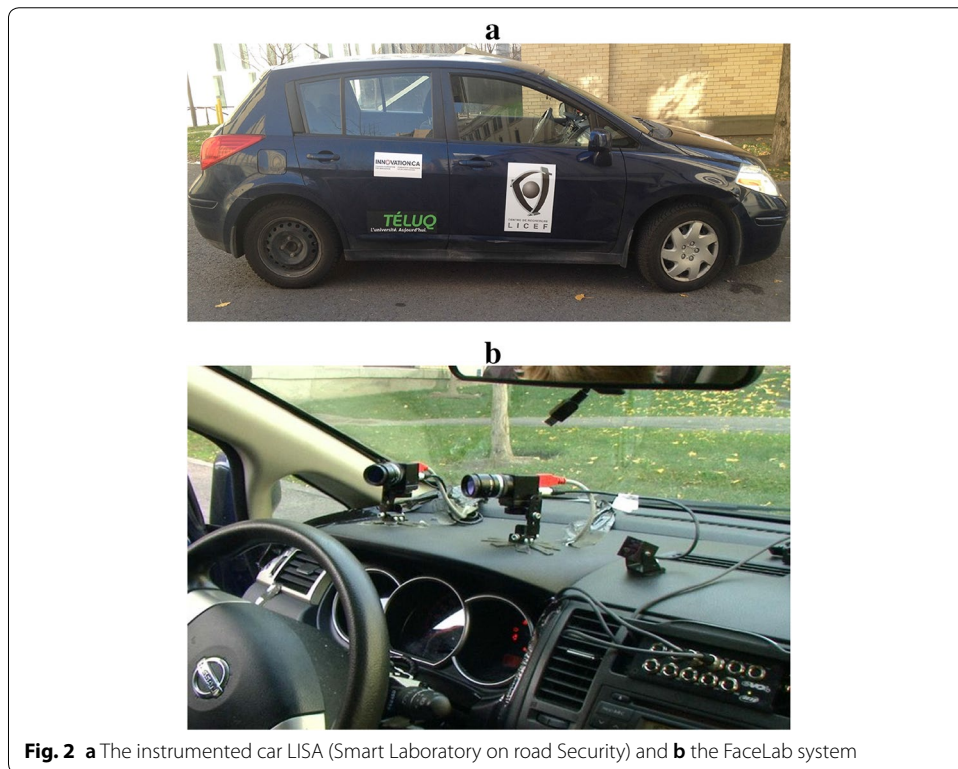
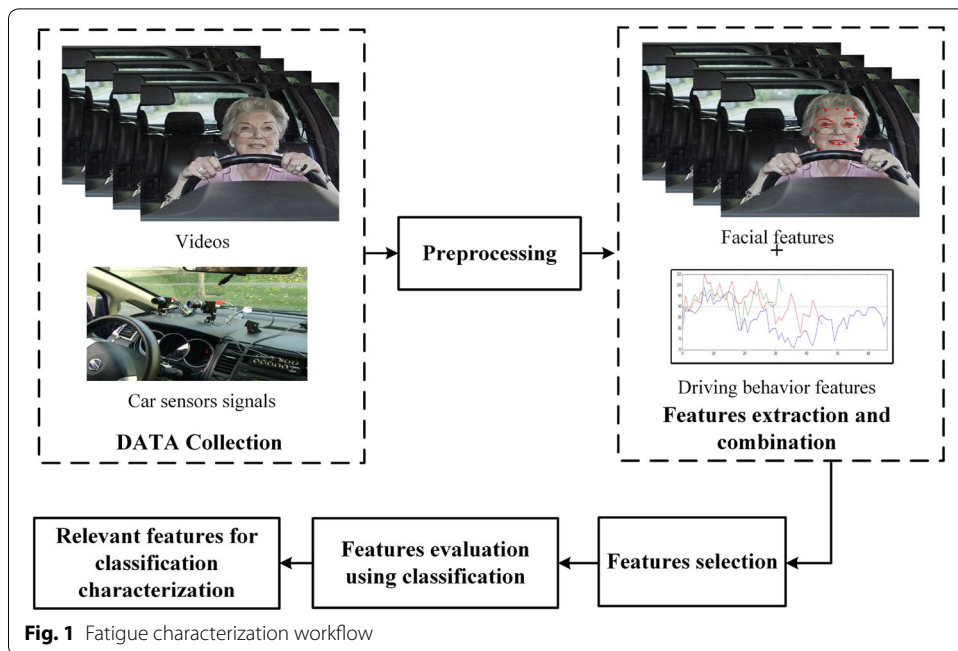
In this paper, we characterize driving fatigue using feature selection and classification algorithms. The used tools as well as the processes of database collection and preprocessing are presented in this section. Figure 1 illustrates the workflow followed in this analysis study.

