

Towards plug & play c-VEP-based BCIs: a pilot study on zero-calibration adaptive decoding using deep learning

Eduardo Santamaría-Vázquez^{1,2,3}, Víctor Martínez-Cagigal^{1,2,3}, Ana Martín-Fernández^{1,2}, and Roberto Hornero^{1,2,3}

¹Biomedical Engineering Group, University of Valladolid, Spain

²Centro de Investigación Biomedica en Red de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Valladolid, Spain

³Valladolid Health Research Institute (IBioVALL), Valladolid, Spain

E-mail: eduardo.santamaria.vazquez@uva.es

ABSTRACT: Brain-computer interfaces (BCI) based on code-modulated visual evoked potential (c-VEP) provide fast and robust control, but most high-performance systems still require user-specific calibration before use. This work evaluates a zero-calibration decoding strategy for a 16-command c-VEP speller encoded with binary *m*-sequences, aiming to reduce preparation time while preserving the possibility of online subject personalization. The proposed model is based on the original EEG-Inception implementation, with task-specific adaptations. We also developed a novel adaptive strategy that simulates online operation, in which high-confidence predictions are converted into pseudo-labels to enable conservative unsupervised adaptation of the batch-normalization layers and the classification head. We evaluated the performance of our approach on 15 subjects using a leave-one-subject-out protocol. Results showed that the zero-calibration approach without adaptation achieved a mean command accuracy of 84.79% after 10 stimulation cycles (5.25s). When two outlier subjects were excluded, this value increased to 95.91%. The adaptive strategy further improved performance, particularly during the early stimulation cycles, with a mean improvement of 5.5%. These findings support the feasibility of subject-independent c-VEP decoding using deep learning and suggest that conservative test-time adaptation can improve robustness in practical plug-and-play spelling scenarios.

INTRODUCTION

Brain-computer interfaces (BCIs) based on code-modulated visual evoked potentials (c-VEP) are among the most accurate and fastest non-invasive paradigms for communication and control [1]. In these systems, each command is encoded by a pseudo-random visual sequence, and the attended target is inferred from the corresponding electroencephalographic (EEG) response. In comparison with more classical evoked-potential paradigms, c-VEPs provide a favorable compromise between number of commands, decoding speed, and robustness, which has enabled spelling applications with very high information transfer rates and reliable operation in both healthy and motor-impaired users [2–5].

However, despite these advances, most high-performance c-VEP systems still depend on subject-specific calibration sessions to estimate user-adapted templates or train subject-dependent decoders [1, 6].

This calibration requirement remains one of the main barriers to the practical deployment of c-VEP-based BCIs. From a translational perspective, the need to record labeled data before each use increases setup time, reduces usability, and limits the development of plug-and-play systems. Moreover, calibration is especially problematic in applications where rapid access, repeated sessions, or user fatigue are relevant concerns. Consequently, reducing the amount of subject-specific data required by c-VEP decoders has become an important research objective in recent years [6–9]. Existing strategies include transfer learning across subjects, multi-session training, semi-supervised updates, and fully unsupervised calibration procedures [6, 10–12]. Although these studies have shown that calibration can be substantially reduced, achieving robust subject-independent performance in unseen users remains challenging because of the marked inter-subject variability of VEP responses.

In parallel with these advances, alternative decoding frameworks have expanded the methodological landscape of c-VEP BCIs. Besides classical template matching, event-wise response modeling approaches have shown that the stimulation pattern can be reconstructed directly from short EEG epochs and later recombined to recover the attended command [13]. This bit-wise reconstruction (BWR) approach decouples event classification from the commands codebook, facilitating transfer across stimulation schemes, and naturally lending itself to modern deep-learning architectures. For instance, our previous work showed that EEG-Inception can accurately reconstruct binary and non-binary stimulation events from EEG under subject-specific calibration, offering a competitive alternative to conventional c-VEP decoding pipelines [14]. In principle, this framework is also well suited to zero-calibration scenarios, as models can be trained across subjects and transferred directly to unseen users. However, studies exploring this possibility are still scarce, and the performance reported so far remains well below that obtained with calibration [15].

