Biomedical Signal Processing and Control 68 (2021) 102746

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Localizing focal brain injury via EEG spectral variance

Sina Khanmohammadi [a,b,*], Osvaldo Laurido-Soto [b], Lawrence N. Eisenman [b], Terrance T. Kummer [b], ShiNung Ching [a,c,d]

[a] Department of Electrical & Systems Engineering, Washington University in St. Louis, St. Louis, MO 63130, USA
[b] Department of Neurology, Washington University School of Medicine, St. Louis, MO 63110, USA
[c] Department of Biomedical Engineering, Washington University in St. Louis, St. Louis, MO 63130, USA
[d] Division of Biology and Biomedical Science, Washington University in St. Louis, St. Louis, MO 63130, USA

ARTICLE INFO

Keywords:
Brain injury
Injury location
Electroencephalography
Spectral variance

ABSTRACT

In this study, we consider the problem of localizing focal brain injuries from surface electroencephalogram (EEG) recordings. To this end, we introduce a new analysis technique termed frequency-based intrinsic network dynamic reactivity (FINDR), which quantifies the extent to which different brain regions (defined in EEG channel space) are responsive to each other in terms of their frequency-domain activity. The technique generalizes the idea of EEG reactivity, a measure of how well EEG signals react/respond to exogenous stimuli. In the present work we generalize this notion to endogenous 'stimuli,' defined as short-time window frequency domain motifs that are most predominant on a per channel basis. For each of these predominant motifs, we quantify the variance of the activity in all other channels as a measure of 'intrinsic reactivity', under the hypothesis that channels proximal to injured regions will be systematically disassociated from other brain areas. We use this method as a front-end to a neural network classifier to predict injury location in a cohort of etiologically heterogeneous comatose patients. We achieve a 0.6 correlation between the predicted injury location and the actual brain injury. These results suggest a possibility of precise localization of brain injury using EEG.

1. Introduction

Brain injury is a leading cause of death, long-term disability, and cognitive impairment [1–3]. Because the cause of brain injuries can be quite varied, diagnosis, including mapping the spatial extent of injury, is a paramount clinical issue. Among various neuroimaging techniques used for diagnosis and prognosis of brain injury, the electroencephalogram (EEG) has the advantage of high temporal resolution, lower cost, and ease of use in clinical settings. The cost factor is especially relevant for critical care settings in smaller hospitals and developing countries where access to expensive imaging equipment such as computerized tomography (CT) or magnetic resonance imaging (MRI) is limited. In these settings, having access to a low-cost and easily deployable diagnostic technology could greatly enhance standard of care. Relatedly, for certain patients, including those who suffer from severe brain injury, it is usually preferable to minimize the transportation of patients from the ICU setting. In these situations, having bedside EEG-based diagnostic could offer considerable advantages relative to neuroimaging that would require moving the patient to a specialized facility. However, because EEG is recorded at the scalp, it offers only modest spatial resolution, which makes it challenging to use this recording modality for mapping the spatial details of brain injury.

Our central hypothesis is that deploying new advancements in signal processing and classification can maximize the explanatory power of this modality, despite the physical limitations of EEG. In so doing, EEG may provide a useful and advantageous means of assaying the spatial extent of injury, a key step in effective treatment and diagnosis [4]. In particular, we posit that EEG-based diagnostics may be able to detect aberrations that are less clearly delineated in structural imaging. Indeed, certain injuries – especially less severe ones – may be harder to detect via structural imaging, but could potentially lead to electrophysiological alterations that are detectable by means of EEG [5,6], given their finer temporal resolution. The overall goal of our study is to evaluate the possibility of new EEG processing techniques to delineate spatial information about brain injury. In order to promote validity testing, we focus here on injuries where confirmatory neuroimaging is available, though

* Corresponding author at: Department of Electrical & Systems Engineering, Washington University in St. Louis, St. Louis, MO 63130, USA.
E-mail addresses: s.khanmohammadi@wustl.edu (S. Khanmohammadi), ojlaurido-soto@wustl.edu (O. Laurido-Soto), leisenman@wustl.edu (L.N. Eisenman), kummert@wustl.edu (T.T. Kummer), shinung@wustl.edu (S. Ching).

https://doi.org/10.1016/j.bspc.2021.102746
Received 27 May 2020; Received in revised form 10 April 2021; Accepted 7 May 2021
Available online 25 May 2021

our long-term goal is to develop this technology as a complement to such imaging.

Previous attempts to extract spatial properties of brain lesions from EEG signals have used the concept of hemispherically asymmetric power between EEG channels. The asymmetry index [7] relies on the hypothesis that the relative difference of average power spectral density of the left–right hemispheres in the healthy resting state should be close to zero, but that this symmetry gets disrupted due to abnormalities such as brain injury. For example, Yi and colleagues proposed symmetrical channel EEG signal analysis (SESA) that aims to map the location of brain injury using spectral density values of left versus right hemispheres [8]. A challenge with such methods is that they require a predefined threshold regarding how much asymmetry is abnormal.

An alternative approach to infer spatial information from EEG is to estimate the current density of sources that generate measured electrical activity at the scalp [9]. This is known as the inverse localization problem, which involves modeling the observed measurements as a linear combination of underlying dipolar time series and then estimating the inverse solution matrix (spatial filters) from the data. This inverse problem is mathematically ill-posed and has an infinite number of solutions. Therefore, an *a priori* constraint such as the number of sources is typically imposed to limit the solution space, allowing for the deployment of numerical inference techniques [10,11]. These methods have been previously used to localize post-traumatic epilepsy (PTE) in acute TBI patients [12,13]. However, they involve secondary magnetic resonance imaging (MRI) to determine the a priori constraints needed for solving the inverse problem. This step mitigates the ease-of-deployment advantage of EEG alone.

