Validation of Cross-Individual Pain Assessment with Individual Recognition Model from Electroencephalogram

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

Abstract— **Cross-individual pain assessment models based on electroencephalography (EEG) allow pain assessment in individuals who cannot report pain (e.g., unresponsive patients). The main obstacle to the generalisation of pain assessment models is the individual variation of brain responses to pain. Hence, we took the individual variation into account in cross-individual model development. We developed two convolutional neural networks (CNN) sharing an encoder architecture. One CNN predicts pain, while the other predicts the identity of an individual. We performed a leave-one-out (LOO) test with the exclusion of each subject and applied evidence accumulation to it for validating the pain prediction model's performance, where the binary classifier involves the states of pain (Hot) and resting state (Eyes-open). The mean accuracy produced by the LOO tests was** 57.81% **(max:** 73.33%**), and the mean accuracy of evidence accumulation achieved** 69.75% **(max:** 100.00%**). The individual recognition model achieved an accuracy of** 99.63%**. Nevertheless, when we acquired the most similar subject to a novel subject using the individual recognition model, where the most similar subject was used to train a subject-wise pain prediction model. The accuracy of predicting the pain-related conditions of the novel subject by the subject-wise model was only** 53.73% **(max:** 79.50%**). Therefore, the approach to utilising the features related to individual variation extracted by the CNN model needs more investigation for improving cross-individual pain assessment.**

Clinical relevance— **This model can be applied to assess pain from EEG signals at the bedside with future improvement, which can help caretakers of unresponsive patients.**

I. INTRODUCTION

Pain is defined as the sensory or emotional unpleasantness associated with actual or potential tissue damage [1]. In keeping with the protective function of pain, the clinicians or patients' caretakers can rely on the patient's pain reports when attending to the patient's health. However, unresponsive patients cannot report pain through communication. Technical advances in machine learning allow us to decode brain activity recorded through electroencephalography (EEG), thus enabling us to identify the experience of pain in unresponsive patients using EEG signals at the bedside [2]. A recent review established that machine learning models can achieve prediction accuracy between 62% and 100% [3]. Among pain assessment models, most were developed for subjects who also provided labelled data for training. However, unresponsive patients are unable to provide labels for training. This bottleneck means that we can only utilise the labelled data from responsive individuals to train the model, and generalise such models to predict pain in novel individuals not involved in the training phase. Importantly, the past research rarely reported effective generalisation of pain prediction models [3]. Yet, in some cases researchers reported optimal generalisation across individuals, e.g. 83% accuracy for three-class classification in [4] and 95.33% accuracy for binary classification in [5].

Cross-individual pain assessment is a cross-domain transfer learning application, in which each individual is a domain [6]. Although many cross-domain transfer learning approaches for various purposes were proposed, the lack of robust cross-individual models is not surprising in EEG-based research. Such a limitation is not specific to pain assessment - unfortunately, the individual variation negates EEG-based machine learning models' generalisation in many brain states [7]. Pain is even more challenging because the neural responses to pain can be very different across individuals, so it is difficult to expose those features that are unspecific to single individuals [8]. Schulz et al. suggested that the individual sensitivity to pain is correlated with the alpha ($8-10Hz$) and gamma ($\leq 80Hz$) bands, both of which have strong cross-individual variation [9]. Therefore, we can utilise such features related to individual variation to build the models, which is suggested key in the cross-domain transfer learning, i.e., cross-individual pain assessment model in this work [10].

In most cases, transfer learning requires deep learning architectures, but deep learning has not been widely implemented yet in the context of pain assessment [3]. Here we aimed to apply deep learning advancements, like the AutoTransfer framework [11], to develop a model sensitive to pain-related conditions as well as to individual variation.

We implemented a convolutional neural network (CNN) to predict pain using phase-based functional connectivity (FC) features in the alpha band, which fit the nature of FC represented by two-dimensional measures [12]. This model yielded ideal accuracy above 90% when applied to the subjects involved in model training, but did not achieve high accuracy when applied to novel subjects (61%). To develop a cross-individual pain assessment model, at least two methodological features must be considered: pain specificity and individual differences. Here, we aimed to test whether FC patterns within the alpha frequency band could be combined to quantify individual specificity [8], [13]. Accordingly, our study had two goals: improving CNN model generalisation through architecture or evidence accumulation, and exploring the use of alpha-phase FC in individual recognition. If

[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

successful, the individual recognition model would identify the most similar subject to a novel subject, and predict the pain status of the novel subject using labelled data from the most similar subject.

