Drowsiness Detection Using Joint EEG-ECG Data With Deep Learning

Guillaume Geoffroy, Lotfi Chaari, Jean-Yves Tourneret, Herwig Wendt

University of Toulouse, IRIT-ENSEEIHT, France
firstname.lastname@irit.fr

Abstract—**Drowsiness detection is still an open issue, especially when detection is based on physiological signals. In this sense, light non-invasive modalities such as electroencephalography (EEG) are usually considered. EEG data provides informations about the physiological brain state, directly linked to the drowsy state. Electrocardiograms (ECG) signals can also be considered to involve informations related to the heart state. In this study, we propose a method for drowsiness detection using joint EEG and ECG data. The proposed method is based on a deep learning architecture involving convolutional neural networks (CNN) and recurrent neural networks (RNN). High efficiency level is obtained with accuracy scores up to 97% on validation set. We also demonstrate that a modification of the proposed architecture by adding autoencoders helps to compensate the performance drop when analysing subjects whose data is not presented during the learning step.**

I. INTRODUCTION

The drowsiness state [1] can be defined as an intermediate state between awake and sleep. Detecting this state is a real and open challenge since high vehicle accident rates are due to drowsiness. This state particularly involves lack of concentration, partial eye closure and relaxation. Generally speaking, drowsiness is linked to reaction heaviness, which means that the brain is altered by this state. In this sense, several studies have investigated the use of biomedical signals to detect the drowsiness state [2–5]. Specifically, electroencephalography (EEG), Electrocardiography (ECG), electromyography (EMG), and electrooculography (EOG) data have been explored for this detection. However, to our knowledge, all these modalities have been analysed separately.

Analysing the spectral content of EEG signals allows different waves of interest to be identified. For example, the Alpha wave (8-13Hz) is linked to relaxation and creativity, while the Theta (4-8Hz) and Beta (13-25Hz) waves indicate the concentration and alertness levels [6, 7], respectively. Regarding the analysis of ECG data, it has been reported that the Heart Rate Variation (HRV) helps to characterize drowsiness [8].

In this paper, we investigate the joint analysis of EEG and ECG data by proposing a deep learning architecture based on both convolutional and recurrent neural networks. Such a combination is expected to improve drowsiness state characterization by providing complementary informations from the ECG and EEG modalities. However, combining different signals raises several problems such as the temporal resolution heterogeneity and the relevance of the used features.

Due to the difficulty to define a relevant analytical model for the drowsy state, machine learning tools have been investigated in the literature. In this sense, support vector machines (SVM) have been proposed to analyse EEG signals [9, 10].

With the emergence of deep CNN, many studies have been conducted by proposing different architectures for EEG analysis [11–13]. In [11], a deep CNN with residual learning has been proposed for driver fatigue detection. In [12], the authors proposed a drowsiness detection method analysing differential entropy with a deep CNN.

Regarding the use of ECG data, a feature extraction step is often used. In [14], heart rate and frequency properties are extracted in a feature vector for both SVM and random forest classifiers. In [15], the same temporal and frequency descriptors are extracted for both EEG and ECG data to feed an SVM classifier. The authors highlight the benefits of adding the ECG signal to the analysis since its use increases the accuracy from 76.36% to 80.9%.

In the proposed study, EEG and ECG are combined using implicit features by resorting to deep neural networks. Specifically, an architecture combining a CNN with two RNN is proposed. The input data is handled differently by each network resulting in two levels of information fusion for the final classification (drowsiness detection). Our motivation for using a deep neural network is that due to the complexity of the drowsy state, extracting features directly from the EEG and ECG signals has its own limits, which will be confirmed by our simulation.

The rest of the paper is organized as follows. Section II describes the used data for drowsiness detection. The deep learning architecture proposed for data analysis is detailed in Section III. Experimental validation is presented in Section IV, while conclusions and perspectives are drawn in Section V.

II. DATA

A. MIT-BIH dataset

The dataset used for this study is extracted from the MIT-BIH Polysomnographic database[1], which involves a collection of recordings of multiple physiological signals acquired during sleep including EEG and ECG recordings [16]. Subjects were monitored in the Boston's Beth Israel Hospital Sleep Laboratory for evaluation of chronic obstructive sleep apnea

---

[1]https://physionet.org/content/slpdb/1.0.0/

syndrome. The database contains over 80 hours of polysomno-graphic recordings. Only ECG and EEG data are considered in this study, both sampled at 250 Hz. The ECG signal has been annotated beat-by-beat with sleep stage annotation every 30 seconds: 0 for awake, 1 for the first sleep stage, 2 for the second sleep stage until "REM" for the final sleep stage. As a drowsy driver state lies in transition between awake and sleep states, the goal will be to discriminate the the two first states "0" and "1", resulting in a binary classification. The following subjects have been chosen: "slp02a", "slp02b", "slp03", "slp04", "slp14", "slp16", "slp45", "slp48", "slp59", "slp60", "slp61", "slp66", "slp67x".