In this paper, we introduce a different approach to identify the location of brain injury from EEG that is based on our previous work [14]. In that prior work, we developed a method called the Intrinsic Network Reactivity Index (INRI) that aims to capture the traditional notion of EEG reactivity [15] without relying on exogenous stimulation. More specifically, INRI measures how rare electrophysiological events (defined channel-wise) impact brain electrical activity in both proximal and distal channels. We have shown that INRI is correlated with the severity of injury (in terms of the Glasgow coma scale). Here, we introduce a variation on INRI that we call frequency-based intrinsic network dynamic reactivity (FINDR) that is developed under the premise that a lesioned area of the brain is limited in its responsiveness. Succinctly, FINDR seeks to capture whether most dominant spectral events in a channel are systematically associated with those of other channels, thus suggesting whether that channel is (or is not) engaged with other brain regions. We formally define this technique and show that it indeed is informative with respect to brain injury location by using it as a front-end feature to a neural network classifier that we evaluate using a formal validation experiment.

2. Materials and methods

2.1. Data description and preprocessing

The retrospective EEG data, including complete medical records, were collected for 15 comatose patients with focal brain injury during three years (2013–2016). The patients underwent EEG recording for routine monitoring purposes in the Neurological and Neurosurgical Intensive Care Unit (NNICU) at Barnes-Jewish Hospital, which is affiliated with the Washington University School of Medicine in St. Louis. Trained neurointensivists administered the clinical ratings and Glasgow coma scale evaluation. The study was approved by Washington University Institutional Review Board. Table 1 provides a summary of the patient information including age, gender, and injury location.

Head CT or MRI images were examined to identify dominant focal and supratentorial lesions. The patients history and clinical images were examined and only patients that had one clearly dominant lesion (radiologically-presumed to be the etiological cause of the coma) with no other confounders were included (confounders that might mask/make the EEG not reliable, like sedation, hx suggestive of seizures). Exclusion criteria were multiple lesions of similar size, injuries expected to result in diffuse pathology based on clinical/radiological characteristic, and severe slowing of EEG signals marked by epileptologists indicative of severe global dysfunction.

Table 1
Summary of the study population.

Variable	Distribution
Age	62 ± 17
Gender	Male (7) and female (8)
Injury location	Left (8), right (7)
Injury type[a]	Ischemic stroke (6), intraparenchymal hemorrhage (8), subdural hematoma (2), aneurysmal subarachnoid hemorrhage (1), traumatic brain injury (1), brain tumor (1)

[a] Some patients had more than one type of injury.

Lesion laterality and the lobar location were defined for each patient. The size of the lesion was measured using the *abc/2* score [16] and the radius calculated as the average of four images. The coordinates in 3D space were determined with the center of the brain serving as a reference. For the *X*-axis, the radiologic left was defined as a positive and radiologic right was defined as negative. For *Y*-axis, anterior was defined as positive, and posterior was defined as negative. The *Z* coordinate was normalized as a distance measured from the vertex. A 2D topographical map was made from this information, which was then matched to the images.

The retrospective data was recorded using 19 electrodes positioned according to the standard 10–20 system of electrode placement and included at least 15 minutes of recordings for each patient. The signals were recorded at either 250 or 500 Hz against a common reference electrode and re-referenced to 18 bipolar channels (FP1-F7, F7-T7, T7-P7, P7-O1, Fp1-F3, F3-C3, C3-P3, P3-O1, Fz-Cz, Cz-Pz, Fp2-F4, F4-C4, C4-P4, P4-O2, Fp2-F8, F8-T8, T8-P8, and P8-O2) to reduce the sensitivity of EEG signals to external noise [17].

EEG preprocessing steps included manual artifact removal by visual inspection, downsampling, and bandpass filtering. More specifically, we looked for eye movement, muscle movement, and electrode artifacts patterns according to reference patterns in [18] and then removed sections of the data that was contaminated with these artifacts. We then downsampled the data that was initially recorded at 500 Hz to 250 Hz, in order to make the sampling rate of all the recordings consistent. Next, as a first step of the proposed framework we calculated the power spectrum of EEG signals in the 1–35 Hz frequency range using a three second long Hanning window with no overlap. The Hanning window was used to take care of discontinuities in the signal introduced by artifact removal, and the non-overlapping sliding windows were used to avoid clustering of events in the first step of the proposed framework. We set the lower cutoff frequency to 1 Hz to avoid potential muscle movement and skin artifacts [19,20]. It should be noted that the proposed approach does not directly use statistics derived from the power spectrum and hence is robust to power spectrum parameter settings such as the window size as shown in Section 3.3 [18,21].

2.2. Localizing injury using FINDR features

The FINDR framework for injury localization involves four main steps: (i) identifying intrinsic spectral events, (ii) obtaining proximal and distal 'responses' to these identified events, (iii) quantifying the divergence of these responses in channel-space, and (iv) feeding the divergence measure into a classifier to determine the final spatial location of injury (Fig. 1).

