Localization of epileptic seizure focus by computerized analysis of fMRI recordings

Rasoul Hekmati[1*] , Robert Azencott[1*], Wei Zhang[2,3], Zili D. Chu[2,4] and Michael J. Paldino[2]

Abstract

By computerized analysis of cortical activity recorded via fMRI for pediatric epilepsy patients, we implement algorithmic localization of epileptic seizure focus within one of eight cortical lobes. Our innovative machine learning techniques involve intensive analysis of large matrices of mutual information coefficients between pairs of anatomically identified cortical regions. Drastic selection of pairs of regions with biologically significant inter-connectivity provides efficient inputs for our multi-layer perceptron (MLP) classifier. By imposing rigorous parameter parsimony to avoid overfitting, we construct a small-size MLP with very good percentages of successful classification.

Keywords: Time series, Deep learning, Mutual information, fMRI, Epilepsy, Seizure focus

1 Introduction

Epilepsy is a common brain disorder that affects 6 out of 1000 children in the United States [36, 52]. Although medication is the mainstay of treatment for these patients, seizures are resistant to anti-epileptogenic drugs in approximately 30% of children. Such medically refractory epilepsies are frequently "focal", defined as originating habitually from a relatively localized region of the brain. In appropriately selected patients, resection of this localized seizure origin can be curative. Despite the promise of this approach, surgical outcomes remain highly variable, even in optimal candidates [6, 7, 17, 24]. Emerging studies motivated by these inconsistencies have demonstrated that pediatric focal epilepsy is not a localized disorder, but rather a disease of large-scale, distributed neural networks [40]. Furthermore, rather than resulting from abnormal activity generated by one or more abnormal foci, the primary organization of seizure origin occurs within a functionally and anatomically connected set of brain regions—a seizure network [2]. Although the exact relationship between ictal dynamics and the inter-ictal seizure network remains to be defined, pathologic connectivity within the seizure network likely contributes to the transition between inter-ictal and ictal states [40, 43]. In other words, the capacity of this network for seizure generation results not only from abnormal hyperactivity within intrinsically dysplastic neural elements, but also from abnormal interactions between otherwise "normal" brain predisposed by pathologic *interictal* connectivity within the seizure network [38, 43]. In some cases this network may be relatively restricted in extent; at other times, it encompasses neuronal populations dispersed throughout multiple distant brain structures [40, 50]. Regardless of extent, failure to sufficiently resect or disconnect the seizure network may be associated with continued seizures after removal of the seizure focus, thereby contributing to surgical failure [1, 18]. At the current time, there is no modality that can accurately map the extent of the seizure network.

Advances in MRI combined with mathematical network approaches have demonstrated great potential to study brain networks non-invasively in a variety of populations, including children with focal epilepsy [4, 5, 20, 21, 41, 45–47]. Resting-state functional MRI is one method by which connectivity in the brain can be measured. This sequence quantifies the blood oxygen level dependent (BOLD) signal over time as an indirect measure of neuronal activity [3, 25]. Functionally connected elements of the brain exhibit similar spontaneous BOLD signal fluctuations at rest [3], which enables

*Correspondence: rhekmati@uh.edu; razencott@uh.edu
[1] Department of Mathematics, University of Houston, Houston, TX, USA
Full list of author information is available at the end of the article

Hekmati *et al. Brain Inf.*   (2020) 7:13

Page 2 of 13

one to capture interactions between brain regions in vivo. Within the network framework, the brain is represented as a graph, a mathematical construct consisting of a set of nodes connected by edges [12, 14, 26, 28, 30, 41, 42]. The nodes are defined as anatomical regions of the brain; the edge between each pair of nodes is estimated as the magnitude of association between their BOLD time series [14, 23, 29, 34, 53]. Studies based on such resting state functional network constructs have consistently shown global aberrations of the epileptic brain from normal; they have also demonstrated that network features have the potential to predict clinically relevant aspects of brain function in children with epilepsy [22, 26, 28, 30, 32, 38, 44, 55]. However, little data exist regarding the use of neuroimaging-defined network models to identify the seizure network in an individual child.

