2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

A Study of Visual Search based Calibration Protocol for EEG Attention Detection

Aung Aung Phyo Wai*, Jee Ern Tchen* and Cuntai Guan#

Abstract— **Attention, a multi-faceted cognitive process, is essential in our daily lives. We can measure visual attention using an EEG Brain-Computer Interface for detecting different levels of attention in gaming, performance training, and clinical applications. In attention calibration, we use Flanker task to capture EEG data for attentive class. For EEG data belonging to inattentive class calibration, we instruct subject not focusing on a specific position on screen. We then classify attention levels using binary classifier trained with these surrogate ground-truth classes. However, subjects may not be in desirable attention conditions when performing repetitive boring activities over a long experiment duration. We propose attention calibration protocols in this paper that use simultaneous visual search with an audio directional change paradigm and static white noise as 'attentive' and 'inattentive' conditions, respectively. To compare the performance of proposed calibrations against baselines, we collected data from sixteen healthy subjects. For a fair comparison of classification performance; we used six basic EEG band-power features with a standard binary classifier. With the new calibration protocol, we achieved $74.37 \pm 6.56\%$ mean subject accuracy, which is about $3.73 \pm 2.49\%$ higher than the baseline, but there were no statistically significant differences. According to post-experiment survey results, new calibrations are more effective in inducing desired perceived attention levels. We will improve calibration protocols with reliable attention classifier modeling to enable better attention recognition based on these promising results.**

I. INTRODUCTION

Attention, multi-faced cognition components, play an crucial role in our daily life for every task performed to achieve a desirable task performance outcome. Attention directly influences task performance, and some applications demand higher attention and longer attention span [1]. Although attention can be triggered by different senses such as auditory, visual, etc, the neural processes behind visual and auditory attention are similar [2]. Besides subjective attention qualification through reaction time in response to the stimulus [3], the use of physiological measurements like EEG enable objective quantification of attention. Different cognitive and psychology tasks activate different attention types such as selective, sustained, divided, etc [4].

Conventionally, attention can be assessed in terms of behavioral responses (reaction time) while performing relevant cognitive tasks. But brain sensing based attention recognition is favored due to its objectivity and quantitative evaluation [4]. With advancements in sensing, signal processing and data analytic fields, EEG-based Brain Computer Interface can be found in applications ranging from performance training, clinical diagnosis to rehabilitation [5]. Attention, motivation and memory processing can be indexed using EEG band-power ratios without model training through attention calibration protocols [5]. However, the accuracy performance of model-free attention indexing is not as good as model-based attention detection [6]. Aside from accuracy performance, subject-specific attention classification is more effective in attention assessment due to non-stationary and time-varying EEG characteristics besides data variability among subjects [4]. A comparison of different standard cognitive tasks with similar 'inattentive task' for attention models training revealed no clear winners among three tasks used in previous study [7]. In this paper, we want to focus on proposing both attentive and inattentive calibration protocols to improve attention modeling in terms of accuracy performance as well as better user perception and usability by leveraging the visual search and white noise paradigm [8].

II. ATTENTION CALIBRATION PROTOCOL

The link between visual search and attention are complexly intertwined and attention is probably guided by five factors [9]. Visual attention, in general, activates selective attention [10], whereas multi-tasking modulates either divided or alternating attention owing to pure visual/audio or combined stimuli [11]. Stimulus complexity does not improve attention demanded in a visual search task, but rather results in faster visual overload [12]. However, conjunction display of simple shape and color differences can achieve parallel processing of stimuli, aiding in the reduction of visual overload according to [13]. As a result, our new attention calibration protocol induces only desirable attentive states while reducing unnecessary visual load and visual processing.