2.2.1. Intrinsic event identification

The first step of the FINDR calculation involves identifying the events

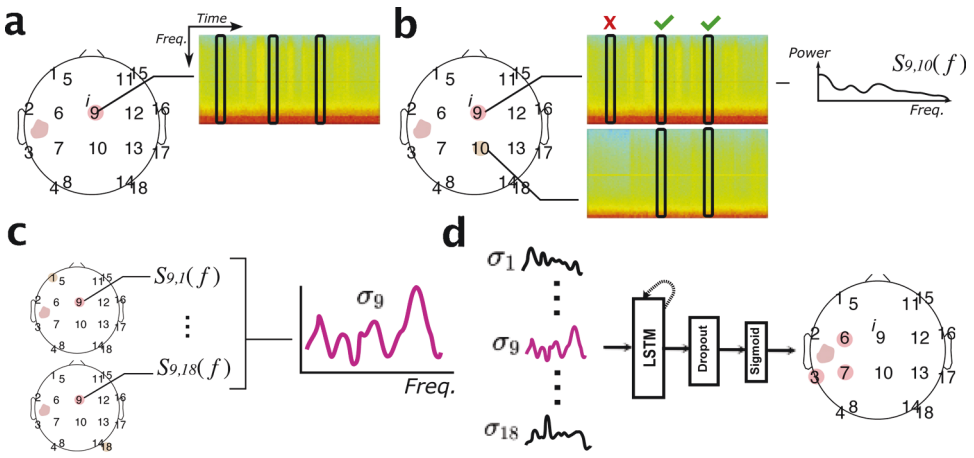

Fig. 1. Frequency-based intrinsic network dynamic reactivity (FINDR). (a) For a given seed channel i (here, $i = 9$), we obtain a time-frequency spectrogram. We then identify 'events' as being those temporal windows whose spectra explain the most variance (over all windows). (b) The procedure for identifying windows is repeated for a non-seed channel j (here, $j = 10$). The average spectrum $S_{i,j}(f)$ is then obtained by averaging all spectra from channel i from event windows that are common to channels i, j. (c) Step (b) is repeated for all non-seed channels, resulting in a set of average spectra $S_{i,1}$, $S_{i,2}$, etc. We then obtain the *spectral variance*, $\sigma_i(f)$, which is the variance across these spectra. This quantity summarizes the heterogeneity in the frequency content at the seed channel during the significant events elsewhere in the EEG montage. (d) Procedure (a)–(c) are iterated so that each channel is treated as a seed, resulting in $\sigma_1(f)$, …, $\sigma_{18}(f)$. These spectral variances are fed into a long short-term memory (LSTM) recurrent neural network to classify each channel as injury or no injury. The classification step is expanded in Fig. 2.

in each channel that are most significant insofar as they account for the most variance of spectral content across time. For a given channel, we calculate the spectral density over sliding windows (Hanning window of length three seconds with no overlap). Then, we apply principal component analysis (PCA) to identify the component that explains the greatest variance. This component is a linear combination of constituent windowed spectra. We then pick the 100 windows with the highest magnitude weights contributing to the first principal component. These selected windows have the most influence on power spectral variation (across windows) and therefore we consider them as the dominant intrinsic windows, or events (Fig. 1 Part (a)).

2.2.2. Calculation of spectral 'responses'

The next step of the FINDR framework is to ascertain the activity in other EEG channels during these intrinsic events. This step begins by performing the above event identification for each channel in the EEG montage. We then use an iterative process to 'scan' the montage to look at how events in one channel are associated with events in other channels. We first fix a seed channel i, then determine the events in a different channel j that temporally overlap with those of the seed. We define the 'response spectrum' $S_{ij}(f)$ as the average spectrum of seed channel i over these common windows (Fig. 1 Part (b)). Hence, the response signal $S_{ij}(f)$ captures the spectral content in the seed channel during the dominant events of all other channels $j = 1, \ldots, 18$. This process is iterated over seed channels until the response value is calculated for each channel pair i, j.

2.2.3. Variance of spectral responses

After obtaining $S_{ij}(f)$ for each i, j, we proceed to calculate the variance of these spectra across the j dimension. That is, for each seed channel, we want to understand how its responses vary relative to the dominant events across the rest of the EEG montage. In order to focus on only the dominant axes of variation, we first use PCA to identify the three most significant spectral components, denoted $\widehat{S}_{i,1}(f)$, $\widehat{S}_{i,2}(f)$ and $\widehat{S}_{i,3}(f)$, respectively. We then calculate the variance as a function of frequency across these components, i.e.,

$$\sigma_i(f) = \frac{1}{3}\sum_{z=1}^{3}\left(\widehat{S}_{i,z}(f) - \widehat{\mu}(f)\right)^2, \tag{1}$$

where $\widehat{\mu}(f)$ denotes the mean of the three component spectra.

Thus, $\sigma_i(f)$ summarizes the heterogeneity present in the power spectrum of channel i at times when other channels are producing likewise significant activity. A key premise of this work is that such heterogeneity can inform the extent to which a given channel is (or is not) associated with others, and therefore, act as a surrogate for the extent of injury proximal to that channel.

2.2.4. Spatial classification

To build on the above premise, we feed the spectral variance $\sigma_i(f)$ into a neural network classification architecture (Fig. 2). We specifically chose a long-short term memory (LSTM) classifier design which processes the matrix

$$\Gamma_f = [\sigma_1(f), \sigma_2(f), \ldots, \sigma_{18}(f)]^T \tag{2}$$

sequentially along the f dimension. Using an LSTM in this way, while nonstandard, allows for the classifier to build associations across both space and frequency by leveraging the memory (wherein the classifier retains information about each sequentially fed frequency components). More specifically, for each patient, we have $N \times M$ matrix Γ_f where N is the number of channels and M is the maximum sample frequency. Each element of Γ_f represents the spectral variance at channel i in frequency f calculated using the FINDR framework. Next we have an N dimensional binary vector $\Theta = [\theta_1, \ldots, \theta_i]$ that represents the location of the brain injury inferred from the MRI/CT scans. Each element of vector Θ is either zero or one depending on whether it is aligned with the location of the brain injury.

The Γ_f and Θ are then used to train the LSTM network model, which consists of an LSTM layer with 50 hidden units, a 20% Dropout layer, and a dense layer with Softmax activation function. In order to train the model, we used the five-fold cross validation scheme. During each iteration of cross validation, we divided the data into five mutually exclusive groups of recordings where one group was used for testing and the rest for training. Hence, at each iteration of the cross validation we train the LSTM model using a $12 \times N \times M$ tensor of spectral variances (σ) and a $12 \times N$ matrix of class labels (θ). The trained model is then tested using the $3 \times N \times M$ tensor of spectral variances for the remaining subjects.

