

Attention Estimation in Virtual Reality with EEG based Image Regression

Victor Delvigne
ISIA Lab, Faculty of Engineering
University of Mons
Mons, Belgium
IMT Lille Douai
CRISTAL UMR CNRS 9189
Villeneuve d'Ascq, France
victor.delvigne@umons.ac.be

Hazem Wannous, Jean-Philippe Vandeborre
IMT Lille Douai
CRISTAL UMR CNRS 9189
Villeneuve d'Ascq, France

Laurence Ris
Neuroscience Department Lab
Faculty of Medicine and Pharmacy
University of Mons
Mons, Belgium

Thierry Dutoit
ISIA Lab
Faculty of Engineering
University of Mons
Mons, Belgium

Abstract—Attention Deficit Hyperactivity Disorder (ADHD) is a neurodevelopmental disorder affecting a certain amount of children and their way of living. A novel method to treat this disorder is to use Brain-Computer Interfaces (BCI) throughout the patient learns to self-regulate his symptoms by herself. In this context, researches have led to tools aiming to estimate the attention toward these interfaces. In parallel, the democratization of virtual reality (VR) headset, and the fact that it produces valid environments for several aspects: safe, flexible and ecologically valid have led to an increase of its use for BCI application. Another point is that Artificial Intelligence (AI) is more and more developed in different domain among which medical application. In this paper, we present an innovative method aiming to estimate attention from the measurement of physiological signals: Electroencephalogram (EEG), gaze direction and head movement. This framework is developed to assess attention in VR environments. We propose a novel approach for feature extraction and a dedicated Machine Learning model. The pilot study has been applied on a set of volunteer and our approach presents a lower error rate in comparison with the state of the art methods.

Keywords—Virtual Reality, Machine Learning, Brain-Computer Interface, Eye-tracking

I. INTRODUCTION

Attention Deficit Hyperactivity Disorder (ADHD) is a neurological disease affecting around 5% of children in the world according to Raman et al. 2018 [26], these figures varying with the location. The syndromes vary in function of the patient subgroup: Hyperactive (ADHD-H), Inattentive (ADHD-I) or Combined (ADHD-C) as explained in the DSM-V [2]. A higher representation of ADHD-I and ADHD-C compared to ADHD-H is reported in Wilcutt 2013 [32]. Among the

different treatments considered the more often preconised is methylphenidate daily intake. Nevertheless, children medication may present controversial aspects and medication is not effective for 20 to 30% of the patient according to Cueli et al. 2019 [10]. A novel method to reduce the symptoms induced by ADHD is the use of Neurofeedback (NF) training. NF consists of a real-time representation of the cerebral activity via an understandable form, e.g. video-game or animation evolution according to inhibition or activation of certain brain regions.

Nowadays, an increasing amount of researches considering the use of Brain-Computer Interfaces (BCI) has been reported. BCI are used in different applications and contexts: health [21], entertainment [1], (neuro)marketing [12], etc. During the last ten years, an increasing number of BCI has been developed in virtual reality (VR) environments. Bashiri et al. 2017 [3] made a review of all the research project using VR in the context of ADHD. As reported in their review, the use of VR presents several motivations including larger freedom for environments creation, combined to higher security in a more ecologically valid environment for researches as explained by Pollak et al. 2009 [25]. Moreover, the recent VR headset allows recording of several physiological signals from the participant, e.g. head position or gaze information.

In another side, AI and dedicated algorithms, e.g. Machine-Learning (ML), are more and more considered in everyday life. In particular, in medical applications where ML can help to detect diseases from biomedical signals [15]; to help in the understanding of genomes [17] and human behaviours [24]; etc. Moreover, researches using ML in VR context have also been made as in movement generation and prediction as

presented in J. Dehesa et al. 2020 [11] or Natural Language Processing in VR as shown in E. Kolve et al. 2017 [16].

In this context, we propose in this paper a BCI application aiming to automatically estimate the visual attention from (neuro)physiological acquisition: Electroencephalogram (EEG); eye-tracking signal (i.e. gaze direction and pupils diameter) and head linear and rotational position in VR environment. The framework developed for this purpose aiming to help to have a better understanding of the attention mechanism in the human brain.

II. RELATED WORK

Different works aiming to estimate attention in VR environments have already been made. In the work proposed by Blume et al. [7], they proposed an environment representing a VR classroom where EEG recordings have been made. They propose an attention assessment based on frequential features from EEG, the considered feature being the ratio between two specific frequency bands: θ and β bands. They show that the use of VR instead of the 2D screen may increase the results of the tasks reflecting attention after multiple training overtimes. Other recordings have also been considered to characterize the attention state: functional near-infrared spectroscopy (fNIRS) representing the cerebral activity and electromyogram (EMG) representing the muscular activity of the in the upper back region.