The goal of this study is twofold: (1) To evaluate the operational feasibility of a zero-calibration c-VEP speller driven by a cross-subject deep-learning mode; and (2) To design, develop, and validate a novel adaptive strategy capable of enhancing decoding performance in real-time by progressively integrating new online data through an unsupervised mechanism. To this end, we combine a leave-one-subject-out (LOSO) training strategy with conservative unsupervised test-time adaptation driven by high-confidence pseudo-labels, enabling the model to be applied directly to unseen users and updated online without supervised calibration. Beyond overall spelling performance, the study also examines inter-subject variability and the influence of outlier users, with the aim of better characterizing the practical limitations of zero-calibration c-VEP decoding.

MATERIALS AND METHODS

Dataset and experimental setting:

The experiments were conducted using the open-source non-binary c-VEP database presented by Martinez-Cagigal *et al.* [16]. This dataset includes several stimulation conditions based on p -ary maximal-length sequences and a 4×4 speller with 16 commands (letters A–P). In this work, only the binary condition was considered, corresponding to the black–white stimulation setting. Although the original database included 16 healthy participants (28.8 ± 5.0 years), the analyses reported here were conducted on 15 subjects, as one participant was excluded due to the use of a different sampling rate during the experiments. Each subject completed 30 calibration trials of 10 stimulation cycles each with an m-sequence of 63 bits and a monitor refresh rate of 120 Hz. The online spelling phase comprised 32 trials per subject, with 10 stimulation cycles per trial. Specifically, participants were tasked with selecting each of the 16 commands in alphabetical order across two full iterations [16]. EEG was recorded from 16 electrodes placed according to the international 10–10 system: F3, Fz, F4, C3, Cz, C4, CPz, P3, Pz, P4, PO7, POz, PO8, Oz, I1, and I2. The ground electrode was located at AFz and the reference at the earlobe. All experiments were performed with the open source "P-ary c-VEP Speller" application of MEDUSA [17]. More information about the experimental protocol can be found in [18].

Preprocessing and feature extraction:

The preprocessing parameters are described in our previous work [14]. EEG signals were filtered using a 7th-order Butterworth band-pass filter between 1 and 60 Hz and a 7th-order notch filter between 49 and 51 Hz. Signal epochs were extracted from 0 to 500 ms after each stimulation event and resampled to 200 Hz, yielding 100 temporal samples per observation. Feature extraction followed a bit-wise reconstruction approach, where each stimulation event was segmented into an EEG epoch and labeled according to its corresponding binary stimulation value. Consequently, each stimulation cycle yielded 63

epochs. In total, the calibration and test datasets contained 18,900 ($63 \times 10 \times 30$) and 20,160 ($63 \times 10 \times 32$) observations per subject, respectively. The pipeline used z-normalization within the analysis window and generated event-wise samples that were later recombined at decoding time to recover the attended command.

Deep learning model:

The classifier was a modified EEG-Inception architecture specifically adapted for bit-wise c-VEP decoding [19]. The backbone started with a temporal inception block composed of three parallel convolutional branches with kernel sizes of 50, 25, and 15 samples, respectively, each containing 12 filters to capture temporal patterns at different scales. This stage was followed by a depthwise spatial filtering block with a multiplier of 2, which learned channel combinations while collapsing the spatial dimension. A second temporal stage then applied a dilated inception block with four parallel branches, all using kernel size 5 but with dilation rates of 1, 5, 10, and 15, respectively, again with 12 filters per branch. The resulting feature maps were progressively compressed by the output block, and finally mapped to the binary classification head. Batch normalization was applied after every convolutional layer throughout the network to stabilize internal activations and accelerate convergence. The network used exponential linear unit (ELU) activations and standard dropout with rate 0.2. Overall, this design preserves the main rationale of EEG-Inception [19] (i.e., multi-scale temporal feature extraction, efficient spatial filtering, and progressive feature compression) while adapting it to the subject-independent, event-wise decoding problem considered here.

