Accident Analysis and Prevention 156 (2021) 106107

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Using long short term memory and convolutional neural networks for driver drowsiness detection

Azhar Quddus[a], Ali Shahidi Zandi[b,*], Laura Prest[b], Felix J.E. Comeau[b]

[a] Au-Zone Technologies Inc., Calgary, AB, Canada
[b] Alcohol Countermeasure Systems Corp. (ACS), 60 International Boulevard, Toronto, ON, Canada

ARTICLE INFO

Keywords:
Driver drowsiness
Fatigue
Long short term memory
LSTM
Convolutional LSTM
Eye detection
Electroencephalogram (EEG)
Non-invasive
Random forest (RF)
Support vector machine (SVM)

ABSTRACT

Fatigue negatively affects the safety and performance of drivers on the road. In fact, drowsiness and fatigue are the cause of a substantial number of motor vehicle accidents. Drowsiness among the drivers can be detected using variety of modalities, including electroencephalogram (EEG), eye movement, and vehicle driving dynamics. Among these EEG is highly accurate but very intrusive and cumbersome. On the other hand, vehicle driving dynamics are very easy to acquire but accuracy is not very high. Eye movement based approach is very attractive in terms of balance between these two extremes. However, eye movement based techniques normally require an eye tracking device which consists of high speed camera with sophisticated algorithm to extract eye movement related parameters such as blinking, eye closure, saccades, fixation etc. This makes eye tracking based drowsiness detection difficult to implement as a practical system, especially on an embedded platform.

In this paper, authors propose to use eye images from camera directly without the need for expensive eye-tracking system. Here, eye related movements are captured by Recurrent Neural Network (RNN) to detect the drowsiness. *Long Short Term Memory* (LSTM) is a class of RNN which has several advantages over vanilla RNNs. In this work an array of LSTM cells are utilized to model the eye movements. Two types of LSTMs were employed: 1-D LSTM (R-LSTM) which is used as baseline and the convolutional LSTM (C-LSTM) which facilitates using 2-D images directly. Patches of size 48×48 around each eye were extracted from 38 subjects, participating in a simulated driving experiment. The state of vigilance among the subjects were independently assessed by power spectral analysis of multichannel electroencephalogram (EEG) signals, recorded simultaneously, and binary labels of *alert* and *drowsy* (baseline) were generated.

Results show high efficacy of the proposed system. R-LSTM based approach resulted in accuracy around 82 % and C-LSTM based approach resulted in accuracy in the range of 95%–97%. Comparison is also provided with a recently published eye-tracking based approach, showing the proposed LSTM technique outperform with a wide margin.

1. Introduction

Now a days people are more susceptible to fatigue due to life style and work requirements. In fact, one gets 20 % less sleep, on average, comparing to a century ago (NCSDR (National Commission on Sleep Disorders Research), 1994). It is estimated that about 50–70 million Americans suffer from sleep disorder (Tjepkema, 2005). Driver performance depreciates significantly under the influence of fatigue, which is more pronounced in the presence of sleep restriction (Philip et al., 2003; Perrier et al., 2016), resulting in higher risk of motor vehicle collisions. In the US, the National Highway Traffic Safety Administration (NHTSA) estimates that 72,000 police-reported motor vehicle accidents in 2015 involved drowsy driving resulting in 41,000 injuries and 800 deaths (NHTSA, 2017). In Canada, the Traffic Injury Research Foundation reported that 6.4 % of motor vehicle fatalities were drowsy-related in 2013 (TIRF (Traffic Injury Research Foundation), 2016). Therefore, monitoring of driver vigilance as a countermeasure for managing fatigue is critical in order to reduce the risk of motor vehicle collisions.

* Corresponding author.
E-mail addresses: azhar@au-zone.com (A. Quddus), aszandi@acs-corp.com (A. Shahidi Zandi), lprest@acs-corp.com (L. Prest), fjecomeau@acs-corp.com (F.J.E. Comeau).

https://doi.org/10.1016/j.aap.2021.106107
Received 23 August 2019; Received in revised form 19 July 2020; Accepted 27 March 2021
Available online 10 April 2021
0001-4575/