B. Windowing

Online drowsiness detection requires analysing a recorded data flux. This flux has to be divided into frames with enough samples so that the analysis is consistent from a statistical point of view.

Denote as $\mathbf{x}^1$ and $\mathbf{x}^2$ the vectors related to EEG and ECG signals, respectively. Let us also assume that the recorded data have a unique duration T. Since EEG and ECG data may be acquired with different sampling rates, the windowig process has to provide frames with two channels of different sizes. The ith frame f_i is made up of EEG and ECG signal vectors $\mathbf{x}_i^1 \in \mathbb{R}^{s_1}$ and $\mathbf{x}_i^2 \in \mathbb{R}^{s_2}$ concatenated to $f_i = \{\mathbf{x}_i^1, \mathbf{x}_i^2\}$.

For our data, as there is one annotation of the sleep stage every 30 seconds, EEG and ECG recordings are split into frames with duration $d = 30$ seconds in order to ensure annotation consistency, with an overlap of 2 seconds. Moreover, frame overlapping is allowed in order to adapt the proposed model to online drowsiness detection.

C. Unbalanced data

To bypass the unbalanced dataset (more awakeness than sleep stages), we resort to data augmentation. Indeed, several works have integrated data augmentation techniques [17, 18] in the training step in order to avoid over-fitting and improve the performance of the networks in terms of classification accuracy. We use the Random Over-sampling technique [19]: random data duplication is performed from the minority class.

III. DETECTION MODEL

The proposed model for joint EEG and ECG data analysis relies on two different analyses, both performed using artificial neural networks: frame and temporal analyses.

A. Frame analysis

This analysis is designed to characterize the drowsy state using data frames passed through a CNN. Each frame $f_i = \{\mathbf{x}_i^1, \mathbf{x}_i^2\}$ is analysed independently from the others. However, $\mathbf{x}_i^1$ and $\mathbf{x}_i^2$ are analysed simultaneously within each frame using a CNN with two channels. Convolution with 1D filters is performed over each channel. Inspired by classics CNN [20], the used CNN is made up of a succession of 6 layers, each of them involving 1 Convolution, 1 MaxPooling and 1 Batch-normalization layer. Finally, a Dropout and a GlobalAverage-Pooling layers are applied.

B. Temporal analysis

In addition to the frame analysis, drowsiness is supposed to be a state that occurs after a progressive attention decrease. As a consequence, analysing segments resulting from successive frames seems to be relevant. Each frame f_i is split into n blocks of dimension m. A short-term Fourier transform (STFT) is applied to each frame signal. Only the positive spectrum part is retained due to the symmetry property. According to [21], build a time-frequency representation simplifies the learning task and in their case increases performances. As will be illustrated in Section IV, this representation allows to reach high accuracy rates.

Finally, a recurrent neural network (RNN) is used to characterize the temporal evolution of different frame. In order to capture the memory effect between successive frames, a Long short-term memory (LSTM) network is used [22]. LSTMs have been widely used in the literature especially for their ability to overcome the vanishing gradient limitation by adding weights managing the memory. Inspired by [23], the LSTM architecture is made up with two layers of n LSTM cells. The complexity of this architecture depends on two parameters: the output y of the LSTM network and the input of the cells from the first layer m. It is worth noting that the complexity of the network decreases as long as n increases.

C. Proposed architecture

Based on the frame and temporal analysis, the proposed architecture is illustrated in Fig. 1. This architecture is based on three parallel networks. The first network is the one ensuring the frame analysis on both EEG and ECG channels. After the windowing step, the two-channels data are introduced to a CNN. An additional 10-neurons layer is applied (the number of neurons is determined by cross-validation).

Two other parallel networks are used to perform the temporal analysis. EEG and ECG data are processed separately using an LSTM with additional 5-neurons output layer, determined too with cross-validation. The outputs of the three networks are presented to a dense layer leading to the final binary classification.