The output of the LSTM network for each patient is the N dimensional probability vector $P = [p_1, \ldots, p_i]$, where p_i represents the probability of channel i belonging to class label one (injury). Since we are dealing with a multi-class problem and to avoid thresholding we assess the performance of this classifier by calculating the correlation (r) between the predicted probability vector P and the ground truth class labels Θ. Hence, the correlation value shows how much the predicted injury locations coincide with the actual injury location. This process is

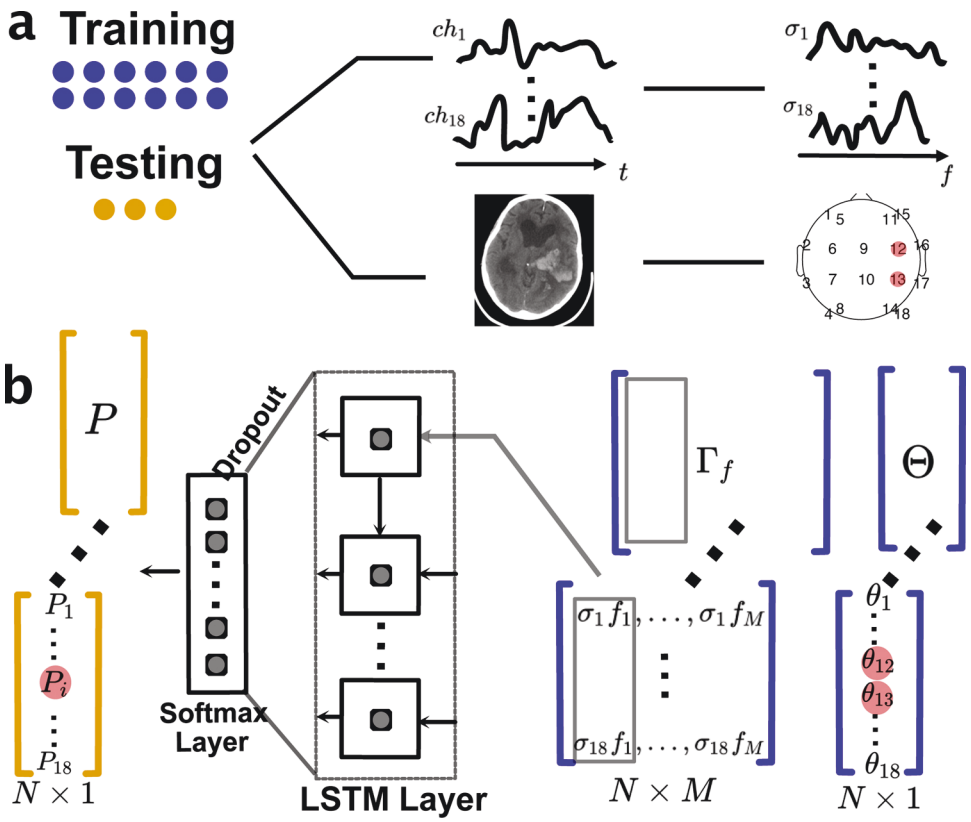

Fig. 2. Predicting location of injury using spectral variance features. (a) For each subject, the spectral response variances $\Gamma_f = [\sigma_1(f), \sigma_2(f), \ldots, \sigma_{18}(f)]$ and the injury location in channel space $\Theta = [\theta_1, \ldots, \theta_{18}]$ are extracted using the EEG signals and MRI/CT scans, respectively. (b) At each iteration of the five-fold cross validation scheme, the LSTM network is trained using a $12 \times N \times M$ tensor of spectral variances extracted from 12 training subjects and then tested using the $3 \times N \times M$ tensor of spectral variances for the remaining three testing subjects. The output of the LSTM network for each patient is an N dimensional vector $P = [p_1, \ldots, p_i]$, where p_i represents the probability that channel i coincides with the injury location.

repeated five times till each group of subjects are tested using the trained model based on data from other groups of patients. The LSTM network was implemented using the TensorFlow backend in a standard Windows 10 PC with Intel Core i7-6700 CPU and 16 Gigabytes of DDR3 RAM.

