Decoding Neural Activity Preceding Balance Loss During Standing with a Lower-limb Exoskeleton using an Interpretable Deep Learning Model

Akshay S Ravindran [1], Christopher A. Malaya[2], Isaac John[2], Gerard E. Francisco[3], Charles Layne[2], Jose L. Contreras-Vidal[1]

[1] Noninvasive Brain-Machine Interface System Laboratory, Department of Electrical and Computer Engineering, University of Houston, 77204, USA

[2] Center for Neuromotor and Biomechanics Research, Department of Health and Human Performance, University of Houston, Houston, 77204, USA

[3] TIRR Memorial Hermann and Department of PMR, University of Texas Health Sciences Center, Houston, Texas, 77204, USA

Abstract.

Falls are a leading cause of death in adults 65 and older in the United States. Recent efforts to restore lower-limb function in these populations have seen an increase in the use of wearable robotic systems; however, fall prevention measures in wearable robots require early detection of balance loss to be effective. Prior studies have investigated whether kinematic variables contain information about an impending fall, but few have examined the potential of using scalp electroencephalography (EEG) as a fall-predicting signal. Moreover, there exists a major gap in our understanding of not only how the brain detects and processes balance perturbations, but also on how it responds to avoid a fall. To address these knowledge gaps, we decoded neural activity in a balance perturbation task during standing with and without an exoskeletal suit. Specifically, we acquired high-density EEG, electromyography (EMG), and center of pressure (COP) data from 7 healthy participants during both fixed and random parameter mechanical perturbations while standing. The timing of the perturbations was randomized in all trials. We found perturbations evoked potentials (PEP) components as early as 75-134 ms after the onset of the external perturbation, which preceded both the peak in EMG activity (˜180 ms) and the COP (˜350 ms). We then trained a convolutional neural network (CNN) to predict the balance perturbations from single-trial EEG. The CNN model, had a mean F score of 75.0 ± 4.3 %. Using a novel approach of clustering GradCAM based model explanations, we demonstrated that the model utilized relevant components in the PEP to infer the predictions and was not driven by artifacts. The model's explanations further aligned with a dynamic functional connectivity measure estimated from the phase difference derivative. Specifically, the nodal connectivity was higher in the occipital-parietal region in the early stage of the perturbations, before shifting to the parietal, motor, and back to the frontal-parietal channels, suggesting a closed-loop network. Continuous-time decoding of COP trajectories from EEG, using a gated recurrent unit model, achieved a mean Pearson's correlation coefficient of 0.7 ± 0.06. Overall, our findings suggest that the EEG signals contain short-latency neural information related to an impending fall, which may be useful for developing brain-machine interface (BMI) systems for fall prevention in robotic exoskeletons.

1. Introduction

The World Health Organization (WHO) reported that over 37 million falls require medical attention each year worldwide [1]. Indeed, falls are a leading cause of injury, loss of independence, hospital admission, and even death. While conventional therapies have been successful in fall reduction and prevention, many individuals with severe illness or injury remain unable to participate in activities of daily living (ADLs) or complete standard care protocols. Recent efforts to aid these populations have utilized wearable robotic systems and, in particular, powered robotic orthoses (i.e., exoskeletons) [2],[3].

The U.S. Food and Drug Administration (FDA) classifies powered exoskeletons as Class 2 medical devices with special controls. They are used frequently for rehabilitation applications due to their ability to provide active, assistive support for walking, sitting, and standing [4],[5]. When compared to traditional therapies, these devices provide intense training in an active and stimulating environment while providing quantifiable markers of progression [6],[7]. In addition to rehabilitation, exoskeletons can also be purposed to reduce the risk of falling and/or aid in fall prevention.

However, falls while wearing the exoskeletons are a significant risk in using these devices [8]. Current FDA-cleared exoskeletons use different strategies for dealing with potential falls and are indicated for use with a trained companion. The effectiveness of these strategies is not studied and is still unclear. Some systems utilize kinematic response assessments to detect fall events based on accelerometers, magnetometers, or joint angles. The Indego and Ekso exoskeleton systems detect falls in real-time by checking for excursions in kinematic variability beyond certain limits. In the case of the Indego device, movements beyond a set threshold will trigger corrective postural movements to reduce the risk of injury [8]. However, while other studies have examined fall risk and incidence [9], tested exoskeletons during perturbations [10],[11], or even developed positioning algorithms to promote safer falls [12], very few appear to both detect and respond to these falls or perturbations. There was only one study that was identified which detailed an exoskeleton system with built-in perturbation or fall detection and response. In this study, Monaco et al. utilized a micro-controller to compare real-time kinematics with predicted walking values. Threshold reaching discrepancies between the predictions and real values were used to apply corrective hip torques to restore balance. Their detection algorithm was able to identify the lack of balance resulting from slippages within about 350 ms of the event [13]. Nevertheless, there are still drawbacks to this mechanism of fall detection; kinematic measures leave minimal time between detection and the fall event. In these systems, given that the use of electric motors with large gear reductions will have reduced response speed, early detection of balance loss is critical. With this in mind, approaches that can identify and act to correct balance loss earlier would be extremely beneficial.

Kinematic measures are not the only way of detecting fall events. Multi-sensory information from visual, somatosensory, and vestibular systems acting on the cerebral cortex, cerebellum, and brainstem have a significant role in postural corrections [14].

These sensory signals might precede the latency of kinematic responses and could offer a longer stimulus to fall interval within which to respond. Physical balance perturbations elicit cortical responses called Perturbation Evoked Potentials (PEP). These PEP can be detected using electroencephalography (EEG). A PEP generally consists of 3 components. The first component is a small positive wave (P1) at approximately 30-90 ms. This is followed by a negative peak at around 90-160 ms with a final, late response (P2 and N2) around 200-400 ms [15]. These PEPs are typically observed by averaging waveforms across many trials. However, if PEP could be detected from a single trial, balance perturbations could be identified much earlier. This would afford considerable, additional time to initiate preventative movements.

Studies examining perturbations during exoskeleton use with an EEG paradigm, as well as the temporal relationship between signal modalities, are rare [16]. More importantly, to our knowledge, no previous studies have evaluated the influence of balance perturbations on EEG during exoskeletal suit use. Further understanding of the influence of exoskeletons on physiological responses observed with EEG as well as physical responses to perturbations is important. In this study, how different perturbations during standing conditions modulated the brain activity was evaluated and tested the possibility of detecting physical perturbations from single-trial EEG in individuals wearing an exoskeleton.

2. Methods

2.1. Participants

Seven healthy participants (5 male) aged 18-32 participated in the study. The experimental protocol was approved by the Institutional Review Board (IRB) at the University of Houston, in accordance with the Declaration of Helsinki. The written informed consent form was collected from each of the participants before the start of the experiment.

2.2. Experimental Setup

Participants were fitted with a 64-channel EEG cap (ActiCap, Brain Products, GmbH, Morrisville, NC) referenced to the ear lobes. 60 active AG/AgCL electrodes were placed in the cap according to the modified 10-20 international system to record EEG signals. Electrodes normally positioned at FT9 and FT10 were moved to replace the AFz and FCz electrodes on the cap (ground and reference, respectively). In addition, electrodes that were to be placed at TP9, TP10, PO9, and PO10 were instead used to measure electro-oculography signals (EOG). Two electrodes were placed above and below the right eye with the remaining two electrodes placed at the lateral canthus of each eye to extract the eye-related artifacts. EEG/EOG data were recorded wirelessly using the MOVE system at 250 Hz and amplified using the BrainAmp DC amplifier (Brain Products, GmbH, Morrisville, NC).

Surface electromyography (EMG) sensors were placed over the tibialis anterior (TA), Medial Gastrocnemius (MG), Lateral Gastrocnemius (LG), and Soleus (S) muscles of both legs, along with one sensor on the forehead and torso. EMG data were collected wirelessly using the Delsys Trigno system (Delsys Inc., Boston, MA).

After set-up and electrode impedance measurements, participants were asked to stand comfortably on a balance platform (Neurocom Balance Manager platform, (NeuroCom, Clackamas, OR) for 2 minutes to acquire eyes open resting-state activity. At the end of 2 minutes, subjects received a series of postural perturbations. This consisted of a series of 32 constant (duration, period, and velocity) perturbations where the platform generated maximal backward translations (displacement of 6.35 cm in 400 ms, i.e. velocity of 15.875 cm/s). This condition is referred to as the Random Timing Condition (RTC), as the timing alone was randomized. The second postural task consisted of 33 random/unexpected perturbations where the platform generated forward/backward/tilted perturbations in a random order (Random Timing and Type Condition - RTTC). Individual trials with the same parameters as the RTC trials were embedded randomly into the RTTC condition. After 16 trials of RTC and RTTC, respectively, a break of approximately 2-5 minutes was given to avoid fatigue. Each trial lasted five seconds and the timing to perturbation onset was randomized in all trials to avoid anticipation of when the perturbation would occur. All conditions were repeated with and without the H2 exoskeleton (in passive mode with the joints decoupled) to evaluate if PEPs would be altered in the presence of the mechanical constraints introduced by wearing the exoskeleton. For every other participant, the order of trials with and without H2 was reversed. The protocol is summarized in Fig 1.

