Classification of Tonic Pain Experience based on Phase Connectivity in the Alpha Frequency Band of the Electroencephalogram using Convolutional Neural Networks

Yiyuan Han[1], Elia Valentini[2] and Sebastian Halder[1]

***Abstract*— The complexity of brain activity involved in the generation of the experience of pain makes it hard to identify neural markers able to predict pain states. The within and between subjects variability of pain hinders the predictive potential of machine learning models trained across participants. This challenge can be tackled by implementing deep learning classifiers based on convolutional neural networks (CNNs). We targeted phase-based connectivity in the alpha band recorded with electroencephalography (EEG) during resting states and sensory conditions (eyes open [O] and closed [C] as resting states, and warm [W] and hot [H] water as sensory conditions). Connectivity features were extracted and re-organized as square matrices, because CNNs are effective in detecting the patterns from 2D data. To assess the classifier performance we implemented two complementary approaches: we 1) trained and tested the classifier with data from all participants, and 2) using a leave-one-out approach, that is excluding one participant at a time during training while using their data as a test set. The accuracy of binary classification between pain condition (H) and eyes open resting state (O) was 94.16% with the first approach, and 61.01% with the leave-one-out approach.**

***Clinical relevance*— Further validation of the CNN classifier may help caregivers track the rehabilitation of chronic pain patients and dynamically modify the therapy. Further refinement of the model may allow its application in critical care setting with unresponsive patients to identify pain-like states otherwise incommunicable to medical personnel.**

I. INTRODUCTION

Pain has become a global health concern. Around 40% of adults are affected by chronic pain in Europe [1]. There is an unmet need for technological advancement to help caregivers detect and monitor patients' pain in both clinical and non-clinical setting [2]. This advance will be beneficial to unresponsive patients such as those with disorders of consciousness, who cannot voluntarily report pain. To achieve the objective of detecting and monitoring pain, scientists are utilising neurophysiological signals, and particularly brain signals recorded through electroencephalography (EEG) [3]. Here, we aimed to develop an EEG-based model for pain prediction rooted in the recent progresses in the field.

The first step in building a machine learning model is to investigate the optimal feature to be used as input. The quest for a specific EEG signature of pain has been mostly unsuccessful over the years because most of the brain responses are similar across sensory modalities and affective states [4] However, recent research indicates the peak of alpha (8-12 Hz) frequency (PAF) as a predictive index of pain sensitivity [5], [6]. Similarly, functional brain connectivity, the representation of integration among brain regions, has also been proposed as a metric of pain intensity in the brain [7]. According to these findings, alpha oscillations and functional connectivity seem to qualify as ideal candidates to validate a brain index for pain detection and monitoring. By combining this evidence, our previous work demonstrated that alpha-phase connectivity behaves as a promising neural marker for pain prediction, thus the machine learning model developed in the current study utilised phase connectivity in the alpha frequency range as classification feature [8].

Recent studies reported high classification accuracy for pain states with convolutional neural network (CNN) models, even above 95% [9]. However, it is currently unknown whether the CNN algorithms previously reported, while effective within the individual participant, can generalise to new individuals. This gap in CNN technology reduces its potential applications in the clinical setting. To address such limitation, we sat out to develop a model able to both predict pain states across individuals and generalise to new individuals. Due to the technical specificity of CNN models whereby one can extract the spatial patterns from 2D data, we organized alpha-phase connectivity as a 2D matrix and inputted it to a CNN pain classifier.

Hence, our model aims to overcome current limitations in the field with a twofold approach for responsive and unresponsive patients, respectively: 1) train (and test) the model on the data obtained from all participants; 2) train the model with a leave-one-out approach, that is excluding one participant at a time from training while using their data as a test set. We submit that by means of the leave-one-out approach we can validate the generalisability of the model to unresponsive patients (who cannot provide labelled data for training).