Data collection

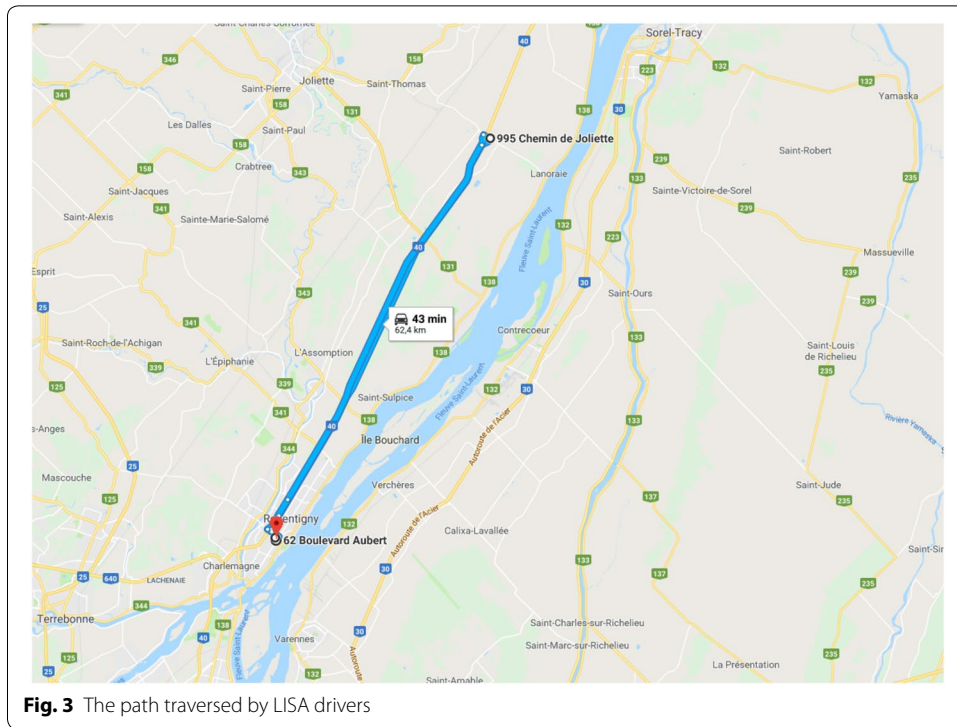
Sixty six participants (30 women and 36 men) have been recruited to drive for about 1 h, they are over 55 years old and have their driving license for 30–47 years. Participants should drive an instrumented car (LISA: Smart Laboratory on road Security, Fig. 2), on a highway in Montreal city in Canada (Fig. 3), the experiments were spread in different hours of the day (10 a.m.–3 p.m.) and over several months (from June to November 2016), which have led to climatic conditions diversities.

LISA (Fig. 2) is a Nissan Versa 2008, 1.8 S, with automatic transmission, equipped by FaceLAB 5.0 system (2 IR-camera) and car sensors such as speed, temperature, pressure, and suspension, which are connected to Evo4 DataLogger. All data are synchronized and recorded by a computer installed on board. FaceLAB captures 60 frames/s and represents outputs in separated files: eyes, head, timing, etc. Car sensors generate vehicle speed and steering wheel movements signals. Data are captured with 10 frames/s frequency.

Driving fatigue is evaluated based on the self-reported fatigue given five times by drivers, twice before and after driving, and three times during experiments. The perceived fatigue scale was inspired by the Karolinska sleepiness scale (Kaida et al. 2006), it is



ranging from 0 not at all fatigued to 10 very much fatigued. Participants are asked only three times while driving, to avoid distracting the driver, that seems sufficient to have an idea about the driver-state changes, and to classify the resulting instances, we use the linear spline interpolation (Habermann and Kindermann 2007) to obtain the whole



data labels. We define two classes: (1) non-fatigued driver, if its self-ratings < 3 and (2) fatigued driver otherwise.

All drivers gave their consent to participate, which was approved by institutional ethics committees.

Data preprocessing

The collected data are multi-modal, coming from different sensors and having different frequencies. Therefore, signals are treated separately and then merged.

First, the FaceLab signals are converted, synchronized and the signals represented in Table 1 are retrieved. Eyes tacking confidence is given by FaceLAB, this information allowed us to filter the eye and head rotation signals as follows: (1) the frames where

Table 1 Signal extracted from FaceLab

Signal	Description
$e_{l,r}$	Eye closure (left and right)
$c_{l,r}$	Eye closure confidence (left and right)
Blink-freq	Blink frequency
Blink-dur	Average blink duration
PERCLOS	Percentage eyelid closure at 75%
φ, ψ	Gaze rotation: pitch, yaw angle
Sac	Saccade
$P_diam_{l,r}$	Pupil diameter (left and right)
$Q_{0,1,2,3}$	Head rotation: quaternions

$(c_L + c_R)/2 < 55\%$ are considered as noise, (2) the head signal is filtered through the head yaw angle $\beta > 15^\circ$. These thresholds are well documented in the literature (Friedrichs and Yang 2010).