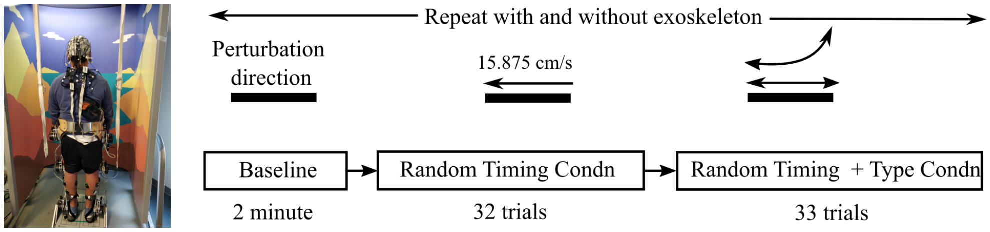


Figure 1: Experimental protocol: the two conditions were repeated with and without the exoskeleton. A 2-5 minute break was provided in between each of the blocks (RTC RTTC).

2.3. Signal pre-processing

The pre-processing steps used to process EEG, EMG, and the Neurocom data are summarized in figure 2.

Both the EEG and EOG signals were bandpass filtered between 0.2 to 50 Hz to remove low-frequency drift and minimize muscular artifacts. A 4th order zero-phase Butterworth filter was used to avoid phase distortion. The high pass cut-off of 0.2 Hz was selected from Tanner et al., which suggested high pass filtering above 0.3 Hz will distort the ERP components [17]. Ocular artifacts were removed using the H-infinity-based

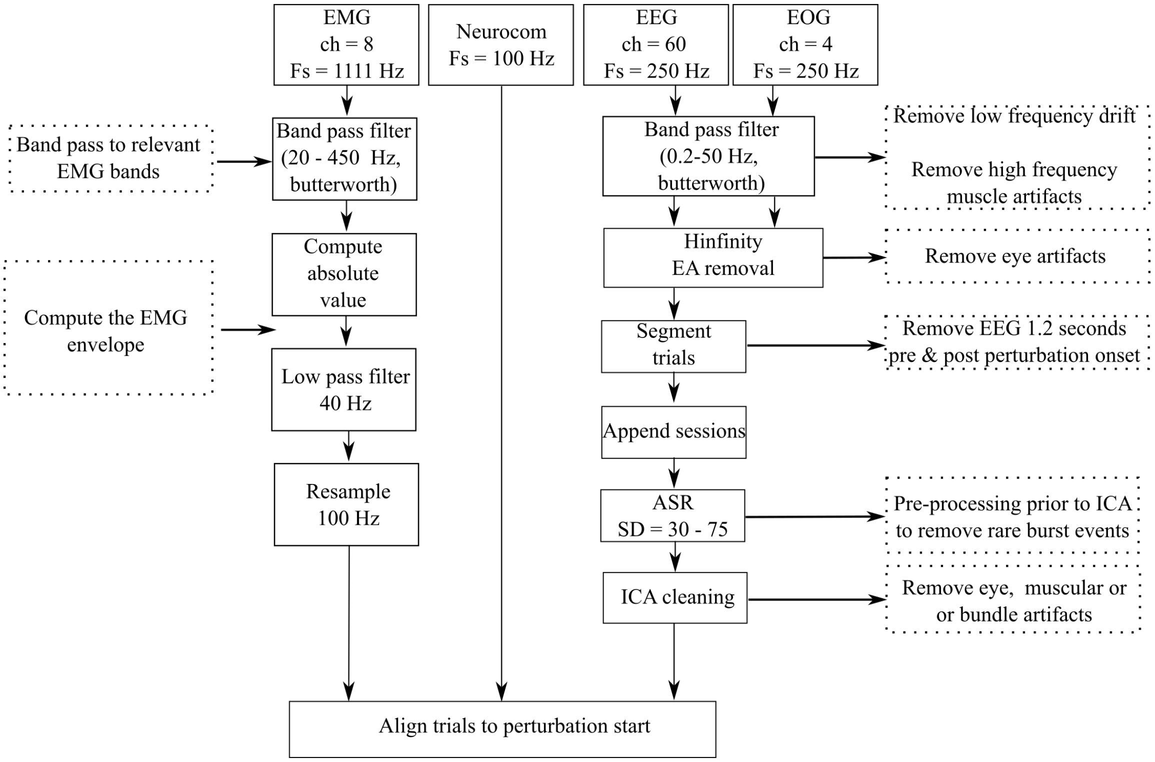

EMG
ch = 8
Fs = 1111 Hz
Neurocom
Fs = 100 Hz
EEG
ch = 60
Fs = 250 Hz
EOG
ch = 4
Fs = 250 Hz
Band pass to relevant EMG bands
Band pass filter (20 - 450 Hz, butterworth)
Compute absolute value
Compute the EMG envelope
Low pass filter 40 Hz
Resample 100 Hz
Band pass filter (0.2-50 Hz, butterworth)
Remove low frequency drift
Remove high frequency muscle artifacts
Hinfinity EA removal
Remove eye artifacts
Segment trials
Remove EEG 1.2 seconds pre & post perturbation onset
Append sessions
ASR SD = 30 - 75
Pre-processing prior to ICA to remove rare burst events
ICA cleaning
Remove eye, muscular or or bundle artifacts
Align trials to perturbation start

Figure 2: Flowchart detailing the different pre-processing steps performed for each of the signal modalities.

adaptive filter [18]. The *gamma* parameter was set to 1.1 and the *q* parameter used was 1e-11 from empirical testing. Data 1.2 seconds before and after the perturbations were discarded and individual trials were concatenated together. Later, to remove any sudden spikes in the EEG and improve Independent Component Analysis (ICA) decomposition, Artifact Subspace Reconstruction (ASR) [19] with less conservative thresholds of 30-75 were used to reconstruct poor components in artifactual windows. The thresholds were selected based on empirical evaluation and also by recommendations from Chang et al. [20]. ICA decomposition was then performed using the Infomax algorithm to identify and remove ocular, muscular, or bundle artifacts (artifacts caused by the physical pulling of cable bundles). Here, a more conservative cleaning is performed to remove 26-44 ICs across subjects. Ocular artifacts were identified by looking at topographical distributions, power spectra with power localized in the delta/theta bands, as well as the time-series data for repeatable ocular artifacts. Muscular ICs were identified by examining the spatial weighting of the IC (localized in the temporal channels), power spectra (looking at the increasing power in 30+ Hz) as well as time-series data for spiking activity. The bundle artifacts were identified by the spatial weight of the IC (alternating pattern for the 2 bundles). Any additional ICs (indicating electrode shifts) were identified and removed. Representative examples of the ICs removed are provided in the supplementary materials. All the pre-processing steps were implemented using the EEGLAB library[21].

EMG data were bandpass filtered with a passband frequency of 20 - 450 Hz using a 4th order Butterworth zero-phase filter. Later, to extract the envelope, data were rectified by computing the absolute value and passing through a second low pass filter at 40 Hz. The envelope of the EMG was then resampled to 100 Hz to match with the sampling rate of the kinematic data from the Neurocom. All three modalities were then aligned to the perturbation onset in each of the trials.

2.4. Latency relationship between the signals

To study how electrophysiological and kinematic responses varied in response to the perturbation, all the signals after baseline correction were trial averaged. This also increased the signal-to-noise ratio. Averaging was done separately for each of the conditions. The period between -500 ms to -200 ms was used to estimate baseline correction values. Perturbation response in the first trial was consistently, significantly larger than the succeeding trials, and thus were removed before averaging. The trial averaged physiological and kinematic signals were aligned to the perturbation onset to evaluate the latency difference between the signals. In the end, the grand average response was computed by averaging the time series across all subjects and trials.