Another methodology to assess visual attention is presented by García-Baos et al. [14]. In their work, the attention is assessed with gaze information. As presented by Varela Casal et al. [30], biomarkers can be deduced from eye-tracking information. From the attention computed from the eye-tracking signal, a tool promoting attention has been developed. The attention computation is based on the measurement of the fixation time to a target in the centre of the screen. The framework has been tested on a population diagnosed with ADHD and shows an improvement in impulsivity and reaction time in comparison with a control group.

In another domain, different research projects aiming to estimate attention state from neurophysiological records have been made to assess vigilance during a driving task as presented by Zheng and Lu 2017 [34] and Cao et al. 2019 [8] in their works. Both of them present a database composed of EEG recordings during attentional tasks.

For all works presented above, the signal processing pipeline considered to estimate attention from the physiological recording is composed of the following key steps:

- Signal acquisition and pre-processing: in this step the signal is recorded during the proceeding of the task. The noise and artefacts, e.g. muscular movements, eye blinks, electrical noise, etc, are removed with frequency filtering and/or independent component analysis (ICA) as shown in Winkler et al. 2011 [33].
- Feature extraction: during this step, the goal is to extract the most discriminant information from the signals to reduce the dimension and to help the signal classifier and/or regressor. From EEG, three types of feature can

be extracted: temporal, spectral (i.e. frequential) and spatial features as presented in Lotte 2014 [19]. A similar approach can also be considered for physiological signals by computing, for instance, the mean variation or acceleration of the signals. To cover as much information as possible the feature can be combined. Other methods independent from the signal nature can also be applied, e.g. principal component analysis (PCA), Riemannian space transformation, t-distributed stochastic neighbour embedding (t-SNE), etc as shown in the works of Lotte et al. 2018 [18]. In recent works, feature arrays have also been considered under different forms. For instance, in their works, Bashivan et al. 2016 [4] considered EEG feature as images with images pixels corresponding to electrodes locations in 3D. Another feature representation method is based on the graph where the nodes are corresponding to the electrodes and edges are proportional to their distance as presented by Zhong et al. 2020 [35].

- Classification/Regression: the goal is to estimate the feature vector described above. For this purpose, it is possible to consider the use of different *traditional* ML algorithms: linear regression, Gaussian classifier, support vector machines (SVM) [29] or random forest (RF) [19]. In the last ten years, other models more robust and complex have been developed but needed a larger amount of data: deep learning (DL) models. DL models have already been used in the context of emotion estimation as shown in the works of Gao et al. 2019 [13] where they consider convolutional neural networks (CNN) to extract feature from EEG to estimate attention state in driving task. Other DL models have also been used with BCI application, e.g. Graph Convolution Network [35] for emotion estimation, Recurrent Neural Network (RNN) for motor activity estimation [4] or Autoencoder for epileptic seizure detection [31].

III. METHODOLOGY

The pipeline of our approach is presented in Fig. 1.

A. EEG Signal Processing Pipeline

As presented in the project pipeline, one of the main steps is the signal acquisition and processing, the considered signals being EEG, eye-tracking and head position. The EEG has been recorded on actiCHamp medical recorder from Brainvision with 32 electrodes headset following the 10/20 electrodes disposition [20]. The VR headset taking space on the back of the scalp, the configuration had to be modified by replacing some electrodes position: P3 \rightarrow AF3; Pz \rightarrow FCz; P4 \rightarrow AF4. A conductive gel has been applied to the participant's scalp to favourite the conduction between electrodes and skin. The recording has proceeded with Brain Vision Recorder software provided with EEG headset. The signal has been automatically filtered between 0-140 Hz during the recording, moreover, FCz electrode has been considered as the reference electrodes. The synchronisation between the stimuli apparition in the VR environment and EEG recording is ensured by automatic

annotations. Before extracting the feature, we proceed to mean filtering on EEG to remove artefacts (e.g. electrical noises and muscular artefacts).

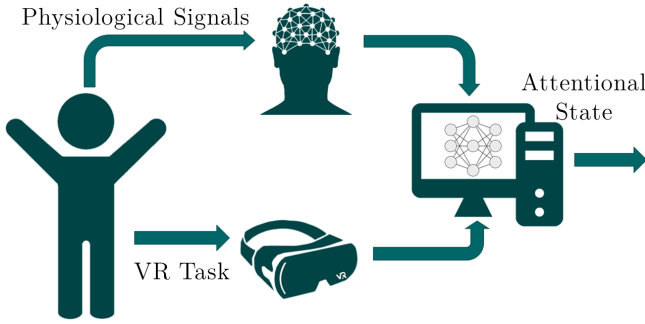


Fig. 1. Framework pipeline. It describes a summary of the framework used in this work with the different steps: VR task, Signal Acquisition and Attentional State computation.