Second, vehicle speed and steering wheel movements signals are extracted from the acceleration pedal and the steering wheel sensors, then they are converted and prepared to be merged with FaceLab data.

Third, data are aggregated in windows of 30 s and the considered features are computed (more details about features are presented in “[Driver’s behavior features extraction](#)” section).

Driver’s behavior feature extraction

In this study, we use non-intrusive sensors to identify the most relevant features for driver’s fatigue characterization. Therefore, we focus on facial expressions and driving behavior features.

Several features have been quoted in the previous researches, especially those related to eye and head behaviors. Table 2 summarizes the most quoted features as well as gives their significance and corresponding formulas.

Note that, N is the number of valid frames in a window of 30 s, $N_{\text{closure75}}$ is the number of frames where the eyelid distance is equal to 75% of the total eye size. $N_{\text{Total closure}}$ and $N_{\text{Total opening}}$ are the number of frames where the eye is totally closed or opened, respectively, i.e, if its closure is less than 20% of the eye size, or more than 80% of the eye size. The $N_{\text{Transition}}$ is the number of frames where the eye passes from opened/closed. These measures are estimated based on the eye closure signal e (left e_L and right e_R) normalized by the $c_{L,R}$ and fused. φ_{head} and ψ_{head} are the pitch and yaw angle of head, the results of head rotation conversion from quaternions to Euler angles. $N_{\text{head frontal position}}$ is equal to 1, if the horizontal head angle is less than 15° .

Regarding the investigation involving the effectiveness of vehicle-based features to predict driving fatigue in real-world, we test the hypotheses which consists of the existence of relationships between the changes in driving speed (the use of acceleration pedal) and driver fatigue, which increases over the time. In fact, it has been demonstrated that drivers can become fatigued within 20–25 min of driving (Otmami et al. 2005; Sahayadhas et al. 2012).

Individual differences were analyzed separately and have lead to these hypotheses. Therefore, we have attempted to verify our hypothesis about the effect of fatigue on driving behaviors (the speed use) on all drivers, which have driven for about 1h. Thus, two driving periods are considered : (1) outward step and (2) return step. Then, Student t test is applied to verify the existence of significant differences in driving behaviors between the outward and return periods. Several driving behavior variables have been considered and are derived from the speed signal, such as minimum, maximum, |minimum–maximum|, mean, standard deviation, variance, norm, norm1, MaxFFT, Skewness, and Kurtosis of speed. Then, other features have been investigated according to the peaks in this signal, such as minimum, maximum, |minimum–maximum|, mean, standard deviation and variance of peak extent, and the number of peaks.

In addition to the features proposed in literature (summarized above), the variables that demonstrated significant differences, with p value less than 0.05 (summarized on Table 3), have been also considered in our driving fatigue characterization investigations.

