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

Fine-tuning and Personalization of EEG-based Neglect Detection in Stroke Patients

Deniz Kocanaogullari[1], Xiaofei Huang[2], Jennifer Mak[3], Minmei Shih[4],
Elizabeth Skidmore[4], George F. Wittenberg[5], Sarah Ostadabbas[2], Murat Akcakaya[1]

Abstract— **Spatial neglect (SN) is a neurological disorder that causes inattention to visual stimuli in the contralesional visual field, stemming from unilateral brain injury such as stroke. The current gold standard method of SN assessment, the conventional Behavioral Inattention Test (BIT-C), is highly variable and inconsistent in its results. In our previous work, we built an augmented reality (AR)-based BCI to overcome the limitations of the BIT-C and classified between neglected and non-neglected targets with high accuracy. Our previous approach included personalization of the neglect detection classifier but the process required rigorous retraining from scratch and time-consuming feature selection for each participant. Future steps of our work will require rapid personalization of the neglect classifier; therefore, in this paper, we investigate fine-tuning of a neural network model to hasten the personalization process.**

Clinical relevance— **The proposed approach will utilize EEG data from multiple individuals, and enable rapid adaptation of the neglect classifier to each specific participant's EEG that could be collected over multiple days. Further research will investigate important EEG channels and it will provide a robust modality for online EEG-guided neglect detection and rehabilitation.**

I. Introduction and Related Work

Visual spatial neglect (SN) is a stroke-related condition afflicting 20-43% of stroke patients [1]. Prevalence of left-sided neglect due to a right hemisphere stroke is more than twice than that of right-sided neglect due to a left hemisphere stroke [2]. Left-sided neglect is usually more severe as well due to to the allocation of attentional processes to the right hemisphere [3]. The condition manifests as inattention to visual stimuli appearing contralesionally, such as missing food on one side of the plate, brushing only one side of the head, or bumping into objects on the contralesional side. There exists no common therapy for addressing SN and rehabilitation of attentional deficits is complicated by the difficulty in identifying and assessing SN.

One of the most common form of SN assessments are pencil-and-paper test. The conventional Behavioral Inattention Tests (BIT-C) is classically comprised of 6 subtests: line crossing, line bisection, letter cancellation, star cancellation, figure and shape copying, and representational drawing [4]. These tests require the assessor to score each subtest and sum the scores. If more than one subtest scores below a cutoff or if the total sum scores below the overall cutoff, the patient has SN. There are several issues with this kind of assessment. First, the result is a binary decision that provides scaling for the severity of neglect. Second, although this test requires that the papers are centered with the patient's body, the patient could change their field of view by simply moving their head, a common compensatory strategy of patients with SN. Lastly, less than half of patients with mild or moderate cases of SN may spontaneously recover partially in the acute phase (2 weeks post-stroke) [5]. It is unclear how sensitive the BIT-C overall is to these rapid changes in the patient's condition. Previous studies have found that just the line bisection subtest performance correlates with the severity of SN seen in clinically-observed behaviors [6], [7]. More reliable methods of measuring and assessing SN are important for providing timely intervention, and therefore, improving rehabilitation outcomes.

In order to overcome the shortcomings of the existing neglect detection methods, we designed an Augmented reality based and electroencephalography (EEG) guided neglect detection and rehabilitation system called AREEN [8]. AREEN uses EEG as the imaging modality of neural signals during a Starry Night Test [9]. The Starry Night Test in AREEN is adapted for augmented reality (AR) [8] (See Figure 1 for the AREEN system) . We already recruited stroke patients and collected EEG data, and performed Phase I analyses of the AREEN system. Phase I analyses investigated power in different frequency bands for SN detection across participant groups and temporal features for detection of potentially neglected targets. We showed that the AREEN system was both generalizable and personalizable in an across-participant setting [8]. However, personalization step included a rigorous retraining and exhaustive search for feature selection for each participant. This personalization step is computationally complex and time-consuming. Phase II of our work will include an EEG-guided neglect rehabilitation over multiple days, and each day will include a calibration session to refine the EEG-based neglected target classifier for each

[1]Deniz Kocanaogullari and Murat Akcakaya are with the Department of Electrical and Computer Engineering, University of Pittsburgh, 3700 O'Hara St, Pittsburgh, PA 15213, USA. dek107@pitt.edu, akcakaya@pitt.edu