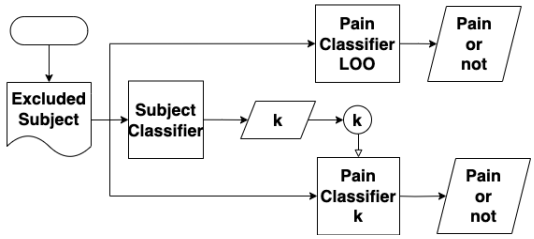

Fig. 1: **The pipeline of subject-transfer pain prediction model.** The testing set of both classifiers on the figure is all the data of the excluded subject. The individual recognition model predicted the 'most similar' subject to the input. Then the pain-related labels were predicted by the subject-wise model trained with the most similar subject (i.e., Pain Classifier k). In the meanwhile, the data from the excluded subject was input into the LOO pain classifier, which also predicted the pain-related labels

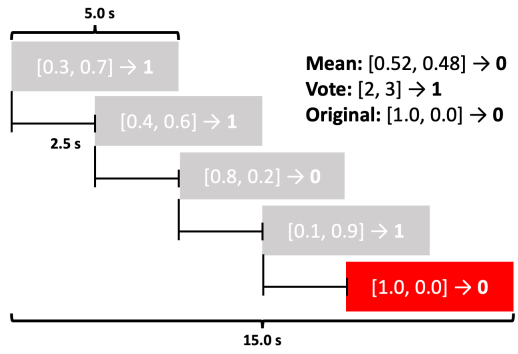

Fig. 2: **An example of the evidence accumulation involving 5 trials. (The values are only examples instead of the real predictions)** The red trial represents the target. The numbers 0 and 1 represent two classes, and [x, y] mean the prediction scores of classes 0 and 1. The prediction based on each paradigm was shown at top right. The original prediction means the prediction from the target trial only.

II. METHODS

A. Subjects and Experimental Paradigm

This study was approved by the ethics committee of the University of Essex. Forty-three healthy individuals participated in the experiment (21 males, mean age=25.36). We excluded seven subjects' data from the analysis due to procedural or technical issues, so the final sample contained 36 subjects.

We involved five conditions in the study, including two thermal stimulus conditions (hot and warm), two resting states (eyes-closed and eyes-open), and one reference sound condition. The experimental procedure can be found in [14]. Toward the aim to classify pain and non-pain conditions, since the subjects kept their eyes open in the thermal conditions, we only involved the thermal stimulus inducing pain (Hot [H]) and the controlled resting state (eyes-open [O]). The recording lasted 5 minutes for each condition in each subject. The detailed experimental progress can be found in [14].

The signals were recorded with a 62-channel EEG system (Easycap, BrainProducts GmbH, Gliching, Germany). We only used 32 channels for feature extraction, which highly contributed to pain assessment declared in [12].

B. Data Pre-processing and Feature Extraction

The sampling rate in EEG recording was 1000 Hz, and it was down-sampled to 500 Hz before further processing. Independent component analysis (ICA) and current source density (CSD) were applied to reduce artefacts and volume conduction, respectively (See details in [14] for ICA).

For feature extraction, the signals were filtered into the alpha band (8-12 Hz). Fraschini et al. suggested that the phase-based FC measures tend to converge to a fixed value with the trial-length above 4 seconds, so we segmented the filtered signals into 5-second trials with 50% overlap between the neighbouring trials [15].

We extracted the inter-site phase clustering (ISPC) as the representation of phase-based FC [16]. Each ISPC value represents the FC between two channels within the single trial, which follows the formula:

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

where C1 and C2 represent two EEG channels, ϕ_{C1} is the phase produced by Hilbert transform from the signals recorded in channel C1 in the trial. Then the ISPC values extracted from the same trial were re-organised into a 32×32 square matrix, whose rows and columns represent the EEG channels in the same order.

C. Classification Model and Validation

1) Convolutional Neural Network (CNN) Model: We built two CNN models, one binary classifier for pain classification, and one multi-classification model for recognising the individual. The two models shared an encoder architecture for feature extraction, containing hidden layers and batch normalisation (Table Ia). The design of dropout layers, activation functions, and fully connected layers differed for each purpose, their architectures are shown in Table Ib and Table Ic.