2. Background and motivation

Physiological responses (Jap et al., 2009; Sun and Lu, 2012; Eoh et al., 2005; Khushaba et al., 2011; Shuyan and Gangtie, 2009; Yang et al., 2010; Chuang et al., 2015), such as electroencephalogram (EEG), are among the most objective measure of fatigue. They can produce reliable measures with very high temporal resolution to detect subtle changes in vigilance well in advance of behavioral lapses. However, the application of such techniques are limited due to their intrusive nature for long-term driver monitoring in practical applications. On the other hand, measures relying on driver behavioral patterns (Jackson et al., 2016a, b; Wang and Xu, 2016; Azim et al., 2014; García et al., 2012; Bergasa et al., 2006; Dinges et al., 1998; Wang et al., 2017), such as eye movements and facial expressions, are non-intrusive and are more suitable for practical reasons. However, accuracy and reliability are the main challenge for these types of measures. For example, percentage of eyelid closure or PERCLOS (Dinges et al., 1998; Wierwille et al., 1994; Dinges and Grace, 1998) has been shown to be sensitive to sleep loss and fatigue. Studies have shown noticeably lower accuracy for PERCLOS as compared to techniques based on biological signals (Sommer and Golz, 2010). PERCLOS also requires computations on a longer time interval (around 1 min) to get an acceptable accuracy (Dong et al., 2011; Sommer and Golz, 2010), resulting in time lag for drowsiness detection (Bergasa et al., 2006), and needs accurate eye opening and closing events for reliable measurements. Driving performance indicators, such as steering wheel patterns, lateral position, or braking patterns have also been used to assess drowsiness in drivers (Perrier et al., 2016; Wang and Xu, 2016; Wierwille et al., 1994; Krajewski et al., 2009; Wang et al., 2016; Zhang et al., 2016a). However, the accuracy of such methods can be affected by several factors including road and weather conditions, driver experience level as well as type of vehicle. Adopting a hybrid approach by combining various physiological, behavioral and driving performance variables, Jacobé de Naurois et al. recently proposed deep learning techniques for multi-level classification of drowsiness (Jacobé de Naurois et al., 2018, 2019). Driver fatigue evaluation using advanced deep learning techniques have been also reported recently in some studies based on physiological responses, e.g. (Gao et al., 2019a, b). In particular, convolutional neural networks (CNNs) and recurrent neural networks (RNNs) have been adopted to gain high accuracy in fatigue detection among drivers; however, the issue of intrusiveness would still limit applications of these methodologies.

Authors have recently proposed advanced machine learning techniques relying on a specific set of features extracted from raw data provided by an advanced eye tracking system in order to address this problem (Shahidi Zandi et al., 2019). Despite high accurate results, however, implementation of eye tracking modules on embedded platform remains a challenge which requires high speed ($\geq$60 Hz.) camera along with high speed processing capabilities.

In this paper, we propose a novel drowsiness detection approach by directly feeding eye images from video frames to specific types of RNNs without explicitly extracting eye tracking features. That is, the objective of this research is to show that by bypassing the eye tracking module and eliminating the related computational costs, which will facilitate the deployment of drowsiness monitoring system in the vehicle environment, we can maintain and even improve the performance of drowsiness detection. Here, we propose to utilize two types of *Long Short Term Memory* (LSTM) to capture the eye movements and subsequently detect drowsiness: regular and convolutional LSTM networks. The drowsiness label was extracted from simultaneously recorded EEG.

3. Introduction to Long short term memory (LSTM)

In this section, a brief introduction of 1-D and 2-D LSTMs is provided for the sake of completeness. The main feature of LSTM (Hochreiter and Schmidhuber, 1997; Graves, 2013; Shi et al., 2015) is its memory cell which is essentially an accumulator and c_t indicates the current status of the memory cell. The cell is modified by several controlling gates. With each input, the cell information will be accumulated if the input gate i_t is on. The past cell state c_{t-1} will be forgotten when forget gate f_t is on. The cell output will be propagated to the final state h_t if output gate o_t is on. The main advantage of using the memory cell and gates to control information flow is that the gradient will be prevented from vanishing too quickly. A multivariate version of LSTM where the input, cell output and states are all 1-D vectors. In this paper we refer it as R-LSTM and its formulation is same as in (Graves, 2013). The key equations are shown below.

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i)$$

$$f_t = \sigma\left(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f\right)$$

$$c_t = f_t\circ c_t + i_t\circ tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}\circ c_{t-1} + b_o)$$

$$h_t = o_t\circ\tanh(c_t)$$

Where '∘' denotes Hadamard product. Multiple LSTMs can be temporally concatenated to form more complex structures.