The goal of this study was to assess the potential of MR imaging-defined networks to depict seizure networks. Specifically, we trained an artificial neural network (ANN) to identify the location of seizure origin given access only to features derived from a resting-state functional network. We measured the accuracy of the learning algorithm against the lobe of seizure origin determined at multidisciplinary epilepsy surgery conference. We selected this reference standard—which relies on the consensus of multiple objective tests as well as the clinical expertise from multiple disciplines—as an accurate and reliable assessment of the lobe of seizure origin [18, 56]. Identification of the lobe of seizure origin can be considered an initial step toward the ultimate validation of the potential for resting-state functional networks to visualize seizure networks in children with focal epilepsy.

2 Methods
2.1 Patients
The HIPAA-compliant study was approved by the local institutional review board. Informed consent was waived for this retrospective study. Patient medical records were retrospectively reviewed to identify patients with the following inclusion criteria: (1) pediatric age group (21 years of age or younger); (2) a clinical diagnosis of focal epilepsy; (3) an available 3-T MR imaging of brain, including a resting-state fMRI sequence; and (4) lobe of seizure origin identified at multidisciplinary epilepsy conference. This determination relies on the consensus of multiple objective tests (including MRI and EEG) as well as the clinical expertise from multiple disciplines (including Neurology, Neurosurgery, and Neuroradiology) [19, 20]. Images were performed from January 2012 to December 2017. Forty-six patients were included. Lobe diagnoses for the cohort are presented in Table 1.

Table 1 Lobe diagnosis of the cohort

Lobe	Frequency
Group 1—right frontal (RF)	10
Group 2—left frontal (LF)	8
Group 3—right temporal (RT)	7
Group 4—left temporal (LT)	11
Group 5—others: right and left parietal + right and left occipital (C5)	6

The 5 classes are not perfectly disjoint because 4 patients were ambiguously diagnosed by clinicians and thus belong to 2 classes simultaneously (RT/C5, RF/RT, LF/LT, LT/RF)

2.2 MR imaging
All imaging was performed on a 3-T Achieva system (Philips, Andover, Massachusetts) with a 32-channel phased array coil. The following sequences were obtained: 1. Structural images: sagittal volumetric T1-weighted images [repetition time (TR)/echo time (TE): 7.2/2.9 ms; 1 acquisition; flip angle: 7°, inversion time: 1100 ms; field of view (FOV): 22 cm; voxel size (mm): $1 \times 1 \times 1$)]. 2. Resting state fMRI: axial single-shot echo planar imaging (EPI) fMRI [TR/TE (ms): 2000/30; flip angle: 80°; 1 acquisition; FOV: 24 cm; voxel (mm): $3 \times 3 \times 3.75$; 300 volumes (duration: 10 min)] performed in the resting state. All images were visually inspected for artifacts, including susceptibility and subject motion.

2.3 Image processing and analysis
The processing pipeline was implemented using MATLAB scripts (version 7.13, MathWorks, Inc) in which adapter functions were embedded to execute FreeSurfer reconstruction (version 5.3.0) and several FMRIB Software Library (FSL) suite tools [39]. Details regarding this pathway have been previously described [26, 28, 30]. A brief summary is provided here.

The reference space was created from images of one patient in our database, who had no visible abnormality and with optimal registration to MNI space [10]. Structural imaging data for each patient were aligned to a standard reference template (MNI152) using FSL's non-linear registration algorithm [37, 39]. Nodes in the network were defined on the template according to parcellation of whole-brain gray matter as follows: First, FreeSurfer reconstruction of cerebral cortical surfaces was performed on the T1 structural image. This processing stream includes motion correction, skull stripping, intensity normalization, segmentation of white matter and gray matter structures, parcellation of the gray matter and white matter boundary, and surface deformation following intensity gradients which

optimally place the gray matter/white matter and gray matter/cerebrospinal fluid borders [8, 9]. The pial and gray-white surfaces were visually inspected using the Freeview software for accurate placement.