Table 2 Features characterizing driving fatigue in literature

Id		Feature descriptions	
F_1	PERCLOS (F_1) (Bergasa and Nuevo 2006; Ji et al. 2004; Jin et al. 2013; Kong et al. 2015; Friedrichs and Yang 2010; Darshana et al. 2015; Sun et al. 2017): represents the percentage of eyelid closure at 75%: $PERCLOS_{75} = \frac{N_{closure75}}{N}$		
F_2	Eye closure duration (ECD) (F_2) (Azim et al. 2014; Bergasa and Nuevo 2006; Kong et al. 2015; Sun et al. 2017, 2015): represents the number of frames where the eye is totally closed. $ECD = \sum N_{Total\ closure}$		
F_3	Pupil diameter (PupilDiam) (F_3) (Miyaji and Kawanaka 2010): represents the diameter of eye iris contour. $PupilDiam = average(P_diam_l, P_diam_r)$		
F_4	Percentage of saccade (Persac) (F_4) (Ji et al. 2004; Liang et al. 2007): represents the percentage of eye pupil saccade		
F_5	Average eye closure speed (AECS) (F_5) (Bergasa and Nuevo 2006; Ji et al. 2004; Friedrichs and Yang 2010): represents the number of frames where the eye changed its state (opened/closed) $AECS = \frac{N_{Total\ closure} + N_{Total\ opening}}{N_{Transition}}$		
F_6	Microsleep (F_6) (Friedrichs and Yang 2010): Microsleep during 30 s $Microsleep = \sum N_{Total\ closure} / N$		
F_7	Eye opened (EOP) (F_7) (Sun et al. 2017, 2015): eye opened during 30 s $EOP = \sum N_{Total\ opening} / N$		
F_8	Gaze angle distribution (Gaz_{dis}) (Ji et al. 2004; Engström et al. 2005; Jin et al. 2013; Darshana et al. 2015; Miyaji and Kawanaka 2010): $Gaze = \sqrt{\varphi^2 + \psi^2}$		
F_9	Gaze angle variation (Gaz_{var}) (Ji et al. 2004): $Gaz_{var} = \sigma(Gaze_{dis})$		
F_{10}	FixedGaze (Bergasa and Nuevo 2006; Liang et al. 2007): it is a Boolean variable, if $Gaze_{dis}$ is stable during 30 s $FixedGaze = 1$, otherwise $FixedGaze = 0$		
F_{11}	Blink frequency ($Blink_{freq}$) (Bergasa and Nuevo 2006; Jin et al. 2013; Friedrichs and Yang 2010; Sun et al. 2017, 2015): $Blink_{freq} = \frac{N_{Total\ closure}}{N}$		
F_{12}	Blink duration (Blink-dur) (Friedrichs and Yang 2010; Hu and Zheng 2009): represents the blink time amount		
F_{13}	Distribution of head tilt angle (Head-tilts-dis) (Bergasa and Nuevo 2006; Ji et al. 2004): $HeadTilts = \sqrt{\varphi_{head}^2 + \psi_{head}^2}$		
F_{14}	Variation of head tilt angle (Head-tilts-var) (Bergasa and Nuevo 2006): $HeadTiltsVar = \sigma(HeadTilts)$		
F_{15}	Head nodding frequency (Head-nod-freq) (Bergasa and Nuevo 2006; Ji et al. 2004): represents the number of peaks in the head tilt angle distribution		
F_{16}	FacePose (Bergasa and Nuevo 2006): represents the number of frames where driver head is in frontal position $\sum N_{head\ frontal\ position}$		
F_{17}	Steering wheel movement (SWM) (Engström et al. 2005; Joly et al. 2018; Thiffault and Bergeron 2003; Sahayadhas et al. 2012; Ingre et al. 2006): represents the variability in steering wheel movements		

Features selection by ensemble learning

The FS field is motivated by the well-known phenomenon of the ‘curse of dimensionality’ which is very recurrent in recent data, as well as by the enhancement of learning generalization over the whole test set. It consists of ranking features according to their relevance and discarding the redundant ones.

Table 3 The proposed driving behavior features

Id	Features	Outward phase		Return phase		t test p value
		Mean	Std	Mean	Std	
F_{18}	Mean (speed)	95.591	4.603	93.386	7.428	0.041
F_{19}	MaxFFT (speed)	132.813	35.179	168.795	43.354	5.28e-07
F_{20}	Skewness (speed)	- 2.042	0.473	- 1.822	0.618	0.022
F_{21}	Number of peaks	9.731	2.171	10.925	2.59	0.004
F_{22}	STD peaks extent	19.56	10.37	16.309	7.801	0.042

Besides improving accuracy and efficiency of the machine-learning algorithms, FS techniques allow understanding, reducing and visualizing data as well as reducing storage and computational costs. Three main FS methods exist: (1) wrapper uses a learning algorithm; they are slow because they scan the whole space, searching the features subset which offers the highest learning performance, (2) Filter methods rank features independently of classifiers. Based on statistics (correlations, maximum variations, etc.) or on information theory (mutual information, entropy, etc.), features are selected depending on their abilities to differentiate data samples belonging to different clusters. Filters are fast and do not require great computational resources, and (3) embedded methods are hybrids. They inject the feature selection technique into the learning algorithms as filtering step.

Combining different methods in decision-making supports systems to be more robust against the individual methods weaknesses. This technique, ensemble learning, has been widely used in pattern recognition, especially by combining classifiers and then voting ensembles of classifiers to improve the classification accuracy. The main idea in designing an ensemble learning is the diversity of combined methods, while respecting both outputs and structure (Duin 2002; Guan et al. 2014).

To investigate the relevance of features for the driving fatigue characterization by ensemble learning, we propose to combine the following ranking methods, which are based on different ranking criteria.

1. Feature Selection and Kernel Learning for Local Learning-Based Clustering (LLCFS) (Zeng and Cheung 2010): it associates a weight to each feature, and incorporates it into the built-in regularization of the LLC algorithm to take into account the relevance of each feature for clustering.
2. Correlation-based Feature selection (CFS): it is an intuitive filter FS method, which sorts features according to pairwise correlations.
3. Laplacian Score (LS) (He et al. 2005): it uses the nearest neighbor graphs to evaluate the feature importance according to its ability for locality preserving.
4. Infinite Feature Selection (Inf-FS) (Roffo et al. 2015): it is composed of: (1) unsupervised feature ranking step, where feature subset is a path connecting them into an affinity graph, and the score of each feature depends on all feature sets, (2) Supervised feature selection step, based on cross-validation to select the best m features.
5. Eigenvector Centrality Feature Selection (ECFS) (Roffo and Melzi 2017): it maps features in an affinity graph by means of mutual information, Fisher's criteria and maximum standard deviation; then Eigenvector centrality measure is used to rank features.
6. Spectral Feature Selection (SFS) (Zhao and Liu 2007): it uses a similarity matrix and spectral graph theory to evaluate and rank features.