R-LSTM is designed for 1-D input data which is not optimal for image inputs. Therefore, convolutional LSTM (C-LSTM) was proposed in (Shi et al., 2015). For purpose of brevity it referred as C-LSTM in this paper. The main feature of this design is that all the inputs $\mathscr{X}_1, \mathscr{X}_2, \ldots \mathscr{X}_t$, cell outputs $\mathscr{C}_1, \mathscr{C}_2, \ldots \mathscr{C}_t$, hidden states $\mathscr{H}_1, \mathscr{H}_2, \ldots \mathscr{H}_t$, and gates i_t, f_t, o_t of the C-LSTM are 3D tensors with last two dimensions are rows and columns. In this paper, the formulation of C-LSTM is the same as in (Shi et al., 2015) and the key equations are shown below.

$$i_t = \sigma(W_{xi}*\mathscr{X}_t + W_{hi}*\mathscr{H}_{t-1} + W_{ci}\circ\mathscr{C}_{t-1} + b_i)$$

$$f_t = \sigma\left(W_{xf}*\mathscr{X}_t + W_{hf}*\mathscr{H}_{t-1} + W_{cf}\circ\mathscr{C}_{t-1} + b_f\right)$$

$$\mathscr{C}_t = f_t\circ\mathscr{C}_t + i_t\circ tanh(W_{xc}*\mathscr{X}_t + W_{hc}*\mathscr{H}_{t-1} + b_c)$$

$$o_t = \sigma(W_{xo}*\mathscr{X}_t + W_{ho}*\mathscr{H}_{t-1} + W_{co}\circ\mathscr{C}_{t-1} + b_o)$$

$$\mathscr{H}_t = o_t\circ\tanh(\mathscr{C}_t)$$

Where '∘' denotes Hadamard product and '∗' denotes Convolution operator. Here cell states can be viewed as the hidden representations of moving objects - a larger kernel should be able to capture faster movements while a smaller kernel should capture slower movements.

4. Proposed system

The proposed system is a two-camera system in which both cameras are directed towards the driver's face. Face detection and eye detection is performed for each frame. There are several off-the-shelf solutions available (such as OpenCV, dlib and MTCNN). MTCNN (Zhang et al., 2016b) based technique was employed due to its superior performance. A 48 × 48 pixel size area around each eye is cropped in both the camera frames. In order to establish baseline, R-LSTM (Hochreiter and Schmidhuber, 1997; Graves, 2013) based approach is presented in Fig. 1. Here, each 48 × 48 input patch is flattened to 2304 (48 × 48) sized array.

The numbers in the parenthesis show the size of data in each path. Each row represents processing for patches belonging to right-camera-right-eye, right-camera-left-eye, left-camera-right-eye and left-camera-left-eye, respectively. In each row LSTM cells (with 128 hidden nodes) are unrolled with n time steps. The output of the final step (in each layer) is fed to two fully connected layers (FC) producing 2-class hot-encoded outputs belonging to "*drowsy*" and "*alert*" states.

Fig. 2 shows the C-LSTM (Shi et al., 2015) based implementation

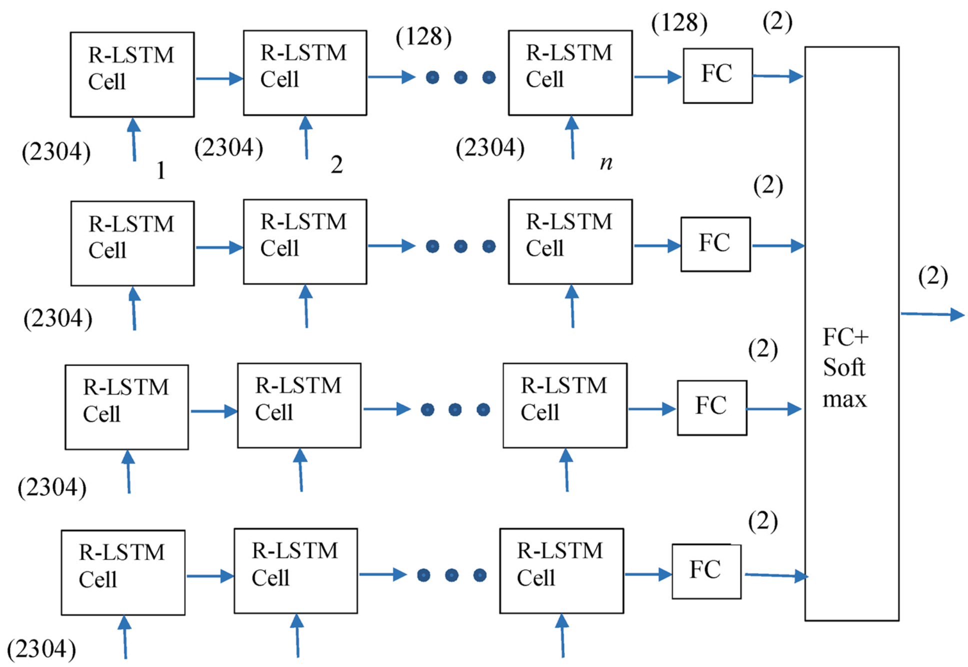

Fig. 1. Baseline R-LSTM implementation. Rows top to bottom represents processing for patches belonging to right-camera-right-eye, right-camera-left-eye, left-camera-right-eye and left-camera-left-eye, respectively.

