

Automatic Identification of Successful Memory Encoding in Stereo EEG of Refractory, Mesial Temporal Lobe Epilepsy

Afarin Famili¹, Gowtham Krishnan¹, Elizabeth Davenport¹, James Germi²,
Ben Wagner¹, Bradley Lega², Albert Montillo¹

¹Department of Radiology, University of Texas Southwestern, Dallas, TX, USA

²Department of Neurosurgery, University of Texas Southwestern, Dallas, TX, USA

ABSTRACT

Surgical resection of portions of the temporal lobe is the standard of care for patients with refractory mesial temporal lobe epilepsy. While this reduces seizures, it often results in an inability to form new memories, which leads to difficulties in social situations, learning, and suboptimal quality of life. Learning about the success or failure to form new memory in such patients is critical if we are to generate neuromodulation-based therapies. To this end, we tackle the many challenges in analyzing memory formation when their brains are recorded using stereoencephalography (sEEG) in a Free Recall task. Our contributions are threefold. First, we compute a rich measure of brain connectivity by computing the phase locking value statistic (synchrony) between pairs of regions, over hundreds of word memorization trials. Second, we leverage the rich information (over 400 values per pair of probed brain regions) to form consistent length feature vectors for classifier training. Third, we train and evaluate seven different types of classifier models and identify which ones achieve the highest accuracy and which brain features are most important for high accuracy. We assess our approach on data from 37 patients pre-resection surgery. We achieve up to 73% accuracy distinguishing successful from unsuccessful memory formation in the human brain from just 1.6 sec epochs of sEEG data.

Index Terms— Epilepsy, stereo EEG, phase synchrony, classification

1. INTRODUCTION

For drug-resistant mesial temporal lobe epilepsy (MTLE) patients, surgical resection of a portion of the temporal lobe containing the seizure foci is often quite effective at eliminating or greatly reducing the frequency of subsequent seizures. Unfortunately, a common side effect is significant loss in the ability to form new memories causing difficulty in social situations, learning, and suboptimal overall quality of life. There is growing interest in medical research to elucidate neuronal correlates that discriminate successful new memory formation, or *encoding* of new information, from the failure to form new memories or *unsuccessful encoding*. The study of human memory in epileptic patients presents unique opportunities. Neurosurgeons utilize invasive stereoencephalography

(sEEG) to fine-tune localization of the seizure foci. This technique entails the insertion of contact-laden probes into the patient's brain to directly record neuronal electrical activity. The probes are left in the subject's brain for up to two weeks, to give sufficient time for a seizure to occur. This also provides ample time to have the subjects perform memory tasks (e.g. Free-Recall task) to gain new insights into how new memory formation can succeed or fail. Meanwhile the sEEG probes provide unparalleled ability to resolve neuronal activation in both space and time.

In this study, we focus on memory encoding by building classifiers with machine learning to predict whether new information (a word) presented to the subject was successfully encoded or not. To do this we utilize features derived from sEEG signals during brief epochs just after the words are presented. Training such classifiers requires constructing features of a fixed length and this is challenging when processing sEEG data for two reasons. First the voltage over time recordings from the probed regions must be converted into a form that characterizes information processing and communication between regions to better capture memory formation. Second, the data must be put into a consistent length since most commonly used and optimized classifiers require training feature vectors of fixed length. However, it is seldom the case in current clinical practices, that the exact same set of neuroanatomical regions will be probed in two different patients. Electrodes are inserted so as to probe the suspected seizure foci whilst minimizing damage to eloquent areas. Our first contribution is to address these challenges. We transform the raw voltage over time recording, $v(t)$, for each pair of regions into a measure of inter-regional brain communication. We utilize statistical measure of synchrony known as the phase locking value (PLV) statistic. To address the variable nature of the probed locations we adopt a strategy of characterizing the most common inter-regional communication channels. Our second contribution is to systematically test and evaluate seven widely used classifiers for suitability to predict memory encoding success or failure in a rigorous framework using nested cross validation, which has been shown to minimize bias in accuracy estimates, as explained in [1].