[2]Xiaofei Huang and Sarah Ostadabbas are with the Department of Electrical and Computer Engineering, Northeastern University, 360 Huntington Ave, Boston, Massachusetts 02115, USA. xhuang@ece.neu.edu, ostadabbas@ece.neu.edu

[3]Jennifer Mak is with the Department of Bioengineering, University of Pittsburgh, 3700 O'Hara St, Pittsburgh, PA 15213, USA. jem356@pitt.edu

[4]Minmei Shih and Elizabeth Skidmore are with the Department of Occupational Therapy, University of Pittsburgh, 100 Technology Drive, Suite 350, Pittsburgh, PA 15219, USA. mis235@pitt.edu, jmk286@pitt.edu, skidmore@pitt.edu

[5]George Wittenberg is with the Department of Neurology, University of Pittsburgh, 3471 Fifth Ave, Pittsburgh, PA 15213, USA. GeoWitt@Pitt.edu

This work is supported by NSF Grant #1915083 and #1915065

participant. Therefore, computationally complex and time-consuming classifier personalization methods cannot be used in Phase II. Methods that rely on fine-tuning rather than retrain entire classifiers from scratch will be more suitable for the AREEN system in a rehabilitation setting.

In this paper, we investigate fine-tuning for a popular convolutional neural network based model called EEGNet [10] for neglected versus non-neglected target classification. This work will build on our previous study that also utilized EEGNet for neglect detection among stroke patients[11]. This investigation will provide a foundation for upcoming phases in our study and the short fine-tuning times will make the AREEN system viable for domain adaptation for online EEG-guided rehabilitation, which will require rapid classifier design, and highly accurate neglected target detection.

II. Datasets

A. Data Collection

In this paper, we use the same data utilized in Phase I analyses of our study to design a fine-tuning and personalization method for the detection of neglected targets [8]. Specifically, we utilize a combination of two EEG datasets in this paper to provide preliminary results towards the above mentioned methods for fine-tuning the classifier design: One dataset is collected using our AREEN system and the other one is collected using the same paradigm as the AREEN system but the presentation was on a computer screen rather than through an AR headset. We denote this system as computer-based BCI (CBBCI). Both datasets are collected through research procedures that were approved by local Institutional Review Board (IRB) under the University of Pittsburgh IRB numbers PRO15020115 and STUDY19060390.

Data is collected using a Starry Night Test through the AREEN system: canvas is divided into 6×12 grids. The entirety of the AR canvas is of size 0.564m $\cdot$ 0.288m and the depth is 1.14m. A red star, or a target, randomly appears every 1.2s to 2.5s. A random number of green stars (between 30 and 35), or distractors, (Figure 1) appears for 0.05s to 0.25s. Targets are shown 216 times total, 3 times in each cell. These stay on the canvas for a maximum of 3s during the clicker-based assessment and FOV (field of view) test, and 0.066s in EEG-based assessment. EEG data, with AREEN system, are collected through 16 electrodes located at Fp1, Fp2, F3, F4, Fz, Fc1, Fc2, Cz, P1, P2, C1, C2, Cp3, Cp4, O1 and O2 according to 10-20 system with sampling frequency of 256Hz. Each experimental session consists of a (1) Signal Check to inspect signal quality, (2) FOV Test to ensure proper mounting and positioning of the HoloLens, (3) Clicker-Based Assessment to gather the groundtruth for each position, and (4) EEG-Based Assessment for EEG data acquisition.

The CBBCI system we developed for neglect detection also utilizes the Starry Night Test paradigm [12]. For time-response data, participants clicked a button on a keyboard for the ground-truth and EEG data was collected using the current method. There are 192 trials instead of 216 presented in our novel AREEN system. It covers the same visual field

