References

[AOG+12] C. Aydin, O. Oktay, A.U. Gunebakan, R.K. Ciftci, and A. Ademoglu. Functional parcellation of memory related brain networks by spectral clustering of Electroencephalography (EEG) data. In *Telecommunications and Signal Processing (TSP), 2012 35th International Conference on*, pages 581–585, 2012.

Aydin et al. investigated the relationship between alpha-band EEG activity and memory function. Clustering, both group and individual wise, was used to classify electrodes for the visuospatial tasks.

[BC96] G.A. Babich and O.I. Camps. Weighted parzen windows for pattern classification. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 18(5):567–570, 1996.

Parzen windows are nonparametric classifiers that estimate a density function for data provided, as many data sets do not follow common distributions. Babich and Camps approach this by clustering training data to find the centers (reference vectors) and weights (window weights) to be used in a kernel estimator. This method makes use of a Parzen Window classifier estimation during training so long as these estimates do not exceed a certain distance from each other. Overall, this weighted Parzen-window classifier is less computational, requires less room, and has little error difference from the full classifier.

[BLUv13] X. Bresson, T. Laurent, D. Uminsky, and J. H. von Brecht. An Adaptive Total Variation Algorithm for Computing the Balanced Cut of a Graph. *ArXiv e-prints*, February 2013.

This article proposes a new method of Total Variation (TV) Clustering. Bresson et al. offer a novel method of choosing an adaptive tolerance such that the convergence of the relaxed Balance Cut Problem (BCP) to a solution vector is monotonic. Monotonicity of convergence is desirable because it leads to a more efficient, more direct convergence. The relaxed BCP is a minimization problem whose objective function uses a TV norm to measure each iteration of the minimization. A threshold must first be defined, and then all points can then be clustered into the two classes by reading indices from the solution vector.

[BWEC13] David A. Bridwell, Lei Wu, Tom Eichele, and Vince D. Calhoun. The spatiospectral characterization of brain networks: Fusing concurrent {EEG} spectra and Functional Magnetic Resonance Imaging (fMRI) maps. *NeuroImage*, 69(0):101 – 111, 2013.

Birdwell et al. applied group Independent Component Analysis (ICA) separately to EEG spectral signals and fMRI Blood-Oxygen-Level-Dependent (BOLD) signals, acquired simultaneously. Selected components of the ICA were deconvolved together to estimate an Impulse Response Function (IRF) for component pairs of each fMRI and EEG components. This method treats the fMRI component time-course as the result of the convolution of the EEG component time-course and the IRF. The deviations in the IRF are indicative of EEG responses being associated with subsequent fMRI BOLD signal. These results were analyzed for similarities in associations in each of the different frequencies, observing positive associations between EEG and fMRI in lower and upper frequencies and negative associations in various alpha sources.

[CV95] Corinna Cortes and Vladimir Vapnik. Support-vector networks. In *Machine Learning*, pages 273–297, 1995.

Support vector networks are a method of classifying data based on clustering of training data–requires a grand truth as well as binary outcomes. The support-vector machine relies on the support vectors given by the training data to determine an optimal hyperplane to separate the clusters. Larger margins usually lead to fewer errors in classification. If the data set is not separable, the kernel trick can be used to convert the problem into a comparable, although higher-dimensional, separable problem.

[EHM12] Andrew D. Engell, Scott Huettel, and Gregory McCarthy. The fMRI {BOLD} signal tracks electrophysiological spectral perturbations, not event-related potentials. *NeuroImage*, 59(3):2600 – 2606, 2012.

Varying the duration of an external stimulus application results in a modulation of γ-Event Related Spectral Perturbations (ERSP) power that is similar to modulation of the BOLD response in certain regions of the brain (peri-calcarine cortex, fusiform gyrus, and lateral-temporal-occipital cortex). Lower-frequency ERSP power is not modulated to the same extent as are the γ-ERSP power and the BOLD signal. This observation suggests that the change in BOLD signal could be indicative of a corresponding change in γ-ERSP power. The methods used (wavelet decomposition with a sliding window over time) are relatively novel and may have yielded better temporal resolution than previous work. This improvement in resolution may have allowed for detection of more minute, temporary changes than was possible in earlier studies. Further, the inclusion of frequencies larger than 45 Hz into the γ-band allows for increased sensitivity to high-frequency signals.

[FGK+12] Nikdokht Farid, Holly M. Girard, Nobuko Kemmotsu, Michael E. Smith, Sebastian W. Magda, Wei Y. Lim, Roland R. Lee, and Carrie R. McDonald. Temporal lobe epilepsy: Quantitative mr volumetry in detection of hippocampal atrophy. *Radiology*, 264(2):542–550, 2012.