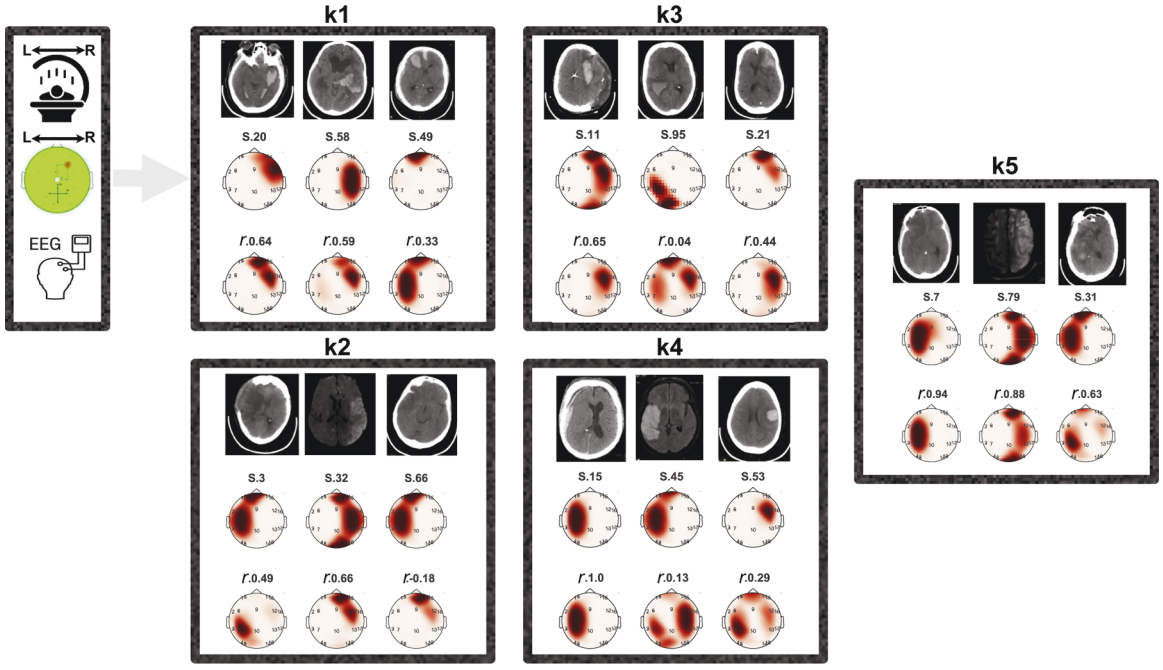

Fig. 3. Correlation between predicted injury locations using FINDR features and the actual brain injury location inferred from MRI/CT scans. Each box represents one fold of the k-fold cross validation scheme. The top row in each box shows the actual MRI/CT scan of the testing subjects, the second row shows the inferred injury location from the corresponding MRI/CT scans, and the third row shows the predicted injury location from trained LSTM networks. For each fold, the LSTM network is tested on the subjects denoted by (S) and trained by the rest of the subjects. The correlation values between the actual brain injury locations in the second row and the predicted injury locations in the third row are denoted by (r).

3. Results

3.1. FINDR spectral variance predicts injury location

Fig. 3 shows the correspondence between the predicted injury location by the FINDR framework and the actual injury location inferred from the MRI/CT scans. Each box represents the testing set in one fold of the five-fold cross validation process. Within each fold, the top row depicts the actual MRI scan from patients, the second row shows a 2D topographical map of injury location inferred from the MRI/CT scans (see Section 2.1), and the final row shows the predicted injury location by the trained LSTM network. For each pair of topographical maps, the subject number (S) and corresponding correlation (r) are presented. In 8/15 test trials we observe $r > 0.5$, with only one case resulting in a negative correlation.

3.2. The results of prediction are not due to random chance

To validate the above finding, we performed a series of control bootstrapping experiments. Overall, these results support our hypothesis that the spectral variance Γ_f obtained through the proposed FINDR method can specifically inform the adjudication of injury location.

3.2.1. The predictions are specific to testing subject

First, we calculated the correlation between the injury location predicted by the FINDR method and that of a different, randomly selected patient within the dataset. That is, we tried to classify the location of injury in patient A by feeding the data from patient B. As shown in Fig. 4, such randomization results is a significant loss of performance. This baseline sanity check establishes that the identified correlations are not due simply to random chance and suggests that the proposed method can indeed inform on injury location.

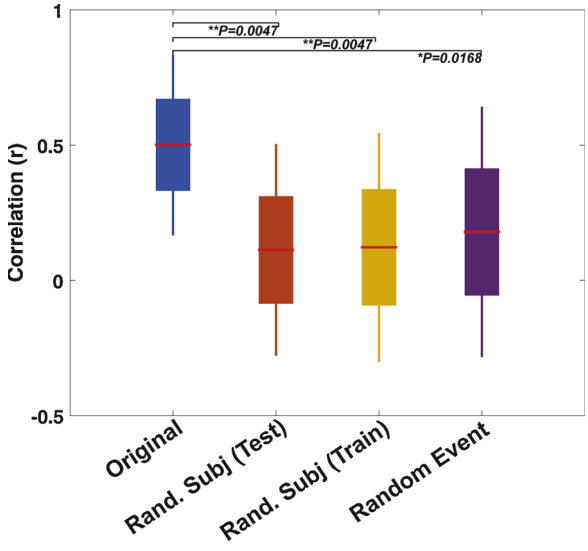

Fig. 4. Bootstrapping analysis of the FINDR framework. The baseline bar plot (Original) shows the distribution of correlation values between the predicted injury location and the actual injury location when the LSTM network is trained using the proposed method. The second bar plot (Random Test Subjects) shows the correlation between the predicted injury locations and a randomly selected patient within the dataset. The third bar plot (Random Train Subjects) shows the correlation between the predicted injury locations and actual injury locations when the LSTM network is trained using random injury locations instead of the actual corresponding injury locations. Finally, the last bar plot (Random Events) shows the correlation between the predicted injury locations and actual injury locations when we use random sections of the data as events instead of the spectral events used in the FINDR framework. p-values are based on a two-sample Kolmogorov–Smirnov test.

3.2.2. FINDR spectral variance is specific to classification performance

Next, to understand the specificity of the FINDR pipeline to the resultant classification performance, we *re-trained* the classifier using random injury locations (injury locations from randomly selected patients). The goal of this test was to assess whether the spectral variance Γ_f of one patient could be associated with the injury location in another patient. As shown in Fig. 4, such re-training results in a lower level of classification accuracy. This is re-assuring as it suggests that the LSTM cannot 'overfit' the injury location based on an incongruent Γ_f.

3.2.3. Event detection is critical to the mechanism of classification

Finally, to delve further into the mechanism of classification, we focused on the step described in Section 2.2. Instead of using the described procedure to identify salient frequency-based events within each EEG channel, we simply picked random windows from which to calculate the subsequent spectral variance. We then re-trained the classifier based on these random events (with the correct patient used for training). As shown in Fig. 4, such randomization and re-training (*random events*) also leads to significant reduction is classification performance. This finding implies that the identification of events through the dimensionality reduction procedure in Section 2.2 is indeed a key step that is meaningful and relevant to injury location.