2) Training and Validation: We used Adam optimiser in the training, with a learning rate of 10^{-3}. The decay rates were set with $\beta_1 = 0.9$ and $\beta_2 = 0.99$, the L2 penalty was 0.01. The batch size was 256 and 100 epochs were trained in each classification. In the validation, the pain assessment models were the concatenation of the encoder and the architecture in Table Ib, and the individual recognition models contained the encoder and the layers in Table Ic. The mean accuracy among all the testing sets in each validation

TABLE I: Architecture of the two CNN models

(a) The encoder's architecture for feature extraction

No.	Layer	Size/Parameter	Output
1	2D Convolution 1	$(7\times7)\times128$	$(26\times26)\times128$
2	Batch Normalisation 1	128	$(26\times26)\times128$
3	2D Convolution 2	$(5\times5)\times64$	$(22\times22)\times64$
4	Batch Normalisation 2	64	$(22\times22)\times64$
5	2D Convolution 3	$(3\times3)\times32$	$(20\times20)\times32$
6	Batch Normalisation 3	32	$(20\times20)\times32$

(b) The architecture for the classifying pain and non-pain conditions

No.	Layer	Size/Parameter	Output
7-1	2D Dropout	0.2	$(20\times20)\times32$
8-1	Flatten 1	-	12800
9-1	Fully Connected 1	100	100
10-1	Activation (ReLU)	-	100
11-1	Flatten 2	-	100
12-1	Activation (sigmoid)	-	100
13-1	Fully Connected 2	2	2
14-1	Softmax	-	2

(c) The architecture for recognising individuals (Number of subjects: N)

No.	Layer	Size/Parameter	Output
8-2	Flatten 1	-	12800
9-2	Fully Connected 1	N	N
10-2	Softmax	-	N

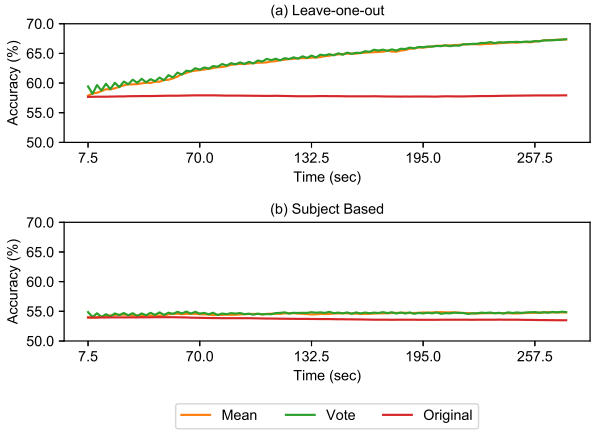

Fig. 3: **The curves of mean accuracy versus the time duration of classification between pain and non-pain with evidence accumulation.** The curves represent the mean accuracy across excluded subjects from the two paradigms based on evidence accumulation, and the original curves represent the accuracy produced by only the target trial in the corresponding test.

mode was analysed: a) in the leave-one-out (LOO) validation of pain assessment, each testing set represents the data from one excluded subject, b) and in the subject-mixture validation, each testing set contained the data from one fold of the cross-validation. The maximum individual accuracy of the test from an excluded subject was also analysed in the LOO tests.

- Individual Recognition
  - *All-mixed:* Data from all subjects were mixed. The data were split into training and testing sets, with a ratio of 35:1 in each fold of cross-validation.
  - *LOO training:* Thirty-six independent models were trained and validated with the exclusion of each subject. Within the data from 35 remained subjects for each model, 80% data from the remained subjects were used for training while the other 20% were the testing set.
- Pain Assessment
  - *Within-subject assessment:* We used the data from each individual separately to train 36 individual-wise classifiers. The data of each subject was split into the training and testing sets, with the ratio of 80% and 20%.
  - *Subject-mixed:* To make the size of data consistent with the further LOO test, 36-fold cross-validation was applied to the mixture of all the data for training and evaluation.
  - *LOO testing:* In each test, one subject was excluded, and the classification model was trained by the remaining 35 subjects'data. Then we tested each model with the excluded subject.
  - *Subject-based transfer assessment:* In each test, we input the data from each excluded subject to the LOO subject recognition model trained with the other subjects. The individual recognition model's most common prediction (i.e., subject k) was identified as the most similar subject to the excluded subject. Following this, we exclusively used the model that was trained on data from subject k, to predict the pain condition of the individual who was excluded. The pipeline is shown in Fig. 1.