Next, a self-developed MATLAB program was applied to the FreeSurfer output to further subdivide the 148 standard gray matter regions according to their surface area [26–30]. During this process, each region was iteratively divided into two new regions of equal size until the surface area of each parcel (as defined on the FreeSurfer gray-white surface mesh) was less than a size threshold of 350 mm^2. The final parcellation contained 780 nodes. Each surface parcel was then converted into a volume mask of gray matter at that region to form a node on the network. All nodes defined in reference space were transformed into each individual patient's space by applying the non-linear transformation matrix (12 degrees-of-freedom) obtained during registration.

The first 5 volumes in each resting state functional data were removed to allow magnetization to reach equilibrium. Standard preprocessing and independent component analysis (ICA) of the functional data sets was performed using FSL MELODIC [39], consisting of motion correction, interleaved slice timing correction, brain extraction, spatial smoothing with a Gaussian kernel full width at half maximum of 5 mm, and high pass temporal filtering equivalent to 100 s (0.01 Hz). Affine boundary-based registration as implemented in FSL FLIRT was then used to align the pre-processed functional image volumes for each patient to that individual's structural T1 dataset using linear registration [19]. The inverse transformation matrix was calculated in this step and subsequently used to transform all masks from structural to functional space. Mean BOLD-signal time series were then computed for each node.

2.3.1 Mutual information between discrete random variables

Mutual information $MI(X, Y)$ between two random variables X and Y quantifies the amount of information brought by the knowledge of X for the prediction of Y. This quantification is symmetric in X and Y, and is based on the concept of conditional entropy. When X and Y can take only a finite number of values denoted x_k and y_m, their marginal and joint probability distributions are:

$$p_k = Pty(X = x_k); q_m = Pty(Y = y_m);$$
$$r_{k,m} = Pty(X = x_k, Y = y_m).$$

Then the entropies of X, Y, and $Z = (X, Y)$ are given by:

$$H(X) = -\sum_k p_k \log(p_k) \geq 0; H(Y) = -\sum_m q_m \log(q_m);$$

$$H(X, Y) = -\sum_{k,m} r_{k,m} \log(r_{k,m}).$$

The mutual information between X and Y is then given by:

$$MI(X, Y) = MI(Y, X) = H(X) + H(Y) - H(X, Y) \geq 0,$$

where $MI(X, Y) = 0$ if and only if X and Y are independent. When X and Y have a bivariate Gaussian distribution, the mutual information $MI(X, Y)$ is also computable as an explicit increasing function of the squared correlation:

$$MI(X, Y) = -\frac{1}{2} \log\left(1 - \text{cor}(X, Y)^2\right).$$

However, in the non-Gaussian case, this formula is not valid. Mutual information can offer major advantages over correlation, especially in the pathologic brain [33, 35, 54]. Mutual information has remarkable invariance property which we now recall. For any random variables (U, V) and any *strictly increasing functions* f and g, the random variables $X = f(U)$ and $Y = g(V)$ verify:

$$MI(X, Y) = MI(f(U), g(V)) = MI(U, V).$$

In this formula, f and g can be *non-linear* functions. This MI invariance property is not satisfied by correlations, since $Cor(U, V)$ and $Cor(f(U), g(V))$ are generally different unless f and g are both linear functions. In the study of fMRI recordings of brain activity, the strong functional invariance of mutual information offers an interesting advantage, which we now outline.

2.3.2 Advantages of mutual information to estimate cortex connectivity

Denote our $N = 780$ cortex parcels by $CP_1, CP_2, \ldots, CP_N$. For any two parcels CP_i, and CP_j, brain activity recordings by fMRI provide, after standard pre-treatments, two sequences of $n = 295$ recordings:

$S(i) = [Y_i(1), Y_i(2), \ldots, Y_i(n)]$ and
$S(j) = [Y_j(1), Y_j(2), \ldots, Y_j(n)]$,

which can be considered as random samples of n observations for two random variables Y_i and Y_j. To evaluate the level of neural interaction between the two parcels i and j, many publications use the correlations $Cor(Y_i, Y_j)$ or their absolute values, which can easily be estimated from the data sequences $S(i)$ and $S(j)$. In this paper, we have preferred instead to use the same data $S(i)$ and $S(j)$ to estimate the mutual information $MI(Y_i, Y_j)$. We now indicate why this choice has a strong pragmatic and