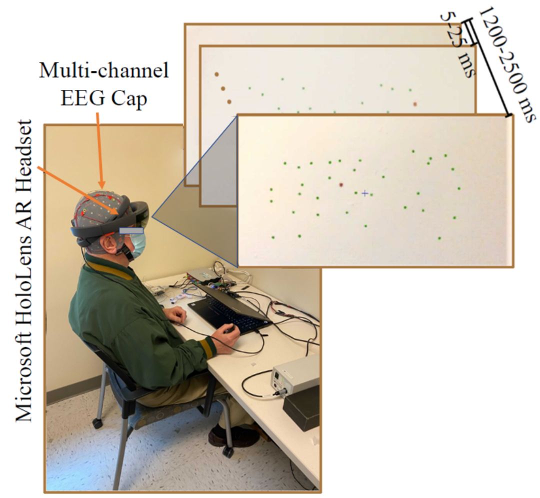

Fig. 1. Overview of the AREEN detection, assessment and rehabilitation system. A Starry Night scheme (green and red dots will appear and disappear at random times in random locations of the visual field in order to assess the region and extent of visual neglect) is presented to the patient while the patient's EEG signal is recorded. When visual neglect is detected using EEG-driven features, the multimodal feedback is triggered for rehabilitation treatment.

TABLE I
PARTICIPANT CHARACTERISTICS

ID	Age	Sex	Stroke Hemisphere	Days Since Stroke	BIT Total	BIT subtests below cutoff (/6)
SN01	76	M	Right	115	44	6
SN02	51	M	Right	9	25	5
SN04	72	F	Right	-	130	2
SN05	57	F	Left	7	134	2
SN101	81	F	Right	701	107	3
WSN01	68	F	Right	17	139	1
WSN101	35	M	Left	2404	138	0
WSN102	57	F	Left	2466	145	0
WSN103	80	M	Right	823	142	0
WSN104	27	M	Left	483	146	0

as our AREEN system but the last two channels differ. Thus, in our analyses, we use 14 channels of EEG data which are the same with AREEN except for O1 and O2 channels. In both systems, stimuli corresponding to each target are marked with trigger values that correspond to 700ms long EEG. As the sampling frequency in both systems is 256Hz, this corresponds to 180 datapoints for each target.

B. Participants

Above mentioned two datasets include EEG data from a total of 10 participants: 5 stroke patients with SN and 5 stroke patients without SN. The initial evaluation of the existence of SN is done with the BIT-C. A diagnosis of neglect was established by either a total BIT score lower than the established cutoff (<129), or a score lower than the cutoff score on more than one subtest. Participants are named with respect to their diagnosis from the BIT-C. If a participant has SN, they are denoted as *SN*. Otherwise, they are denoted as *WSN*. More specifically, we recruited 1 SN participant 4 WSN participants to use our AREEN system in Phase I for neglect detection and assessment. The remaining 5 participants (4 SN and 1 WSN) used the CBBCI system during EEG data collection. Patients with recent seizures were excluded from this study. The details of the participants are provided in Table I. Note that the numeral codes denote whether a participant used CBBCI or AREEN: those which start with one (also shown in bold) used the AREEN system and those starting with 0 used CBBCI.

III. Methodology

Time-domain EEG data is preprocessed and fed through the neural network model for classifying slow- versus fast-response targets (representing potentially neglected and non-neglected targets, respectively). As mentioned in Section I, considering the variability of EEG signals across different individuals, we develop an EEG-based personalized neglect detector. In order to overcome the inability to train an independent individual detection model from scratch due to insufficient individual data, and to minimize individual model training time, we have trained the same model multiple times from scratch by leaving a participant's data out and getting metrics from that specific data before and after fine-tuning.

A. Preprocessing

After collecting ground-truth time-response data corresponding to each target, target locations corresponding to slow-response and fast-response targets are identified. Slow-response targets correspond to potentially neglected targets whereas fast-response targets correspond to potentially observed targets. In order to achieve this separation, Otsu's method [13] is used to create a time-threshold for every patient: if a target's corresponding response time is greater than the threshold, the target is considered as a slow-response target; if it is smaller than the threshold, then it is considered as a fast-response target. To suppress outlier activity, we first get the median of the three time-response data in each cell, as each cell represents part of a visual field of a participant, and get the threshold value from resulting group of median values. After calculating the threshold, each time-point is then compared with the threshold value for initial labeling. To further suppress outlier activity, we follow a majority-voting procedure for the targets in their respective cells. This procedure provides information about potentially perceived or potentially neglected locations on the visual field and it is used as ground-truth to label EEG data. EEG data are first put through an 8^{th} order Butterworth filter with corner frequencies of 2Hz and 60Hz, and then through a 4^{th} order notch filter with corner frequencies of 58Hz and 62Hz.