Zero-calibration approach:

The proposed framework follows a strictly subject-independent training strategy based on leave-one-subject-out (LOSO) cross-validation. For each fold, one participant was held out for testing and the model was trained exclusively with the preprocessed calibration data from the remaining subjects (i.e., 30 trials/subject). Therefore, no labeled EEG data from the test user were available during model training. Consequently, the evaluated system simulates a zero-calibration approach from the user perspective, since decoding starts without any supervised subject-specific adjustment. The training stage for each LOSO fold used a learning rate of 0.01, batch size of 256, validation split of 0.1, early stopping with patience of 5 epochs, and a maximum of 100 epochs.

During testing, the network estimated the binary label of each stimulation event independently [14]. These event-wise predictions were then recombined to reconstruct the stimulation sequence associated with each trial. Command decoding was finally performed by comparing the reconstructed sequence with the 16 binary m-sequences assigned to the speller commands and selecting the command with maximum Pearson correlation. This procedure produced a command estimate after each stimulation cycle, thus allowing performance to be analyzed as a function of stimulation duration.

Unsupervised online adaptation:

In addition to the zero-calibration baseline, we implemented an unsupervised adaptive strategy that simulates progressive online personalization. Starting from the cross-subject model, adaptation was applied trial by trial in causal order. After each prediction, the confidence margin, defined as the difference between the highest and second-highest command-correlation scores at the last available stimulation cycle, was computed. When this margin exceeded a threshold of 0.1, the trial was deemed reliable and used to update the model via pseudo-labeling: the predicted command was mapped back to its event-wise binary sequence, and the corresponding EEG epochs were added to an adaptation buffer.

Model updating was deliberately conservative. Only the batch-normalization layers and the classification head were fine-tuned, while the convolutional backbone remained frozen, thereby mitigating overfitting and catastrophic forgetting while allowing the model to recalibrate its feature normalization and decision boundaries to the new user. Adaptation used a learning rate of 0.001, batch size of 256, validation split of 0.2, early stopping with patience of 3 epochs, minimum improvement of 0.001, and a maximum of 10 epochs per update

RESULTS

We compared our zero-calibration approach with and without adaptation. Performance was assessed in terms of command accuracy as a function of stimulation cycle, which provides a direct measure of the speller's speed-accuracy trade-off.

Table 1 summarizes the results obtained with the subject-independent model without adaptation. Mean accuracy increased progressively from 52.71% at cycle 1 to 84.79% at cycle 10. However, performance was strongly affected by two outlier subjects (S10 and S11), whose accuracies remained close to chance level across the whole session. When these subjects were excluded, mean accuracy increased from 59.38% at cycle 1 to 95.91% at cycle 10, with a clear reduction in inter-subject variability.

Table 2 shows the results obtained with the unsupervised adaptive strategy. In this case, mean accuracy improved from 63.12% at cycle 1 to 87.50% at cycle 10 across all subjects. After excluding the two outliers, performance increased from 72.36% to 98.08%. Overall, the adaptive approach yielded consistently higher accuracies than the non-adaptive baseline across all stimulation cycles.

The magnitude of this improvement is displayed in Figure 1. Across all subjects, the adaptive strategy provided a mean gain of 10.42% at cycle 1, with the benefit gradually decreasing as more stimulation cycles were accumulated, reaching 2.71% at cycle 10. On average, the adaptive strategy improved accuracy by 5.50% across all cycles.

Analysis of outlier subjects:

Given the marked inter-subject variability observed in our experiments, a complementary analysis was conducted to

better characterize the two outlier subjects (S10 and S11). Interestingly, these subjects did not show atypical performance in our previous within-subject frameworks, where their accuracies fell within the normal range [14, 20]. In contrast, in the present cross-subject setting they exhibited substantially lower accuracies than the rest of the cohort under both the baseline and adaptive conditions, remaining close to chance level even after 10 stimulation cycles. Their inclusion therefore had a pronounced effect on the overall mean performance and variability estimates, warranting a more detailed inspection of their electrophysiological patterns.