2.5. Detecting perturbations from single trials

A CNN was implemented to detect the presence of perturbations from 200 ms long windows of single-trial EEG. Class 1 was composed of individual trial windows during the baseline period (1200 to 500 ms) prior to the onset of perturbation. Class 2 consisted of EEG segments between -200 ms until +500 ms post perturbation onset. Windows of 200 ms from each of these classes were extracted in a sliding window manner with a one-sample difference. The data were scaled by dividing by a value of 100 (μV). The baseline period per trial was selected as class 1, instead of the resting state, to avoid the model prediction being confounded by impedance change between the two segments. It further ensures that internal states unrelated to the perturbations are comparable across the classes.

To increase the sample size for the classifier, trials not involving the exoskeleton were also included. Therefore, a total of 60 trials of RTC trials were used for training the model. Trials were randomized and divided into train, test, and validation sets. 15% of trials were divided into validation and 15% into the test set. The data was divided based on trials and not by random sampling of all the windows. This was done to avoid any potential data leakage due to the high level of overlap. This ensured that there was no shared information between the three sets. A total of 5 such held out sets were created for cross-validation to evaluate the generalizability of the model. In addition to test accuracy, F-score was also computed for each of the folds.

The architecture for the model is summarized in Fig. 3. The input to the model is the 200 ms EEG window (batch size x 50 samples x 60 channels). The model consisted of 5 temporal convolution layers of 8 units each (3 x 1 kernel size with a stride length of

1). A temporal pooling layer of 2x1 pooling dimension with a stride length of 2 was also used after every pair of convolutional filter layers except the last block. These should help with the trial-by-trial translational variance of the PEP components. The output from these convolutional layers was flattened and fed into a dense, fully connected layer of 16 hidden units followed by an output layer with softmax activation.

A dropout layer with alpha = 0.5 was added in between the dense layer and the output layer to reduce overfitting. Except for the output layer, the model utilized ReLU as the activation function. An Adam optimizer [22] with a learning rate of 0.0001 was used to train the model. A batch size of 32 and epoch length of 100 was set. An early stopping condition was set to avoid the model from being overfitted. This stopped the training if the validation loss did not improve in 5 consecutive epochs. A re-initialized independent copy of the same model architecture was used for each fold and subject. The proposed model was implemented in python 3.6 using keras 2.15 [23] wrapper using Tensorflow [24] backend.

The model architecture was selected to better facilitate the GradCAM algorithm in identifying relevant channels. Most currently available models use a spatial filter in the early stage of the architecture. If spatial filters are used early on, the deeper layers can only see a mixed channel (time x number of filters dimension) representation. GradCAM will not be able to identify the relevant channel distribution. Here, the emphasis was put on explaining the model decision to ensure the model is indeed learning from relevant components and not driven by irrelevant signals. To ensure that prioritizing explainability during architecture selection did not impair decoding performance, the performance of the model was compared with the DeepConvNet architecture [25]. The original paper that proposed the DeepConvNet architecture used a 2-second long EEG, sampled at 256 Hz as input. For the DeepConvNet, to account for the difference in dimensions the architecture hyper-parameters were modified to make it compatible with our data. In this study, three blocks were used instead of four as the window size is not long enough to accommodate the 4th block. Additionally, to evaluate the impact of denoising, the process was repeated by training the model used in this study on EEG data prior to ICA cleaning instead of the denoised EEG.

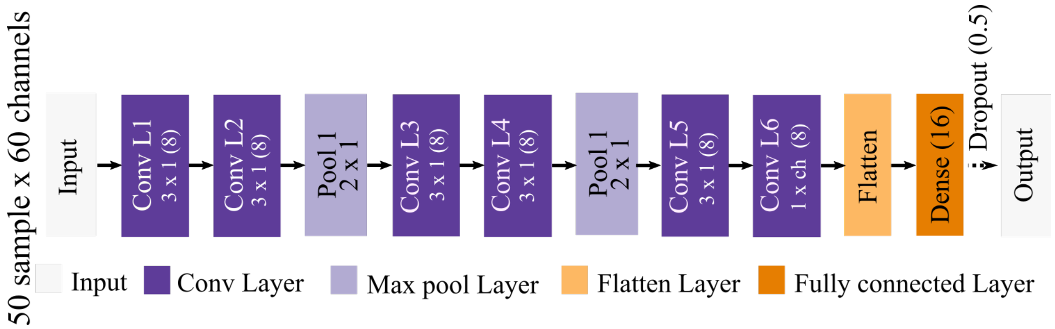


Figure 3: Model architecture: Each block correspond to different types of layers in the model. The dotted line is to illustrate the dropout operation during the training phase aimed at reducing overfit. During inference, all units were retained.

2.6. Explaining the CNN model decision

The model decision explanation was carried out using the GradCAM method [26]. GradCAM is a class-specific explanation technique that identifies relevant regions in the input that the model used to make the prediction pertaining to a specific class. The algorithm is explained in Selvaraju et al. 2017 [26]. GradCAM is a generalization for Class Activation Map (CAM) as CAM limits the CNN to require a global average pooling layer at the end of the convolutional blocks. GradCAM on the other hand does not require this. GradCAM computes the gradient of the score of the class of interest with respect to each of the feature map activations of the penultimate layer being considered. These gradients are then global average pooled to serve as weights for the particular feature map. A weighted sum of the feature map activations with respect to these weights is then computed. These are then are passed through a ReLU operation to consider only positive values as they contribute to making the correct prediction. Here, the penultimate layer used is the convolutional layer L5 to learn channel relevancy. From the model explanations, time-averaged GradCAM is computed to identify the relevant channels per window.

Next, k-means clustering was performed on the model decisions. All the correctly predicted data points across all subjects from the best performing fold (combined validation and test set) were fed into the clustering algorithm instead of visualizing hand-selected examples to avoid bias. The distance measure used was squared euclidean with the maximum number of iterations allowed set to the total number of samples present. The optimal cluster number was selected using the elbow method. K-means was evaluated for a variable number of clusters ranging from one to 100. The total within-cluster sums of point-to-centroid distances were computed for each of the K values. The K values that corresponded to the knee of the curve were selected. Instead of manually selecting the knee point which could be subjective, the Kneedle algorithm was used to detect the knee [27]. The parameter S was set to 0 as recommended in the offline setting in the original paper[27]. The process was repeated 5 times and the average K values were chosen for the final k-means clustering. The cluster results were then evaluated to assess whether the model was learning from the PEP components and not being driven by artifacts. The process was repeated on separate models trained on pre-processed EEG as well as raw EEG without ICA cleaning.

2.7. Post-hoc test to evaluate model explanation with traditional signal processing approaches

To evaluate how the network dynamics evolve with time during the PEP, a measure of dynamic functional connectivity called phase difference derivative (PDD) [28] was calculated for each trial. One of the more prominent models of the origin of ERP is the phase reset model wherein resetting of phase of the ongoing oscillations cause the generation of ERP [29]. We expect the PDD to quantify the dynamics of the various PEP components across trials. PDD is a measure of the stability of phase difference between two signals. It computes the instantaneous phase of the signal based on the analytic signal extracted from the Hilbert transform of each of the signals. For phase-locked signals, the difference in phase remains constant across time, in which case the derivative of that should be approximately zero. Taking the negative exponent of the derivative further ensures that it is bounded between 0 and 1 with a value of 0 meaning no coupling between the signals. The equation to estimate PDD is

$$PDD_{ij}(t) = \exp(-|\frac{d\Delta\Phi_{ij}(t)}{dt}|). \tag{1}$$

Here, $\Delta\Phi_{ij}$ is the phase difference between signals i and j at time t. The PDD in the alpha band was calculated by initially band-pass filtering the signal using a 4th order zero-phase butter worth filter in the band (8-13 Hz). The PDD was estimated with a center frequency of 10 Hz and a window size of 128 ms. The window size was selected such that it contains at least one cycle of the lowest frequency of interest (8 Hz). The measure was estimated from seven channels. Six of the channels were relevant to the task (based on model explanations from CNN). A seventh channel, which we expected to be task-independent (TP7) was also evaluated. The PDD was baseline corrected (w.r.t. -500 ms to -200 ms) to further remove any residual connectivity across channels that are not task-dependent. The grand average ERP and PDD were estimated from each of these channels using the same procedure as described in the section above.