Hekmati *et al. Brain Inf.* (2020) 7:13

Page 4 of 13

theoretical advantage over the use of $Cor(Y_i, Y_j)$. Indeed, the actual values $Y_i(n)$ and $Y_j(n)$ generated by fMRI recordings after pretreatment are essentially of the form:

$Y_i(n) = f[BOL_i(n)]$ and

$$Y_j(n) = g[BOL_j(n)],$$

where at time *n*, the numbers $BOL_i(n)$ and $BOL_j(n)$ represent the BOLD (blood oxygen level dependent) signals for the two parcels CP_i and CP_j. Here, the functions $f(u)$ and $g(v)$ are two *increasing* functions of *u* and *v*, which are essentially *unknown* to the experimenter because they are determined by many variable factors such as the actual hardware settings of the fMRI acquisition modalities, the meta-parameters of the fMRI signal reconstruction software, the geometric positions of CP_i and CP_j within the cortex, etc. Hence, contrary to the actually observed data sequences $S(i)$ and $S(j)$, the two BOLD signal sequences $BS(i) = [BOL_i(1), \ldots, BOL_i(n)]$ and $BS(j) = [BOL_j(1), \ldots, BOL_j(n)]$ are in fact not directly observable. Consider the sequences $BS(i)$ and $BS(j)$ as random samples from two random variables BOL_i and BOL_j. Since the sequences $S(j)$ and $BS(j)$ are linked by the function f, the observable random variable Y_i is of the form $Y_i = f(BOL_i)$. Similarly, Y_j is of the form $Y_i = g(BOL_j)$. Most changes in the instrumental settings of the fMRI hardware may modify f and g in an unknown way. Moreover, if one changes the parcels, the functions f and g will also be modified. Fortunately, due to the functional invariance of mutual information mentioned above, we have:

$$MI(BOL_i, BOL_j) = MI(f(BOL_i), g(BOL_j)) = MI(Y_i, Y_j)$$

for arbitrary increasing functions f and g. By contrast, $Cor(Y_i, Y_j)$ and $Cor(BOL_i, BOL_j)$ can generally be quite different unless f and g are linear functions, which is quite unlikely in this context. So to estimate cortex connectivity by quantifying the strength of pairwise interactions between the BOLD signals BOL_i, BOL_j at parcels CP_i and CP_j, the mutual information $MI(Y_i, Y_j)$ will be more stable and more relevant than the correlation $Cor(Y_i, Y_j)$. Consistent with theoretical superiority, the real-world advantages of MI have been documented in this setting [35, 54].

2.3.3 Numerical computation of mutual information matrix

As above, we expect strong values of $MI_{i,j} = MI(Y_i, Y_j)$ to detect the frequent presence of simultaneous high blood oxygen levels BOL_i and BOL_j within cortex parcels CP_i and CP_j, a phenomenon indicative of strong neuronal interactivity between CP_i and CP_j. Each observed time series $Y_i(t)$ involves $n = 295$ observations of the