As shown in Fig. $2(c.1)$, the subject looks at the center of screen and only needs to press a left arrow key while spotting a target among non-target stimuli randomly displayed throughout the trials. To increase the difficulty, we add an auditory direction task to detect changes in direction of 'beep' while continuously 'beeping' on one side of the speaker as secondary task in visual search task. To respond to auditory change stimulus, the user presses the right arrow key for each direction change. The simultaneous visual and auditory stimuli induce divided and selective attention in a single task, forcing subjects to be in a state of high attention. We use TV static noise visual stimuli, with and without static noise sounds, for two designs similar to 'attentive'

*Authors with equal contributions. All authors are with the School of Computer Science and Engineering, Nanyang Technological Univeristy, 50 Nanyang Ave, Singapore 639798.

#Corresponding Author. ctguan@ntu.edu.sg

calibration. White noise is used as an 'inattention' task because it reduces voluntary eye movements when compared to looking around the screen border while still allowing the subject to focus on one spatial point, according to [14]. We introduce a 'resting task' before each 'attention' task pair to perform visual and audio guided deep breathing exercises to counter mental fatigue from performing repetitive and high workload mental tasks. The deep breathing relaxation does not only reset the user's fatigue but also improves task performance due to better exploitation of alpha band activities [15].

III. EXPERIMENT DESIGN AND DATA COLLECTION

In addition to the intrinsic EEG signal issues, the sub-optimal EEG-based attention classification performance is complicated by the inability to induce attentive or inattentive states in the subjects during calibration. To avoid mental and physical fatigue from repetitive and monotonous calibration tasks, we incorporate a 'resting task' in which the user performs a deep breathing exercise guided by visual and audio instructions, as shown in Fig. 1. The 180-second rest period before each attention calibration task also helps in performing attention task better. Each calibration task is made up of five activity blocks where each block include 30 seconds each for 'attentive' (ATT) and 'inattentive' (INA) task sequence. Each calibration task T_i takes approximately 5 minutes to complete. To understand how users perceived their mental states in the experiment, we asked short online survey questions that reflected perceived users' mental states and subjective experience of the experiment, as shown in Fig. 1.

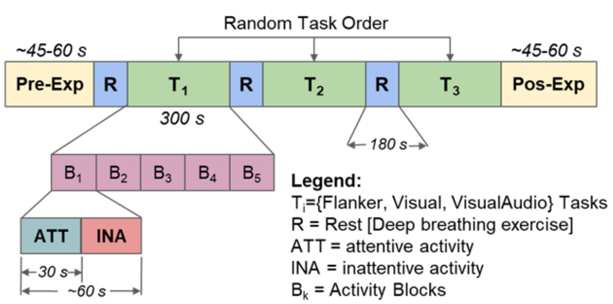

Fig. 1. Experiment design with multiple tasks involved in data collection.

We simultaneously collected data from two EEG amplifiers: a clinical-grade Nuamp (Neuroscan Inc.) with four EEG and four EOG channels and a consumer-grade Muse-2 (Interaxon Inc.) with four EEG channels. We down sampled raw EEG signals to 250 Hz that is the most commonly used sampling rate for data analysis. We compute basic spectral band-power features from the down sampled EEG signals to evaluate the attention classification performance using a binary Support Vector Machine (SVM) classifier. Both bipolar and re-referenced EEG data are derived from physical unipolar EEG channels during pre-processing.

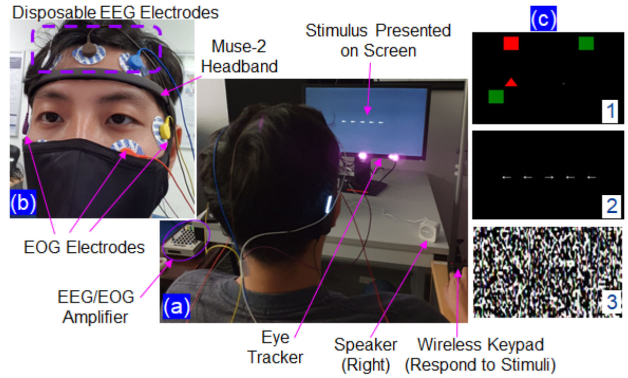

Fig. 2. Description of experiment screen (a) Subject performed tasks in the experiment (b) Close-up view of EEG and EOG electrodes with Muse-2 setup (c) Screenshots of experiment tasks (1) proposed 'attentive task' (2) arrow Flanker 'baseline' attentive task (3) proposed 'inattentive task'