IV. RESULTS

A. Validation with a frame leave-p-out strategy

In the first experiments, we asses the performance of the proposed architecture for drowsy and awake frames classifi-cation. EEG and ECG recordings from the database described in Section II-A are processed as indicated in Section II. Frames from the different subjects are mixed and split into two groups: 70% for learning and 30 % for testing. With a final accuracy score of 97%, Fig. 2 displays the accuracy and loss curves with respect to (w.r.t.) the number of epochs. The curves confirm the good convergence properties after approximately 300 epochs using the "Adam" optimizer. Note that the similarity between the values of the learning and
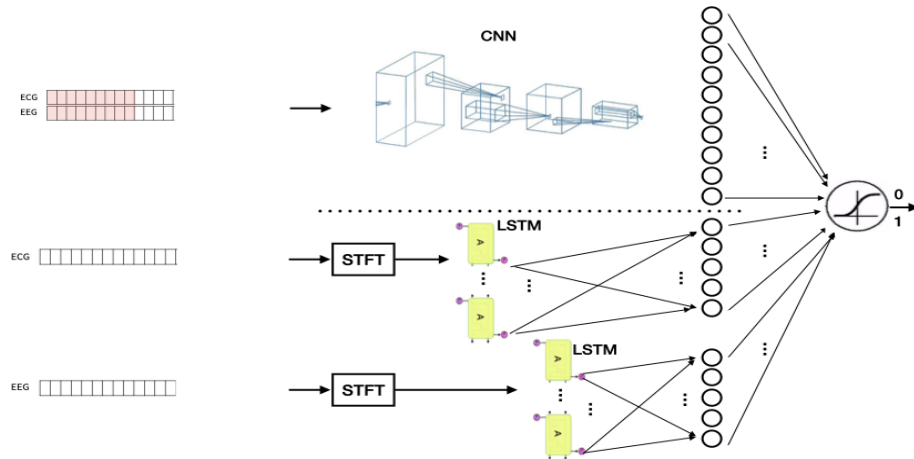
Fig. 1. Proposed architecture for drowsiness detection.

validation accuracy and the loss function confirm the good generalization of the proposed architecture. This is partly due to the use of a high dropout level. Based on multiple experimental settings, the best configuration was obtained with a dropout of 0.85 for the convolutional part and 0.5 before the final classification. The obtained accuracy rate is higher that the one reported in [10] (93%) where the authors showed that their approach outperforms other state-of-the-art methods.
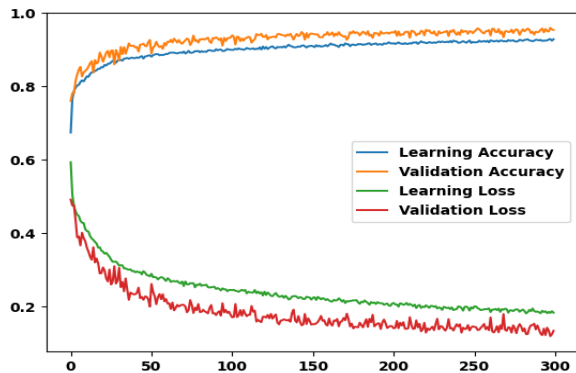

Fig. 2. Accuracy and loss curves w.r.t. epochs.

Fig. 3 displays the confusion matrix where the "0" label corresponds to the drowsy state. From this matrix, sensitivity (the ratio of True Positive to the number of element of the considered class) and specificity (the ratio of True Negative to the number of element of the other class) scores considering Drowsy class can be calculated, and are respectively 97.3% and 96.6%.
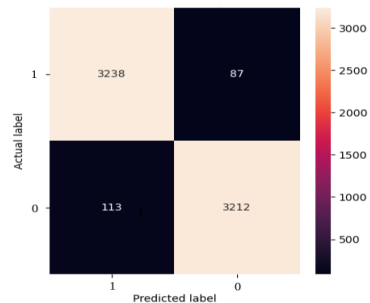

Fig. 3. Confusion Matrix.