3.3. FINDR Performance is robust to window size and number of events

Because FINDR relies on the choice of a moving window (to compute short-time power spectra, i.e., in Section 2.2.1), we assessed the sensitivity of the method to window size. As shown in Fig. 5 Part (a), there is no significant difference in correlation performance for window lengths ranging from 1 to 5s, suggesting robustness over this range. Interestingly, increasing the window size did not improve the performance of our framework. This is partly because by increasing the window size we are potentially decreasing the variability across different windows, which ultimately affects the spectral variance of the signal. As decreasing the window size reduces the frequency resolution of the signal, we selected a window size of three in this study as a compromise between frequency resolution and spectral variability. We also assessed the sensitivity of our framework to the number of selected windows (events) as seen in Part (b) of Fig. 5. According to the results, setting the number of windows below 100 degrades the overall performance. This is expected, as lowering the number of windows makes it more difficult to estimate the spectral responses in step 2 of the FINDR framework. Finally, we assessed the sensitivity of FINDR framework to average (CART), average of A1&A2, and single channel Cz re-referencing methods in Part (c) of Fig. 5. Based on the results, both bipolar and CART referencing methods performed well. In comparison, the re-referencing to Cz and Average A1/A2 did not perform as well as the other methods. A plausible explanation for this might be the higher sensitivity of these two approaches to noise, as well as different spatial specificity associated with the referencing schemes. For a more detailed comparison of different EEG re-referencing methods, the readers can refer to [17,22,23].

4. Discussion

In this paper we adapt the classical idea of EEG reactivity (how the EEG reacts to exogenous stimuli) and translate it into a paradigm for injury localization that relies only on passive EEG measurement. The paradigm proceeds by quantifying a 'reaction' of a channel in terms of how heterogeneous its activity is during periods of significant activity in other channels. Here, significant events are defined using a systematic procedure of windowed power spectral density estimation followed by dimensionality reduction. These quantification of spectral heterogeneity are then fed to a machine learning classifier that predicts actual injury locations. In a majority of tested patients, the predicted injury location correlated to a significant level relative to chance and other validation

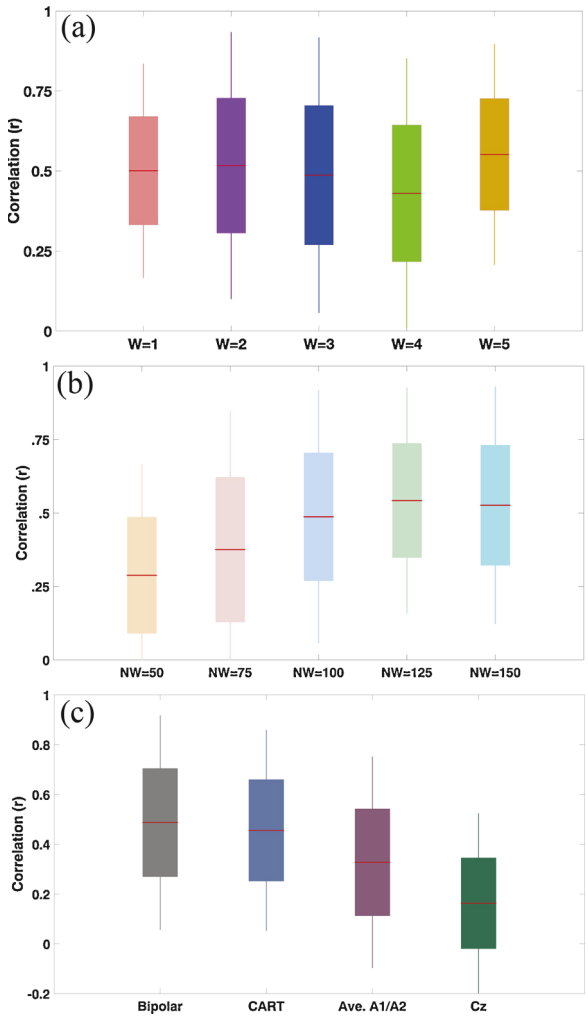

Fig. 5. Sensitivity analysis of the FINDR framework. (a) The correlation between predicted injury location and actual injury location is calculated for different window lengths. (b) The correlation between predicted injury location and actual injury location is calculated for different number of windows (events). (c) The correlation between predicted injury location and actual injury location is calculated for different EEG re-referencing techniques.

tests, indicating that the proposed method indeed carries information regarding injury location.

4.1. Interpretation of FINDR quantification of spectral variance

This study extends our prior work on intrinsic network reactivity (INRI) [14], which discussed how to glean information on injury severity and etiology from EEG recordings. More specifically, The INRI approach was developed to characterize the reactivity of the EEG signal to intrinsic events, thus generalizing the classical notion of EEG reactivity to exogenous stimuli. The underlying hypothesis of the INRI approach was that such "intrinsic reactivity" can be informative with respect to extent of injury in comatose patients. From a technical perspective, the INRI looks holistically at signals across the entirety of the EEG montage relative to the intrinsic events in question. The FINDR approach, on the other hand, is designed to quantify the interaction of different brain regions with each other in a dynamical systems sense, which we hypothesized would be informative with respect to spatial properties (in this case location) of a brain injury. The overarching premise of these approaches involves characterization of the 'lability' of brain regions, i.e., the extent to which they can and do produce diverse patterns of electrical activation. The spectral variance that underlies the FINDR approach, i.e., Γ_f is designed as a surrogate quantity that captures this idea.

It is important to note that the mapping from the spectral variance Γ_f to actual spatial location of injury is not explicit in our framework, since it is embedded within the LSTM network. Thus, we are not able to make statements about whether particular values of Γ_f are more or less attributed to injury location. However, we speculate that the LSTM may in fact be embedding a 'U' shaped function wherein extremal values within Γ_f are associated with injury. Indeed, we hypothesize that this mapping may encapsulate two scenarios. In the first, an injured region is suppressed on account of injury and thus exhibits very low spectral variance relative to other regions. The second scenario is one in which an injured region becomes completely disassociated from other regions leading to high spectral variance. Clearly, the performance achieved by the trained LSTM is not perfect and for several patients the classifier fails to characterize injury location. The heterogeneity of injury etiologies present in our dataset may mean that some patients fall outside of the above premises regarding spectral variance and an important question moving forward will be to understand when the method does or does not work and to analytically characterize the transformation learned by the LSTM classifier.