IV. DATA ANALYSIS AND RESULTS

Firstly, we remove trials with excessive eye blinks, by utilizing 2D eye gaze points across the stimulus presentation screen synchronized with high amplitude EOG signals. We apply fourth order Butterworth band-pass filters with cut-off frequencies of 0.5 and 45 Hz to raw EEG signals. In this paper, we only use EEG data to compare the classification performance between baseline and proposed protocols. We use Wilcoxon rank sum test in all our statistical hypothesis tests as normality testing with One-sample Kolmogorov-Smirnov test shows some data do not conform to normal distribution. All accuracy results adopt six band-power features $(6BP)$ from six channel inputs except accuracy comparison between different number of band-power features.

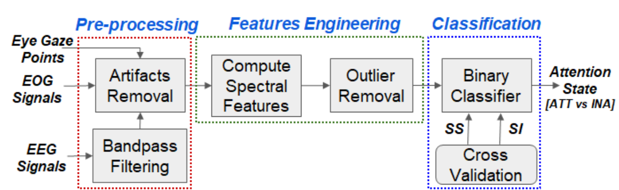

Fig. 3. Data analysis steps from EEG signals to 'attention' classification.

We begin by detecting trials with excessive ocular artifacts using eye gazes from an eye tracker in order to eliminate bad trials from analysis. We then use EOG signal as artifact references to reconstruct ocular artifact-free EEG signals by removing all possible ocular artifacts [16]. We re-referenced all EEG channels to linked average mastoids, resulting in re-referenced unipolar and bipolar channels, in addition to frontal EEG channels recorded directly from amplifiers (physical channels). As a result, we have six channel inputs for feature extraction: unipolars (Fp_1, Fp_2), bipolar $(Fp_1 + Fp_2)$, re-referenced unipolars (L_mFp_1, L_mFp_2), and re-referenced bipolar $(L_mFp_1 + L_mFp_2)$. The figure 4 below shows that mean accuracy for re-reference electrodes (6 Ch) is higher in all tasks compared with only physical electrodes (2 Ch). However, the 'baseline' task comparison reveals a statistically significant difference among channel options ($p < 0.05$). Although the proposed protocol design's accuracy is higher than the baseline, there is no statistically significant difference results between them.

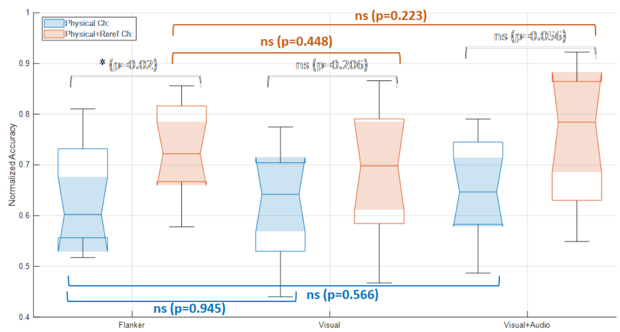

Fig. 4. Mean Accuracy comparison between two channel options: Physical channels only Vs Physical + Re-referenced channels

We use Pwelch power spectral density analysis on pre-processed EEG to extract band-power features in six bands: δ, θ, α, β_{Low}, β_{High} and γ [7]. Due to the lack of true ground-truth labels, we use one-class SVM with RBF Kernel to remove 5% of feature outliers from both classes [17]. Finally, we employ an SVM binary classifier with RBF kernels using the total number of features according to the number of band-powers $(3BP, 4BP, 5BP, 6BP)$ and EEG channels $(2, 6)$.