2.8. Continuous decoding of COP from EEG

The predictive power of EEG to continuously decode the COP variations in response to perturbations was then evaluated. Gated Recurrent Units (GRU) were used to decode the COP values. Considering the perturbations were solely a backward translation, only the y component of COP was decoded as it had the largest modulations. To evaluate the ideal model parameters, a hyperparameter search was performed by varying the number of layers (1 to 3) and the number of units per layer (8, 16, 32, 64, 128, 256, 512, 1024). This was followed by a dense layer with a ReLU activation function and the number of units equal to that of the GRU units. The dense layer was then connected to the output layer with a linear activation function. To evaluate the decoding performance, the coefficient of determination (R2 score), Pearson's correlation coefficient (R-value) and mean squared error (MSE) metrics were used [30]. All of these were implemented in python using the Scikit library [31]. Similar to the classification model, 70% of the trials were divided into training, 15% for validation, and 15% for testing. The GRU model was trained and tested on five such splits to evaluate generalizability. Here, unlike the classification model, EEG from 1.2 seconds prior to perturbation onset until 1-second post perturbation onset was used. Separate models were trained for each combination of participant x number of layers x number of GRU units x folds. Predicted and actual COP values were evaluated using the measures on the validation set across all 5 sets to identify the optimal model hyperparameters. Upon identifying the optimal hyperparameters for the model with minimal computational cost, the optimized model was evaluated on the test set to determine final performance values. The models were trained using the Keras library with the TensorFlow backend. The initial learning rate was set to 0.001 with the model weights optimized using Adam optimizer [22]. The batch size used was 128 and trained for a maximum of 200 epochs with an early stopping condition of stopping the training if the validation loss did not improve in 5 consecutive epochs. The GRU was trained to minimize the mean squared error between the actual and predicted COP values. To further evaluate how the model generalized when the person was not only blind to the timing but also the type of perturbation, trials with the same type of perturbations that were randomly present in the RTTC sessions were also tested.

3. Results

3.1. Latency relationship between the signals

Fig 4 depicts the grand average response across channels during the exoskeleton RTC condition. The top row shows the grand average PEP components in the Cz EEG channel. All the previously reported components of the PEPs including P1, N1, and P2 are retained while wearing the exoskeleton. In addition, the P1 peak (75 ± 8 ms) and N1 peak (137 ± 12 ms) precedes the peak in EMG (MG: 195 ± 27 ms; LG: 182 ± 19 ms; TA: 180 ± 14 ms; S: 181 ± 13 ms) which again precedes the peak in the COP (365 ± 22 ms). The peak of COP is the point at which the participants start initiating the return to the original position. This indicates that EEG contains discriminatory information much earlier than the kinematic response which could be used to detect the balance perturbations.

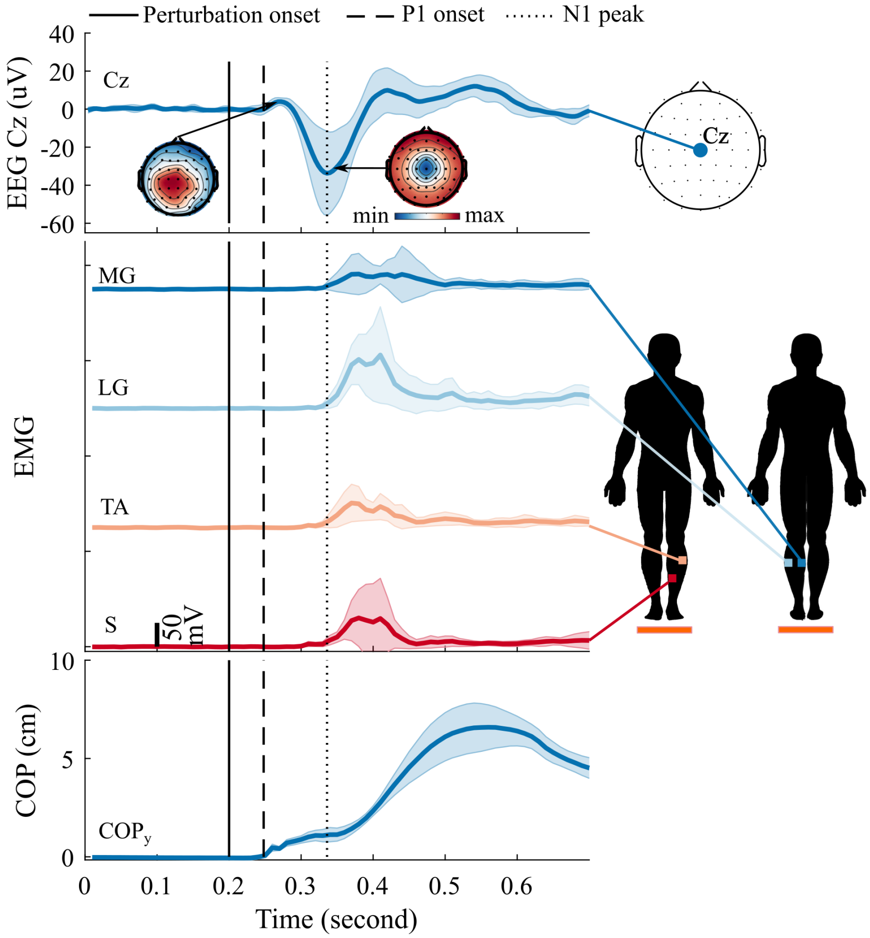


Figure 4: Between subject grand average latency difference between different electrophysiological and kinematic responses associated with balance perturbation while wearing the exoskeleton. The muscles shown are from the left leg with the following abbreviations: MG (medial gastrocnemius), LG (lateral gastrocnemius), TA (Tibialis Anterior), S (Soleus).

Table 1: Cross validated performance metrics evaluated on the test set; all numbers are in percentages; Raw: model trained on EEG without ICA denoising, Clean: model trained on ICA cleaned EEG, DCN: DeepConvNet trained on ICA cleaned EEG.

Sub	**Accuracy**			**F-score**		
	Raw	**Clean**	**DCN**	**Raw**	**Clean**	**DCN**
S1	79.3 ± 4.1	75.2± 3.6	67.3 ± 2.5	79.2 ± 4.0	75.0 ± 3.4	65.4 ± 2.1
S2	72.2 ± 5.4	77.6 ± 6.8	73.3 ± 3.1	72.1 ± 5.4	77.5 ± 6.8	72.7 ± 3.7
S3	77.2 ± 11.5	75.5 ± 3.7	67.0 ± 5.5	77.0 ± 11.6	75.4 ± 3.7	65.3 ± 7.2
S4	84.1 ± 1.7	70.3 ± 5.4	65.0 ± 5.0	84.0 ± 1.6	69.2 ± 7.1	64.0 ± 4.9
S5	74.9 ± 4.2	71.4 ± 4.1	70.5 ± 5.2	74.5 ± 4.2	71.1 ± 4.1	69.9 ± 5.6
S6	81.7 ± 3.9	80.0 ± 1.8	79.8 ± 2.6	81.6 ± 3.9	79.8 ± 1.9	79.6 ± 2.6
S7	76.9 ± 5.3	75.3 ± 4.8	70.7 ± 4.1	76.6 ± 5.2	75.2 ± 4.7	69.7 ± 3.9
Avg	78.0 ± 5.2	75.0 ± 4.3	70.5 ± 4.0	77.9 ± 5.1	74.7 ± 4.5	69.5 ± 4.3

3.2. Detection of balance perturbation using a convolution neural network

The capability for CNN to detect the PEP components and other underlying neural representations from single trials alone in a data-driven manner was tested. The cross-validated results are summarized in Table 1. Overall, all the subjects obtained above chance level ($\sim$ 50%) classification scores. A cross-validated mean test F score of 74.7 ± 4.5 % was obtained. Subject 4 had the lowest F score of 69.2 ± 7.1 % whereas subject 6 obtained the highest F score of 79.8 ± 1.9 %. The same model was tested on EEG without ICA denoising (Raw) and that model achieved a higher decoding accuracy (F score = 78.0 ± 5.2).

DeepConvNet yielded a mean test F score of 69.5 ± 4.3. Compared to DeepConvNet, our model performed better. However, we emphasize that the study do not claim superiority for the architecture. Instead, this is evaluated only to show that focusing on architecture by prioritizing model explanation did not compromise model performance. To make the comparison fairer, randomization of the trials was made consistent for all models by assigning the same seed per fold.

3.3. Explaining the CNN model decision

The optimal K value to perform the k-means on the model explanations was identified as 11 for the model trained on clean EEG and 14 for the model trained on EEG without ICA cleaning. Fig 5 shows the clustering results on the best-performing fold for both cases. Fig 5.a summarizes the clustering performed on the explanations from the model trained on cleaned denoised EEG. Fig 5.b corresponds to the explanations from the model trained on the raw EEG without ICA cleaning. The top row in both cases shows the mean relevancy score for the channels in each of the identified clusters. The middle row represents the distribution of window latency relative to perturbation onset (w.r.t. the last sample in each window). The distribution was normalized for visualization purposes. The third row shows the contribution of the examples in each cluster from each of the 7 participants.