Several related recent studies use machine learning to predict successful versus unsuccessful memory states. Balci

et al. trained linear support vector machine models for memory state prediction from fMRI data [2]. However, fMRI measures the BOLD signal, an *indirect* measure of neuronal activation. This signal has a slow ~ 2 second temporal resolution, while brain activation occurs on the order of several msec. Achieving sufficiently high temporal resolution requires using other modalities that directly measure neuronal activation (e.g. EEG or MEG). Hohne et al. evaluated the prediction accuracy of the linear SVM using EEG phase information [3]. They explored the data from a cortical surface grid EEG, which although simpler to analyze, does not directly measure neuronal activation well from deep memory targeted regions such as the hippocampus. This is due to interference from intervening tissue between the deep region and the cortical surface. Since deep regions such as the hippocampus are a focus of our study due to their involvement in memory, we develop a method to process sEEG data that does not suffer from such signal interference limitations.

2. MATERIALS AND METHODS

2.1 Dataset

Our dataset consists of stereo-EEG recordings from 37 patients with refractory mesial temporal lobe epilepsy who have been selected for temporal lobe resection by neurosurgeons at the UT Southwestern Medical Center, Dallas, TX USA. The recordings were sampled at 1000 Hz continuously (24h/day) for between 7-14 days. Each patient has between 10 and 13 probes, with 10 contacts along the length of each probe, implanted in their lateral and mesial structures. Across these 37 subjects a total of 28 structures were probed in at least one subject. The list of probed regions is: {*anterior hippocampus left/right, anterior temporal left/right, basal temporal lateral left/right, basal temporo-medial left/right, lateral mid temporal left/right, lateral orbit frontal left/right, lateral posterior temporal left/right, lateral temporoparietal left/right, mid orbitofrontal left/right, posterior hippocampus left/right, posterior cingulate left/right, precuneus left/right, superior temporal post left/right, superior temporal med left/right*}.

2.2 Test Episodic Memory with Free Recall Task

During extraoperative recording, each patient performed the Free Recall task (Fig. 1, top). In this task, the subjects are asked to remember a list of words, and then perform a distractor arithmetic task, then given 45 seconds to recall as many words as they can from the previously presented word list [4]. Words that were presented and subsequently recalled, were labeled by the clinician as successfully encoded (which we represent as a 1 in our training set target vector, y), while words that were presented but not recalled were labeled as unsuccessfully encoded and labeled as a 0 in y [5]. In our task, 16 word encoding trials were performed with 15 words per trial. We use the data from all trials and words. We extract and process the subset of the sEEG data in the 1600ms epoch following every word presentation which ends prior to the

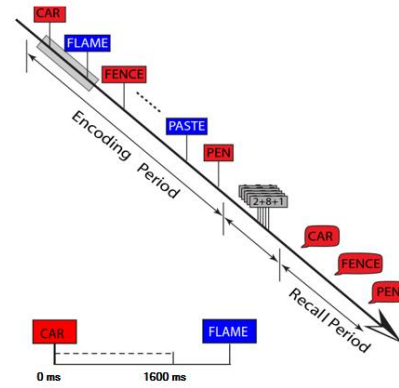


Figure 1: Free-Recall Task

next word. An example of the time-locked epochs that we process is illustrated for the word “CAR” in Fig. 1, bottom.