The detection of hippocampal atrophies in epilepsy patients is crucial in lateralization of Temporal Lobe Epilepsy (TLE). However, slight atrophies present early in the disease may be difficult to identify, even by expert radiologists. Farid et al., implement fully automated volumetric Magnetic Resonance Imaging (MRI) in an attempt to correctly lateralize TLE in their patients. Volumetric analysis can confirm hippocampal cell loss, and thus is a strong indicator of the presence and degree of hippocampal atrophy. Individual patient's brain scans were normalized with respect to age and sex, and then compared to appropriate norms in order to determine if atrophy was present. Although this volumetric approach was successful in determining the presence and laterality of these atrophies, it is best used as an aid to visual analysis.

[FYS+12] Niels K. Focke, Mahinda Yogarajah, Mark R. Symms, Oliver Gruber, Walter Paulus, and John S. Duncan. Automated {MR} image classification in temporal lobe epilepsy. *NeuroImage*, 59(1):356 – 362, 2012. Neuroergonomics: The human brain in action and at work.

Support vector machines can be used to classify between left Mesial Temporal Lobe Epilepsy (mTLE), right mTLE, and the control group (no unilateral mTLE), with or without marked Hippocampal Sclerosis (HS). The classification/diagnosis was particularly accurate when gray matter based segmentation and Diffusor Tensor Imaging (DTI) were implemented. When dealing with mTLE, it can be beneficial to use local weighting.

[HO00] Aapo Hyvrinen and Erkki Oja. Independent component analysis: algorithms and applications. *Neural Networks*, 13:411–430, 2000.

Hyvrinen and Oja give an overview of Independent Component Analysis (ICA) and its implementation in recovering individual data samples from multivariate data. This separation of data is achieved through application dependent preprocessing of the data, such as centering and whitening and utilizing an algorithm that maximizes the contrast function of components. The application of ICA becomes useful in the separation of individual brain activity components from measured potentials in EEG, the separation of brain activity from artifacts in the data (eye movements or blinks, or malfunctions in sensors), and also in the denoising of images.

[HR11] Thomas R. Henry and Deborah D. Roman. Presurgical epilepsy localization with interictal cerebral dysfunction. *Epilepsy & Behavior*, 20(2):194 – 208, 2011.

Accurate localization is crucial to the success of resection aimed towards curing TLE. Fludeoxyglucose Positron Emission Tomography (PDG-PET) is valuable in planning resection surgery and predicting surgical outcomes. Different types of epilepsy can characterized by the apparent hypometabolism observed in regions such as the mesial and lateral temporal lobe.

[KCA+13] W.T. Kerr, A.Y. Cho, A. Anderson, P.K. Douglas, S.T. Nguyen, N.M. Reddy, E.P. Lau, E.S. Hwang, K.R. Raman, A. Trefler, D.H. Silverman, and M.S. Cohen. Balancing clinical and pathological relevance in machine learning diagnosis of epilepsy. In *3rd International Workshop Pattern Recognition in Neuron Aging*, Philadelphia, PA, USA, 2013.

[Filler citation for now–likely contains errors.] It is important to consider when evaluating machine learning used in the diagnosis of epilepsy the control group–whether the comparison group is normal or Persons with Non-Epileptic Seizures (PWN). Kerr et al. argue that because the clinical question in epilepsy diagnosis is whether patients are Persons with Epilepsy (PWE) or PWN given that seizures are already present, the control group should be PWN. That is, the clinical comparison is between patients with epileptic seizures and those with non-epileptic seizures, not between patients with epileptic seizures and those whom not present seizures.

[KNC+13] Wesley T. Kerr, Stefan T. Nguyen, Andrew Y. Cho, Edward P Lau, Daniel H. Silverman, Pamela K. Douglas, Navya M. Reddy, Ariana Anderson, Jennifer Bramen, and Noriko Salamon. Computer-aided diagnosis and localization of lateralized temporal lobe epilepsy using interictal FDG-PET. *Frontiers in Neurology*, 2013.

Kerr et al. focus on the problem of diagnosing PWE verus PWN using a Computer Aided Diagnosis (CAD) tool. The model used involved partitioning PDG-PET data into 47 Region(s) of Interest (ROI) and incorporating the metabolic activity of each ROI in training a Multilayer Perceptron (MLP) evaluated with Cyclical Leave One Out Cross Validation (CL1OCV).

[LS00] Daniel D. Lee and H. Sebastian Seung. Algorithms for non-negative matrix factorization. In *In NIPS*, pages 556–562. MIT Press, 2000.

Many data classification techniques rely on matrix factorizations. When the data set at hand is guaranteed to be nonnegative (i.e. the physical measurements only take on nonnegative values) the nonnegative matrix factorization (NMF) method may be useful. NMF is an alternative strategy to using SVD or PCA. This paper examines different techniques which can be used to compute the NMF which will approximate a general nonnegative matrix.

[ML10] C. Mulert and L. Lemieux. *Eeg - Fmri: Physiological Basis, Technique, and Applications.* Springer Berlin Heidelberg, 2010.