Then, driving fatigue features are ranked based on a 'plurality vote'. This later is a simplest method; it counts the number of decisions for each rank, then assigns features to the rank that obtained the highest number of votes (agreement index) (Guan et al. 2014). If two features have the same agreement index for a given position, the feature,

which already appears in a better rank, is privileged; otherwise, features are classed consecutively.

Evaluation

In this study, we are seeking for features characterizing driving fatigue. We investigated the supervised SVM classifier with RBF kernel and the unsupervised density-based algorithm 'DBSCAN' to evaluate the impact of features and their combinations in detecting driving fatigue.

The SVM (Mountrakis et al. 2011) is a classification algorithm, generally used to validate the results of FS as well as embedded in ranking and subset selection for Wrapper and Embedded methods (Chen and Lin 2006). SVM is an efficient supervised learning algorithm, which is largely used in the detection and the recognition of objects (faces, hand-written characters, text ...) from signals (images, videos ...). During learning, SVM attempts to design hyper planes that leave the maximum margin of classes, to separate and classify data. Several kernel functions exist such as linear, polynomial, and radial basis (RBF). To analyze features characterizing driving fatigue, we use RBF kernel, which classifies data even if classes are non-linearly separable.

We also use the first density-based clustering algorithm DBSCAN (Ester et al. 1996); it was designed to discover clusters of arbitrary shapes in the presence of noise. Based on the ϵ -neighborhood cardinality of data points and some density concepts, it searches high-density clusters. We use this algorithm to evaluate the ability of features to separate data and then, illustrate the distribution of data points in each dimension.

A holdout validation is used to assess feature rankings given by FS. Training and testing sets are sampled, respectively, 60% and 40% of the whole database. All FS methods are applied on training set, and then SVM and DBSCAN are used to classify test set while varying features amount and their ranking. The k-fold validation ($k=50$) has been considered to validate the SVM classification results.

Experimental results

Driving fatigue feature selection

We investigated the combination of several FS methods to select relevant features that characterize driving fatigue from a set of features listed in "[Driver's behavior features extraction](#)" section.

First, we analyze the relevance of facial expression features. Table 4 summarizes their ranking by means of various FS methods. The obtained rankings depend on the FS method and consequently on the used criteria (clustering, the laplacian index or graphs). However, we note that several features occupy the same or roughly the same rank for different FS method.

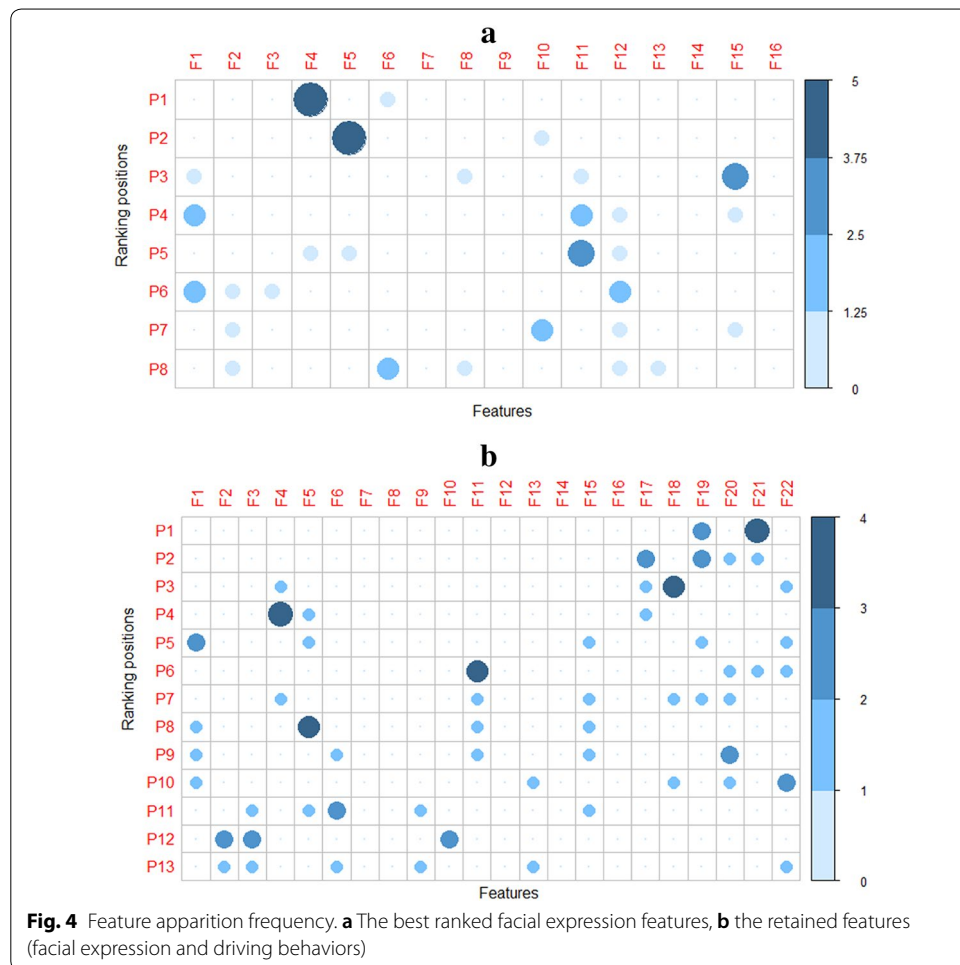
For instance, the feature F_4 which corresponds to 'Persac' tops the list for most methods.