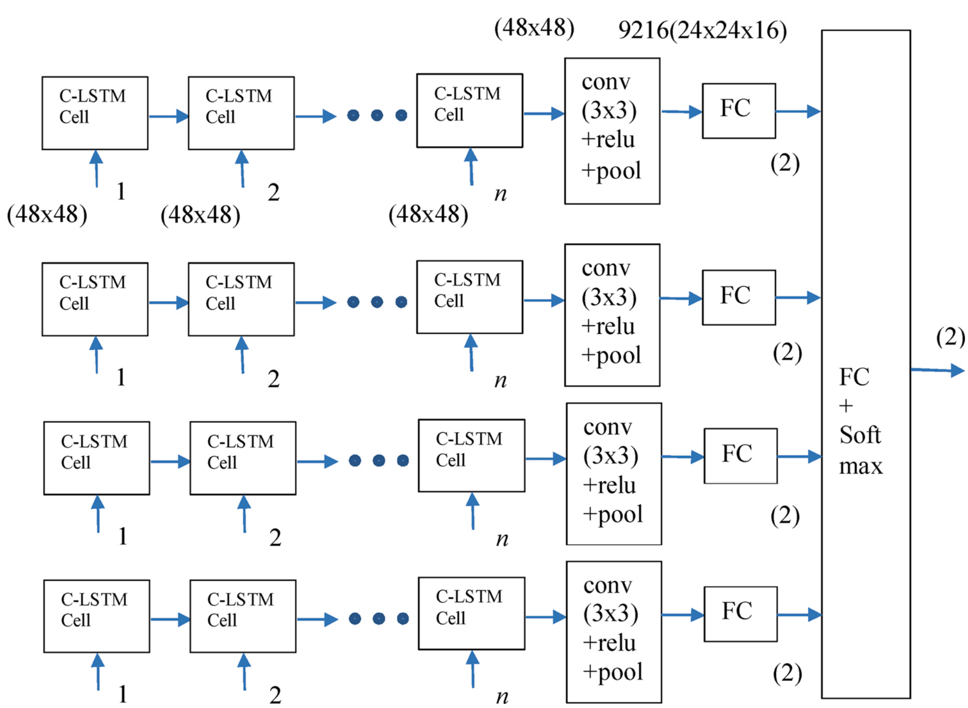

Fig. 2. 2-D C-LSTM based implementation. Rows top to bottom represents processing for patches belonging to right-camera-right-eye, right-camera-left-eye, left-camera-right-eye and left-camera-left-eye, respectively.

where input is 2-D patches of size (48 × 48). Here also each row represents processing for patches belonging to right-camera-right-eye, right-camera-left-eye, left-camera-right-eye and left-camera-left-eye, respectively. In each row C-LSTM cells (with output size of 48 × 48) are unrolled with *n* time steps. The output of the final time step is fed to convolutional neural network (kernel size of 3 × 3) with ReLU and followed by max-pooling layers. The output of pooling is 24 × 24 × 16 which flattened to 9216 (24 × 24 × 16) and fed to fully connected layer (FC). Finally the outputs of all (4) channels (rows) are combined through fully connected layers (FC) producing 2-class hot-encoded outputs belonging to drowsy and alert states, respectively.

5. DATA acquisition and preparation

The experiment was designed and conducted at the Somnolence Laboratory of Alcohol Countermeasure Systems Corp. (ACS), Toronto, Canada. The laboratory was setup with a driving simulator, two 60 Hz video cameras (with infra-red illumination), and EEG recording equipment. More specifically, Simuride pro driving simulator from AplusB software corporation was used with 38 participants (ages 41 ± 9 years) where each participant was asked to drive in the afternoon for approximately 30 min under monotonous driving conditions on a low-traffic highway. In order to increase the chance of drowsiness, the afternoon session was conducted at the mid-afternoon dip of the alertness circadian cycle (Goel et al., 2011) (typically around 2 pm). Moreover, participants were asked not to drink tea or coffee one hour before the driving session. All the subjects were asked to drive on the same set of roads and they were allowed using glasses and contact lenses as they normally do in normal driving conditions.