B. Validation with a subject leave-p-out strategy

In this experiment, drowsiness detection is performed using the architecture in Fig. 1. However, the dataset has first been split into two groups: 10 subjects have been dedicated for the learning phase ("slp02a", "slp02b", "slp03", "slp04", "slp14", "slp48", "slp59", "slp60", "slp61", "slp67x"), while 3 subjects have been kept for test ("slp16", "slp66", "slp45"). During the learning phase, no frames from the second group are presented to the model. Consequently, the model cannot learn properties linked to the subject specificities.
Frame windowing is then performed within each group.
The obtained results show an accuracy decrease to 70%, which means that intrinsic information related to a given subject plays an important role in the drowsy state detection. It is worth noting that in [24] the authors report a sensitivity decrease from 80.3% to 41.4% in this case. **Moreover, in ... []**
For the proposed method, a sensitivity decrease from 97.3% to 82% is reported (see Table I). However, the decrease in specificity is more significant (from 96.6% to 67%), which means that a higher false alarm rate is expected if no frames corresponding to the target subject are provided for the learning phase.
In the following Section we develop an improvement of the proposed method to overcome this limitation related to the training conditions.

3

C. Autoencoded detection

As discussed in the previous Section, it has been reported in the literature that when testing classifiers on data from subjects not present in the learning database, a sensitivity drop is always observed [24, 25]. In order to limit overfitting with intrinsic features related to the subject from EEG and ECG data, we propose the use of autoencoders before feeding the data to our classifier. Autoencoders have been used in the literature [26] as unsupervised deep learning models where a first group of weights is trained to encode the input data into a representation vector of lower dimension, and a second group is trained to decode this new representation and reconstruct the original input. After the encoding/decoding steps, our expectation is to make the model forget intrinsic features of the subject. By reducing the dimensionality, autoencoders provide a better generalization for the input data. Two autoencoders handling EEG and ECG data are therefore introduced in the proposed architecture as illustrated in Fig. 4. Both autoencoders have the same architecture: three layers of Convolution-MaxPooling-BatchNormalization to encode the signal into a dimension of 625 features, and three symmetrics Convolution-UpSampling-BatchNormalization layers to decode the signal. Once these two autoencoders have been trained, the ECG and EEG data are first encoded into the smaller dimension and then fed to the neural networks for the frame and temporal analyses.

The proposed model is trained as for the experiment in Section IV-B. The obtained accuracy and loss curves are displayed in Fig. 5 with a best accuracy score equal to 84%.

This result confirms the importance of the autoencoding step, and shows that we improved classification performance (in terms of accuracy score) compared to [24] (direct comparison is not possible due to data availability issues).

To summarize the quantitative evaluation of the proposed architectures performance, Table I reports sensitivity and specificity scores for the proposed architectures with and without autoencoders, and this for both frame and subject leave-p-out strategies. According to the reported scores, a sensitivity and specificity drop is reported with the autoencoded architecture for the frame leave-p-out strategy, as expected. Indeed, the autoencoders effect is useless in this case due to the learning procedure where the model has the possibility to learn from the target subjects. As regards the subject leave-p-out detection, the reported gain in both sensitivity and specificity clearly shows the effectiveness of the proposed architecture with scores higher than those reported in the recent literature (for example [25] with accuracy score of 76% with a leave-one-out strategy).

On the other hand, we can notice that the sensitivity score is stable with and without autoencoding if the subject leave-p-out strategy is considered. However, autoencoding clearly helps improving specificity performance.

V. CONCLUSION

This paper studied a new method for drowsiness detection using EEG and ECG data jointly. The proposed method is

TABLE I
SENSITIVITY AND SPECIFICITY SCORES WITH AND WITHOUT AUTOENCODERS, FOR BOTH FRAME AND SUBJECT LEAVE-P-OUT STRATEGIES.

		Without autoencoder	With autoencoder
Sensitivity	Frame p-out	97.3%	91.6%
	Subj. p-out	82%	84%
Specificity	Frame p-out	96.6%	90.5
	Subj. p-out	67%	81%

based on a combination of convolutional and recurrent neural networks. A very accurate detection was obtained, especially when the learning step involves frames from the target subjects. When learning from the target subject is impossible, the accuracy of the algorithm drops to 70%. To overcome this performance loss, an autoencoded version is proposed to reach an accuracy level of 83.38%. The proposed final architecture is therefore very efficient, enjoys high generalization level, and can be used for applications where calibration using data from the target subject is impossible.

Future work will focus on investigating the use of fractal features to improve the performance of the proposed model, as well as large scale validation with data collected using connected EEG and ECG devices.

REFERENCES

[1] J. Yu, S. Park, S. Lee, and M. Jeon, "Driver drowsiness detection using condition-adaptive representation learning framework," *IEEE Trans. on Intell. Transp. Syst.*, pp. 1–13, December 2018.