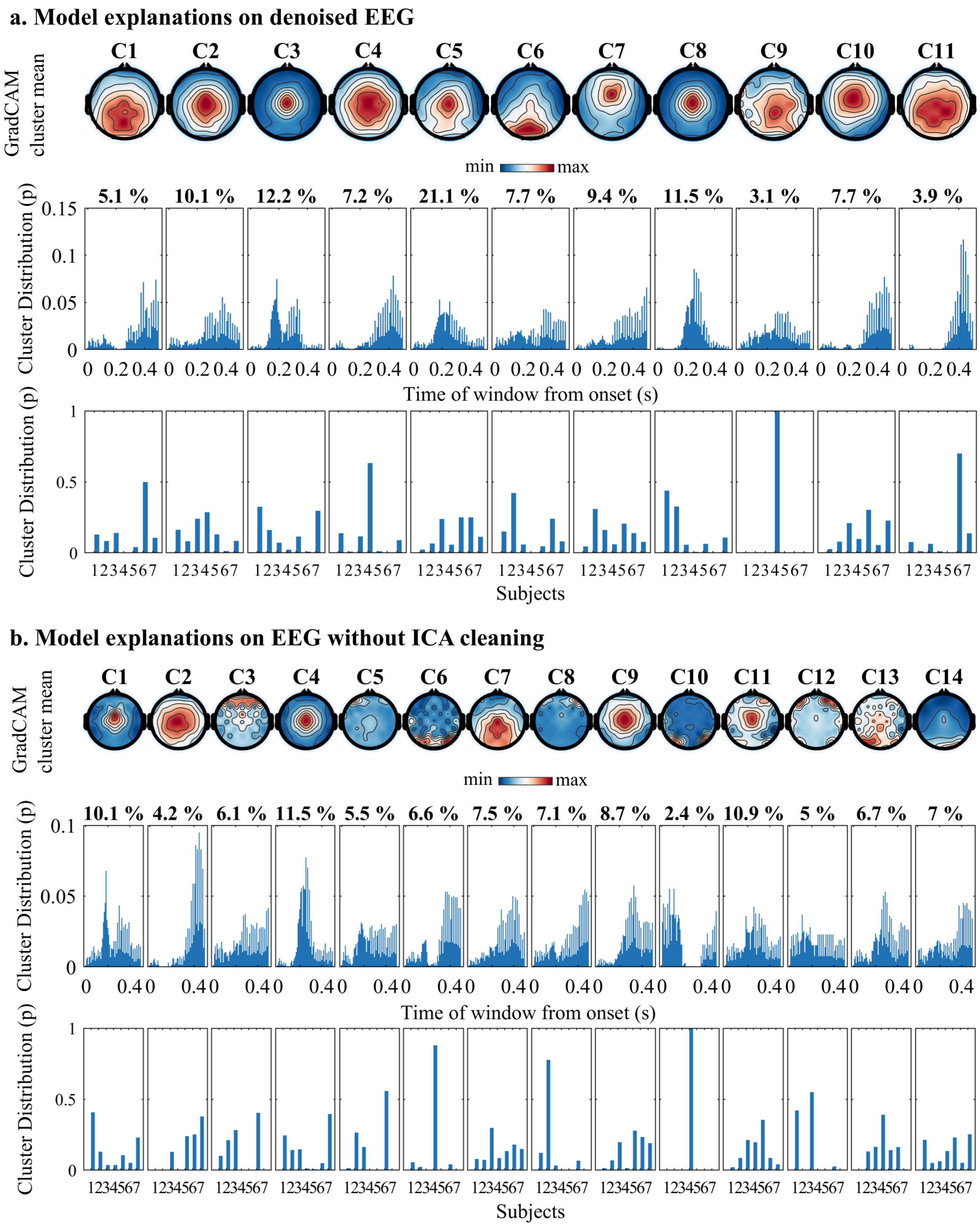
From Fig 5.a, it can be seen that none of the clusters were weighing in on the periphery channels, which are often strongest if driven by artifacts. Almost all clusters were focusing on the channels in the motor, parietal and pre-motor regions to arrive at the decisions. From these, clusters C3 and C8 are localized in the Cz channel and are centered around the time when N1 peaks. Similarly, the parietal channels become more relevant both in the early and late stages of the perturbations (C1, C6). The clusters localizing in the frontal channels (C7, C10) are centered in the latter half of the perturbation. In multiple clusters, the model is focusing on a broader range of channels but is still centered around the motor regions (C1, C2, C4, C11). The largest cluster, C5 had contributions from both the central as well as the parietal channels. Overall, in evaluating the spatial map distribution, the response is found to be highly dynamic, involving multiple brain regions varying over time.

To further verify that the model explanation was not biased against detecting artifacts and that the pre-processing was reliable and significant, the process of training and explaining the model decisions was repeated on EEG without ICA cleaning. The clustering results of data with artifacts are summarized in Fig 5.b. Even though the model trained on data without ICA cleaning achieved higher performance (F-score: 78 ± 5.2), evaluating the model explanations, it was observed that the model was learning the artifacts for decoding purposes. The model learned to detect the bundle artifacts indicated by alternative localized channel relevancy (C3, C6, C13) as well as started giving more relevance to the peripheral channels (C3, C5, C8, C10, C11, C12). These were absent in our pre-processed data.

3.4. Post-hoc test to evaluate model explanation with traditional signal processing approaches

The variability of the dynamic measure of the functional connectivity ΔPDD is shown in Fig 6. The channels over the parietal and parietal-occipital lobe that are heavily reported to be involved with sensory processing have increased connectivity in both the start and end of the perturbations. The variability in the motor channels particularly the Cz is centered around the N1 peak. The FCz on the other hand has an increase in connectivity relative to other channels soon after the N1 peak as well.

In addition, the connectivity strength of the Cz, C2, and FCz channels is high w.r.t. the frontal and parietal channels prior to the perturbations suggesting anticipatory mechanisms. TP7 which is task-irrelevant does not appear to have significant activity throughout the duration of interest.



a. Model explanations on denoised EEG

b. Model explanations on EEG without ICA cleaning

Figure 5: Clustering result of the model explanations from the highest performing fold.