continuous random variable Y_i, we discretize each Y_i to transform it into a random variable taking only 5 values. To this end we first compute 6 quantiles of the *n* observations $Y_i(t)$, at levels 0, 20, 40, 60, 80, 100%. This subdivides the range of Y_i into 5 intervals $[Z_1, \ldots, Z_5]$ with midpoints $[z_1, \ldots, z_5]$, and whenever $Y_i(t)$ falls in Z_k, we replace $Y_i(t)$ by $\widehat{Y}_i(t) = z_k$. Then we approximate $MI_{i,j}$ by $MI(\widehat{Y}_i, \widehat{Y}_j)$.
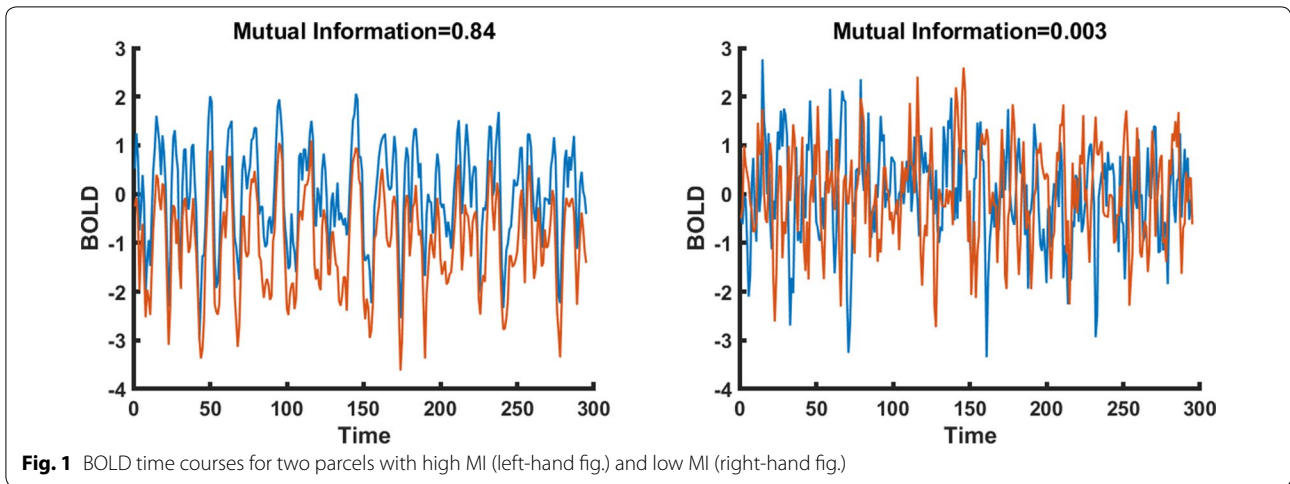
To evaluate the accuracy of our mutual information estimates, we have applied a variant of the "Jackknife resampling" technique by randomly taking out 10% of time points (29 time points) and re-calculating MI on the new data set. This process has been repeated 1000 times and the absolute error of MI estimation has been calculated. For 60% of the mutual information matrix coefficients $MI_{i,j}$, the relative estimation error on $MI_{i,j}$ is inferior to 10%. The use of 20 bins instead of 5 bins to generate the discretized $Y_i(t)$ tends to double the average errors of estimation for the $MI_{i,j}$, as can be expected from theory. Indeed, our exploration of more detailed discretization, and even optimized discretization, has shown that using 5 bins of equal frequency 20% for each Y_i was quite effective in our context. Figure 1 displays two nodes whose BOLD time courses have high MI and another two nodes whose time courses have low MI.

2.3.4 Mutual information between pairs of cortex subregions

For each patient, there are roughly 780 parcels of size 350 mm^2 and for each parcel one time series with 295 recordings. We characterize the cortex connectivity of each patient by the symmetric 780 × 780 mutual information matrix $MI_{i,j}$. The coefficients of this MI matrix provide a very large number of features $\cong 780^2/2$ for each patient [31]. To reduce the number of features we have developed and implemented an innovative multiscale analysis of these MI matrices.