Same for feature F_5 , which corresponds to the average eye closure speed 'AECS' that occupied the second rank. In contrast, the feature F_7 is ranked usually last among the 16 features.

Moreover, the ranking agreement between the several FS methods is more important for the first eight ranks. Figure 4a illustrates the FS methods agreement index of

Table 4 Features ranking according the FS methods (numerical values of ranking represent the index of feature F_i)

FS methods	Ranking															
Laplacian	4	5	15	11	12	1	2	6	8	13	10	14	9	16	3	7
CFS	4	10	8	11	5	3	6	12	1	16	13	14	2	9	7	15
LLCFS	4	5	15	1	11	12	10	8	16	2	9	13	6	14	3	7
Inf-FS	4	5	15	1	11	12	10	13	14	2	8	16	9	6	3	7
ECFS	4	5	1	15	11	2	12	6	10	8	13	3	14	9	16	7
SFS	6	5	11	12	4	1	15	2	3	13	16	8	10	7	14	9



each rank computed based on the feature occurrence frequency. For instance, features F_4 , F_5 and F_{15} have a very high agreement index for their corresponding positions 1, 2 and 3. The agreement between FS methods in ranking features confirms the relevance of features, especially those considered in first ranks.

Second, we combined all feature sets (facial expression and driving behaviors features) and evaluate their relevance by FS methods. A pronounced agreement is noted for the 13 first ranks. Therefore, we retain the intersection between the features in the 13 first ranks for classification and the detection of driver’s fatigue (see Fig. 4b).

Feature evaluation

The feature evaluation has been performed using two classifiers i.e., SVM and DBSCAN, and classification results are validated by *k*-fold method, with *k* = 50 (see "Evaluation" section). The retained feature subset is composed of : {*F*₁, *F*₂, *F*₄, *F*₅, *F*₁₀, *F*₁₁, *F*₁₅, *F*₁₇, *F*₁₈, *F*₁₉, *F*₂₀, *F*₂₁, *F*₂₂}.

Table 5 illustrates comparisons of classification rates, given by SVM and DBSCAN, while varying features subsets.

The classification accuracies using the SVM are almost higher due to its supervised aspects, and more consistent which can be explained by the data nature.

Both SVM and DBSCAN achieved their best classification rate with the retained feature subset; this subset is composed of all driving behaviors features and the most relevant facial expressions ones (See Table 5).

Facial expression features are less relevant than driving behaviors. Based on these features, SVM detects fatigue at 83.14% and DBSCAN accuracy is 78.5%.

The combination of facial expression features and driving behaviors enhances the classification rate (All features) at 85.53% for SVM and to 79.23% for DBSCAN. Although, the use of only relevant features as ranked by FS methods enhances algorithm accuracies (retained features). These results demonstrate the consistence of the retained features to characterize driving fatigue.

This study demonstrates the relevance of driving behavior features in driving fatigue detection; they illustrate a high relevance justified by the classification accuracies and by their rank given by FS methods. Their combination with some facial expression feature such as Perclos (*F*₁), ECS (*F*₂), Persac (*F*₄), and AECS (*F*₅), and increases the detection accuracy.