3) Evidence Accumulation: The time duration required for assessing pain should be considered for bed-side applications[17]. Accordingly, we tested the accuracy associated with evidence accumulation over time. Because the trial length was 5 seconds and the overlapping ratio was 50 %, the involved time duration increased by 2.5 seconds with one more trial included in the assessment. We tested the effects of trial numbers from 2 to 109 (7.5 to 275.0 seconds). Two paradigms were used for the predictions. One of them is based on the mean prediction scores of each class across the involved trials. And the other is based on a voting system that will identify the most frequent class among the trial-based predictions. Fig. 2 shows an example involving five trials.

TABLE II: Accuracy produced by the individual recognition and pain assessment models under different validation modes

Model	Validation	Mean Accuracy (%)	Max Individual Accuracy (%)
Individual recognition	All-mixed	99.63 ± 0.23	
	Leave-one-out	99.55 ± 0.22	
Pain assessment	Within-subject	91.86 ± 1.30	100.00
	Subject-mixed	96.52 ± 1.02	
	Leave-one-out (original)	57.81 ± 7.48	73.33
	Leave-one-out (max., mean)	69.56 ± 14.72	100.00
	Leave-one-out (max., vote)	69.75 ± 14.59	99.43
	Subject-based transfer (original)	53.73 ± 10.59	79.50
	Subject-based transfer (max., mean)	58.41 ± 15.32	100.00
	Subject-based transfer (max., vote)	58.83 ± 14.92	100.00

III. RESULTS AND DISCUSSION

The mean and maximum accuracy produced by each validation mode is shown in Table II, where the mean

accuracy (column 'Mean Accuracy') produced with all the data is from the folds in cross-validation, and the LOO results are across excluded subjects. Since the performance was seriously diverse across subjects, the maximum accuracy of a single excluded subject (column 'Max Individual Accuracy' in Table II) from the LOO tests was also shown as the highest level achieved by an individual, which can show the potentially best accuracy achieved by the model.

The within-subject (91.86%) and subject-mixed (96.52%) validations suggested ideal performances of this model in the labelled data. This equated the benchmark across several frequency bands, for example, max accuracy of 97.37% in [18]. For generalisation, some work revealed that specific measures representing the integration across different brain regions or frequency bands are correlated to pain [2]. However, the cross-individual models typically used time-frequency features within specific channel or frequency band, but did not sufficiently utilise measures of integrations, such as FC [4], [5], [19]. Although we could not improve the accuracy of predicting pain with a single trial compared to our previous machine learning implementation [12], the use of evidence accumulation optimised the performance (69.75%). Ten of thirty-six subjects produced accuracy above 80% with the evidence accumulation, and the best subject reached 100%. Despite the average accuracy of cross-individual prediction was worse than the benchmark (89.45%; [5]), the large standard deviation suggested some of them can be predicted with ideal accuracy. The important difference between the mean and maximum accuracy also suggests the relevance of the effects associated with inter-individual variation.

The FC features extracted from the alpha band showed good performances in the recognition task, which is consistent with the correlation between individual pain sensitivity and the alpha band. The accuracy in classifying the 36 subjects reached 99.63%, and the mean accuracy for each participant subset was 99.55%. With such robustness, we can theoretically rely on the individuals with labelled data to predict their similar novel subject's pain. Therefore, we implemented the subject-based transfer paradigm as described in Section II-C.2 and Fig. 1. Nevertheless, the accuracy of the subject-based transfer test was worse than the models trained with 35 subjects (Fig. 3a). This may be expected due to the number of samples used in training each individual-wise model (i.e., $1/35$ of the data used for training the respective LOO model). One strategy to capitalise on the similar subjects to predict the pain in a novel individual may entail increasing the number of samples within a generative adversarial network (GAN) [20], or building an adversarial learning model that will contrast a pain assessment network and an individual recognition network [21].

REFERENCES

[1] S. N. Raja, D. B. Carr, M. Cohen, N. B. Finnerup, H. Flor, S. Gibson, F. Keefe, J. S. Mogil, M. Ringkamp, K. A. Sluka *et al.*, "The revised iasp definition of pain: concepts, challenges, and compromises," *Pain*, vol. 161, no. 9, p. 1976, 2020.