For any cortex subregion A, denote $L(A)$ the list of n_A cortex parcels contained in A. For any two cortex subregions A and B generate the set $S(A, B)$ of all n_A by n_B mutual information coefficients $MI_{i,j}$ with i in $L(A)$ and j in $L(B)$. Let $L(A, B)$ to be the list of all elements of $S(A, B)$. We then define the mutual information $MI(A, B) = MI(B, A)$ as the 75% quantile of the list $L(A, B)$. When $MI(A, B)$ is high, at least 25% of cortex parcels pairs CP_i in $L(A)$ and CP_j in $L(B)$ have high $MI_{i,j}$ and hence are expected to have strong neural interaction.
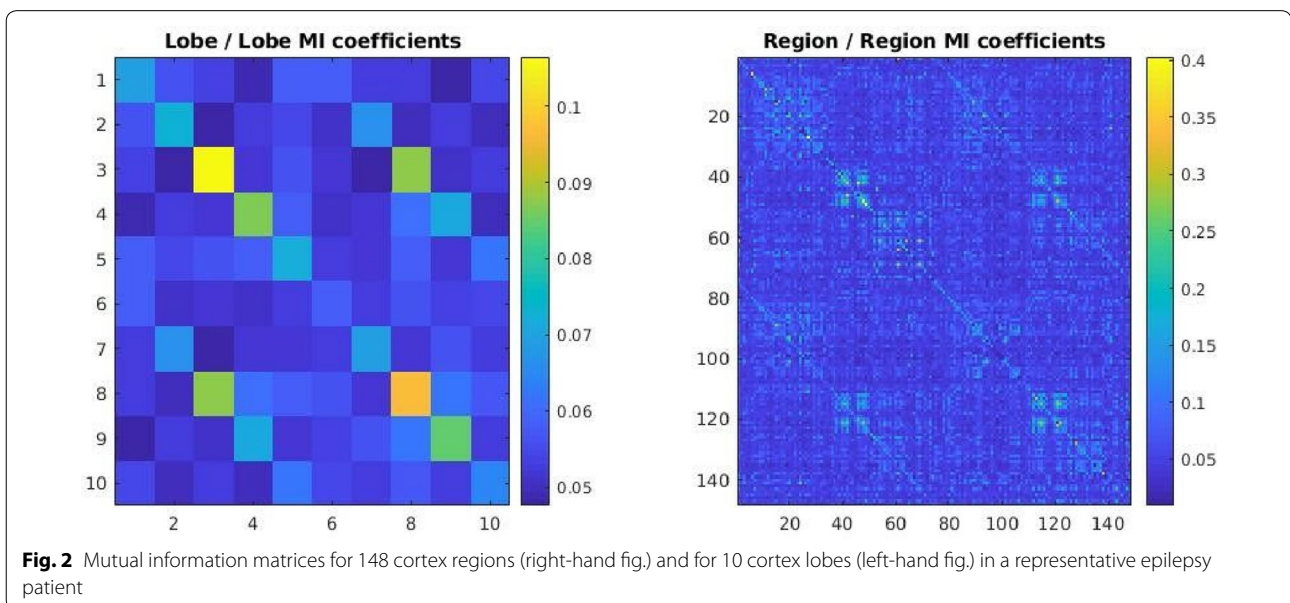
In our definition of $MI(A, B)$, instead of the 75% quantile, we could a priori have used other quantiles of $S(A, B)$. To explore these possibilities, for each patient and each mutual information coefficient $MI_{i,j}$, we have computed the Relative Standard Error of estimation $RES(A, B)$ for $MI(A, B)$. These results show that $RES(A, B)$ becomes fairly small when $MI_{i,j}$ lies between

Fig. 1 BOLD time courses for two parcels with high MI (left-hand fig.) and low MI (right-hand fig.)

the 70%-quantile and 90%-quantile of the patient's MI matrix coefficients, so that these $MI(A, B)$ can be considered as quite robustly estimated. Moreover, for regions A and B of small or only moderate sizes, the high quantiles of all the $MI_{i,j}$ in the list $L(A, B)$ are of course very sensitive to possible outliers. This has led us to explore defining $MI(A, B)$ by quantiles of $L(A, B)$ between 70 and 90% to define $MI(A, B)$. We have thus separately tested 4 quantiles namely 70, 75, 80, 85%. These tests showed that, using any one of these 4 possible quantiles to implement our features selection algorithm, we ended up with fairly similar performances for our neural network classifier, with a small qualitative

advantage for the 75% quantile. This explains our choice for the definition of the regional MI coefficients $MI(A, B)$.