In order to compare the performance of the proposed methodologies with eye tracking based techniques, eye tracking data was also simultaneously collected using Smart Eye Pro software. The gaze accuracy of this system is 0.5 degrees and it provides output for the head, left and right eyes including eye gaze, head gaze, eyelid opening, pupilometry, blinks, fixations, and saccades among other measures. At the start of each session eye tracking system was recalibrated.

5.1. EEG-based drowsiness labeling

In order to assess the level of drowsiness, the characteristic changes in the EEG power spectrum were analyzed (Jap et al., 2009; Bonnet, 2011; Lal and Craig, 2001), and subsequently generate binary labels (i.e., "*alert*" vs. "*drowsy*"). EEG waveforms in four distinct frequency bands were analyzed (Sanei and Chambers, 2007) viz. the δ-band (0.5–4 Hz), the θ-band (4–7 Hz), the α-band (8–12 Hz), the β-band (13–30 Hz). A 60-Hz notch filter was first applied to the multichannel EEG signal, and the signal was band-passed filtered between 0.2 and 45 Hz. Bipolar EEG channels (AF_3-AF_4, F_7-F_8, F_3-F_4, FC_5-FC_6, T_7-T_8,P_7-P_8,O_1-O_2) were then used for power spectral analysis.

Here, short-time Fourier transform was applied to EEG channels and the β/α and $\alpha/(\delta+\theta)$ power ratios were computed for each EEG window in every channel. Intuitively, the higher these power ratios, the higher the level of the alertness. The EEG-based algorithm uses an "adaptive" reference calculated based on a moving EEG window to quantify the amount of relative changes in power ratios at each point in time and estimate a quantitative alertness level. The algorithm also learns a subject-specific threshold for the level of alertness based on the changes in power ratios in a driving *control session*, which was conducted in late morning when the participants were highly alert and engaged. This threshold is then used to create binary labels for the video frames in this work. If alertness level is inferior to the threshold, the "*drowsy*" state is declared; otherwise, the state of the video frame is considered as "*alert*". Since the sampling rate of EEG recording equipment is not the same as video frame rate, precise time stamps were used to label the video frames.

5.2. Data extraction and setup

At the start of the driving session, subjects are alert and drowsiness increases as driving progresses. In order to extract balanced data, the labels associated with each frames are analyzed and approximately twenty (20) minutes of session is identified with roughly ten minutes of *alert* state and ten minutes of *drowsy* state. In order to detect facial landmarks on each frame, MTCNN (Zhang et al., 2016b) based technique was employed for the twenty minutes of the driving session. A 48 × 48 pixel size patch around each eye was cropped in both the camera frames leading to around 288 K (20 × 60 × 60 × 4) patches per subject.

This data was split for training and testing for each subject. Approximately 20 % (14,400) frames worth of data was taken randomly for testing purpose (with roughly equal alert and drowsy states) and the rest was used for training.

6. Training and results

A desktop with NVIDIA GTX 1070 GPU (with 8 GB RAM) was used for training. The two cameras were operating at 60 Hz and training was done for specific window sizes with 30 time steps ($n = 30$ in all the LSTM simulations). Here, window size of t second means down sampling with a factor of $t \times 60/n$. For example, window size of 1 s leads to down sampling factor of 2. This results in effective sample rate of n/t.The batch size was set at 90 which was decided empirically based on the available memory and resources.

At each iteration, training was done by randomly applying a batch worth of data (of size × 48 × 48) from each eye to the network for each subject sequentially. With two camera system data from each eye were fed in parallel to the networks shown in Figs. 1 and 2. One epoch was defined by training one random batch from all the subjects. Average training loss and test accuracy was recorded over each epoch. Adam optimizer was used with initial learning rate at 0.0001 for all the LSTM training setups. Training with regular R-LSTM require 1-D input; therefore, input patches were flattened to 2304 (48 × 48) array. Training was run for 2000 epochs. Training with C-LSTM (Graves, 2013) require 2-D inputs; hence, input patches (48 × 48) were used and the training was run for 1000 epochs.