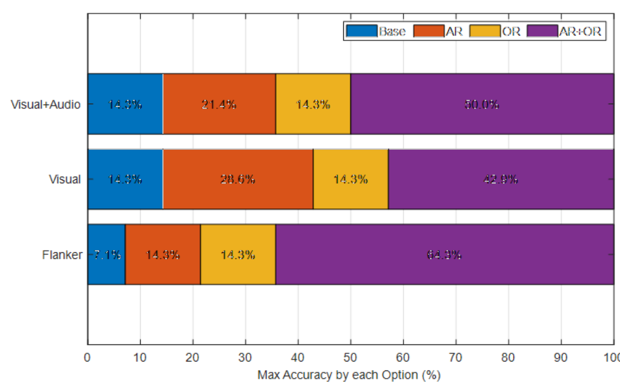

Fig. 5. Contributions of maximum accuracy per subject by four analysis options for different calibration protocols [Basic: Baseline, AR: Artifact Removal, OR: Outlier Removal, AR+OR: both AR and OR steps]

As shown in Fig. 5, we added OR and AR steps to basic analysis pipeline for accuracy comparison with different feature options for all calibration tasks. The 'Basic' option is standard EEG analysis steps similar to [7] that only achieve mean subject accuracy of $52.4 \pm 10.9\%$. The 'AR' option removes and reconstructs artifact-free EEG signals by utilizing synchronized eye gazes and EOG based ocular artifacts signatures [16]. As seen in Fig. 6, the reason for subpar mean accuracy of 74.37% is primarily due to three poor subjects whose accuracy is close to chance level, 50%.

A. DISCUSSION

We highlight the importance of an effective calibration protocol that stimulates user attention states, resulting in more representative surrogate measures, based on empirical evaluation results. Another step forward will be to conduct a feature-level assessment of poor accuracy subjects in relation to their reaction time and subject profile. Tab. I points out that the proposed calibration design improves accuracy of all analysis options compared to the baseline 'Flanker task'

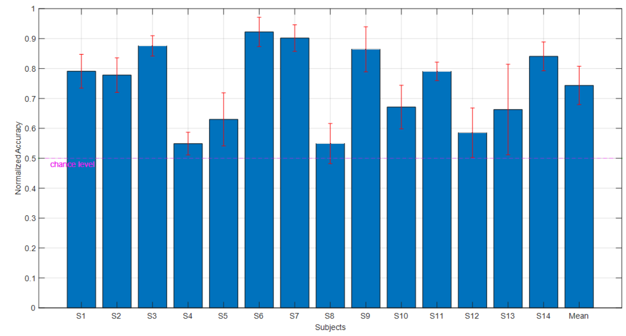

Fig. 6. Mean accuracy of each subject for proposed calibration protocol (Visual+Audio) using 6 BP features with AR+OR analysis option and Physical+Re-referenced channels

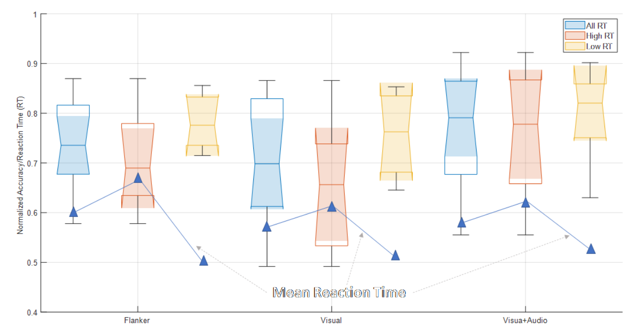

Fig. 7. Differences in maximum accuracy of subjects among different reaction-time (RT) based subject groups proposed calibration protocol (Visual+Audio) using 6 BP features with Physical+Rereferenced channels

A possible explanation for the lack of a significant difference in accuracy between baseline and the proposed solution is that there were fewer training data due to the small number of subjects and the short calibration time per task. Another question is whether accuracy alone is a good indicator or not for comparison among different calibration protocols.