Then, the EEG signals have been separated in windows of four seconds length around the stimuli apparition (-1 second before and +3 second after). Each window corresponding to one stimulus, various overlapping can be considered in the function of the stimulus apparition time.

To analyse the signal, a feature extraction step has been made. The considered features are power spectral density (PSD) of specific frequency bands. In the context of EEG study, different rhythms have been scientifically defined each of them related to cognitive function as explained in Collura and Siever 2009 [9]:

- delta band (0 - 4 Hz) - deep sleep state.
- theta band (4 - 8 Hz) - task related to memory.
- alpha band (8 - 12 Hz) - relaxation eye open and closed.
- beta bands (12 - 30 Hz) - physical activity, movement, excitation.
- gamma band (30 - 45 Hz) - nervous state, high activity and excitation.

It is important to note that the frequency bands limits and their corresponding state can vary from a participant to another as shown in Bazanova and Aftanas 2010 [5]. A method based on the recording of the resting state eye open and closed has been presented in their study.

As said earlier, it has already been reported that the attentional state can be observed in theta and beta bands or their ratio as shown in the work of Bioulac et al. 2019 [6]. The PSD for each frequency band has been computed from the raw EEG. From the five different frequency bands, we consider only three bands: theta, alpha and beta. This choice has been motivated by the fact that delta bands are more presented in the deep sleeping state [9] which is not our case here and that the gamma bands tend to be more affected by noise (e.g. electrical noise having a frequency of 50 Hz) and are also merge with beta band in several studies, their contribution being minors.

From the feature array representing the spectral information with the dimension $dim = [n_{sample} \times n_{electrodes} \times n_{frequency\ bands}]$ we construct an image as described in the work of Bashivan et al. 2016 [4]:

- azimuthal projection from 3D electrodes Cartesian coordinates system in 2D coordinate projection.
- interpolation of the discrete point in 2D array to a continuous image representation with bicubic interpolation.

B. VR Environments

Five environments have been designed in the framework as shown in Fig. 2.

Each of them corresponding to a life scene: bedroom; birthday party; amusement park; forest; gym hall. It was chosen to work with five environments to allow the participant to choose the place where he feels the most comfortable in. In each of these environments three tasks of five minutes duration are performed:

- relaxation
- selective attention
- sustained attention

To assess attention, visual stimuli related to the environment appear (e.g. butterflies in the forest) and it is asked to the participant to look in its direction without moving the head. The difference between the second and third task is the apparition of a combination of right and wrong stimuli (i.e. to look at or not) in the selective attention task, whereas only one type of stimulus appears in the sustained attention task.

During the selective attention task, two stimuli appear with a variable frequency, the time interstimulus interval (ISI) is randomly chosen in an interval: $ISI \in [3000 - 250; 3000 + 250]ms$. It is asked to the participant to look at the direction of the target predefined before the acquisition. We then measure the time elapsed during which the participant looked at the target t_{target} and wrong stimulus t_{wrong} and the apparition duration $t_{apparition}$.

Throughout the sustained attention task, a similar approach is considered with the apparition of only the target and a different $ISI \in [3000 - 500; 3000 + 500]ms$. The time elapsed between the target apparition and the moment at which the participants look at the target $t_{elapsed}$. We noted that the ISI duration parameters have been chosen after different try and that these parameters can be easily tuned in the framework.

The environments have been developed in C# with Unity software, the full recording framework is freely available on github¹.

C. Eye tracking Recording - Analysis

Eye-tracking signals have been recorded with the eye-tracker placed into the HTC Vive Pro Eye VR headset. The physiological signals are recorded two times with different approaches: 1). Constant recording during the three tasks at 5 Hz. The eyes position in 3D VR environment and their direction are registered in an array. 2). The time measurement for the attentional tasks for each stimuli apparition. Simultaneously, other physiological signals are also recorded from the framework with the VR headset at 5 HZ: the rotational and linear head position and the pupils' diameter over time.