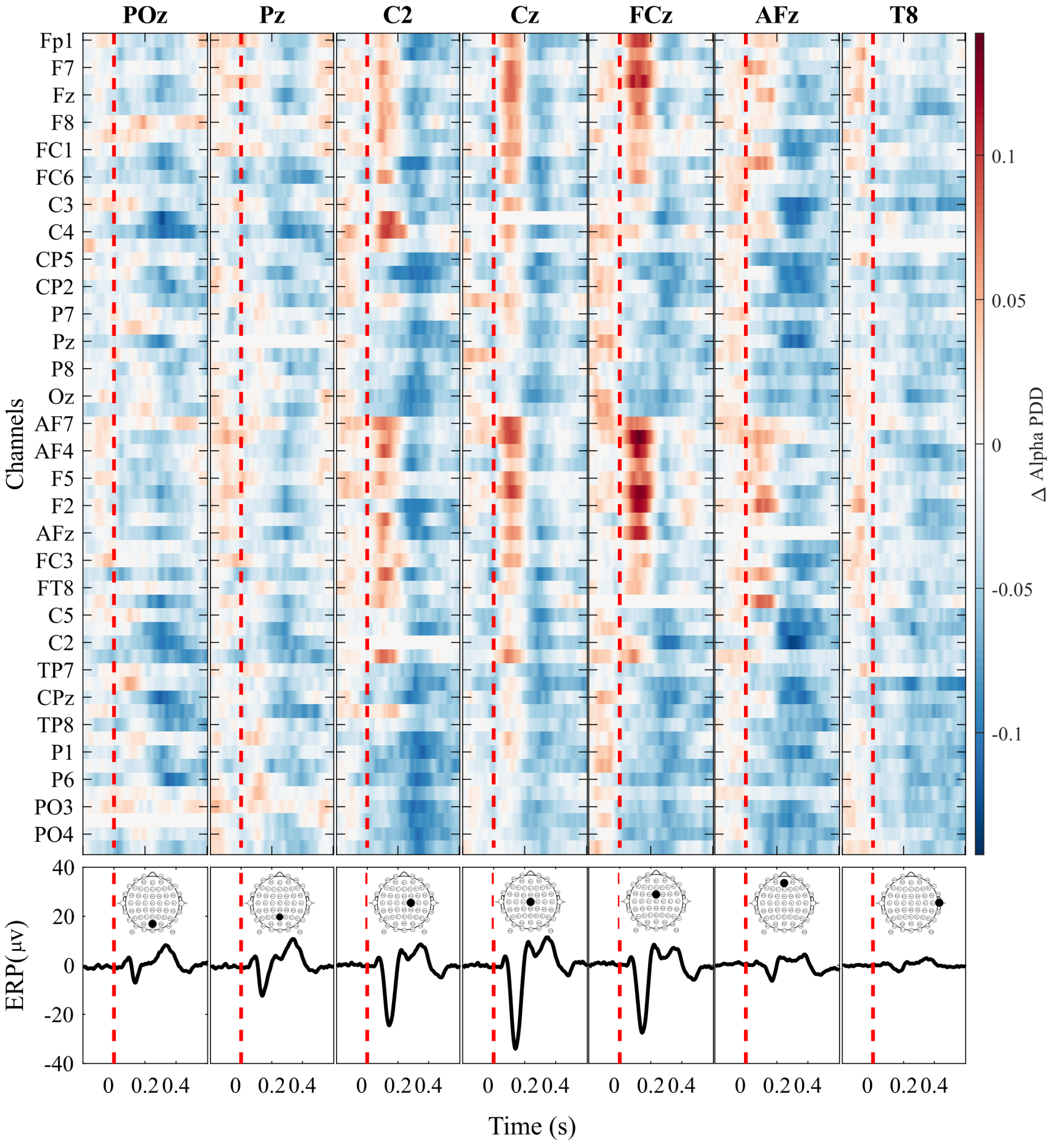


Figure 6: The difference in alpha band PDD w.r.t. -500 to -200 ms prior to the trials. Each column corresponds to connectivity w.r.t. one specific channel. The top row indicates how the alpha band Δ PDD of all other channels w.r.t. the channel of interest changes with time. The bottom row is a grand average PEP for the channel of interest.

Table 2: GRU decoder performance metrics on the test set.

Subject	Correlation (RTC)	R2-Score (RTC)	Correlation (RTTC)	R2-Score (RTTC)
HS1	0.45 ±0.08	0.13 ±0.13	0.33 ±0.06	0.07 ±0.06
HS2	0.76 ±0.03	0.54 ±0.07	0.81 ±0.02	0.65 ±0.03
HS3	0.71 ±0.06	0.47 ±0.14	0.8 ±0.02	0.63 ±0.04
HS4	0.56 ±0.06	0.29 ±0.08	0.37 ±0.06	0.07 ±0.06
HS5	0.81 ±0.05	0.64 ±0.09	0.76 ±0.04	0.56 ±0.06
HS6	0.85 ±0.06	0.7 ±0.1	0.74 ±0.02	0.51 ±0.02
HS7	0.78 ±0.04	0.59 ±0.08	0.64 ±0.03	0.34 ±0.06
mean ±*s.d.*	0.7 ±0.06	0.48 ±0.1	0.64 ±0.03	0.41 ±0.05

3.5. Continuous decoding of COP from EEG

From the model explanation results and the PDD analysis, it was observed that there are dynamical changes in response to perturbations with time. Additionally, from the PEP, it is clear that distinct PEP components exist at varying latencies. With this in mind, the possibility to estimate the variation of COP associated with balance perturbation from EEG was tested. Initially, the cross-validated grid search identified the optimal hyperparameters for the GRU architecture. Fig 7. A) shows the distribution of R-value, R2 value, and MSE losses for all combinations of the hyperparameters used. After the hyperparameters were selected based on the performance metrics evaluated on the validation set, the optimized model was tested on the held-out test set. The performance measures are summarized in table 2. Evaluating the violin plot, the number of layers was found to be not critical here. The performance initially increases with the number of units but starts decreasing/saturating after 256 units. Considering this, the number of layers was chosen as one and the number of units to be 256. The model was then trained using these architectural choices.

The final optimized model yielded an across subject mean R-value of 0.7 ± 0.06, R2 score of 0.48 ± 0.1 on the test set (RTC - random timing alone), and a mean R-value of 0.64 ± 0.03, an R2 score of 0.41 ± 0.05 on the RTTC test set (random timing + type). Participant 6 had the highest decoding performance with an R-value of 0.85 ± 0.06 and an R2 value of 0.7 ± 0.4 on the test set. Participant 1 had the lowest decoding performance with an R-value of 0.45 ± 0.08 and an R2 value of 0.13± 0.13 on the test set. Fig 7b. shows the continuous sample-by-sample decoder results corresponding to the best fold from the worst-performing participant (S1). Fig 7c. shows the continuous point-by-point decoder results corresponding to the best fold from the best performing participant (S6).

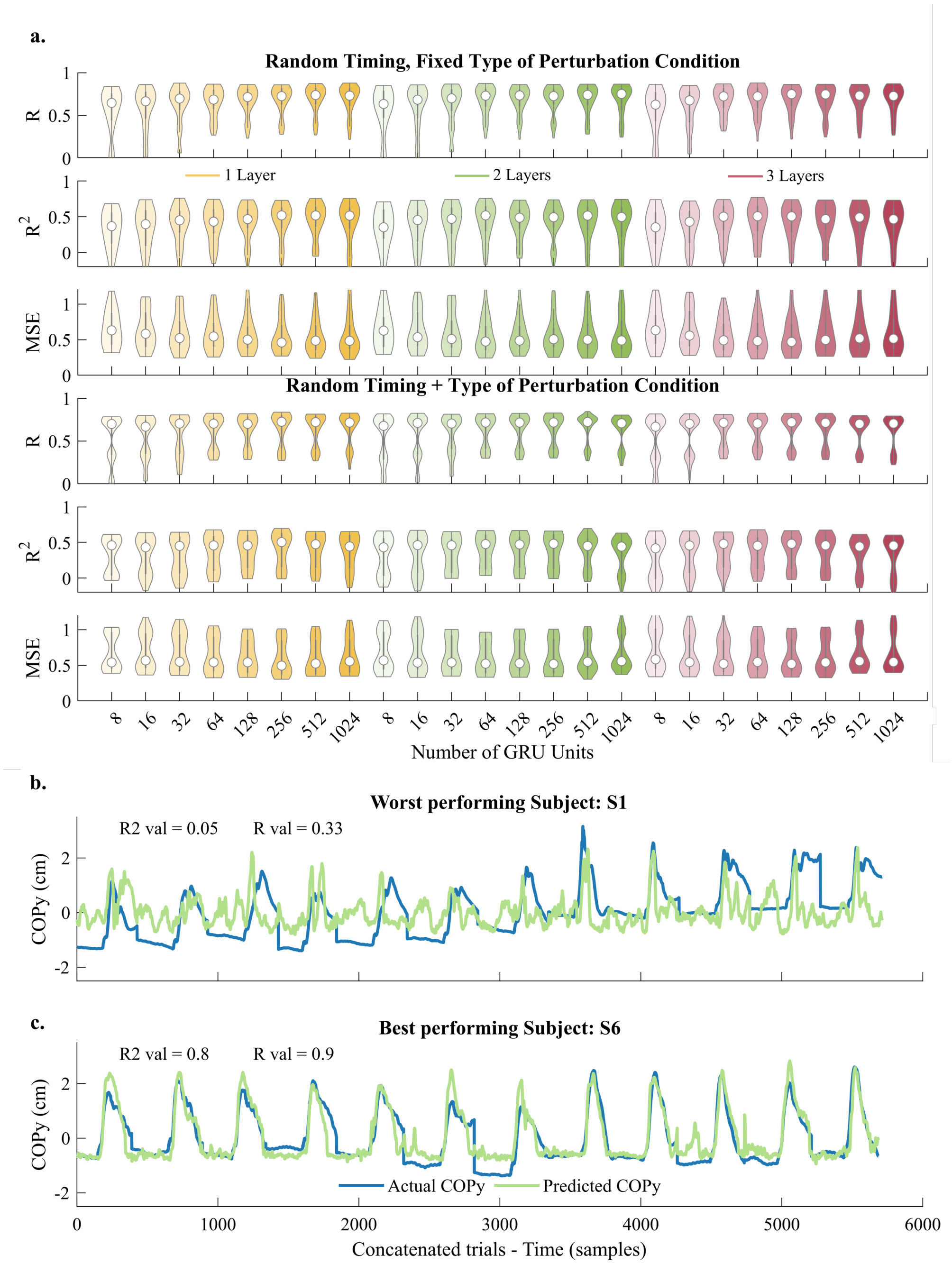


Figure 7: a) Performance measures evaluated on the validation set for varying hyperparameters for the GRU architectures. Each row corresponds to different evaluation metrics; b) decoded COP from the best performing subject (test set, RTC condition); c)decoded COP from the lowest-performing subject (test set, RTC condition).