II. METHODS

A. Participants and Experimental Paradigm

This study was approved by the ethics committee of the University of Essex. Forty-three healthy individuals participated in the experiment (22 females, mean age 25.36 years, range 20-56). Data from seven participants were removed,

[1]Yiyuan Han and Sebastian Halder are with School of Computer Science and Electronic Engineering, University of Essex, Colchester, United Kingdom y.han@essex.ac.uk, s.halder@essex.ac.uk

[2]Elia Valentini is with Department of Psychology and Centre for Brain Science, University of Essex, Colchester, United Kingdom evalent@essex.ac.uk

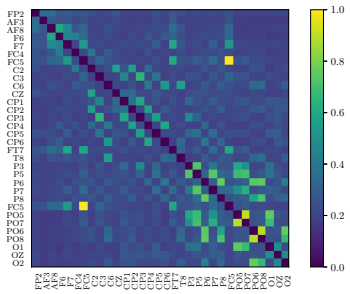


Fig. 1: The mean matrices of ISPCs across all time epochs recorded in the H condition.

because of procedural or technical issues, leaving a final sample of 36 participants.

The study entailed four experimental conditions for classification. Two thermal conditions, hot (H) and warm (W), eliciting pain and innocuous somatosensory sensations respectively, and two resting-state conditions, eyes-open (O) and eyes-closed (C). The thermal stimuli were induced by a prolonged immersion of the participants' left hand in a 30-litre tank (RW-3025P, Medline Scientific). Each condition lasted five minutes. The average temperature of H condition was 44.5 °C, and the temperature of W was 6 lower, the temperature was adjusted according to the individual's unpleasantness ratings delivered through visual-analogue scale (VAS). For more details about the methodology and procedure see [6].

We recorded the EEG with a 62 Ag/AgCl electrodes cap (Easycap, BrainProducts GmbH, Gliching, Germany). According to our previous work, we selected 32 representative electrodes to analyse connectivity (FP2, AF3, AF8, F6, F7, FC4, FC5, C2, C3, C6, Cz, CP1, CP2, CP3, CP4, CP5, CP6, FT7, T8, P3, P5, P6, P7, P8, FC5, PO5, PO6, PO7, PO8, O1, Oz, O2) [8].

TABLE I: Architecture of the CNN model: Three basic structures (layers 1-9) were applied, in which the activation function of each convolutional hidden layer was a rectified linear unit (ReLU) function. The 2D sizes in brackets involved in 'Size/Parameter' represent the kernel size of the corresponding layer, and the parameter multiplied with the kernel size in each hidden layer is the number of filters.

No.	Layer	Size/Parameter	Output
1	2D Convolution 1	$(7\times7)\times128$	$(32\times32)\times128$
2	2D Max-pooling 1	(3×3)	$(10\times10)\times128$
3	Batch Normalization 1	-	$(10\times10)\times128$
4	2D Convolution 2	$(5\times5)\times64$	$(10\times10)\times64$
5	2D Max-pooling 2	(3×3)	$(3\times3)\times64$
6	Batch Normalization 2	-	$(3\times3)\times64$
7	2D Convolution 3	$(3\times3)\times32$	$(3\times3)\times32$
8	2D Max-pooling 3	(3×3)	$(1\times1)\times32$
9	Batch Normalization 3	-	$(1\times1)\times32$
10	2D Dropout	0.2	$(1\times1)\times32$
11	Flatten 1	-	32
12	Fully Connected 1	100	100
13	Activation (ReLU)	-	100
14	Flatten 2	-	100
15	Activation (sigmoid)	-	100
16	Fully Connected 2	2	2
17	Softmax	-	2

B. Data Pre-processing

The EEG signal was down-sampled to 500 Hz from 1000 Hz, then independent component analysis (ICA) was applied and artefactual components were removed. Artefact-reduced data were then transformed into current source density (CSD) to reduce the effects associated with volume conduction [10], [11]. A Butterworth filter was applied to filter the signals into the alpha band from 8 to 12 Hz. After filtering the data, they were segmented into 5-second epochs with 50% overlap between neighbouring epochs.

C. Feature Extraction

We extracted alpha inter-site phase clustering (ISPC) as connectivity feature for classification [12]. The features were extracted from all pairs of the selected channels of each epoch with the formula:

$$ISPC_{C1,C2} = |\frac{1}{n}\sum_{t=1}^{n} e^{i(\phi_{C1}(t)-\phi_{C2}(t))}| \quad (1)$$

in which C1 and C2 represent two channels, ϕ_{C1} is the phase series produced by Hilbert transform from the signals recorded in channel C1 during the corresponding epoch. The range of each ISPC value is between 0 and 1.