As shown in Tables 1 and 2, the baseline model reached only 18.75% and 6.25% accuracy at cycle 10 for S10 and S11, respectively. With adaptation, S10 improved modestly to 31.25%, whereas S11 remained at 6.25%. This indicates that the adaptive strategy was not sufficient to recover normal performance in these cases, although it provided a limited benefit for one of the two subjects.

To further explore this finding, Fig. 2 compares the average evoked-response template at POZ and the spatial distribution of template root mean square (RMS) between outlier and non-outlier subjects. The non-outlier group show clearer evoked-response morphology and a more marked posterior RMS maximum over occipital electrodes, whereas the outlier group tends to exhibit lower amplitudes and less distinctive spatial pattern. Although these observations should be interpreted with caution, they are compatible with the possibility that the outlier subjects presented less typical c-VEP responses and electrophysiological profiles, which may have limited the effectiveness of cross-subject transfer.

DISCUSSION

This study evaluated a zero-calibration binary c-VEP decoding framework based on subject-independent deep learning and assessed whether a novel unsupervised online adaptation strategy could improve performance during use. Overall, the approach was feasible for most subjects, with the baseline model reaching a mean accuracy of 84.79% after 10 stimulation cycles and the adaptive strategy increasing this value to 87.50%. When the two outlier subjects were excluded, performance reached 95.91% and 98.08%, respectively. The main benefit of adaptation was observed during the early stimulation cycles, with mean gains of 10.42% and 8.12% at cycles 1 and 2, indicating that its primary effect was to accelerate convergence toward a correct decision rather than to substantially improve final performance. Taken together, these results suggest that our decoding strategy may provide a useful basis for calibration-free c-VEP spelling, while also highlighting the presence of inter-subject variability and the influence of a small number of difficult cases.

A useful reference for contextualizing these results is our previous within-subject study on the same database [14]. In that calibrated setting, the binary condition achieved

Table 1: Accuracy per cycle (%) — Zero-Calibration (no adaptation)

Cycle	1	2	3	4	5	6	7	8	9	10
Stim. dur. (s)	0.53	1.05	1.58	2.10	2.62	3.15	3.68	4.20	4.73	5.25
S0	46.88	62.50	71.88	78.12	81.25	81.25	81.25	84.38	84.38	87.50
S1	56.25	75.00	81.25	87.50	96.88	96.88	100.00	100.00	100.00	100.00
S2	87.50	90.62	93.75	96.88	100.00	100.00	100.00	100.00	100.00	100.00
S3	68.75	87.50	93.75	93.75	96.88	96.88	96.88	100.00	100.00	96.88
S4	37.50	43.75	56.25	53.12	59.38	71.88	71.88	81.25	84.38	84.38
S5	62.50	81.25	81.25	84.38	90.62	90.62	93.75	93.75	100.00	100.00
S6	71.88	84.38	87.50	90.62	90.62	90.62	90.62	93.75	93.75	90.62
S7	59.38	65.62	75.00	81.25	87.50	87.50	96.88	96.88	100.00	93.75
S8	71.88	90.62	93.75	96.88	96.88	96.88	96.88	96.88	96.88	96.88
S9	62.50	81.25	93.75	100.00	100.00	100.00	100.00	100.00	100.00	100.00
S10 [†]	12.50	12.50	12.50	18.75	18.75	15.62	12.50	15.62	18.75	18.75
S11 [†]	6.25	9.38	9.38	6.25	9.38	12.50	6.25	9.38	6.25	6.25
S12	71.88	90.62	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
S13	28.12	40.62	53.12	59.38	65.62	65.62	71.88	75.00	87.50	96.88
S14	46.88	56.25	78.12	84.38	87.50	93.75	96.88	96.88	100.00	100.00
Mean (all)	52.71	64.79	72.08	75.42	78.75	80.00	81.04	82.92	84.79	84.79
Std (all)	23.14	27.33	28.28	29.02	28.96	28.66	30.65	29.62	30.02	29.85
Mean (w/o outliers)	59.38	73.08	81.49	85.10	88.70	90.14	92.07	93.75	95.91	95.91
Std (w/o outliers)	16.19	17.70	14.63	14.67	13.04	11.04	10.33	8.27	6.30	5.32

[†]Outlier subjects (S10, S11).