4.2. Novelty and significance of the proposed approach

There are few attempts in the literature to localize brain injury using the scalp EEG data, but they usually rely on descriptive statistics on the power spectrum (hence susceptible to frequency resolution) or depend on extrinsic stimulation (hence less practical). Even though a major component of FINDR approach is calculating the spectrum of the signal, it is not as sensitive as other methods to the frequency resolution. This is partly because unlike those methods such as asymmetry index [7], the power spectrum here is just a tool to capture the responsiveness of channels, and any resolution deficiency is globally distributed and canceled out during the process. This could be seen in Fig. 5 where the information content of the FINDR values doesn't seem to depend on the length of windows.

As alluded to in the introduction, the capability of our methodology could become even more powerful for cases of mild traumatic brain injury where the actual brain tissue is not damaged and therefore not detectable by structural imaging techniques [24]. In such cases, it is necessary to analyze the functional properties of the brain, which generally involve capturing the brain's response to an external stimuli. These responses can be captured using functional imaging techniques such as fMRI [25], but they suffer from the same limitations as structural imaging techniques. They are time-consuming, costly, and not suitable for rapid deployment. The rapid deployment is especially important for at risk population such as athletes and military service members who might not have immediate access to functional imaging equipment. Hence, EEG is an important asset to study the functional properties of the brain, including disruptions of interactions between different brain regions caused by injury.

Perhaps a more significant advantage of the FINDR method is that it does not depend on any external stimuli and the responsiveness of brain is quantified in terms of intrinsic events. The application of such an approach goes beyond localizing brain injury and includes a wider set of problems that involve understanding the underlying mechanism of the brain in resting state. For example, recently there has been a lot of interest in studying large-scale brain networks called default mode networks which are known to be active during the resting state and relate to various neurological disorders [26]. The FINDR approach is a great candidate to study such networks and understand the underlying mechanism that drive such networks. Finally, unlike source localization methods, the FINDR approach does not require a priori information such as the geometrical model of the head. The only parameters used in this approach are the typical power spectrum parameters including the number of windows, amount of overlap, and the length of each window,

which are easy to optimize.

4.3. Limitations and future directions

Our results establish that information regarding injury location may be embedded in certain dynamical features of the EEG. Nonetheless, the average correlation performance of 0.6 that we achieved clearly needs to be improved to entertain the possibility of deploying this technique in a real world setting. Thus, the current results should be viewed as an initial proof of concept study toward the overall problem of brain injury characterization from surface EEG recordings.

First and foremost, as noted above, while we have shown that the spectral variance is informative with respect to injury location, more is required to study the specific way that the machine learning algorithm uses this quantity. Occlusion methods [27,28] and other emerging techniques for interrogating the mechanisms of supervised classification methods in healthcare applications [29,30] could be useful in this regard.

In addition, there are a number of basic data processing issues that need to be addressed before this method could be translated to a clinical setting. The data preprocessing and artifact rejection steps were performed offline and required manual inspection of the EEG traces. For a fully automated system that can be used in remote locations (the most likely application of this approach), an automated pre-processing pipeline needs to be developed.

Another limitation to the current study is that it only considers injury location in two-dimensional space (x and y-axis), whereas the depth of the injury is also clearly important for diagnosis and prognosis of brain injury.

4.4. Conclusion

In summary, we developed a new method termed frequency-based intrinsic network dynamic reactivity (FINDR) that quantifies the variance of neural response at different brain regions to intrinsic events in other regions. We showed that despite some limitations, this technique could be used to localize brain injury. The proposed method is a passive technique that does not require administration of external stimuli, making it potentially suitable for cases of brain injury where patients do not respond to external stimuli or where there is no immediate access to advanced neuroimaging equipment. The current results contribute to a growing set of techniques for characterizing coma from brain electrical dynamics and provide a foundation for automated and easily deployable EEG-based injury diagnostics.

Authors' contribution

The study was conceived by SK, TTK, and SC. TTK, LE, and OL-S acquired the data. Analysis and statistics were designed and performed by SK and SC. All authors aided in analysis and interpretation of data. SK, LE, TTK, and SC drafted the manuscript.