Consequently, a 32×32 square matrix could be generated with these ISPCs, in which each element represents the functional connectivity between two channels in the alpha range within one time epoch, Fig. 1 shows the mean ISPC matrices across all epochs in the H condition as an example.

D. Classification

1) CNN model: Table I shows the architecture of the CNN classifier. The ISPC matrix was used as input to the CNN-based classifier. Due to its 2D nature and the fixed value range between 0 and 1, the ISPC matrix is a perfect input to a CNN model. Consequently, no further processing was necessary. At last, the softmax layer produced the prediction of each epoch.

We utilised Gradient-weighted Class Activation Mapping (Grad-CAM) to identify the activated regions within the connectivity graph and differentiate the conditions [13]. We fed the classifier with the mean ISPC matrix from each condition to detect the corresponding 2D pattern generated by the last hidden convolutional layer (layer 7 in Table I), then we resized the pattern to fit the original size of the input, thus the target connectivity can be indicated by the activation in the generated pattern. Finally, we computed the absolute values of arithmetic differences between the activated patterns from all pairs of conditions to expose the cogent connectivity in the corresponding binary classification. The difference matrices are represented in Fig. 2.

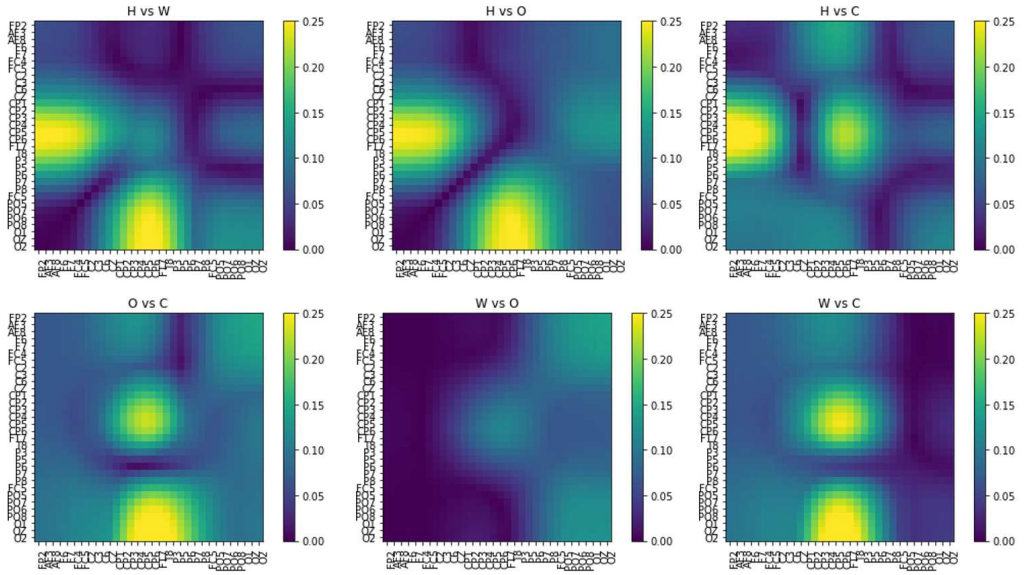
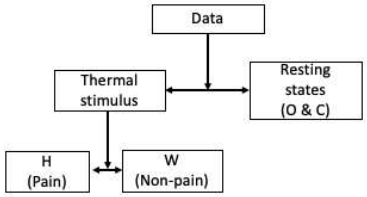
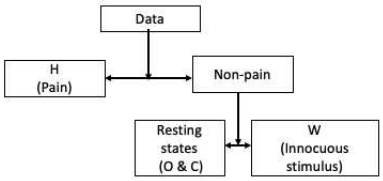


Fig. 2: The absolute values of arithmetic differences among the activated patterns out of layer 7 in Table I between each pair of conditions. These differences can show the cogent regions reflecting the connectivity able to significantly classify the corresponding conditions: Hot [H], Warm [W], Eyes-open [O], and Eyes-closed [C].



(a) Decision tree 1



(b) Decision tree 2

Fig. 3: Two types of decision trees in pain prediction: (a) classifying if thermal stimulus was induced before distinguishing the intensity of the thermal stimulus, (b) targeting at recognizing the pain, then classifying the non-pain conditions, respectively.