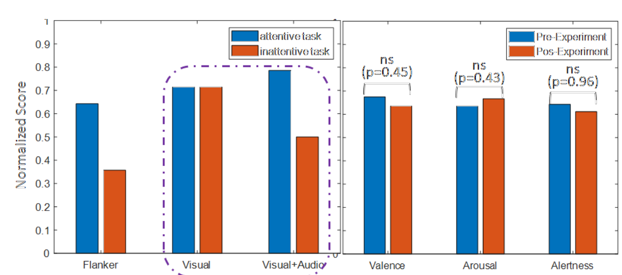

Fig. 8. Subject's feed-backs on effectiveness of calibration protocols and their perceptive emotional states before and after the experiment

Surprisingly, with the exception of one subject, most of the subjects outperform basic analysis steps such as AR, OR, and AR+OR in terms of accuracy. This means the EEG spectral features we used are generalizable and free from ocular artifacts. This strengthens the comparison of accuracy differences between three band-power options, such as 4, 5, and 6 bands per channel.

TABLE I
ATTENTION CLASSIFICATION IMPROVEMENT OF PROPOSED SOLUTION OVER BASELINE AMONG DIFFERENT NUMBER OF INPUT FEATURES.

Analysis Opt:	$3BP$	$4BP$	$5BP$	$6BP$
Base	$+5.3$	$+3.91$	$+4.30$	$+5.94$
AR	$+5.4$	$+2.11$	$+1.24$	$+1.46$
OR	$+4.86$	$+5.6$	$+4.63$	$+5.81$
AR+OR	$+4.52$	$+1.49$	$+2.51$	$+1.69$

We conclude that the attention classification with all six band-power features resulted the highest accuracy for the majority of the subjects. However, we still do not know how much each band contributed to the overall classification performance. We will look into features-level understanding of how each band-power influences on classification accuracy. It is worthwhile to investigate dimensional reduction approach for better features discrimination. In addition, we will compare the generalizability of attention classification using L1SO cross validation compared to L1BO subject-dependent CV. The reasons are due to EEG characteristics of time-varying and subject-dependent as well as how calibration task reliability induces desirable mental states.

Recently, deep learning methods show better performance for different EEG signals, we would like to apply deep learning methods for evaluating classification performance as well as automated learned EEG features directly from either raw or pre-processed data [18]. In address with subject-related modeling issues, transfer learning approach may be the next step in improving attention modeling performance. We still need to compare and evaluate the similarity and difference in performance between consumer EEG headbands with clinical EEG unit simultaneously collected data in the experiment. Currently, we only consider the 'attentive-inattentive' task-pair and it may be interesting to investigate what performance gains or losses can occur when data from cross task pairs among different calibration paradigms are used. Finally, we could use a multi-task learning to discover distinct profiles of attentive and inattentive states with task-agnostic features representations.

V. CONCLUSIONS

Attention is important in our daily life in order to achieve optimal task fulfillment. In contrast to modeling with real ground truths, EEG-based attention classification only has surrogate ground truths. As a result, we attempt to reduce unreliable modeling issues by developing new calibration protocols that can stimulate desirable states of attention. We use a visual search-like paradigm and random white noise to induce attentive and inattentive states respectively. We devised a study that included subjective questionnaires, a baseline, relaxation, and two variations of the proposed calibration protocols. Our analysis of sixteen subjects' data shows that our proposed solutions outperform the baseline task pairs. We can further improve current work in areas such as design, feature engineering, modeling, and so on to achieve more robust and reliable attention detection.

ACKNOWLEDGMENT

The study was reviewed and approved by the Institutional Review Board at Nanyang Technological University under reference number, IRB-2018-12-011. We would like to acknowledge all subjects for their participation in the experiment.

REFERENCES

[1] J. R. McDaniel, S. M. Gordon, and V. J. Lawhern, "Modeling the relationship between cognitive state and task performance in passive bcis using cross-dataset learning," in *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2020, pp. 4020–4025.