[2] M. Ploner, C. Sorg, and J. Gross, "Brain Rhythms of Pain," *Trends in Cognitive Sciences*, vol. 21, no. 2, pp. 100–110, 2017. [Online]. Available: http://dx.doi.org/10.1016/j.tics.2016.12.001

[3] T. Mari, J. Henderson, M. Maden, S. Nevitt, R. Duarte, and N. Fallon, "Systematic review of the effectiveness of machine learning algorithms for classifying pain intensity, phenotype or treatment outcomes using electroencephalogram data," *The Journal of Pain*, vol. 23, no. 3, pp. 349–369, 2022.

[4] T. Nezam, R. Boostani, V. Abootalebi, and K. Rastegar, "A novel classification strategy to distinguish five levels of pain using the eeg signal features," *IEEE Transactions on Affective Computing*, vol. 12, no. 1, pp. 131–140, 2021.

[5] V. Vijayakumar, M. Case, S. Shirinpour, and B. He, "Quantifying and characterizing tonic thermal pain across subjects from eeg data using random forest models," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 12, pp. 2988–2996, 2017.

[6] Q. Yang, Y. Zhang, W. Dai, and S. J. Pan, *Transfer learning*. Cambridge University Press, 2020.

[7] W.-L. Zheng and B.-L. Lu, "Personalizing eeg-based affective models with transfer learning," in *Proceedings of the twenty-fifth international joint conference on artificial intelligence*, 2016, pp. 2732–2738.

[8] A. J. Furman, T. J. Meeker, J. C. Rietschel, S. Yoo, J. Muthulingam, M. Prokhorenko, M. L. Keaser, R. N. Goodman, A. Mazaheri, and D. A. Seminowicz, "Cerebral peak alpha frequency predicts individual differences in pain sensitivity," *Neuroimage*, vol. 167, pp. 203–210, 2018.

[9] E. Schulz, A. Zherdin, L. Tiemann, C. Plant, and M. Ploner, "Decoding an individual's sensitivity to pain from the multivariate analysis of eeg data," *Cerebral cortex*, vol. 22, no. 5, pp. 1118–1123, 2012.

[10] K. Weiss, T. M. Khoshgoftaar, and D. Wang, "A survey of transfer learning," *Journal of Big data*, vol. 3, no. 1, pp. 1–40, 2016.

[11] N. Smedemark-Margulies, Y. Wang, T. Koike-Akino, and D. Erdogmus, "Autotransfer: Subject transfer learning with censored representations on biosignals data," in *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2022, pp. 3159–3165.

[12] Y. Han, E. Valentini, and S. Halder, "Classification of tonic pain experience based on phase connectivity in the alpha frequency band of the electroencephalogram using convolutional neural networks," in *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2022, pp. 3542–3545.

[13] J. E. Letzen, B. Remeniuk, M. T. Smith, M. R. Irwin, P. H. Finan, and D. A. Seminowicz, "Individual differences in pain sensitivity are associated with cognitive network functional connectivity following one night of experimental sleep disruption," *Human brain mapping*, vol. 41, no. 3, pp. 581–593, 2020.

[14] 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.

[15] M. Fraschini, M. Demuru, A. Crobe, F. Marrosu, C. J. Stam, and A. Hillebrand, "The effect of epoch length on estimated eeg functional connectivity and brain network organisation," *Journal of neural engineering*, vol. 13, no. 3, p. 036015, 2016.

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

[17] A. C. Chen, S. F. Dworkin, J. Haug, and J. Gehrig, "Human pain responsivity in a tonic pain model: psychological determinants," *Pain*, vol. 37, no. 2, pp. 143–160, 1989.

[18] 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. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231219313980

[19] Y. Bai, G. Huang, Y. Tu, A. Tan, Y. S. Hung, and Z. Zhang, "Normalization of pain-evoked neural responses using spontaneous eeg improves the performance of eeg-based cross-individual pain prediction," *Frontiers in computational neuroscience*, vol. 10, p. 31, 2016.

[20] K. G. Hartmann, R. T. Schirrmeister, and T. Ball, "Eeg-gan: Generative adversarial networks for electroencephalograhic (eeg) brain signals," *arXiv preprint arXiv:1806.01875*, 2018.

[21] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. Marchand, and V. Lempitsky, "Domain-adversarial training of neural networks," *The journal of machine learning research*, vol. 17, no. 1, pp. 2096–2030, 2016.