Test performance was computed with varying window sizes (0.5 s, 1 s, 2 s, 5 s, and 10 s). Table 1 shows the *False Accept Rate* (FAR), *False Reject Rate* (FRR) and *Accuracy* (ACC) for both with R-LSTM and C-LSTM. It can be observed here that best R-LSTM is around 88 % with window sizes of 5 and 10 s whereas C-LSTM performs best with window sizes of 1 s. and 2 s. Effective sampling rate is also provided which indicates that the R-LSTM achieves best performance at lower sampling rate resulting in computationally less expensive for real-time implementation. On the hand, C-LSTM would require higher sampling rate which will be more computationally expensive for real-time implementation. The CNNs perform very well on 2-D inputs; therefore, C-LSTM noticeably outperforms R-LSTM.

7. Comparison with eye tracking based system

In this section, a brief comparison of accuracy is provided with a state-of-the-art eye-tracking based drowsiness detection system (Shahidi Zandi et al., 2019), where 34 distinct eye tracking features are extracted from four major categories of eye-tracking measurements, namely, Gaze, Blink, Pupil and Eyelid are extracted and fed to random forest (RF) and nonlinear support vector machine (SVM) classifiers. Eye-tracking data for the same subjects was acquired simultaneously as explained in section 5. The features (detailed in (Shahidi Zandi et al., 2019)) were computed over epochs of few seconds with 50 % overlap. Table 2 shows

Table 1
Test FAR, FRR and Accuracy with R-LSTM and C-LSTM with various window sizes.

Window Size (t) ⟶ (Effective sampling rate = n/t)		0.5 s. (60 Hz.)	1 s.(30 Hz.)	2 s.(15 Hz.)	5 s.(6 Hz)	10 s. (3 Hz.)
R-LSTM	FAR %	9.385	5.718	11.487	**7.333**	6.410
	FRR %	8.538	9.846	7.897	**4.692**	5.769
	ACC %	82.077	84.436	80.615	**87.974**	87.821
C-LSTM	FAR %	3.769	**1.538**	1.128	3.949	5.462
	FRR %	1.692	**0.590**	2.077	1.769	2.897
	ACC %	94.538	**97.872**	96.795	94.282	91.641

Table 2
Test Accuracy of Eye tracking based system with SVM and Random Forest classifiers (Shahidi Zandi et al., 2019).

Epoch Size ⟶	5 s.	10 s.	15 s.	20 s.
Support Vector Machines (SVM)	83.6 %	82.7 %	84 %	83.8 %
Random Forest (RF)	84.1 %	84.5 %	85.1 %	85.9 %

performance of these traditional machine learning techniques (SVM and RF) with different epoch sizes.

From Tables 1 and 2 it can be easily observed that R-LSTM outperforms the eye-tracking based classifiers marginally, whereas C-LSTM outperforms them with a wide margin.

8. Conclusion

In this paper, a novel technique for driver drowsiness detection is presented in which eye movements and characteristics are captured by directly feeding eye images from video frames to RNNs, without the need of a separate eye-tracking module. Although we have shown results with two cameras, the architecture is flexible to any number of cameras. The eye detection was achieved using already published works, and 48 $\times$ 48 pixel size of image patches were extracted around each eye of the driver. These image patches were fed to R-LSTM (Fig. 1) and C-LSTM (Fig. 2) sequentially. The outputs (corresponding to inputs from each eye) from the last stages of LSTMs were combined to produce final prediction about '*alert*' vs. '*drowsy*'. Test accuracy with C-LSTM significantly outperforms the accuracy with R-LSTM. Overall, the results of this invention reveal a high level of correspondence between the eye movements (captured by RNNs) and EEG as an objective physiological measure of vigilance; which in turn shows that the state of vigilance can be non-intrusively classified with very high accuracy. We also compared the accuracies with machine learning based system using eye tracking features and the proposed technique outperforms with a wide margin.

The data used in this work were collected in a laboratory setting, the future studies should consider testing the limits of this technology under real-driving scenarios, in which more challenging conditions, such as using sunglasses and various lighting conditions, will be also considered. This research ultimately will lead to development of an unobtrusive reliable technique for long-term assessment of the state of vigilance, a crucial step towards managing fatigue in drivers and reducing motor vehicle collisions.

CRediT authorship contribution statement

The authors confirm contribution to the paper as follows:
Azhar Quddus: Conceptualization, Methodology, Algorithm Development (software), Writing- Original Draft Preparation, Writing-Reviewing and Editing. **Ali Shahidi Zandi:** Conceptualization, Study Conception and Design, Methodology, Algorithm Development, Data Collection, Writing- Reviewing and Editing. **Laura Prest:** Data Collection. **Felix J.E. Comeau:** Writing- Reviewing and Editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Azim, T., Jaffar, M.A., Mirza, A.M., 2014. Fully automated real time fatigue detection of drivers through fuzzy expert systems. Applied Soft Computing. Elsevier B.V., pp. 25–38. May;18.