¹<https://github.com/VDeIv/VERA>

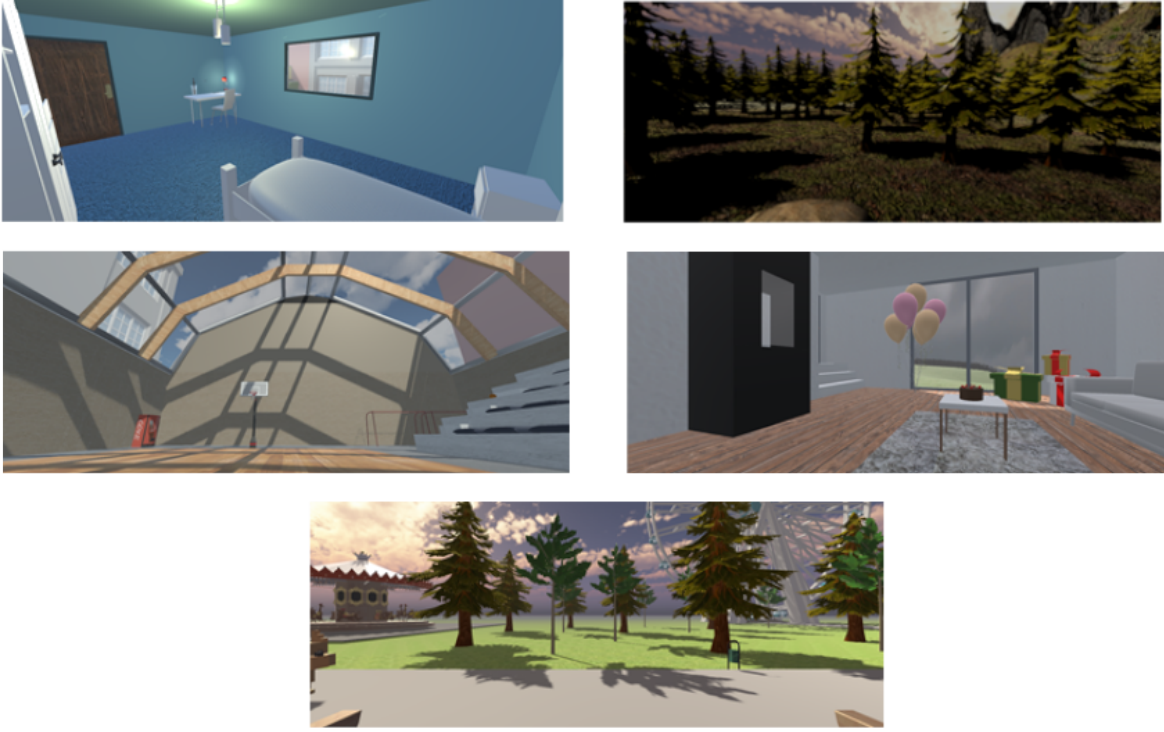


Fig. 2. Virtual Environments. The five environments representing lifespan scenes are displayed: bedroom, walk in the forest, gymhall, birthday party and amusement park.

From the head position and pupils diameter, we compute the derivative to study the effect of parameters evolution over time.

From the time measurement t_{target} , t_{wrong} , $t_{elapsed}$ we deduced an attention score for the second and third task at the time t :

$$score_{t2}(t) = \text{sigm} \left(\frac{t_{target}(t) - t_{wrong}(t)}{t_{apparition}(t)} \right) \quad (1)$$

with $score_{t2}(t)$ the score at moment t for the second task, $\text{sigm}(x) = \frac{1}{1+e^{-k*x}}$ the sigmoid function (k being a scaling parameter).

$$score_{t3}(t) = \text{sigm} \left(t_{elapsed}(t) - \overline{t_{elapsed}} \right) \quad (2)$$

With $\overline{t_{elapsed}}$ the mean time per participant for all the stimuli of the third task, this normalization helping to have a better score generalization.

After the following steps, a score representing the visual attention between 0 (totally not focus) and 1 (totally focus) has been computed for each stimuli apparition. For each stimulus, a corresponding image representing the spectral information from the EEG and the variation of the physiological signals are available.

D. Signal Classification Method

The goal of this step is to retrieve the attentional score with the EEG spectral feature and the physiological signals. Different approaches have been studied: considering as input the images created with EEG and/or the physiological signals.

The model presented in this paper is a convolutional neural network (CNN) followed by fully-connected layers. This approach is inspired by the VGG network presented in Simonyan and Zisserman 2015 [28], Bashivan et al. 2016 [4] showed a similar approach for motor imagery in the context of electroencephalogram signals processing.

The network configuration is shown in Fig. 3 and described in Table I. The activation function considered is the ReLU except for the output layer where a sigmoid function is used to have output value in range $[0, 1]$.

The dataset was composed by two inputs of a respective dimension of $[n_{sample} \times 7]$ for the physiological signals (with \ddot{x} , \ddot{y} , \ddot{z} , $\ddot{\theta}$, $\ddot{\varphi}$, $\ddot{\rho}$ for the linear and rotational acceleration and Δd for the pupils diameter evolution) and $[n_{sample} \times 3 \times 32 \times 32]$ for the *EEG images*.