Conclusion

In this study, we investigated several feature selection methods to extract relevant features that characterize driving fatigue. These features are derived from facial expression which measure eyes and head behaviors as well as from vehicle speed and steering wheel movements which reflect driving behavior. Features based on driving behavior illustrate for the first time their efficiency in real driving environment. The relevance of retained features was evaluated using classification methods, they well discriminate fatigued driver and no fatigued driver. As future works, we plan to develop an on-line fatigue detection prototype based on data stream analysis methods and the obtained

Table 5 Classification rates using a SVM and a DBSCAN classifiers on the different feature set

Features sets	Number of features	SVM		DBSCAN
		ACC (%)	Std	ACC (%)
All features	22	85.53	1.02	79.23
Facial expression features	16	83.14	1.16	78.5
Eye-based features	12	82.65	1.1	78
Head-based features	4	81.30	0.33	75.63
Driving behaviors features	5	84.14	0.66	78
Retained features	13	89.13	1.01	80.91

conclusions about the effectiveness of the retained features in driving fatigue detection on a real setting. Such prototypes would activate systems of autonomous helper cars when the drivers' states are unsafe.

Authors' contributions

The authors discussed the problem and the solutions were proposed all together. All authors participated in collecting dataset, drafting and/or revising the final manuscript. All authors read and approved the final manuscript.

Author details

¹ Centre de recherche LICEF, université TELUQ, 5800 Rue St-Denis, Montréal, QC, Canada. ² Laboratoire de recherche en imagerie et orthopédie (LIO), Centre de recherche du CHUM, 900 Rue St-Denis, Montréal, QC, Canada.

Competing interests

The authors declare that they have no competing interests.

Funding

This research have been funded by the Foundation CAA- Québec, the Canadian Foundation for Innovation (CFI) - Leader funds, the program MITACS, accelerate FR20930 and by the Canada research chair on biomedical data mining (950-231214).

Availability of data and materials

The dataset will be posted online after anonymization of drivers' personal information.

Ethics approval and consent to participate

All drivers gave their consent to participate, which was approved by institutional ethics committees.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 19 July 2018 Accepted: 20 October 2018

Published online: 01 November 2018

References

- Abraham H, Lee C, Brady S, Fitzgerald C, Mehler B, Reimer B, Coughlin JF (2016) Autonomous vehicles, trust, and driving alternatives: a survey of consumer preferences. *Massachusetts Inst. Technol, AgeLab, Cambridge*, pp 1–16
- Ahlstrom C, Kircher K, Kircher A (2013) A gaze-based driver distraction warning system and its effect on visual behavior. *IEEE Trans Intell Transp Syst* 14(2):965–973
- Awake Consortium (IST 2000-28062) System for effective assessment of driver vigilance and warning according to traffic risk estimation (AWAKE). <http://www.awake-eu.org>. Accessed Sept 2001–2004
- Azim T, Jaffar MA, Mirza AM (2014) Fully automated real time fatigue detection of drivers through Fuzzy Expert Systems. *Appl Soft Comput J* 18:25–38
- Bergasa LM, Nuevo J (2006) Real-time system for monitoring driver vigilance. *IEEE Int Symp Ind Electron III(1)*:1303–1308
- Chen Y, Lin C (2006) Combining SVMs with various feature selection strategies. *Featur Extr* 324(1):315–324
- Darshana S, Fernando D, Jayawardena S, Wickramanayake S, Silva C D (2015) Efficient PERCLOS and gaze measurement methodologies to estimate driver attention in real time. In: *Proc. - int. conf. intelligent systems, modelling and simulation, ISMS. IEEE, Piscataway*, pp 289–294
- Duin R (2002) The combining classifier: to train or not to train? In: *Proc. of the international conference on pattern recognition ICPR 2002, Quebec (Canada)*, pp 765–770
- Engström J, Johansson E, Östlund J (2005) Effects of visual and cognitive load in real and simulated motorway driving. *Transp Res Part F Traffic Psychol Behav* 8(2):97–120
- Ester M, Kriegel HP, Sander J, Xu X (1996) A density-based algorithm for discovering clusters in large spatial databases with noise. In: *Proc. 2nd int. conf. knowl. discov. data min.* pp 226–231
- Friedman JH, Bias On (1997) Variance, 0/1- loss, and the curse-of-dimensionality. *Data Min Knowl Discov* 77:55–77
- Friedrichs F, Yang B (2010) Camera-based drowsiness reference for driver state classification under real driving conditions. In: *Proc. IEEE Intell. Veh. Symp. IEEE, Piscataway*, pp 101–106
- Guan D, Yuan W, Lee Y, Najeebullah K, Rasel MK (2014) A review of ensemble learning based feature selection. *IETE Tech Rev* 31(3):190–198
- Habermann C, Kindermann F (2007) Multidimensional spline interpolation: theory and applications. *Comput Econ* 30(2):153–169
- He X, Cai D, Niyogi P (2005) Laplacian score for feature selection. *Adv Neural Inf Process Syst* 18:507–514
- Hu S, Zheng G (2009) Driver drowsiness detection with eyelid related parameters by Support Vector Machine. *Expert Syst Appl* 36(4):7651–7658
- Ingre M, Akerstedt T, Peters B, Anund A, Kecklund G (2006) Subjective sleepiness, simulated driving performance and blink duration: examining individual differences. *J Sleep Res* 15:47–53
- Ji Q, Zhu Z, Lan P (2004) Real-time non-intrusive monitoring and prediction of driver fatigue. *IEEE Trans Veh Technol* 53(4):1052–1068
- Jin L, Niu Q, Jiang Y, Xian H, Qin Y, Xu M (2013) Driver sleepiness detection system based on eye movements variables. *Adv Mech Eng* 5:64843