2.3 Measure Inter-regional, Functional Brain Connectivity with Phase-locking Value Statistic

Since sEEG directly measures neuronal activity within the brain, it permits us to derive estimates of functional connectivity. In this research, we focus on measuring the connectivity between pairs of probed brain regions. Several bi-variate estimators have been proposed, including classical *linear* methods to measure the directionality of interactions (correlation) or coherence (phase). *Non-linear* methods can model both linear relationships as well as non-linear ones. Since the nature of the memory success or failure in epileptics is not well understood, we choose a flexible non-linear phase synchronization method. This method assumes that two brain regions are connected functionally when they are phase locked. We compute the phase locking value statistic which measures whether the voltage recordings in probed regions 1 and 2, $v_1(t)$ and $v_2(t)$, are statistically significantly phase locked over multiple trials over a time window within the encoding epoch. We compute the phase locking value statistic for specific frequency components within the recordings and for multiple epoch time windows. To calculate these PLV statistics, each $v_i(t)$ recording at each region i , is filtered using Morlet wavelets to 53 logarithmically-spaced frequencies ranging from 2 Hz to 181 Hz. The appearance of locally generated oscillations at such frequencies in the human brain is widely appreciated [6]. We then apply a Hilbert transform to compute the instantaneous phase, $\phi(t)$, of the signal, where $\phi(t) \in [-\pi, \dots, +\pi]$. The difference between the phase time course of two electrodes, $\theta(t) = \phi_1(t) - \phi_2(t)$ quantifies the locking between the phases. If a stimulus or activity (e.g. memory) causes the two regions to rise and fall together or with a fixed time lag, then $\theta(t)$ will be consistent across the trials. If there is no relationship, $\theta(t)$ will be random. To quantify the randomness in $\theta(t)$ and compute PLV we use:

$$PLV_t = \frac{1}{N} \left| \sum_{n=1}^N \exp(j\theta_n(t)) \right|$$

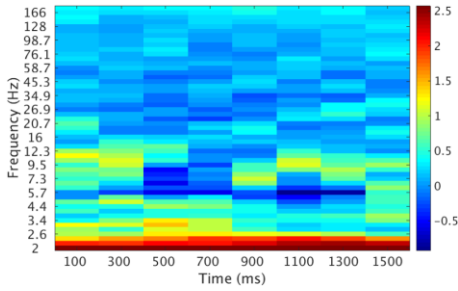


Figure 2: Phase locking value statistic over frequency and time forms our region pair (edge) feature.

where N is the number of trials and $\theta_n(t)$ is the phase difference at trial n . The PLV values are then Z-score normalized. Since there can be multiple electrodes (probe contacts) within a given region, the z-statistic of the synchrony values are averaged over all pairs of electrodes from a region pair. Electrodes are localized to a neuroanatomical region by co-registering pre-surgical CT and MRI. For further detail see [4,7].

This calculation is computed separately for each of 53 frequencies and 8 linearly-spaced time windows from 0ms to 1600ms post word presentation. This yields an edge feature matrix with 53 frequency bins and 8 time-bins of phase locking values for every pair of probed regions (Fig. 2), that we linearize columnwise to form a 424-length feature vector.

2.3 Identify Common Pairs of Probed Regions among Subjects

We formed a graph consisting of one vertex for each region and one edge between each region pair. We hypothesized that because the edge features are information rich, the individual edges would be sufficient to detect the success or failure of new memory formation, particularly when one or both regions perform a memory related task. We also found that several edges exist in common across our subjects (Table 1, bottom). The location of these edges is illustrated in Fig. 3. Utilizing common edges allows to extract features of fixed length across subjects, a requirement of several powerful classifiers.

We also hypothesized using two edges might also aid in characterizing memory formation and enable the construction of a classifier capable of discriminating successful from unsuccessful encoding. In our dataset the edge pair between anterior hippocampus left and lateral mid temporal left (AH_L and LMT_L) and between anterior temporal left and lateral mid temporal left (AT_L and LMT_L), is the most common (occurring in 16 subjects), while six other edge pairs (listed in the columns heading of table 3) occur in 15 subjects. We train and evaluate seven (7) types of two-category classifiers to discriminate successful and unsuccessful encoding. We do this *and* select optimal hyperparameters using a *nested* cross validation approach with 10 folds for each of the two levels of cross validation.

Table 1: The most common edges, delineated by probed region pairs, in our dataset. (top) Region names and their abbreviations. (bottom) Common edges and the number of subjects with each edge.