Other ML-based approaches have also been considered: SVM, Random Forest Regressor (RF) and Multi-Layer Perceptron (MLP). For these classical ML algorithm, the EEG bandwidth have directly been considered as array of dimension $[n_{sample} \times n_{channel} \times 3]$.

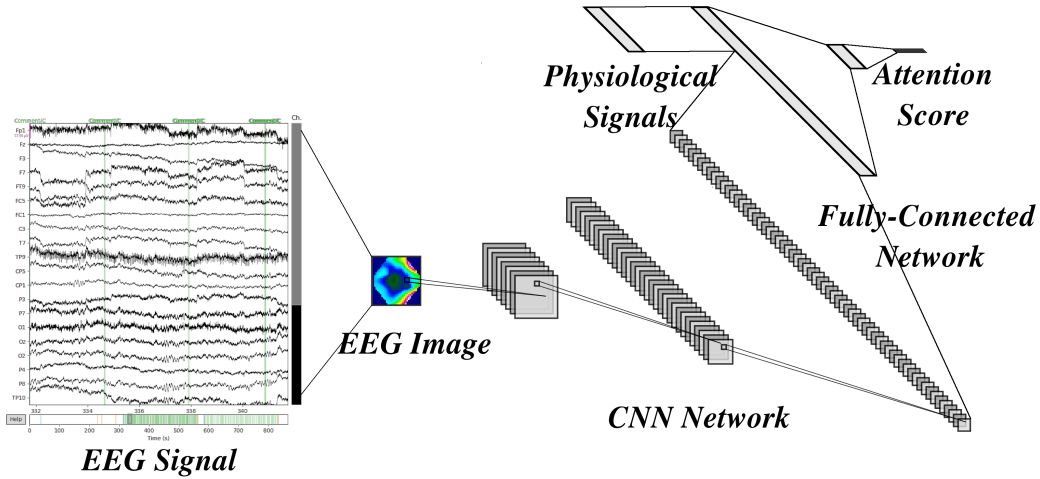


Fig. 3. CNN Architecture. Overview of our approach with the different steps: transformation of the EEG into images; feature extraction into convolutional network; concatenation with physiological signals and regression to estimate the attention score.

CNN Configuration	
input image	$[batchsize \times 3 \times 32 \times 32]$
Conv2d	(out _{channel} = 16)
Conv2d	(out _{channel} = 64)
Conv2d	(out _{channel} = 64)
BatchNormalization	
MaxPool2d	(2 × 2)
Conv2d	(out _{channel} = 64)
Conv2d	(out _{channel} = 128)
BatchNormalization	
MaxPool2d	(2 × 2)
Conv2d	(out _{channel} = 128)
MaxPool2d	(2 × 2)
FC	- 16
Sigmoid	

TABLE I
MODEL CONFIGURATION.

Architecture proposed for the final CNN model inspired by the VGG architecture from Simonyan and Zisserman 2015 [28].

IV. EXPERIMENTS

A. Study Design

The study has been elaborated in collaboration with the psychologist in the context of a multi-disciplinary research project. Moreover, it has been approved by the related ethical board.

Five volunteers have participated in the fifteen minutes study, as presented in Roy et al. 2019 [27] the dataset length can be considered as sufficient for BCI application. The restriction for the participation of the study was not to present any neurological issues and difficulties with VR. No issues were reported from the participants before or after the acquisition. The full experimentation follows the procedure:

- Briefing about the study and short explanation about electroencephalogram.

- Preparation of the acquisition: gel and electrodes application and devices calibration.
- Signal acquisition.
- Study debriefing, the possibility of signal rejection in case of a bad understanding of instructions or acquisition issues.

B. Results

The mean score for the two tasks of each participant are listed in Table II. As shown, the score varied from a participant to another after considering the subject based normalisation (see Equation 1 and 2) which showing a correct score distribution.

# Subject	2 nd Task	3 rd Task
1	64.86/16.34	80.36/11.89
2	63.36/16.88	83.49/10.51
3	68.67/5.41	51.22/6.98
4	69.44/3.50	42.08/4.92
5	70.29/3.86	51.87/4.59

TABLE II
MEAN ATTENTION SCORE AND STD FOR PARTICIPANTS.

The average score for the five participant to the pilot study at the second and third attentional task and the respective standard deviation.

After the signal acquisition, we preprocess the signal to extract the feature and adapt the data in a correct format depending on the model.

The architecture considered for the CNN model is presented in Fig. 3. Different approaches have been considered: only EEG images or physiological signals and the combination of both. The results of the different models are presented in Table III.