Bergasa, L.M., Nuevo, J., Sotelo, M.A., Barea, R., Lopez, M.E., 2006. Real-time system for monitoring driver vigilance. IEEE Trans. Intell. Transp. Syst. 7 (1), 63–77.

Bonnet, M.H., 2011. Acute sleep deprivation. In: Kryger, M., Roth, T., Dement, W. (Eds.), Principles and Practice of Sleep Medicine, 5th ed. Saunders: Elsevier, pp. 54–66.

Chuang, C.-H., Huang, C.-S., Ko, L.-W., Lin, C.-T., 2015. An EEG-based perceptual function integration network for application to drowsy driving. Knowledge-based systems. Elsevier B.V. 80, 143–152.

Dinges, D.F., Grace, R., 1998. PERCLOS: A Valid Psychophysiological Measure of Alertness As Assessed by Psychomotor Vigilance. Fed. Highway Admin., Office Motor Carriers, Washington, DC.

Dinges, D.F., Mallis, M.M., Mailslim, G., Powell, J.W., 1998. Evaluation of Techniques for Ocular Measurement As an Index of Fatigue and the Basis for Alertness Management. U.S. Dept. Transp., NHTSA, Washington, DC.

Dong, Y., Hu, Z., Uchimura, K., Murayama, N., 2011. Driver inattention monitoring system for intelligent vehicles: a review. IEEE Trans. Intell. Transp. Syst. 12 (June (2)), 596–614.

Eoh, H.J., Chung, M.K., Kim, S.H., 2005. Electroencephalographic study of drowsiness in simulated driving with sleep deprivation. Int. J. Ind. Ergon. 35 (4), 307–320.

Gao, Z., Li, Y., Yang, Y., Ma, C., 2019a. A recurrence network-based convolutional neural network for fatigue driving detection from EEG. Chaos. 29, 113126.

Gao, Z., Wang, X., Yang, Y., Mu, C., Cai, Q., 2019b. W. EEG-Based spatio-temporal convolutional neural network for driver fatigue evaluation. IEEE Trans. Neural Netw. Learn. Syst. 30 (9), 2755–2763.

García, I., Bronte, S., Bergasa, L.M., Almazán, J., Yebes, J., 2012. Vision-based drowsiness detector for real driving conditions. IEEE Intelligent Vehicles Symposium, Proceedings 618–623.

Goel, N., Van Dongen, H.P.A., Dinges, D.F., 2011. Circadian rhythms in sleepiness, alertness, and performance. In: Kryger, M.H., Roth, T., Dement, W.C. (Eds.), Principles and Practice of Sleep Medicine, 5th ed. Saunders: Elsevier, pp. 445–455.

Graves, A., 2013. Generating sequences with recurrent neural networks. arXiv preprint arXiv, 1308.0850.

Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9 (8), 1735–1780.

Jackson, M.L., Kennedy, G.A., Clarke, C., Gullo, M., Swann, P., Downey, L.A., et al., 2016a. The utility of automated measures of ocular metrics for detecting driver drowsiness during extended wakefulness. Accid. Anal. Prevent. Elsev. Ltd. 87, 127–133.

Jackson, M.L., Raj, S., Croft, R.J., Hayley, A.C., Downey, L.A., Kennedy, G.A., et al., 2016b. Slow eyelid closure as a measure of driver drowsiness and its relationship to performance. Traffic Inj. Prev. 17 (3), 251–257.

Jacobé de Naurois, C., Bourdin, C., Bougard, C., Vercher, J.L., 2018. Adapting artificial neural networks to a specific driver enhances detection and prediction of drowsiness. Accid. Anal. Prev. 121, 118–128.

Jacobé de Naurois, C., Bourdin, C., Stratulat, A., Diaz, E., Vercher, J.L., 2019. Detection and prediction of driver drowsiness using artifical neural network models. Accid. Anal. Prev. 126, 95–104.

Jap, B.T., Lal, S., Fischer, P., Bekiaris, E., 2009. Using EEG spectral components to assess algorithms for detecting fatigue. Exp. Syst. Appl. Elsev. Ltd. 36 (March(2)), 2352–2359.

Khushaba, R.N., Kodagoda, S., Lal, S., Dissanayake, G., 2011. Driver drowsiness classification using fuzzywavelet-packet-based feature-extraction algorithm. IEEE Trans. Biomed. Eng. 58 (1), 121–131.