After filtering, EEG data that are 700ms long and time-locked to the presented targets are extracted from each participant's recorded EEG data, for a total of 216 EEG segments. As EEG is very person-specific and we are aiming at a classification across individuals, the 200ms before a target is presented is used for baseline correction. Baseline correction is done in the time domain such that the mean of amplitude values of the baseline segment is calculated and the actual segment is corrected.

TABLE II
Neural Network Architecture, based on EEGNet

Layer	# Layers	Size	Info	# Params
conv2d	8	$(30, 1)$		256
batchnorm	-	-		32
depthwise_conv2d	16	$(1, 16)$		224
batchnorm	-	-		64
activation, avgpool	-	$(4, 1)$	ELU	-
dropout	-	-	$p = 0.5$	-
separable_conv2d	16	$(1, 16)$	512	
batchnorm	-	-	64	
activation, avgpool	-	$(8, 1)$	ELU	-
dropout	-	-	$p = 0.5$	-
flatten	-	-		162
dense, activation	1	2	softmax	-
Total				1.314

B. Classification Algorithm

We have utilized a convolutional neural network-based classifier called EEGNet[10]. EEGNet has been demonstrated to classify multi-channel EEG data with high metrics. The model has back-to-back convolutional layers to learn spatial and temporal features. To meet our data's needs, we have modified the hyperparameters; thus, our model's structure is explained below.

The model begins with a $(30, 1)$ 2D convolutional layer, which can be seen as a temporal filter. It is followed by a depthwise convolutional layer [14] of size $(1, 16)$, which is a spatial filter. A batch normalization layer is then used along the resulting spatiotemporal features. To introduce non-linearity, we use an exponential linear unit (ELU) and to reduce computational complexity, we use an average pooling layer of size $(4, 1)$. For more generalizability, we also use a dropout layer with a rate of 0.25.

These layers are continued with a separable convolution layer [14] of size $(1, 16)$ to further reduce the number of parameters and to combine the extracted spatiotemporal features. Finally, the output is put through another average pooling layer of size $(8, 1)$ and flattened. The flattened input goes through a dense layer with 2 for classification using softmax.

A brief description of the model can be seen in Table II. As the input shape in our modality is $(14, 180)$, 180 timepoints from 14 channels, the number of parameters with respect to that is correlated with the input shape.

For each participant, the general model is trained from scratch on the rest of participants' data. During this training stage, we adopt an Adam optimizer with a learning rate of $1e - 4$ for 500 epochs. The next fine-tuning stage sets the same learning rate but for 100 epochs. The model is written in Tensorflow[15] on Python and all the models are trained with CPU.

IV. Results

The experimental results with our proposed methodology are given in Table III.

Results show that fine-tuning the model to a participant's data improves the accuracy of detecting slow-response targets in a very short amount of time with a small amount of data.

TABLE III
EXPERIMENTAL RESULTS

Participant	Accuracy before Fine-tuning (%)	Accuracy after Fine-tuning (%)	Fine-tuning time (sec)
SN01	64.33%	82.3%	92.3
SN02	18.7%	81.6%	128.6
SN04	43.2%	65.4%	94.5
SN05	41.4%	75.8%	96.8
SN101	52.8%	76.8%	59.8
WSN01	23.8%	86.7%	94.7
WSN101	39.3%	97.2%	52.5
WSN102	15.9%	90.2%	56.5
WSN103	9.6%	91.8%	54.3
WSN104	6.5%	98.6%	55.8

This can be seen as domain adaptation, as EEG is very person-specific and we are adapting the general EEG model to each participant's EEG. This is also viable in Phase II as in each session, data will be collected before the rehabilitation sessions. The rapid computation time of fine-tuning shows that fine-tuning before each rehabilitation session is feasible.

The low metrics before fine-tuning shows the shortcomings of our project for the time being as Phase I is still ongoing. Deep learning methods usually need large amounts of data and our current dataset is comprised of ten participants. Thus, fine-tuning is crucial for both domain adaptation and addressing the dataset shortcomings.