The implementation of the CNN model has been made with PyTorch 1.5.1 [22] and perform our training on NVIDIA RTX2060 GPU with 6GB GPU memory. The optimizer chosen is Adam with a learning rate of 0.001 and a mean absolute

Model	Dataset	Cross-Validation Error [%]
CNN	EEG Images	5.93
CNN	Combined Dataset	4.56
MLP	EEG Frequency Bands	8.44
MLP	Phy. Signal	8.53
MLP	Combined Dataset	7.47
SVM	EEG Frequency Bands	6.98
SVM	Phy. Signal	6.16
SVM	Combined Dataset	7.5
RF	EEG Frequency Bands	7.49
RF	Phy. Signal	6.68
RF	Combined Dataset	6.72

TABLE III
ERROR FOR DIFFERENT MODELS.

error loss (i.e. L1Loss), a batch size of 64 and 150 epochs. For the other ML models, the sci-kit-learn library [23] has been used and the corresponding hyperparameters have been fine-tuned. The documented code is available on github². Different model combinations (e.g. the number of layer, kernel size, activation function, etc), loss function or hyperparameters have been tried but we keep only the ones with the lowest error rate.

As shown in Table III, the model with the lowest error rate is our approach considering EEG as images with the combined dataset (i.e. EEG+phy signals). The considered error is the leave-one-out subject cross-validation error, i.e. the model is trained with the $N - 1$ patient signals and test with the remaining one. In this case, our model can estimate the correct score with a margin of $\pm 4.5\%$ that outperform the other ML-based approaches.

Moreover in Table IV we can see the error rate per participant.

From this result we can note that the error rate is rising for some participant compared to other (e.g. S1 and S2), this could be explained by the fact that the model has not enough data to converge to the global minimum. This issue could be solved in future works with more recorded signals.

V. CONCLUSION

In this work, we propose a novel framework aiming to assess attention in VR environments from various physiological signals: EEG, eye-tracking signals and head position. To avoid possible parasitic effects induce by stress or nervousness, the environments have been designed to allow the participant to choose the place he feels the most comfortable in. The framework purpose is to estimate the evolution of visual attention computed from attentional tasks, to implement the attention detector in neurofeedback aiming to reduce the symptoms of ADHD. The method presented in this work outperforms the

²<https://github.com/VDelv/Attention-EEG>

# Subject	S1	S2	S3	S4	S5
Error [%]	7.5	7.52	3.16	2.25	2.38

TABLE IV
ERROR PER PARTICIPANT WITH CNN MODEL.

basic approaches proposed in previous studies to estimate the attentional state. Moreover, the signals recorded will help to have a better understanding of the mechanism behind attention in the brain.

In the future works, it can be interesting to consider other recent ML-based approaches and to pursue the data collection with children diagnosed with ADHD.

For the future, we hope that virtual reality will continue to help for the elaboration of innovative assessments, as in our framework, and treatments for neurological disease.

ACKNOWLEDGMENT

The authors would like to thank the volunteers for their participation to the study and the research group of Cognitive Psychology and Neuropsychology of the Faculty of Psychology (University of Mons, Belgium) for their help in the definition of the attentional tasks. This collaboration has been made in the context of the CiPsE research centre³.

REFERENCES

- [1] Minkyu Ahn, Mijin Lee, Jinyoung Choi, and Sung Chan Jun. A Review of Brain-Computer Interface Games and an Opinion Survey from Researchers, Developers and Users. *Sensors (Basel, Switzerland)*, 14(8):14601–14633, August 2014.
- [2] American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. American Psychiatric Association, fifth edition edition, May 2013.
- [3] Azadeh Bashiri, Marjan Ghazisaeedi, and Leila Shahmoradi. The opportunities of virtual reality in the rehabilitation of children with attention deficit hyperactivity disorder: a literature review. *Korean Journal of Pediatrics*, 60(11):337–343, November 2017.
- [4] Pouya Bashivan, Irina Rish, Mohammed Yeasin, and Noel Codella. Learning Representations from EEG with Deep Recurrent-Convolutional Neural Networks. February 2016. _eprint: 1511.06448.
- [5] O. M. Bazanova and L. I. Aftanas. Individual EEG Alpha Activity Analysis for Enhancement Neurofeedback Efficiency: Two Case Studies. *Journal of Neurotherapy*, 14(3):244–253, August 2010.
- [6] Stéphanie Bioulac, Diane Purper-Ouakil, Tomas Ros, Hilario Blasco-Fontecilla, Marie Prats, Louis Mayaud, and Daniel Brandeis. Personalized at-home neurofeedback compared with long-acting methylphenidate in an european non-inferiority randomized trial in children with ADHD. *BMC Psychiatry*, 19(1):237, August 2019.
- [7] Friederike Blume, Justin Hudak, Thomas Dresler, Ann-Christine Ehlis, Jan Kühnhausen, Tobias J. Renner, and Caterina Gawrilow. NIRS-based neurofeedback training in a virtual reality classroom for children with attention-deficit/hyperactivity disorder: study protocol for a randomized controlled trial. *Trials*, 18(1):41, January 2017.
- [8] Zehong Cao, Chun-Hsiang Chuang, Jung-Kai King, and Chin-Teng Lin. Multi-channel EEG recordings during a sustained-attention driving task. *Scientific Data*, 6(1):19, April 2019. Number: 1 Publisher: Nature Publishing Group.
- [9] Thomas F. Collura and David Siever. Chapter 8 - Audio-visual entrainment in relation to mental health and EEG. In Thomas H. Budzynski, Helen Kogan Budzynski, James R. Evans, and Andrew Abarbanel, editors, *Introduction to Quantitative EEG and Neurofeedback (Second Edition)*, pages 195–224. Academic Press, San Diego, January 2009.
- [10] Marisol Cueli, Celestino Rodríguez, Paloma Cabaleiro, Trinidad García, and Paloma González-Castro. Differential Efficacy of Neurofeedback in Children with ADHD Presentations. *Journal of Clinical Medicine*, 8(2), February 2019.
- [11] Javier Dehesa, Andrew Vidler, Christof Lutteroth, and Julian Padget. Touché: Data-Driven Interactive Sword Fighting in Virtual Reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, pages 1–14, Honolulu, HI, USA, April 2020. Association for Computing Machinery.