This chapter presents a discussion of various techniques for integrating EEG and fMRI information. The goal is to formulate an accurate representation of brain activity with the spatial density of fMRI and temporal density of EEG. Asymmetrical approaches use one form of data as a predictor for details of the other form, for example EEG and its temporal density (resp. fMRI and its spatial density) as a predictor for temporal details of fMRI (resp. spatial details of EEG). These models assume a strong level of truth to the predictor data. Symmetrical approaches, information fusion, equally weight the different types of data and attempt to identify complementary features. The goal of these techniques is to invert the data, taking EEG (resp. fMRI) from [scalp-measured voltage] (resp. BOLD signal) to underlying neural activity.

[MPC+07] D. Mantini, M.G. Perrucci, S. Cugini, A. Ferretti, G.L. Romani, and C. Del Gratta. Complete artifact removal for {EEG} recorded during continuous fMRI using independent component analysis. *NeuroImage*, 34(2):598 – 607, 2007.

Simultaneous recording of EEG and fMRI data is useful in combining the strengths of both procedures. However, EEG recordings made in the MRI become contaminated by both physiological and environmental disturbances in the MRI environment. Mantini et al. propose a method of using ICA to remove ballistocardiographic (BCG) and ocular artifacts along with contamination due to MRI from the data. The source signals were separated into brain signals and artifacts through manual and automated classification, and artifacts were subtracted from the data. In doing so, desired signals were recovered and noise free data was reconstructed to be used for later clinical study.

[PLD05] H. Peng, Fulmi Long, and C. Ding. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(8):1226–1238, 2005.

Peng et al. present a method for efficiently selecting the most important factors in a dataset after distinct classes (clusters) have been distingushed. They present the Max-Relevance, Min-Redundancy criteria (*mRMR*), based on the principle of mutual information for bivariate joint probability density functions. This process is presented in contrasted to the Max-Dependency criteria, which requires joint probability functions of higher dimensions (these are characterized as ill-posed problems).

[PM02] RD Pascual-Marqui. Standardized low-resolution brain electromagnetic tomography (sLORETA): Technical details. *METHODS AND FINDINGS IN EXPERIMENTAL AND CLINICAL PHARMACOLOGY*, 24(D):5–12, 2002. 12th Meeting of the International-Pharmaco-EEG-Group, BARCELONA, SPAIN, NOV 21-24, 2002.

[Not 100% citation]sLORETA is a technique for transforming the EEG recordings from spatially diffuse electrode space into MR voxel space. It is currently the best solution to the inverse problem for finding the exact sources of neural activity as it is instantaneous, distributed, discreate and linear. It is also accurate for localizing electrode readings whose sources are deep in the cortex. Importantly, sLORETA computes with a zero localization error.

[Shl03] John Shlens. A tutorial on principal component analysis derivation, discussion and singular value decomposition. Technical report, University of California San Diego, March 2003.

Principal component analysis provides a method for reducing the dimensionality of data–or more importantly, reducing redundancy and noise. It is derived from the assumption that noise and redundant data will typically exhibit low variance, whereas the principal components will exhibit higher variance. It is proven that given a few additional assumptions, the principal components can be obtained from the original data by applying a linear transformation to the original data. This linear transform is obtained from the assumption that we wish to diagonalize the covariance matrix–and therefore reduce the redundancy of the data.

[vL07] Ulrike von Luxburg. A tutorial on spectral clustering. *CoRR*, abs/0711.0189, 2007.

Spectral clustering first involves the establishment of a graph and definition of a pairwise vertex similarity metric. A graph type (ϵ-neighborhood, k-nearest and mutual k-nearest neighborhood, fully-connected) must be specified. The type of Laplacian (normalized random-walk, normalized symmetric, or unnormalized) should be chosen, and the Laplacian matrix computed. A matrix should be created with columns consisting of the k smallest eigenvectors of the Laplacian matrix. The row vectors of this new matrix should be partitioned into k classes, via k-means clustering. If the ith row vector was assigned to the jth cluster, then the ith vertex in the training set is assigned to some jth cluster of vertices. Temporal lobe epilepsy is characterized by newly-presented hypometabolic features in some region of the brain. Spectral clustering could be applied to the diagnostic problem by identifying the patient as having a specific class of epilepsy (bilateral, Left Temporal Lobe Epilepsy (LTLE), Right Temporal Lobe Epilepsy (RTLE), unlateralized) based on which ROI has experienced the hypometabolism. Toward the fMRI/EEG fusion problem: after transforming EEG spectral bands onto fMRI voxel space, different spectral bands (either single or multiple bands at a time) could be clustered with fMRI signal. In this scheme the spectral band data and fMRI data would be vertices on a graph that is weighted based on temporal proximity, similarity in signal strength, or another factor.