4. Discussion

This study investigates whether the PEP components would be preserved when a user wears an unpowered exoskeleton. It was found that all the components of the PEP were preserved and that the latency of the P1 and N1 wave preceded that of EMG and kinematic response peaks. This suggests the P1 and N1 components are a viable signal for fall prediction and prevention in exoskeletons. Fall detection in exoskeleton systems is limited and latencies are often too long to be pragmatic in real-world applications. A system detailed in Monaco et al. [13] identified balance perturbation while walking at 350 ms based on hip joint angles. It was also observed that the kinematic response from balance perturbation (while standing) peaked at approximately 350 ms. Comparatively, muscular activity peaked earlier than the COP. Also, PEP components appear as early as 75-137 ms in response to the perturbations. This provides us with a much longer window to perform actions to prevent/reduce fall-related injuries than relying exclusively on temporal kinematic features of the perturbation response.

In a recent review, Varghese et al. suggest that P1 is the earliest non-specific cortical response to a perturbation [15]. They argue that the P1 is not related to the context of the balance perturbation task, and does not contain information related to the predictability of the perturbation or whether the perturbations are internally or externally induced. It is the earliest exogenous cortical response driven by the somatosensory input typically in the range of 0.2-12.7 μV [15]. Compared to P1, N1 is a significantly larger component distributed across the central, frontal, and parietal channels at a latency of 100-150 ms [32],[33]. Prior studies have reported the N1 peak to be as high as 60 μV, localizing in the Cz or FCz channels [32]. Unlike P1, N1 potential has been shown to not just be influenced by afferent signals. Instead, it is also influenced by the predictability and difficulty of the balance task, [34],[35] as well as the presence of competing cognitive tasks [36]. This suggests a higher-order cognitive processing role [33]. Typically, EEG data are trial-averaged to improve signal-to-noise ratio from event-related potentials. After confirming that PEP components were preserved while wearing the exoskeleton, it was determined that perturbations can also be detected from single-trial EEG. This is a crucially important step towards the real-time detection of perturbations. In real-world applications, decoding must occur in real-time. Studies decoding PEP components from single trials are rare and only find one study examining the feasibility was identified [16]. However, that particular study was conducted in a seated condition with a whole-body perturbation and did not examine standing or the use of an exoskeleton. No previous studies that target decoding PEP components from single trials in neither standing nor with an exoskeletal suit were found.

Initially, a CNN model was used to check if the presence of balance perturbation could be detected from single trials. The architecture of the CNN-based decoder was selected considering the usability of the GradCAM approach. GradCAM was chosen specifically because many of the other saliency methods were mentioned to be unreliable and GradCAM was known to be one of the most robust model explanation methods [37]. One point to note is that ideally GradCAM should be applied to the final layer. Here, since we are interested in channel relevancy we use the layer before the last convolutional layer. The performance was compared w.r.t the DeepConvNet [25] model. However, this study does not claim the superiority of the used model architecture or the decoder. The optimization of hyperparameters for both models was not performed, as that is outside the scope of this study. Here, the evaluation is done to confirm the existence of predictive power to detect balance perturbation on a single trial basis and further ensure the model architecture used for prioritizing explainability is comparable to existing architectures. Subject S4 had the lowest decoding performance. During the experiment, this participant reported having congenital nystagmus. There exist a possibility that the PEP might have been corrupted by sudden eye movements and gotten removed during the pre-processing or there may be a difference in the PEP response either of which could cause a reduction in decoding performance.

It was also demonstrated that the CNN model used to detect perturbations was primarily driven by PEP components and not by artifacts. Unlike prior studies that reported few hand-selected examples to demonstrate model explanation, a clustering approach was employed in this study to visualize the model explanation of all the windows from the test and validation set. Model explanations in deep learning studies on EEG are rare. There are only a few ($\sim$ 1.5%) studies that explore the interpretability or explainability of the model used [38]. It is very important to assess whether the outputs of deep learning models are driven by artifacts or PEP signals. This is even more critical considering that the majority of the published studies using deep learning methods currently do not handle the artifacts. A recent review by Roy et al. reported that only 23% of studies performed artifact handling [38]. A similar review by Craik et al. [39] reports 63 % of studies did not preprocess the EEG for classification tasks. As seen from this study, even though the prediction score is higher when using the model trained on EEG without ICA cleaning, many of the decisions were driven by artifacts.

Examining the outputs gave further confidence that the artifact handling pipeline was successful. When the model was trained on data that was not pre-processed, it was biased by artifacts as shown in figure 5b. The CNN started learning from bundle artifacts (C3, C6, C8, C13) and also emphasized peripheral channels more prominently (C5, C8, C10, C12). However, these were not present when the pre-processed data was used to train the model. The study thus highlights the need for providing model explanations in deep learning studies involving EEG, as context is important to assess the main factors behind different decisions. This study also shows how using the data-driven approach coupled with model explanations can help reduce the number of channels required for the decoder. Here, the number of channels was reduced from 60 to 8 without compromising the decoding accuracy.

In addition, from the model explanations shown in 5a., it was observed that depending on the position of the window being considered relative to the perturbation onset (middle row), different channel combinations become most relevant. Channels in the parietal, and occipital regions were the most relevant in the earliest and the latter part of the perturbation onset (C1 and C6). Between 100-300 ms, the model shifted relevance to motor channels (C3, C5, C8). From 200-300 ms, the model was prioritizing the parietal and fronto-motor channels (C1, C2, C4, C7, C10, C11). This suggests the dynamic recruitment of different brain regions in response to the balance perturbation. The model explanations are in agreement with prior works that demonstrated the significance for these regions in balance perturbation tasks [40],[15],[41]. Further exploring these dynamics by computing a measure of dynamic functional connectivity (PDD) similar effects were observed. Specifically, the nodal connectivity was higher in the occipital-parietal region in the early stage of the perturbations, shifting to the motor, then to frontal, and back to the parietal channels.

Given these dynamics, it was expected that the EEG would have the information to be able to continuously decode the instantaneous COP variation. This was validated using a GRU model to decode continuous COP responses from single-trial EEG. It was demonstrated that the GRU model was able to decode, on a sample-by-sample basis, the COP variability from EEG alone for all participants. Evaluating the hyperparameters, it was observed that the number of layers did not contribute significantly to the model performance, which is in agreement with prior work [30]. However, the number of hidden layer units does impact the model significantly. This effect appears most noticeably in the variance across different folds. Comparing the three metrics, a U-shaped relationship was observed between the number of units and the decoding measures, with the performance peaking at 128 or 256 units. The variance was higher with a smaller number of units, suggesting lower predictive power in small models yields poor performance on out-of-distribution data. The variance again increased for large values of hidden units, mostly indicating the tendency towards overfitting to the training data. The selection of an appropriate number of hidden units per layer seemed to be the most critical model hyperparameter. Additional tests were conducted on similar types of trials (i.e. the backward perturbations described above) which were randomly introduced in between variable types of perturbations that included toes up, toes down and forward translations. There was a slight reduction in performance in this condition potentially resulting from additional cognitive processes required to anticipate both the timing and the type of perturbations. The decoding score across all participants exhibited good performance (R-value greater than 0.5), except for participant 1. Participant 1 consistently opted for a specific, non-stereotypical strategy to counteract the perturbation. However, it was noticed that the strategy used by this participant was not working effectively as the participant had the greatest difficulty restoring postural equilibrium. It is possible that the strategy chosen by this participant conflicted with the variable nature of the perturbation, and led to poor decoding.

In summary, relevant components in PEPs were detected as early as $\sim$ 75-137 ms after the onset of a mechanical external perturbation. These components preceded both the peak in EMG activity ($\sim$ 180 ms) and the COP data ($\sim$ 350 ms). It was observed that the perturbations could be decoded from single-trial EEG using a CNN model. Also, it has been demonstrated that the model was driven primarily by relevant components in the PEP to infer the predictions and not by artifacts. The model explanations further aligned with the dynamic functional connectivity measure estimated using PDD. Moreover, the feasibility of decoding continuous COP values from the EEG using a GRU model was established. Overall, the findings suggest that the EEG signals contain short-latency neural information related to an incoming fall, which may be useful for developing brain-machine interface (BMI) systems for fall prevention in neurally-controlled robotic exoskeletons.

5. Acknowledgments

This project was sponsored by the NSF IUCRC Building Reliable Advances and Innovation in Neurotechnology (BRAIN) Center Award (NSF Award 1650536). This work was completed in part with resources provided by the Research Computing Data Core at the University of Houston. In addition, we would like to thank Dr. Manuel Cestari for assisting with the data collection and preparing the IRB documents. He also contributed towards writing codes to automate data curation process. We would also like to thank Nina Dorfner, Peijun Zhao, and Chase Philip for their contributions to the project.