Region Name	Abbreviation
Anterior Hippocampus left	AH _L
Anterior Hippocampus right	AH _R
Anterior Temporal left	AT _L
Anterior Temporal right	AT _R
Lateral Mid Temporal left	LMT _L
Lateral Mid Temporal right	LMT _R
Posterior Cingulate right	PC _R

Common edge	# Subjects
AH _L and LMT _L	21
AT _L and LMT _L	17
AH _L and AH _R	16
LMT _L and LTP _L	16
LTP _R and PC _R	16

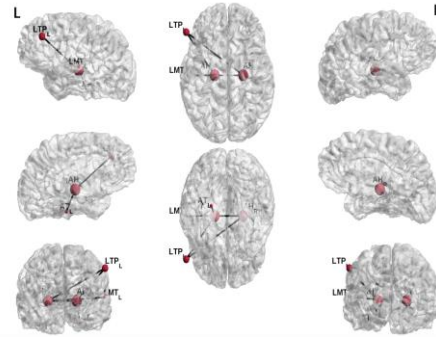


Figure 3. Locations of the 5 most common edges. Visualizations generated using BrainNet Viewer [8].

2.4. Systematically Train and Compare Seven (7) Different Classifiers Using PLV Features

Nesting provides minimal bias in accuracy estimates that are more likely to hold up when we translate our methods to the clinic. The list of classifiers we construct includes: {*logistic regression, linear support vector machine, extremely randomized forest, randomized forest, Adaboost, Gradient boosting (Gradboost), and voting (a combination of the remaining 6 classifiers)*} with implementations from [1].

3. RESULTS

In this section we describe our two experiments. In the first, we evaluate the classifiers on 424 PLV features from common single edges; in the second, we evaluate their performance when trained on the 848 PLV edge pair features.

3.3. Results Using Common Single Edges

When trained on common single edges, we observe that Adaboost, gradient boosting and voting classifiers achieve accuracy well above chance (50%) when using the PLV features from the edge connecting lateral mid temporal left (LMT_L) and lateral temporoparietal left (LTP_L). We also observe elevated standard deviations for some models, which

Table 2. Classifier accuracy using features from common single edges. Gradient boosting and voting classifiers with 73% accuracy values on edge number 4.

Classifier	Most frequent edge				
	AH _L -LMT _L	AH _R -LMT _L	AH _L -AH _R	LMT _L -LTP _L	LTP _R -PC _R
Log. Reg (%)	53 ±26	43 ±11	48 ±21	58 ±0	50 ±20
SVM (%)	52 ±25	45 ±19	53 ±8	45 ±15	55 ±10
ERF (%)	52 ±12	40 ±17	50 ±19	55 ±20	43 ±24
RF (%)	44 ±24	50 ±11	48 ±8	55 ±24	45 ±19
Adaboost (%)	53 ±16	40 ±17	30 ±22	63 ±8	60 ±23
GradBoost (%)	48 ±15	45 ±20	55 ±30	73 ±20	40 ±20
Voting (%)	38 ±13	43 ±19	53 ±21	73 ±21	38 ±19

is expected due to limited cohort size. The full results for all common single edges are shown in Table 2.

3.4. Results Using Common Edge Pairs

We trained the same set of classifiers on the features from the common edge pairs and these results are summarized in Table 3. We observe that most of the ensemble classifiers, and the logistic regression performed better than chance using the edge pairs (AH_L-LMT_L; AH_R-LMT_L) and (AH_L-AH_R; AH_R-LMT_L). The random forest performed better than chance on edges classification accuracies on edge pairs (AH_L-LMT_L; AH_L-AH_R), (AH_L-AH_R; AH_R-LMT_L) and (AH_L-LTP_L; LMT_L-LTP_L).

4. DISCUSSION

We would anticipate that features from regions involved in memory formation would be most important to identify brain states of successful encoding. Our results support that expectation; all classifiers with high accuracy (>60%) include the LMT_L as at least one communicating region and this region is known to be involved in memory formation. In our dataset, 21 of 37 subjects (57%) have at least one of the edges or edge pairs with classifier accuracy better than chance (i.e. >63%).

Fig 4 shows the PLV feature importances for the ERF classifier using the edge pair, AH_L-AH_R; AH_R-LMT_L. We observe the most important, discriminative features lie in the gamma, beta bands (known to be important for memory) and theta bands, and tend to be in the middle of the memory encoding epoch.