V. CONCLUSION

In this paper, we provide a modified version of a deep learning model called EEGNet and conduct a slow-response versus fast-response classification for spatial neglect assessment. For online rehabilitation, as a preliminary step, we also fine-tune the model for each participant in order to address person-specificness of EEG with small amount of data. The experimental results show that with fine-tuning, the model adapts to each specific participant with high detection performance metrics.

Future investigations will use the novel AREEN BCI setup to progress on the rehabilitation procedures. For a better classification, we will provide a four-class classification for slow-response and fast-response targets from both participants with and without SN. Accordingly, our novel BCI and the classifier could improve the assessment of SN and it may result in a more versatile assessment of spatial neglect in stroke patients compared to the BIT-C.

REFERENCES

[1] D. S. Ting, A. Pollock, G. N. Dutton, F. N. Doubal, D. S. Ting, M. Thompson, and B. Dhillon, "Visual neglect following stroke: Current concepts and future focus," *Survey of Ophthalmology*, vol. 56, no. 2, pp. 114–134, 2011.

[2] J. M. Ringman, J. L. Saver, R. F. Woolson, W. R. Clarke, and H. P. Adams, "Frequency, risk factors, anatomy, and course of unilateral neglect in an acute stroke cohort," *Neurology*, vol. 63, no. 3, pp. 468–474, 2004.

[3] K. Li and P. A. Malhotra, "Spatial neglect," *Practical Neurology*, vol. 15, no. 5, pp. 333–339, 2015.

[4] B. Wilson, J. Cockburn, and P. Halligan, "Development of a behavioral test of visuospatial neglect," *Archives of physical medicine and rehabilitation*, vol. 68, p. 98—102, February 1987.

[5] A. Farnè, L. Buxbaum, M. Ferraro, F. Frassinetti, J. Whyte, T. Veramonti, V. Angeli, H. Coslett, and E. Làdavas, "Patterns of spontaneous recovery of neglect and associated disorders in acute right brain-damaged patients," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 75, pp. 1401 – 1410, 2004.

[6] P. Azouvi, C. Samuel, A. Louis-Dreyfus, T. Bernati, P. Bartolomeo, J. Beis, S. Chokron, M. Leclercq, F. Marchal, Y. Martin, G. Montety, S. Olivier, D. Pérennou, P. Pradat-Diehl, C. Prairial, G. Rode, E. Sieroff, L. Wiart, and M. Rousseaux, "Sensitivity of clinical and behavioural tests of spatial neglect after right hemisphere stroke," *Journal of neurology, neurosurgery, and psychiatry*, vol. 73, pp. 160–6, 09 2002.

[7] R. Luukkainen-Markkula, I. Tarkka, K. Pitkanen, J. Sivenius, and H. Hamalainen, "Comparison of the behavioural inattention test and the catherine bergego scale in assessment of hemispatial neglect," *Neuropsychological rehabilitation*, vol. 21, pp. 103–16, 01 2011.

[8] **D. Kocanaogullari, J. Mak, X. Huang**, J. Kersey, M. Shih, E. Skidmore, G. F. Wittenberg, S. Ostadabbas, and M. Akcakaya, "Detection and assesment of stroke-induced visual neglect using augmented reality and electroencephalography," *Journal of Neural Engineering*, Under Review.

[9] L. Deouell, Y. Sacher, and N. Soroker, "Assessment of spatial attention after brain damage with a dynamic reaction time test," *Journal of the International Neuropsychological Society : JINS*, vol. 11, pp. 697–707, 11 2005.

[10] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "EEGNet: A Compact Convolutional Network for EEG-based Brain-Computer Interfaces," *CoRR*, vol. abs/1611.08024, 2016.

[11] D. Kocanaogullari, J. Mak, J. Kersey, A. Khalaf, S. Ostadabbas, G. Wittenberg, E. Skidmore, and M. Akcakaya, "Eeg-based neglect detection for stroke patients," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, pp. 264–267, 2020.

[12] A. Khalaf, J. Kersey, S. Eldeeb, G. Alankus, E. Grattan, L. Waterstram, E. Skidmore, and M. Akcakaya, "Eeg-based neglect assessment: a feasibility study," *Journal of neuroscience methods*, vol. 303, pp. 169–177, 2018.

[13] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, pp. 62–66, Jan 1979.

[14] F. Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions," *CoRR*, vol. abs/1610.02357, 2016.

[15] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015. Software available from tensorflow.org.