Table 2: Accuracy per cycle (%) — Zero-Calibration with adaptive strategy

Cycle	1	2	3	4	5	6	7	8	9	10
Stim. dur. (s)	0.53	1.05	1.58	2.10	2.62	3.15	3.68	4.20	4.73	5.25
S0	65.62	87.50	87.50	90.62	96.88	93.75	96.88	96.88	96.88	96.88
S1	71.88	87.50	87.50	90.62	96.88	100.00	100.00	100.00	100.00	100.00
S2	96.88	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
S3	87.50	93.75	96.88	100.00	100.00	100.00	100.00	100.00	100.00	100.00
S4	40.62	46.88	68.75	81.25	78.12	81.25	87.50	93.75	96.88	96.88
S5	78.12	87.50	90.62	93.75	96.88	96.88	96.88	96.88	100.00	100.00
S6	87.50	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
S7	75.00	81.25	90.62	90.62	93.75	96.88	96.88	96.88	100.00	100.00
S8	93.75	96.88	96.88	96.88	96.88	96.88	96.88	96.88	96.88	96.88
S9	71.88	84.38	87.50	100.00	100.00	100.00	100.00	96.88	100.00	100.00
S10 [†]	3.12	6.25	18.75	15.62	12.50	12.50	25.00	28.12	25.00	31.25
S11 [†]	3.12	3.12	15.62	9.38	9.38	6.25	12.50	9.38	6.25	6.25
S12	84.38	96.88	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
S13	25.00	34.38	46.88	65.62	78.12	81.25	78.12	78.12	81.25	84.38
S14	62.50	87.50	96.88	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Mean (all)	63.12	72.92	78.96	82.29	83.96	84.38	86.04	86.25	86.88	87.50
Std (all)	30.94	33.36	28.73	29.85	30.52	31.12	28.07	28.17	29.53	28.59
Mean (w/o outliers)	72.36	83.41	88.46	93.03	95.19	95.91	96.39	96.63	97.84	98.08
Std (w/o outliers)	20.60	20.07	15.12	10.06	7.82	6.80	6.49	5.91	5.16	4.33

[†]Outlier subjects (S10, S11).

99.8% accuracy after 10 cycles, whereas the adaptive zero-calibration framework evaluated here reached 87.50%, increasing to 98.08% when the two outlier subjects were excluded. Thus, although the adaptive subject-independent approach did not reach the performance of the calibrated within-subject framework in the full cohort, it approached that reference level for most subjects while eliminating the calibration phase.

Two subjects showed markedly lower accuracies than the rest of the cohort despite not being atypical in our previous within-subject framework [14, 20], suggesting reduced cross-subject transfer in these cases. Although their descriptive characterization was inconclusive, some features pointed to a potentially different neurophysiological profile and less reliable c-VEP response. The limited effect of adaptation in these subjects is also consistent with the fact that this strategy depends on sufficiently reliable initial predictions. Overall, these results suggest

that conservative unsupervised adaptation may be useful as a refinement mechanism, but not always sufficient to compensate for poor initial transfer.