5. CONCLUSION

We presented a method for predicting whether information is successfully encoded in the human brain using sEEG data. To accomplish this, we overcame the challenges of heterogenous data where the set of probed regions were highly variable from one patient to the next. By constructing rich PLV feature matrices to characterize every edge, we were able to utilize common edges as input features to our classifiers that yielded high prediction accuracy, up to 73% in distinguishing memory encoding states, in 1.6-second epochs of sEEG data.

Table 3: Classifier accuracy using features from common edge pairs.

Classifier	Most frequent edge pair								
	AH _L -LMT _L	AH _R -LMT _L	AH _L -LMT _L	LMT _L -LTP _L	AH _L -AH _R	AH _L -LMT _L	AH _R -LMT _L	AH _L -LTP _L	AH _R -LTP _L
Log. Reg (%)	45 ^{±1} 9	40 ±17	50 ±16	60 ±16	50 ±16	50 ±16	43 ±11	48 ±8	
SVM (%)	40 ±20	55 ±15	53 ±15	48 ±8	45 ±8	53 ±17	43 ±16		
ERF (%)	58 ±16	48 ±18	53 ±14	65 ±13	68 ±26	53 ±15	35 ±23		
RF (%)	43 ±19	55 ±15	60 ±18	58 ±22	65 ±21	53 ±22	63 ±20		
Adaboost (%)	43 ±22	53 ±23	48 ±17	65 ±26	65 ±20	38 ±11	50 ±19		
GradBoost (%)	45 ±16	43 ±22	40 ±8	63 ±20	55 ±22	38 ±12	40 ±32		
Voting (%)	43 ±13	50 ±21	45 ±12	63 ±17	68 ±20	45 ±13	35 ±22		

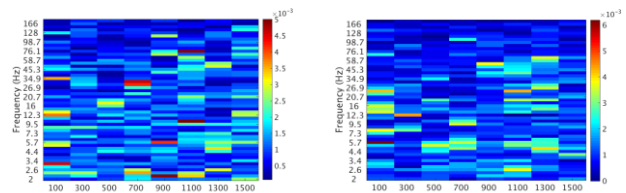


Figure 4: Feature importances for the extremely randomized forest (ERF) using the edges (left) AH_L-AH_R and (right) AH_R-LMT_L.

We extended our analysis to common edge pairs. We trained and evaluated seven classifiers and reported accuracy on all. In the future, we aim to further improve accuracy with additional rich brain connectivity features and increase the number of patients studied. Our work has brain states with high propensity for successful memory formation for which our ultimate goal is to use for targeted therapies, such as deep brain stimulation, to help restore full memory capabilities for the MTLE patients.

6. REFERENCES

- [1] S. Varma and R. Simon. "Bias in Error Estimation When Using Cross-validation for Model Selection." *BMC Bioinformatics*, 7(1):91, 2006.
- [2] Balci, S. K., et al. "Prediction of successful memory encoding from fMRI data." *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2008.
- [3] Höhne, M., et al. "Prediction of successful memory encoding based on single-trial rhinal and hippocampal phase information." *NeuroImage* 139: 127-135, 2016.
- [4] Lega, B., Burke J., Jacobs J., Kahana M.J., "Slow-Theta-to-Gamma Phase-Amplitude Coupling in Human Hippocampus Supports the Formation of New Episodic Memories." *Cereb. Cortex Cerebral Cortex*, 26(1): 268-278, 2014.
- [5] Burke, J. F., Zaghoul, K. A., Jacobs, J., Williams, R. B., Sperling, M. R., Sharan, A. D., & Kahana, M. J. "Synchronous and asynchronous theta and gamma activity during episodic memory formation." *The Journal of Neuroscience*, 33(1), 292-304, 2013.
- [6] Kahana, M.J. "The cognitive correlates of human brain oscillations." *The Journal of Neuroscience* 26(6): 1669-1672, 2006
- [7] Lachaux, J. P., et al. Measuring phase synchrony in brain signals, *Human brain mapping*, 8(4), 194-208, 1999.
- [8] Mingrui X., Jinhui W., and Yong He Y., "BrainNet Viewer: a network visualization tool for human brain connectomics." *PLoS one* 8(7), 2013.