We then compute the coefficients $MIreg(m, n) = MI(REG_m, REG_n)$ for any two regions in our list of 148 anatomically defined cortex regions to obtain for each patient a symmetric matrix $MIreg$ of "inter-region connectivity". $MIreg$ has size 148×148, and each one of its $(148*149)/2 = 11{,}026$ distinct coefficients will provide 11,026 *potential input features* $MIreg(m, n)$ for our patient classification task. We have also studied the symmetric 10×10 symmetric matrices $MIlob$ of mutual information between our 10

Fig. 2 Mutual information matrices for 148 cortex regions (right-hand fig.) and for 10 cortex lobes (left-hand fig.) in a representative epilepsy patient
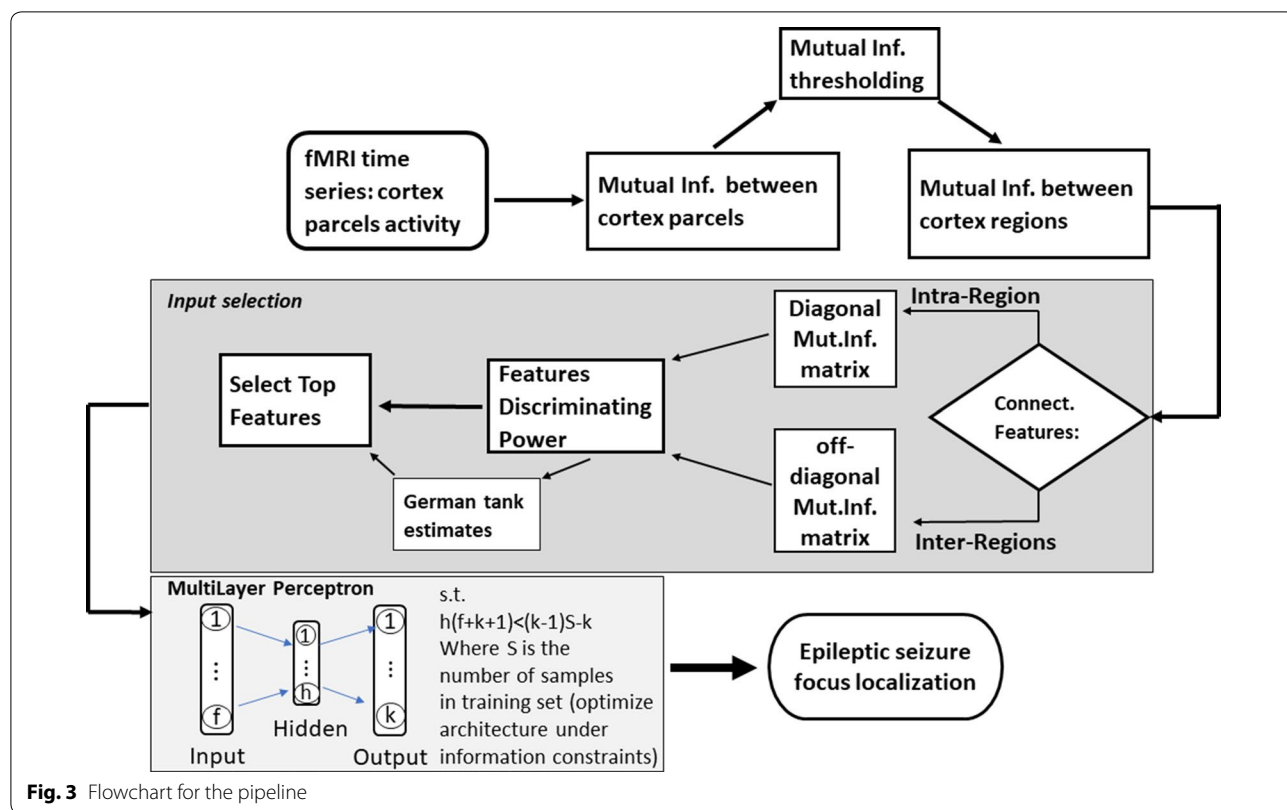
Hekmati *et al. Brain Inf.* (2020) 7:13

Page 6 of 13

Fig. 3 Flowchart for the pipeline

anatomically defined cortex lobes. Figure 2 displays (for only one patient) the matrices *MIreg* and *MIlob*.