2) Training and testing: The primary aim in pain prediction is to distinguish pain from non-pain conditions. We investigated binary classifiers with the aim to ultimately combine them into binary decision trees. Once accurate prediction is obtained, one can attempt to classify the pain intensity. We propose two approaches to designing the decision trees: 1) distinguish resting states from thermal states (both pain and non-pain) (see Fig. 3a); 2) distinguish pain from non-pain states (both resting states and non-painful thermal stimuli) (see Fig. 3b). Based on this reasoning, we trained and tested the binary classification for the class combinations presented in Table II.

For each binary classification test, we drew training data from 75% of all of the participants' data and tested on remaining 25%. For the leave-one-out tests, all the data from the excluded participant composed the test set, while the data from the remaining participants were the training data.

For each training set, 7500 epochs per each class were randomly selected from the training data to balance the training set in the binary classification. If one binary class contained the mixture of several conditions, e.g. in pain (H) vs. non-pain (mixture of W, O and C), an equal number of epochs was selected for each class (i.e. 2500 epochs). For all training runs, we applied 10-fold cross-validation, and the accuracy of each classification output was used as the metric of model performance.

III. RESULTS AND DISCUSSION

A. Performance of CNN model

Table II shows the performance of the CNN model for each binary classification. When using data of all participants the classifier delivered a satisfactory accuracy (mean accuracy: 96.26%, accuracy between pain (H) and eyes open resting state (O): 94.16%), which was better than the performance delivered by the within-subject support-vector machine model proposed in our previous work (mean accuracy: 71.67%) [8]. Nevertheless, the classifier achieved a less satisfactory performance with the leave-one-out training mode (mean accuracy: 63.69%, accuracy between pain (H) and resting state (O): 61.01%), thus suggesting this approach is not currently reliable enough for detecting and monitoring pain in unresponsive patients.

These results suggest that when all participants are involved in training the classifier, the model is able to differentiate between pain and resting states, and between pain and innocuous sensation within each single individual. However, the model does not optimally predict pain states of a novel participant not involved in model training, possibly due to high individual variability. Such a limitation currently prevents us from applying the model to unresponsive patients.

Although Furman et al. proposed that PAF is sensitive enough to account for individual variability [5], current data do not seem to support the generalisability of the CNN classifier in the context of the leave-one-out training mode, thus implying this approach may not be currently extended to unresponsive patients. Future work will have to address this bottleneck and investigate the specificity of pain detection using other sensory conditions bearing comparable emotional effect with tonic experimental pain (e.g. unpleasant prolonged auditory experience).

TABLE II: Performance of binary classification with CNN model: (1) 'All': the accuracies of the models trained and tested with all participants. (2) 'LOO': In each test, one participant was excluded from training and used in testing, the mean accuracies produced by predicting the conditions of the excluded participants are shown here.

Binary classification		All	LOO
Pain (H) vs Non-pain (W+O+C)		94.37%	61.55%
Thermal stimulus (H+W) vs Resting states (O+C)		96.32%	63.28%
Pain (H) vs Resting states	H vs O	94.16%	61.01%
	H vs C	98.85%	77.20%
	H vs RS (O+C)	96.85%	63.87%
Warm (W) vs Resting states	W vs O	98.35%	57.64%
	W vs C	97.76%	72.54%
	W vs RS (O+C)	95.25%	63.94%
H vs W		94.43%	52.21%

B. Activated patterns from hidden convolutional layers

Fig. 2 displays the differences between activated patterns from all experimental conditions. The analysis aimed to determine the features used for recognizing the H condition (shown in the first column). The activated patterns in the H condition distribute over the central-parietal region of the scalp. This finding is compatible with the neurophysiology of somatosensation and somatic pain because the parietal region is assumed to primarily reflect the somatosensory and insular cortices encoding of thermal pain[14], [15]. Though the central-central, frontal-parietal and parietal-parietal connectivity can provide some evidence about the differences between W and O (See 'W vs O' in Fig. 2), they were significantly weaker than the pain-related connectivity, and the patterns to classify them separately with C condition (See 'O vs C' and 'W vs C' in Fig. 2) can also support this trend.