Relative to the state of the art, the present work contributes to the ongoing effort to reduce calibration requirements in c-VEP BCIs. Previous studies, particularly by Thielen *et al.* [12], have shown that calibration-free or zero-training c-VEP decoding is feasible using reconvolution and canonical correlation analysis (CCA) formulations. Other studies have also reported adaptive or training-reduction strategies in c-VEP systems, highlighting both the potential and the difficulty of maintaining robust performance when subject-specific calibration is limited [7, 10]. Although direct comparisons across studies are difficult due to differences in datasets and decoding approaches, the present work presents, to the best of our knowledge, the first unified framework combining subject-independent deep-learning-based c-VEP de-

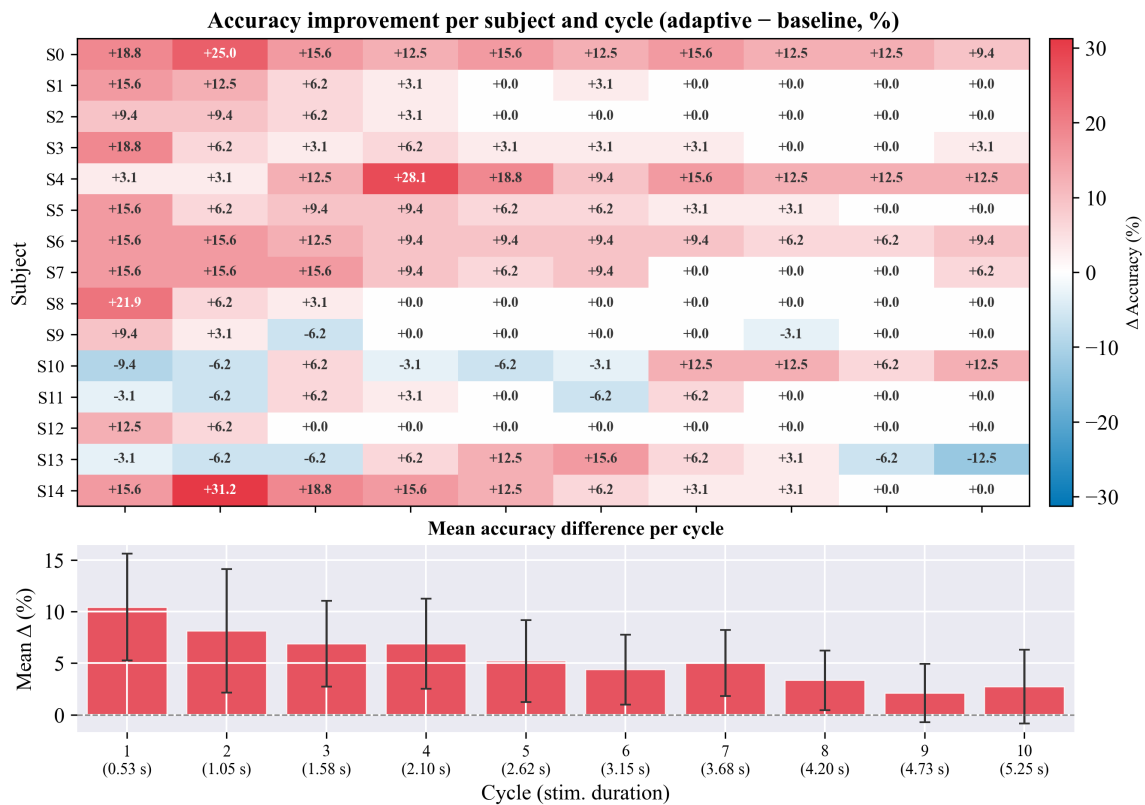


Figure 1: Accuracy improvement per subject and stimulation cycle for the adaptive strategy relative to the zero-calibration baseline. The upper panel shows the subject-wise difference in command accuracy (adaptive – baseline, percentage points) across stimulation cycles, whereas the lower panel shows the corresponding mean improvement across subjects with standard deviation.