[2] A. Sahayadhas, K. Sundaraj, and M. Murugappan, "Electromyogram signal based hypovigilance detection," *Biomedical Research (India)*, vol. 25, pp. 281–288, 2014.

[3] F.Wang, H.Wang, and R.Fu, "Real-time ECG-based detection of fatigue driving using sample entropy," *Entropy*, vol. 20, no. 3, 2018.

[4] S. Ahn, T. Nguyen, H. Jang, J. G. Kim, and S. C. Jun, "Exploring neuro-physiological correlates of drivers' mental fatigue caused by sleep deprivation using simultaneous EEG, ECG, and fNIRS data," *Frontiers in Human Neuroscience*, vol. 10, pp. 219, May 2016.

[5] C. Basri, K .Wan, I. Zunaidi, L. Lee, A. B. Shahriman, M. Razlan, Z. Mustafa, W. Zulkepli, and N. Noriman, "Muscle fatigue detections during arm movement using EMG signal," *IOP Conference Series: Materials Science and Engineering*, vol. 557, pp. 012004, June 2019.

[6] A. Ishiim, T. Ishizuka, M. Tanaka, Y. Muta, Y. Watanabe, and E. Ya-mano, "The neural effects of positively and negatively re-experiencing mental fatigue sensation: a magnetoencephalography study," *Exp. Brain Res.*, , no. 6, pp. 1735–1747, June 2018.

[7] A. Craig, Y. Tran, N. Wijesuriya, and H. T. Nguyen, "Regional brain wave activity changes associated with fatigue," *Psychophysiology*, vol. 49, no. 4, pp. 574–582, Feb. 2012.

[8] M. Gromer, D. Salb, T. Walzer, Natividad M. Madrid, and R. Seepold, "Ecg sensor for detection of driver's drowsiness," *Procedia Comput. Sci.*, vol. 159, pp. 1938–1946, Jan. 2019.

[9] T. Lajnef, S. Chaibi, P. Ruby, P.-E. Aguera, J.-B. Eichen Laub, M. Samet, A. Kachouri, and K. Jerbi, "Learning machines and sleeping brains: Automatic sleep stage classification using decision-tree multi-class support vector machines," *J. of Neuroscience Methods*, vol. 250, pp. 94—105, July 2015.

[10] Y. Abichou, S. Chaabene, and L. Chaari, "A sleep monitoring method with EEG signals," in *Digital Health in Focus of Predictive, Preventive and Personalised Medicine*, L. Chaari, Ed. 2020, pp. 25–32, Springer.

[11] H. Zeng, Y. Chen, D. Guojun, Q. Feiwei, Z. Jianhai, and K. Wanzeng, "EEG classification of driver mental states by deep learning," *Cognitive Neurodynamics*, vol. 12, no. 6, pp. 597–606, July 2018.

[12] K. Wonjun, O. Kwanseok, J. Eunjin, and H. I. Suk, "Vignet: A deep convolutional neural network for EEG-based driver vigilance
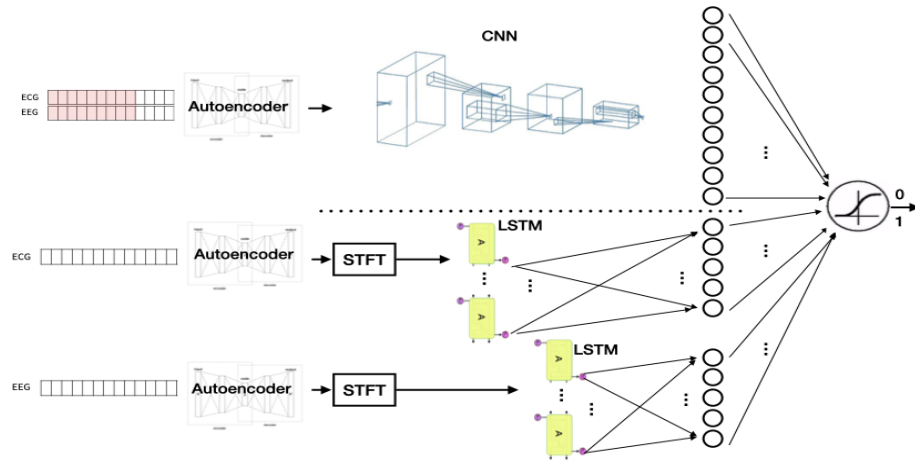
Fig. 4. Proposed architecture for drowsiness detection using autoencoders.
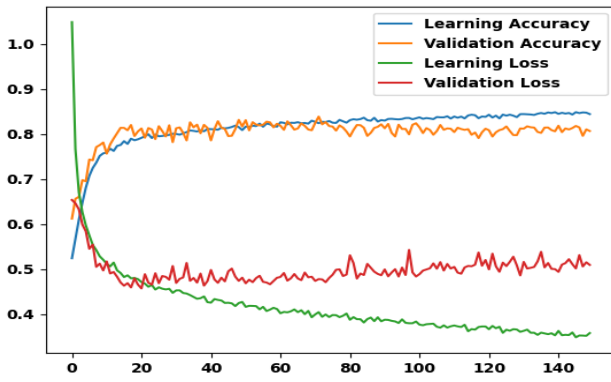