³<https://web.umons.ac.be/sante/fr/centres-de-recherche-associies/cipse/>

- [12] Ya-Nan Feng, Zhao-Hui Wang, Jia-Rong Fan, Ting Fu, and Zhi-Yuan Chen. A Cross-Reference Line Method Based Multiobjective Evolutionary Algorithm to Enhance Population Diversity, July 2020. ISSN: 1687-5265 Library Catalog: www.hindawi.com Pages: e7179647 Publisher: Hindawi Volume: 2020.
- [13] Zhongke Gao, Xinmin Wang, Yuxuan Yang, Chaoxu Mu, Qing Cai, Weidong Dang, and Siyang Zuo. EEG-Based Spatio-Temporal Convolutional Neural Network for Driver Fatigue Evaluation. *IEEE Transactions on Neural Networks and Learning Systems*, 30(9):2755–2763, September 2019. Conference Name: IEEE Transactions on Neural Networks and Learning Systems.
- [14] Alba García-Baos, Tomas D’Amelio, Isadora Oliveira, Peter Collins, Carme Echevarria, Laura Pérez Zapata, Elizabeth Liddle, and Hans Supèr. Novel Interactive Eye-Tracking Game for Training Attention in Children With Attention-Deficit/Hyperactivity Disorder. *The primary care companion for CNS disorders*, 21(4), July 2019.
- [15] Mathieu Hatt, Baptiste Laurent, Anouar Ouahabi, Hadi Fayad, Shan Tan, Laquan Li, Wei Lu, Vincent Jaouen, Clovis Tauber, Jakub Czakon, Filip Drapejkowski, Witold Dyrka, Sorina Camarasu-Pop, Frédéric Cervenansky, Pascal Girard, Tristan Glatard, Michael Kain, Yao Yao, Christian Barillot, Assen Kirov, and Dimitris Visvikis. The first MICCAI challenge on PET tumor segmentation. *Medical Image Analysis*, 44:177–195, February 2018.
- [16] Eric Kolve, Roozbeh Mottaghi, Winson Han, Eli VanderBilt, Luca Weihs, Alvaro Herrasti, Daniel Gordon, Yuke Zhu, Abhinav Gupta, and Ali Farhadi. AI2-THOR: An Interactive 3D Environment for Visual AI. *arXiv*, 2017.
- [17] Yifeng Li, Wenqiang Shi, and Wyeth W. Wasserman. Genome-wide prediction of cis-regulatory regions using supervised deep learning methods. *BMC Bioinformatics*, 19(1):202, May 2018.
- [18] F Lotte, L Bougrain, A Cichocki, M Clerc, M Congedo, A Rakotomamonjy, and F Yger. A review of classification algorithms for EEG-based brain-computer interfaces: a 10 year update. *Journal of Neural Engineering*, 15(3):031005, April 2018. Publisher: IOP Publishing.
- [19] Fabien Lotte. A Tutorial on EEG Signal-processing Techniques for Mental-state Recognition in Brain-Computer Interfaces. In Eduardo Reck Miranda and Julien Castet, editors, *Guide to Brain-Computer Music Interfacing*, pages 133–161. Springer London, London, 2014.
- [20] Robert Oostenveld and Peter Praamstra. The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology*, 112(4):713–719, April 2001.
- [21] Rupert Ortner, Brendan Z. Allison, Gerd Korisek, Herbert Gaggl, and Gert Pfurtscheller. An SSVEP BCI to Control a Hand Orthosis for Persons With Tetraplegia. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 19(1):1–5, February 2011. Conference Name: IEEE Transactions on Neural Systems and Rehabilitation Engineering.
- [22] Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. Automatic differentiation in PyTorch. 2017.
- [23] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.
- [24] Nhatthai Phan, Dejing Dou, Brigitte Piniewski, and David Kil. A deep learning approach for human behavior prediction with explanations in health social networks: social restricted boltzmann machine (srbm+). *Social Network Analysis and Mining*, 6(1):79, Sep 2016.