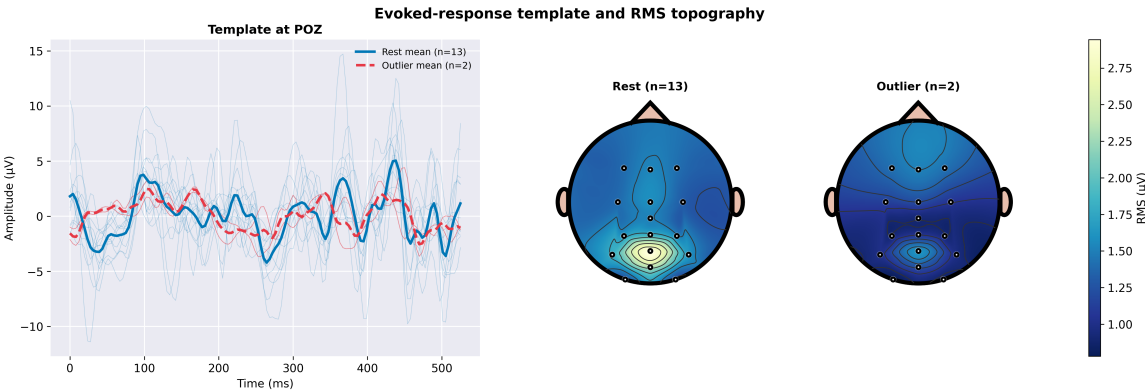


Figure 2: Evoked-response template and RMS topography averaged across regular and outlier subjects. The left panel shows the c-VEP template waveform at electrode POz for each subject (thin lines) together with the group mean (thick lines). The right panels display the mean root mean square (RMS) scalp topography for regular and outlier subjects, with a shared colour scale.

coding with an unsupervised online adaptation strategy.

Several limitations should be considered. The subject sample size was modest, which limits the interpretation of the outlier analyses, and the observed inter-subject variability may partly reflect the limited amount and diversity of data available for subject-independent training. Crucially, the current approach has yet to be validated within populations with severe motor disabilities, such as individuals with spinal cord injury or amyotrophic lateral sclerosis, whose neurophysiological profiles may differ

significantly from the current sample. In this regard, the proposed processing pipeline is expected to benefit from larger and more heterogeneous datasets, which should improve cross-subject robustness and reduce the impact of underrepresented response profiles. In addition, the adaptive strategy was deliberately conservative, which favors stability but may also limit personalization in difficult cases. Future work should therefore explore more flexible adaptation schemes.

CONCLUSION

This paper presents the first zero-calibration deep-learning framework for c-VEP spellers based on subject-independent training and unsupervised online adaptation. The proposed approach avoids supervised fine-tuning on the test subject and achieves performance levels that, for most users, approach those reported in conventional calibrated systems, thereby substantially reducing setup time while preserving high-speed communication capabilities. Nevertheless, a small subset of users remained challenging in the present study, highlighting the need for further work to improve robustness and personalization for underrepresented neurophysiological profiles.

ACKNOWLEDGEMENTS

This research was supported by Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), through Instituto de Salud Carlos III, co-funded by the European Regional Development Fund (ERDF); the Regional Government of Castilla y León (Junta de Castilla y León, Consejería de Educación) and the EU-FEDER under the project VA140P24; and the European Union through the Interreg VI-A Spain-Portugal Program (POCTEP) 2021-2027 project "0124_EUROAGE_MAS_4_E". A. Martín-Fernández was supported by a PIF grant from UVA.