6. Author contributions statement

J.L.C-V, G.E.F., and C.L. conceived the experiment(s), C.L. and J.L.C-V designed the experiment, A.S.R., C.M., I.J. conducted the experiment, A.S.R analyzed the EEG data sets supervised by J.L.C-V. C.M. wrote the code for analyzing the EMG and COP data. All authors reviewed the manuscript.

References

[1] Organization WH. Falls: fact sheet;. Available at https://www.who.int/news-room/fact-sheets/detail/falls.

[2] He Y, Eguren D, Azorín JM, Grossman RG, Luu TP, Contreras-Vidal JL. Brain–machine interfaces for controlling lower-limb powered robotic systems. Journal of neural engineering. 2018;15(2):021004.

[3] Pinto-Fernandez D, Torricelli D, del Carmen Sanchez-Villamanan M, Aller F, Mombaur K, Conti R, et al. Performance evaluation of lower limb exoskeletons: a systematic review. IEEE Transactions on Neural Systems and Rehabilitation Engineering. 2020;28(7):1573-83.

[4] Shi D, Zhang W, Zhang W, Ding X. A review on lower limb rehabilitation exoskeleton robots. Chinese Journal of Mechanical Engineering. 2019;32(1):1-11.

[5] Hong YW, King Y, Yeo W, Ting C, Chuah Y, Lee J, et al. Lower extremity exoskeleton: review and challenges surrounding the technology and its role in rehabilitation of lower limbs. Australian Journal of Basic and Applied Sciences. 2013;7(7):520-4.

[6] Contreras-Vidal JL, Bhagat NA, Brantley J, Cruz-Garza JG, He Y, Manley Q, et al. Powered exoskeletons for bipedal locomotion after spinal cord injury. Journal of neural engineering. 2016;13(3):031001.

[7] Rodríguez-Fernández A, Lobo-Prat J, Font-Llagunes JM. Systematic review on wearable lower-limb exoskeletons for gait training in neuromuscular impairments. Journal of neuroengineering and rehabilitation. 2021;18(1):1-21.

[8] He Y, Eguren D, Luu TP, Contreras-Vidal JL. Risk management and regulations for lower limb medical exoskeletons: a review. Medical devices (Auckland, NZ). 2017;10:89.

[9] Wu CH, Mao HF, Hu JS, Wang TY, Tsai YJ, Hsu WL. The effects of gait training using powered lower limb exoskeleton robot on individuals with complete spinal cord injury. Journal of neuroengineering and rehabilitation. 2018;15(1):1-10.

[10] Ringhof S, Patzer I, Beil J, Asfour T, Stein T. Does a passive unilateral lower limb exoskeleton affect human static and dynamic balance control? Frontiers in Sports and Active Living. 2019;1:22.

[11] Steinhilber B, Seibt R, Rieger MA, Luger T. Postural control when using an industrial lower limb exoskeleton: Impact of reaching for a working tool and external perturbation. Human Factors. 2020:0018720820957466.

[12] Khalili M, Borisoff JF, Van der Loos HM. Developing safe fall strategies for lower limb exoskeletons. In: 2017 International Conference on Rehabilitation Robotics (ICORR). IEEE; 2017. p. 314-9.

[13] Monaco V, Tropea P, Aprigliano F, Martelli D, Parri A, Cortese M, et al. An ecologically-controlled exoskeleton can improve balance recovery after slippage. Scientific reports. 2017;7(1):1-10.

[14] Takakusaki K. Functional neuroanatomy for posture and gait control. Journal of movement disorders. 2017;10(1):1.

[15] Varghese JP, McIlroy RE, Barnett-Cowan M. Perturbation-evoked potentials: Significance and application in balance control research. Neuroscience & Biobehavioral Reviews. 2017;83:267-80.

[16] Ditz JC, Schwarz A, Müller-Putz GR. Perturbation-evoked potentials can be classified from single-trial EEG. Journal of neural engineering. 2020;17(3):036008.

[17] Tanner D, Morgan-Short K, Luck SJ. How inappropriate high-pass filters can produce artifactual effects and incorrect conclusions in ERP studies of language and cognition. Psychophysiology. 2015;52(8):997-1009.

[18] Kilicarslan A, Grossman RG, Contreras-Vidal JL. A robust adaptive denoising framework for real-time artifact removal in scalp EEG measurements. Journal of neural engineering. 2016;13(2):026013.

[19] Mullen TR, Kothe CA, Chi YM, Ojeda A, Kerth T, Makeig S, et al. Real-time neuroimaging and cognitive monitoring using wearable dry EEG. IEEE Transactions on Biomedical Engineering. 2015;62(11):2553-67.

[20] Chang CY, Hsu SH, Pion-Tonachini L, Jung TP. Evaluation of artifact subspace reconstruction for automatic EEG artifact removal. In: 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE; 2018. p. 1242-5.

[21] Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of neuroscience methods. 2004;134(1):9-21.

[22] Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv preprint arXiv:14126980. 2014.

[23] Chollet F, et al.. Keras; 2015. https://keras.io.

[24] Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, et al.. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems; 2015. Software available from tensorflow.org. Available from: https://www.tensorflow.org/.

[25] Schirrmeister RT, Springenberg JT, Fiederer LDJ, Glasstetter M, Eggensperger K, Tangermann M, et al. Deep learning with convolutional neural networks for EEG decoding and visualization. Human brain mapping. 2017;38(11):5391-420.

[26] Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-cam: Visual explanations from deep networks via gradient-based localization. In: Proceedings of the IEEE international conference on computer vision; 2017. p. 618-26.

[27] Satopaa V, Albrecht J, Irwin D, Raghavan B. Finding a" kneedle" in a haystack: Detecting knee points in system behavior. In: 2011 31st international conference on distributed computing systems workshops. IEEE; 2011. p. 166-71.

[28] Breakspear M, Williams LM, Stam CJ. A novel method for the topographic analysis of neural activity reveals formation and dissolution of 'dynamic cell assemblies'. Journal of computational neuroscience. 2004;16(1):49-68.

[29] Sauseng P, Klimesch W, Gruber WR, Hanslmayr S, Freunberger R, Doppelmayr M. Are event-related potential components generated by phase resetting of brain oscillations? A critical

discussion. Neuroscience. 2007;146(4):1435-44.

[30] Nakagome S, Luu TP, He Y, Ravindran AS, Contreras-Vidal JL. An empirical comparison of neural networks and machine learning algorithms for EEG gait decoding. Scientific reports. 2020;10(1):1-17.

[31] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine learning in Python. the Journal of machine Learning research. 2011;12:2825-30.

[32] Marlin A, Mochizuki G, Staines WR, McIlroy WE. Localizing evoked cortical activity associated with balance reactions: does the anterior cingulate play a role? Journal of neurophysiology. 2014;111(12):2634-43.

[33] Mierau A, Hülsdünker T, Strüder HK. Changes in cortical activity associated with adaptive behavior during repeated balance perturbation of unpredictable timing. Frontiers in Behavioral Neuroscience. 2015;9:272.

[34] Payne AM, Hajcak G, Ting LH. Dissociation of muscle and cortical response scaling to balance perturbation acceleration. Journal of neurophysiology. 2019;121(3):867-80.

[35] Goel R, Ozdemir RA, Nakagome S, Contreras-Vidal JL, Paloski WH, Parikh PJ. Effects of speed and direction of perturbation on electroencephalographic and balance responses. Experimental brain research. 2018;236(7):2073-83.

[36] Wittenberg E, Thompson J, Nam CS, Franz JR. Neuroimaging of human balance control: a systematic review. Frontiers in human neuroscience. 2017;11:170.

[37] Adebayo J, Gilmer J, Muelly M, Goodfellow I, Hardt M, Kim B. Sanity checks for saliency maps. arXiv preprint arXiv:181003292. 2018.

[38] Roy Y, Banville H, Albuquerque I, Gramfort A, Falk TH, Faubert J. Deep learning-based electroencephalography analysis: a systematic review. Journal of neural engineering. 2019;16(5):051001.

[39] Craik A, He Y, Contreras-Vidal JL. Deep learning for electroencephalogram (EEG) classification tasks: a review. Journal of neural engineering. 2019;16(3):031001.

[40] Varghese JP, Marlin A, Beyer KB, Staines WR, Mochizuki G, McIlroy WE. Frequency characteristics of cortical activity associated with perturbations to upright stability. Neuroscience letters. 2014;578:33-8.

[41] Goel R, Nakagome S, Rao N, Paloski WH, Contreras-Vidal JL, Parikh PJ. Fronto-parietal brain areas contribute to the online control of posture during a continuous balance task. Neuroscience. 2019;413:135-53.