- [25] Yehuda Pollak, Patricia L. Weiss, Albert A. Rizzo, Merav Weizer, Liron Shriki, Ruth S. Shalev, and Varda Gross-Tsur. The utility of a continuous performance test embedded in virtual reality in measuring ADHD-related deficits. *Journal of developmental and behavioral pediatrics: JDBP*, 30(1):2–6, February 2009.
- [26] Sudha R. Raman, Kenneth K. C. Man, Shahram Bahmanyar, Anick Berard, Scott Bilder, Takoua Boukhris, Greta Bushnell, Stephen Crystal, Kari Furu, Yea-Huei KaoYang, Øystein Karlstad, Helle Kieler, Kiyoshi Kubota, Edward Chia-Cheng Lai, Jaana E. Martikainen, Géric Maura, Nicholas Moore, Dolores Montero, Hidefumi Nakamura, Anke Neumann, Virginia Pate, Anton Pottegård, Nicole L. Pratt, Elizabeth E. Roughead, Diego Macias Saint-Gerons, Til Stürmer, Chien-Chou Su, Helga Zoega, Miriam C. J. M. Sturkenbroom, Esther W. Chan, David Coghill, Patrick Ip, and Ian C. K. Wong. Trends in attention-deficit hyperactivity disorder medication use: a retrospective observational study using population-based databases. *The Lancet. Psychiatry*, 5(10):824–835, 2018.
- [27] Yannick Roy, Hubert Banville, Isabela Albuquerque, Alexandre Gramfort, Tiago H. Falk, and Jocelyn Faubert. Deep learning-based electroencephalography analysis: a systematic review. *Journal of Neural Engineering*, 16(5):051001, August 2019.
- [28] Karen Simonyan and Andrew Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv:1409.1556 [cs]*, April 2015. arXiv: 1409.1556.
- [29] Y. Tan, D. Zhu, H. Gao, T. Lin, H. Wu, S. Yeh, and T. Hsu. Virtual Classroom: An ADHD Assessment and Diagnosis System Based on Virtual Reality. In *2019 IEEE International Conference on Industrial Cyber Physical Systems (ICPS)*, pages 203–208, May 2019.
- [30] Paloma Varela Casal, Flavia Lorena Esposito, Imanol Morata Martínez, Alba Capdevila, Maria Solé Puig, Núria de la Osa, Lourdes Ezpeleta, Alexandre Perera i Lluna, Stephen V. Faraone, Josep Antoni Ramos-Quiroga, Hans Supèr, and Jose Cañete. Clinical Validation of Eye Vergence as an Objective Marker for Diagnosis of ADHD in Children. *Journal of Attention Disorders*, 23(6):599–614, April 2019.
- [31] Tingxi Wen and Zhongnan Zhang. Deep Convolution Neural Network and Autoencoders-Based Unsupervised Feature Learning of EEG Signals. *IEEE Access*, 6:25399–25410, 2018. Conference Name: IEEE Access.
- [32] Erik G. Willcutt. The prevalence of DSM-IV attention-deficit/hyperactivity disorder: a meta-analytic review. *Neurotherapeutics: The Journal of the American Society for Experimental Neurotherapeutics*, 9(3):490–499, July 2012.
- [33] Irene Winkler, Stefan Haufe, and Michael Tangermann. Automatic Classification of Artifactual ICA-Components for Artifact Removal in EEG Signals. *Behavioral and Brain Functions*, 7(1):30, August 2011.
- [34] Wei-Long Zheng and Bao-Liang Lu. A multimodal approach to estimating vigilance using EEG and forehead EOG. *Journal of Neural Engineering*, 14(2):026017, February 2017. Publisher: IOP Publishing.
- [35] Peixiang Zhong, Di Wang, and Chunyan Miao. EEG-Based Emotion Recognition Using Regularized Graph Neural Networks. *IEEE Transactions on Affective Computing*, pages 1–1, 2020. Conference Name: IEEE Transactions on Affective Computing.