Acknowledgements

This work is primarily supported by the grant #1R21-NS096590-01A1 from National Institutes of Health (NIH). Other sources of support include National Science Foundation (NSF CMMI 1537015, and NSF ECCS 1509342), National Institutes of Health (NIH 1R21-EY027590-01, NIH UL1 TR000448), and Career Awards at the Scientific Interface (CASI) from Burroughs Wellcome Fund (BWF).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] J. Bruns, W.A. Hauser, The epidemiology of traumatic brain injury: a review, Epilepsia 44 (2003) 2–10.
[2] J.A. Langlois, W. Rutland-Brown, M.M. Wald, The epidemiology and impact of traumatic brain injury: a brief overview, J. Head Trauma Rehabil. 21 (2006) 375–378.
[3] A.I. Maas, N. Stocchetti, R. Bullock, Moderate and severe traumatic brain injury in adults, Lancet Neurol. 7 (2008) 728–741.
[4] C.-L. Chen, F.-T. Tang, H.-C. Chen, C.-Y. Chung, M.-K. Wong, Brain lesion size and location: effects on motor recovery and functional outcome in stroke patients, Arch. Phys. Med. Rehabil. 81 (2000) 447–452.
[5] M.R. Nuwer, D.A. Hovda, L.M. Schrader, P.M. Vespa, Routine and quantitative eeg in mild traumatic brain injury, Clin. Neurophysiol. 116 (2005) 2001–2025.
[6] J.N. Ianof, R. Anghinah, Traumatic brain injury: an eeg point of view, Dementia Neuropsychol. 11 (2017) 3–5.
[7] M.J. van Putten, J.M. Peters, S.M. Mulder, J.A. de Haas, C.M. Bruijninckx, D. L. Tavy, A brain symmetry index (bsi) for online eeg monitoring in carotid endarterectomy, Clin. Neurophysiol. 115 (2004) 1189–1194.
[8] L. Yi, L. Xiao-Ping, L. Xian-Hong, L. Jing-Qi, Y. Wen-Wei, Z. Dan-Ke, L. Li-Hua, Y. Yong, Mapping brain injury with symmetrical-channels' eeg signal analysis – a pilot study, Sci. Rep. 4 (2014), 5023.
[9] C.M. Michel, M.M. Murray, G. Lantz, S. Gonzalez, L. Spinelli, R.G. de Peralta, Eeg source imaging, Clin. Neurophysiol. 115 (2004) 2195–2222.
[10] Z.J. Koles, Trends in eeg source localization, Electroencephalogr. Clin. Neurophysiol. 106 (1998) 127–137.
[11] R. Grech, T. Cassar, J. Muscat, K.P. Camilleri, S.G. Fabri, M. Zervakis, P. Xanthopoulos, V. Sakkalis, B. Vanrumste, Review on solving the inverse problem in eeg source analysis, J. Neuroeng. Rehabil. 5 (2008) 25.
[12] A. Irimia, S.M. Goh, C.M. Torgerson, M.C. Chambers, R. Kikinis, J.D. Van Horn, Forward and inverse electroencephalographic modeling in health and in acute traumatic brain injury, Clin. Neurophysiol. 124 (2013) 2129–2145.
[13] A. Irimia, S.-Y.M. Goh, C.M. Torgerson, N.R. Stein, M.C. Chambers, P.M. Vespa, J. D. Van Horn, Electroencephalographic inverse localization of brain activity in acute traumatic brain injury as a guide to surgery, monitoring and treatment, Clin. Neurol. Neurosurg. 115 (2013) 2159–2165.
[14] S. Khanmohammadi, O. Laurido-Soto, L.N. Eisenman, T.T. Kummer, S. Ching, Intrinsic network reactivity differentiates levels of consciousness in comatose patients, Clin. Neurophysiol. 129 (2018) 2296–2305.
[15] E. Gütling, A. Gonser, H.-G. Imhof, T. Landis, Eeg reactivity in the prognosis of severe head injury, Neurology 45 (1995) 915–918.
[16] R.U. Kothari, T. Brott, J.P. Broderick, W.G. Barsan, L.R. Sauerbeck, M. Zuccarello, J. Khoury, The abcs of measuring intracerebral hemorrhage volumes, Stroke 27 (1996) 1304–1305.
[17] J. Osselton, Acquisition of eeg data by bipolar unipolar and average reference methods: a theoretical comparison, Clin. Neurophysiol. 19 (1965) 527–528.
[18] J.W. Britton, L.C. Frey, J. Hopp, P. Korb, M. Koubeissi, W. Lievens, E. Pestana-Knight, E.L. St, Electroencephalography (EEG): An Introductory Text and Atlas of Normal and Abnormal Findings in Adults, Children, and Infants, American Epilepsy Society, Chicago, 2016.
[19] X. Chen, X. Xu, A. Liu, S. Lee, X. Chen, X. Zhang, M.J. McKeown, Z.J. Wang, Removal of muscle artifacts from the eeg: a review and recommendations, IEEE Sens. J. 19 (2019) 5353–5368.
[20] M.A. Colombo, M. Napolitani, M. Boly, O. Gosseries, S. Casarotto, M. Rosanova, J.-F. Brichant, P. Boveroux, S. Rex, S. Laureys, et al., The spectral exponent of the resting eeg indexes the presence of consciousness during unresponsiveness induced by propofol, xenon, and ketamine, Neuroimage 189 (2019) 631–644.
[21] J.M. Stern, Atlas of EEG Patterns, Lippincott Williams and Wilkins, 2005.
[22] L. Dong, X. Liu, L. Zhao, Y. Lai, D. Gong, T. Liu, D. Yao, A comparative study of different eeg reference choices for event-related potentials extracted by independent component analysis, Front. Neurosci. 13 (2019) 1068.
[23] D. Yao, Y. Qin, S. Hu, L. Dong, M.L.B. Vega, P.A.V. Sosa, Which reference should we use for eeg and erp practice? Brain Topogr. 32 (2019) 530–549.
[24] S.S. Shin, J.W. Bales, C.E. Dixon, M. Hwang, Structural imaging of mild traumatic brain injury may not be enough: overview of functional and metabolic imaging of mild traumatic brain injury, Brain Imaging Behav. 11 (2017) 591–610.
[25] B.C. McDonald, A.J. Saykin, T.W. McAllister, Functional mri of mild traumatic brain injury (mtbi): progress and perspectives from the first decade of studies, Brain Imaging Behav. 6 (2012) 193–207.
[26] R.L. Buckner, L.M. DiNicola, The brain's default network: updated anatomy, physiology and evolving insights, Nat. Rev. Neurosci. 20 (2019) 593–608.
[27] M.D. Zeiler, R. Fergus, Visualizing and understanding convolutional networks. European Conference on Computer Vision, Springer, 2014, pp. 818–833.
[28] J. Van Der Westhuizen, J. Lasenby, Techniques for Visualizing LSTMS Applied to Electrocardiograms, 2017. arXiv:1705.08153.
[29] H. Suresh, N. Hunt, A. Johnson, L.A. Celi, P. Szolovits, M. Ghassemi, Clinical intervention prediction and understanding with deep neural networks, Machine Learning for Healthcare Conference (2017) 322–337.
[30] R. Miotto, F. Wang, S. Wang, X. Jiang, J.T. Dudley, Deep learning for healthcare: review, opportunities and challenges, Brief. Bioinformatics 19 (2018) 1236–1246.