REFERENCES

- [1] Martínez-Cagigal V, Thielen J, Santamaría-Vázquez E, Pérez-Velasco S, Desain P, Hornero R. Brain-computer interfaces based on code-modulated visual evoked potentials (c-VEP): a literature review. *Journal of Neural Engineering*. 2021;18(6):061002.
- [2] Gembler FW, Rezeika A, Benda M, Volosyak I. Five Shades of Grey: Exploring Quintary m-Sequences for More User-Friendly c-VEP-Based BCIs. *Computational Intelligence and Neuroscience*. 2020;2020.
- [3] Sun Q, Zheng L, Pei W, Gao X, Wang Y. A 120-target brain-computer interface based on code-modulated visual evoked potentials. *Journal of Neuroscience Methods*. 2022;375(109597).
- [4] Moreno-Calderón S, Martínez-Cagigal V, Santamaría-Vázquez E, Pérez-Velasco S, Marcos-Martínez D, Hornero R. Combining brain-computer interfaces and multiplayer video games: an application based on c-VEPs. *Frontiers in Human Neuroscience*. 2023;17.
- [5] Verbaarschot C *et al.* A visual brain-computer interface as communication aid for patients with amyotrophic lateral sclerosis. *Clinical Neurophysiology*. 2021;132(10):2404–2415.
- [6] Thielen J, Marsman P, Farquhar J, Desain P. From full calibration to zero training for a code-modulated visual evoked potentials brain computer interface. *Journal of Neural Engineering*. 2021;18(5):56007.
- [7] Stawicki P, Volosyak I. cVEP Training Data Validation—Towards Optimal Training Set Composition from Multi-Day Data. *Brain Sciences*. 2022;12(2).
- [8] Miao Y *et al.* High-performance c-VEP-BCI under minimal calibration. *Expert Systems with Applications*. 2024;249(PB):123679.
- [9] Zheng L, Dong Y, Tian S, Pei W, Gao X, Wang Y. A calibration-free c-VEP based BCI employing narrow-band random sequences. *Journal of Neural Engineering*. 2024;21(2).
- [10] Spüler M, Rosenstiel W, Bogdan M. Unsupervised online calibration of a c-vep brain-computer interface (bci). In: *Artificial Neural Networks and Machine Learning—ICANN 2013: 23rd International Conference on Artificial Neural Networks Sofia, Bulgaria, September 10-13, 2013. Proceedings 23. 2013*, 224–231.
- [11] Turi F, Gayraud NT, Clerc M. Auto-calibration of c-vep bci by word prediction. 2020.
- [12] Thielen J, Sosulski J, Tangermann M. Exploring new territory: Calibration-free decoding for c-vep bci. *arXiv preprint arXiv:2403.15521*. 2024.
- [13] Nagel S, Dreher W, Rosenstiel W, Spüler M. The effect of monitor raster latency on VEPs, ERPs and Brain-Computer Interface performance. *Journal of Neuroscience Methods*. 2018;295:45–50.
- [14] Santamaría-Vázquez E, Martínez-Cagigal V, Hornero R. Bit-wise reconstruction of non-binary visual stimulation patterns from eeg using deep learning: A promising alternative for user-friendly high-speed c-vep-based bcis. In: *International Work-Conference on Artificial Neural Networks*. 2023, 603–614.
- [15] Behboodi M, Kinney-Lang E, Etemad A, Kirton A, Abou-Zeid H. Leveraging foundation models for calibration-free c-vep bcis. *IEEE*, Oct. 2025, 4564–4571.
- [16] Martínez-Cagigal V. Dataset: Non-binary m-sequences for more comfortable brain-computer interfaces based on c-veps. 2025.
- [17] Santamaría-Vázquez E *et al.* MEDUSA©: A novel Python-based software ecosystem to accelerate brain-computer interface and cognitive neuroscience research. *Computer Methods and Programs in Biomedicine*. 2023;230(107357).
- [18] Martínez-Cagigal V, Santamaría-Vázquez E, Pérez-Velasco S, Marcos-Martínez D, Moreno-Calderón S, Hornero R. Non-binary m-sequences for more comfortable brain-computer interfaces based on c-VEPs. *Expert Systems with Applications*. 2023;232(120815).
- [19] Santamaría-Vázquez E, Martínez-Cagigal V, Vaquerizo-Villar F, Hornero R. Eeg-inception: A novel deep convolutional neural network for assistive erp-based brain-computer interfaces. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*. 2020;28:2773–2782.
- [20] Martínez-Cagigal V, Santamaría-Vázquez E, Pérez-Velasco S, Martín-Fernández A, Hornero R. Reevaluating performance in c-vep bcis: The impact of calibration time. *Biocybernetics and Biomedical Engineering*. 2025;45:685–696.