Krajewski, J., Sommer, D., Trutschel, U., Edwards, D., Golz, M., 2009. Steering wheel behavior based estimation of fatigue. Proceedings of the Fifth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design 118–124.

Lal, S.K.L., Craig, A., 2001. A critical review of the psychophysiology of driver fatigue. Biol. Psychol. 55 (3), 173–194.

NCSDR (National Commission on Sleep Disorders Research), 1994. Wake Up America: A National Sleep Alert, Vol. II. Working Group Reports, Washington, DC.

NHTSA, 2017. Asleep at the Wheel: A National Compendium of Efforts to Eliminate Drowsy Driving. National Highway Traffic Safety Administration (NHTSA). U.S. Department of Transportation, pp. 1–24.

Perrier, J., Jongen, S., Vuurman, E., Bocca, M.L., Ramaekers, J.G., Vermeeren, A., 2016. Driving performance and EEG fluctuations during on-the-road driving following sleep deprivation. Biol. Psychol. Elsev. B.V. 121, 1–11.

Philip, P., Sagaspe, P., Taillard, J., Moore, N., Guilleminault, C., Sanchez-Ortuno, M., et al., 2003. Fatigue, sleep restriction, and performance in automobile drivers: a controlled study in a natural environment. Sleep 26 (3), 277–280.

Sanei, S., Chambers, J., 2007. EEG Signal Processing. John Wiley & Son, Ltd, West Sussex, England, p. 289.

Shahidi Zandi, A., Quddus, A., Prest, L., Comeau, F., 2019. Non-intrusive detection of drowsy driving based on eye tracking data. Transp. Res. Rec. 2673 (6), 247–257.

Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.-K., Woo, W.-C., 2015. Convolutional LSTM network: a machine learning approach for precipitation nowcasting. Twenty-Ninth Annual Conference on Neural Information Processing Systems (NIPS).

Shuyan, H., Gangtie, Z., 2009. Driver drowsiness detection with eyelid related parameters by Support Vector Machine. Exp. Syst. Appl. Elsev. Ltd. 36 (May(4)), 7651–7658.

Sommer, D., Golz, M., 2010. Evaluation of PERCLOS based current fatigue monitoring technologies. In: 32nd Annu Int Conf IEEE Eng Med Biol Soc (EMBS). January; 2010, pp. 4456–4459.

Sun, H., Lu, B., 2012. EEG-based fatigue classification by using parallel hidden Markov model and pattern classifier combination. In: The 19th International Conference on Neural Information Processing - Volume Part IV. Berlin, Heidelberg: Springer-Verlag, pp. 484–491.

TIRF (Traffic Injury Research Foundation), 2016. Fatigure-related Fatal Collisions in Canada. Ottawa.

Tjepkema, M., 2005. Insomnia. Health Rep. 17 (1), 9–25.

Wang, X., Xu, C., 2016. Driver drowsiness detection based on non-intrusive metrics considering individual specifics. Accid. Anal. Prev. Elsevier Ltd 95, 350–357.
Wang, M.S., Jeong, N.T., KIM, S.B., Yang, S.M., You, S., Lee, J.H., et al., 2016. Drowsy behaviour detection based on driving information. Int. J. Automot. Technol. 17 (1), 165–173.
Wang, Y., Xin, M., Bai, H., Zhao, Y., 2017. Can variations in visual behavior measures be good predictors of driver sleepiness? A real driving test study. Traffic Inj. Prev. 18 (2), 132–138.
Wierwille, W., Wreggit, W., Kirn, C., Ellsworth, A., Fairbanks, R., 1994. Research on Vehicle-Based Driver status/performance Monitoring: Development, Validation, and Refinement of Algorithms for Detection of Driver Drowsiness. U.S. Dept. Transp., Washington, DC. NHTSA Final Report: DOT HS 808 247.
Yang, G., Lin, Y., Bhattacharya, P., 2010. A driver fatigue recognition model based on information fusion and dynamic Bayesian network. Inform. Sci. Elsev. Inc. 180 (May (10)), 1942–1954.
Zhang, H., Wu, C., Huang, Z., Yan, X., Qiu, T.Z., 2016a. Sensitivity of lane position and steering angle measurements to driver fatigue. Transp. Res. Rec. 2585, 67–76.
Zhang, K., Zhang, Z., Li, Z., Qiao, Y., 2016b. Joint face detection and alignment using multitask cascaded convolutional networks. IEEE Signal Process. Lett. 23, 1499–1503.