[2] B. G. Shinn-Cunningham, "Object-based auditory and visual attention," *Trends in Cognitive Sciences*, vol. 12, no. 5, pp. 182–186, May 2008.

[3] E. Salo, V. Salmela, J. Salmi, J. Numminen, and K. Alho, "Brain activity associated with selective attention, divided attention and distraction," *Brain Research*, vol. 1664, pp. 25–36, Jun. 2017.

[4] P. Aricò, G. Borghini, G. D. Flumeri, N. Sciaraffa, and F. Babiloni, "Passive BCI beyond the lab: current trends and future directions," *Physiological Measurement*, vol. 39, no. 8, p. 08TR02, aug 2018.

[5] A. Y. Shestyuk, K. Kasinathan, V. Karapoondinott, R. T. Knight, and R. Gurumoorthy, "Individual eeg measures of attention, memory, and motivation predict population level tv viewership and twitter engagement," *PLOS ONE*, vol. 14, no. 3, pp. 1–27, 03 2019.

[6] V. Sandhu, A. A. P. Wai, and C. Y. Ho, "Evaluation of learning performance by quantifying user's engagement investigation through low-cost multi-modal sensors," in *2017 International Conference on Orange Technologies (ICOT)*, Dec 2017, pp. 180–183.

[7] A. A. P. Wai, M. Dou, and C. Guan, "Generalizability of EEG-based mental attention modeling with multiple cognitive tasks," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, Jul. 2020, pp. 2959–2962.

[8] J. Palmer and E. Davis, "Visual search and attention: An overview," *Spatial Vision*, vol. 17, no. 4, pp. 249–255, 2004.

[9] J. M. Wolfe and T. S. Horowitz, "Five factors that guide attention in visual search," *Nature Human Behaviour*, vol. 1, no. 3, Mar. 2017.

[10] H. J. Müller and J. Krummenacher, "Visual search and selective attention," *Visual Cognition*, vol. 14, no. 4-8, pp. 389–410, Aug. 2006.

[11] A.-M. Bonnel and E. R. Haftser, "Divided attention between simultaneous auditory and visual signals," *Perception & Psychophysics*, vol. 60, no. 2, pp. 179–190, Jan. 1998.

[12] E. T. Davis, T. Shikano, S. A. Peterson, and R. K. Michel, "Divided attention and visual search for simple versus complex features," *Vision Research*, vol. 43, no. 21, pp. 2213–2232, Sep. 2003.

[13] S. W. Han, "The involvement of central attention in visual search is determined by task demands," *Attention, Perception, & Psychophysics*, vol. 79, no. 3, pp. 726–737, Jan. 2017.

[14] G. B. Söderlund, S. Sikström, J. M. Loftesnes, and E. J. Sonuga-Barke, "The effects of background white noise on memory performance in inattentive school children," *Behavioral and Brain Functions*, vol. 6, no. 1, p. 55, 2010.

[15] J. R. Stieger, S. Engel, H. Jiang, C. C. Cline, M. J. Kreitzer, and B. He, "Mindfulness improves brain–computer interface performance by increasing control over neural activity in the alpha band," *Cerebral Cortex*, vol. 31, no. 1, pp. 426–438, Sep. 2020.

[16] S. Blum *et al.*, "A riemannian modification of artifact subspace reconstruction for EEG artifact handling," *Frontiers in Human Neuroscience*, vol. 13, Apr. 2019.

[17] L. M. Manevitz and M. Yousef, "One-class svms for document classification," *J. Mach. Learn. Res.*, vol. 2, p. 139–154, Mar. 2002.

[18] A. Craik, Y. He, and J. L. Contreras-Vidal, "Deep learning for electroencephalogram (EEG) classification tasks: a review," *Journal of Neural Engineering*, vol. 16, no. 3, p. 031001, apr 2019.