- Joly A, Zheng R, Kaizuka T, Nakano K (2018) Effect of drowsiness on mechanical arm admittance and driving performances. *IET Intell Transp Syst* 12(3):220–226
- Kaida K, Takahashi M, Akerstedt T, Nakata A, Otsuka Y, Haratani T, Fukasawa K (2006) Validation of the Karolinska sleepiness scale against performance and EEG variables. *Clin Neurophysiol* 117:1574–1581
- Kong W, Zhou L, Wang Y, Zhang J, Liu J, Gao S (2015) A system of driving fatigue detection based on machine vision and its application on smart device. *J Sensors* 548602:11
- Liang Y, Reyes ML, Lee JD (2007) Real-time detection of driver cognitive distraction using support vector machines. *IEEE Trans Intell Transp Syst* 8(2):340–350
- Mitra P, Majumder DD (2004) Feature selection and gene clustering from gene expression data. In: Proc. 17th Int. Conf. Pattern Recognition, IEEE, Piscataway, ICPR pp 1590–1599
- Miyaji M, Kawanaka H (2010) Effect of Pattern Recognition Features on Detection for Drivers Cognitive Distraction. In: Proc. conf. intelligent transportation systems (ITSC). IEEE, Piscataway, pp 605–610
- Mountrakis G, Im J, Ogole C (2011) Support vector machines in remote sensing: a review. *ISPRS J Photogramm Remote Sens* 66(3):247–259
- Of O, Carriers M (1998) PERCLOS: A valid psychophysiological measure of alertness as assessed by psychomotor vigilance. *FEDERAL HIGHWAY ADMINISTRATION, OFFICE OF MOTOR CARRIER RESEARCH & STANDARDS* 31(5):1237–1252
- Otmani S, Pebayle T, Roge J, Muzet A (2005) Effect of driving duration and partial sleep deprivation on subsequent alertness and performance of car drivers. *Physiol Behav* 84(5):715–724
- Rahate RS, Emmanuel M (2013) Feature selection for sentiment analysis by using SVM. *Int J Comput Appl* 84(5):24–32
- Rakoczy M, McGaughey D, Korenberg MJ, Levman J, Martel AL (2013) Feature selection in computer-aided breast cancer diagnosis via dynamic contrast-enhanced magnetic resonance images. *J Digit Imaging* 26(2):198–208
- Roffo G, Melzi S (2017) Ranking to learn: feature ranking and selection via eigenvector centrality. *New Front. Min. Complex Patterns* 1–15
- Roffo G, Melzi S, Cristani M (2015) Infinite feature selection. In: proceedings of the IEEE international conference on computer vision. IEEE, Piscataway, pp 4202–4210
- Sahayadhas A, Sundaraj K, Murugappan M (2012) Detecting driver drowsiness based on sensors: a review. *Sensors* 12(12):16937–16953
- Sigari MH, Fathy M, Soryani M (2013) A driver face monitoring system for fatigue and distraction detection. *Int J Veh Technol* 263983:11
- Sun W, Zhang X, Peeta S, He X, Li Y, Zhu S (2015) A self-adaptive dynamic recognition model for fatigue driving based on multi-source information and two levels of fusion. *Sensors* 15(9):24191–24213
- Sun W, Zhang X, Peeta S, He X, Li Y (2017) A real-time fatigue driving recognition method incorporating contextual features and two fusion levels. *IEEE Trans Intell Transp Syst* 99:1–13
- Sáez-Francàs N, Hernández-Vara J, Corominas-Roso M, Alegre J, Jacas C, Casas M (2014) Relationship between poor decision-making process and fatigue perception in Parkinson's disease patients. *J Neurol Sci* 337(1):167–172
- Thiffault P, Bergeron J (2003) Monotony of road environment and driver fatigue: a simulator study. *Accid Anal Prevent* 35:381–391
- Zeng H, Cheung YM (2010) Feature selection and Kernel learning for local learning based clustering. *IEEE Trans Pattern Anal Mach Intell* 33(8):1532–1547
- Zhao Z, Liu H (2007) Spectral feature selection for supervised and unsupervised learning. In: Proc. 24th Int. Conf. Mach. Learn. - ICML '07. ACM, New York, pp 1151–1157
- Zhao L, Wang Z, Wang X, Liu Q (2018) Detecting driver drowsiness based on sensors: a review. *IET Intell Transp Syst* 12(2):127–133

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