Fig. 5. Accuracy and loss curves using autoencoders.

estimation," in *Proc. Int. Winter Conf. on Brain-Computer Interface*, Gangwon, Korea, Feb. 2020.

[13] A. Boudaya, S. Chaabene, L. Chaari, B. Bouaziz, A. Ammar, and A. Hokelmann, "EEG-based hypo-vigilance detection using convolutional neural network," in *The Impact of Digital Technologies on Public Health in Developed and Developing Countries*. 2020, pp. 69–78, Springer.

[14] F. Bourghelle, "Development of an automatic drowsiness monitoring system using the electrocardiogram," *Master thesis*, pp. 14–46, 2015-2016, University of Liège.

[15] M. Awais, N. Badruddin, and M. Drieberg, "A hybrid approach to detect driver drowsinessutilizing physiological signals to improve systemperformance and wearability," *Sensors (MDPI)*, vol. 17, no. 9, pp. 1–16, August 2017.

[16] A. Goldberger, L. Amaral, L. Glass, Hausdorff, Ivanov J., R. P. C., Mark, J. E. Mietus, G. B. Moody, C. K. Peng, and H. E. Stanley, "Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[17] J. Salamon and J. P. Bello, "Deep convolutional neural networks and data augmentation for environmental sound classification," *IEEE Signal Process. Letters*, vol. 24, no. 3, pp. 279–283, 2017.

[18] Y. Roy, H. Banville, I. Albuquerque, A. Gramfort, T. H. Falk, and J. Faubert, "Deep learning-based electroencephalography analysis: a systematic review," *J Neural Eng.*, vol. 16, pp. 1—38, July 2019.

[19] J. L. Leevy, T. M. Khoshgoftaar, R. A. Bauder, and N. Seliya, "A survey on addressing high-class imbalance in big data," *J. Big Data*, vol. 5, no. 42, pp. 1–30, 2018.

[20] Hinton G LeCun Y, Bengio Y, "Deep learning.," *nature 521:436–444*, 2015.

[21] G. Ruffini, D. Ibañez, M. Castellano, L. Dubreuil-Vall, A. Soria-Frisch, R. Postuma, J. Gagnon, and J. Montplaisir, "Deep learning with eeg spectrograms in rapid eye movement behavior disorder," *Frontiers in Neurology*, vol. 10, pp. 1–9, July 2019.

[22] S. Kamilya and J. Alex Pappachen, "A survey on LSTM memristive neural network architectures and applications," *The European Physical Journal Special Topics*, vol. 228, no. 10, pp. 2313–2324, 2019.

[23] G. Zhang, V. Davoodnia, A. Sepas-Moghaddam, Y. Zhang, and A. Etemad, "Classification of hand movements from EEG usinga deep attention-based LSTM network," *IEEE Sensors Journal*, vol. 20, no. 6, 2019.

[24] H. Martensson, O. Keelan, and C. Ahlström, "Driver sleepiness classification based on physiological data and driving performance from real road driving," *IEEE Trans. on Intell. Transp. Syst.*, vol. 20, no. 2, pp. 421–430, 2019.

[25] J. H. Jeong, B. W. Yu, D. H. Lee, and S. W. Lee, "Classification of drowsiness levels based on a deep spatio-temporal convolutional bidirectional LSTM network using electroencephalography signals," *Brain sciences*, vol. 9, no. 12, pp. 1–18, 2019.

[26] E. Q. Wu, X. Y. Peng, C. Z. Zhang, J. X. Lin, and R. S. F. Sheng, "Pilots' fatigue status recognition using deep contractive autoencoder network," *IEEE Trans. on Instrum. and Meas.*, vol. 68, no. 10, pp. 3907–3919, 2019.