IV. CONCLUSIONS

The satisfactory performance of the proposed CNN classifier with training data selected from all participants suggests the proposed model can generalise across participants. However, the sensitivity to individual differences depended on having data of each participant in the training set, thus casting a shadow on the ability of the model to generalise to novel unknown individuals.

ACKNOWLEDGMENT

The authors acknowledge the use of the High Performance Computing Facility (Ceres) and its associated support services at the University of Essex in the completion of this work.

REFERENCES

[1] H. Breivik, E. Eisenberg, and T. O'Brien, "The individual and societal burden of chronic pain in europe: the case for strategic prioritisation and action to improve knowledge and availability of appropriate care," *BMC public health*, vol. 13, no. 1, pp. 1–14, 2013.

[2] R. Wang, Y. Deng, S. Zhou, and J. Zhang, "Eeg-derived pain threshold index for prediction of postoperative pain in patients undergoing laparoscopic urological surgery: a comparison with surgical pleth index," *Journal of Clinical Monitoring and Computing*, vol. 35, no. 6, pp. 1395–1402, 2021.

[3] D. Lenoir, W. Willaert, I. Coppieters, A. Malfliet, K. Ickmans, J. Nijs, K. Vonck, M. Meeus, and B. Cagnie, "Electroencephalography during nociceptive stimulation in chronic pain patients: A systematic review," *Pain Medicine*, vol. 21, no. 12, pp. 3413–3427, 2020.

[4] I. Timmers, A. L. Park, M. D. Fischer, C. A. Kronman, L. C. Heathcote, J. M. Hernandez, and L. E. Simons, "Is empathy for pain unique in its neural correlates? a meta-analysis of neuroimaging studies of empathy," *Frontiers in behavioral neuroscience*, vol. 12, p. 289, 2018.

[5] A. J. Furman, M. Prokhorenko, M. L. Keaser, J. Zhang, S. Chen, A. Mazaheri, and D. A. Seminowicz, "Sensorimotor peak alpha frequency is a reliable biomarker of prolonged pain sensitivity," *Cerebral Cortex*, vol. 30, no. 12, pp. 6069–6082, 2020.

[6] E. Valentini, S. Halder, D. McInnerney, J. Cooke, I. L. Gyimes, and V. Romei, "Assessing the specificity of the relationship between brain alpha oscillations and tonic pain," *NeuroImage*, p. 119143, 2022.

[7] P. Modares-Haghighi, R. Boostani, M. Nami, and S. Sanei, "Quantification of pain severity using eeg-based functional connectivity," *Biomedical Signal Processing and Control*, vol. 69, p. 102840, 2021.

[8] Y. Han, E. Valentini, and S. Halder, "Prediction of tonic pain using support vector machines with phase-based connectivity features," in *Abstract Book of the 8th International BCI Meeting*. BCI Society, 2021, p. 6.

[9] M. Yu, Y. Sun, B. Zhu, L. Zhu, Y. Lin, X. Tang, Y. Guo, G. Sun, and M. Dong, "Diverse frequency band-based convolutional neural networks for tonic cold pain assessment using eeg," *Neurocomputing*, vol. 378, pp. 270–282, 2020.

[10] T. J. Sejnowski, "Independent component analysis of electroencephalographic data," in *Advances in Neural Information Processing Systems 8: Proceedings of the 1995 Conference*, vol. 8. MIT press, 1996, p. 145.

[11] C. E. Tenke and J. Kayser, "Reference-free quantification of eeg spectra: combining current source density (csd) and frequency principal components analysis (fpca)," *Clinical Neurophysiology*, vol. 116, no. 12, pp. 2826–2846, 2005.

[12] M. X. Cohen, *Analyzing neural time series data: theory and practice*, 1st ed. MIT Press, 2014.

[13] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, Oct 2017.

[14] M. Bushnell, G. Duncan, R. Hofbauer, B. Ha, J.-I. Chen, and B. Carrier, "Pain perception: is there a role for primary somatosensory cortex?" *Proceedings of the National Academy of Sciences*, vol. 96, no. 14, pp. 7705–7709, 1999.

[15] B. Horing, C. Sprenger, and C. Büchel, "The parietal operculum preferentially encodes heat pain and not salience," *PLoS biology*, vol. 17, no. 8, p. e3000205, 2019.