2.4 Automatic classification by multilayer perceptron

The chart in Fig. 3 shows the workflow of the pipeline from fMRI time series to epileptic seizure focus localization.

Here we explain the MLP architecture and its input selection procedure.

2.4.1 Architecture of our multi-layer perceptron (MLP)

To implement automatic classification of seizure origin by computer analysis of fMRI recordings of resting state brain activity, we have decided to use and evaluate Artificial Neural Networks (ANNs). Our ANNs are feed forward MLPs (multi-layer perceptrons), technically similar to the classical type of ANNs used to classify handwritten digits (MNIST data).

We first implement, as detailed further on, our feature selection algorithm which extracts a set of f highly discriminating input features, selected among our 11,026 region/region mutual information coefficients $MIreg(m, n)$. For each patient, the f selected input features are encoded by f "neurons" which constitute the 1st layer of our MLP. The 2nd layer (called the hidden

layer) contains h neurons, with all-to-all "synaptic" links between input layer and hidden layer. Each hidden neuron has a Sigmoid Response Function combining the f input features and using $f + 1$ unknown parameters (f weights and *one* threshold). The 3rd layer of our MLP contains $k = 5$ neurons (1 neuron for each one of our k classes of patients). Each one of these k neurons is linked to all the h hidden neurons and has a Sigmoid Response function, involving $h + 1$ unknown parameters. The 4th MLP layer is the output layer and contains k neurons. Each output neuron is linked to all the k neurons of the 3rd layer. The response function of output neuron j is classically computed by the SoftMax formula, which involves no unknown parameter, and which generates the probability $p(j)$ that the current f inputs come from a patient belonging to class $CL(j)$.

2.4.2 Parameters parsimony constraints for the MLP inputs

Most MLPs trained on very large training data sets, such as the MNIST data, have many layers successively generated by deep learning, and a quite large number of neurons. But here, given the moderate number of diagnosed patients in our data set, we have instead focused on implementing a radical reduction of the number of weights and thresholds to be learned during MLP training, in order to enforce a robust learning of synaptic

Hekmati *et al. Brain Inf.* (2020) 7:13

Page 7 of 13

weights and thresholds. Enforcing a strong parsimony for the number of unknown parameters (weights and thresholds) to be "learned" by an MLP is indeed known to *enhance the robustness* of automatic classifiers [15, 16]. Indeed, deep theoretical results of Vapnik [48] have demonstrated that this is a good recipe to perform stable automatic learning on a moderately large training set.

In such situations, feedforward neural networks (such as MLPs) are known to exhibit better generalization capacity and less overfitting when their architecture involves fewer weights. But one still wants the MLPs to achieve good classification performance, which often requires to increase the number of weights. Our approach to implement an ideal balance between these two criteria has involved a meticulous algorithmic selection of a small number of highly discriminating input features for our MLPs. Let "f" be the number of selected input features and "k" be the number of patient classes. Our MLP classifiers involve 4 successive layers of artificial neurons: an input layer of size f, a hidden layer of size h, a 3rd layer of size $k = 5$, and finally an output layer of size k which computes k probabilities $p(1) \ldots p(k)$ verifying $p(1) + \cdots + p(k) = 1$. The number w of weights and thresholds (also called "unknown parameters") of such an MLP is given by:

$$w = h(f + k + 1) + k.$$

Call W the vector of all w unknown parameters. Each patient provides a vector U of f input values for our MLP, and the outputs $p(1) \ldots p(k)$ computed by the MLP are all of the form $p(j) = G_j(U, W)$ where G_j is an explicit non-linear function of U and W. But these functions are linearly dependent because by construction one always has $G_1 + \cdots + G_k = 1$. For the nth training case, the input U_n is known, as well as the true values $v_1(n) \ldots v_k(n)$ for the k MLP outputs $p(1) \ldots p(k)$. In the coding of true outputs, for a case in class $CL(j)$, the only non-zero true output is $v_j = 1$, so that $v_1 + \cdots + v_k = 1$. Hence the nth training case yields k non-linear equations satisfied by the vector W of unknown parameters, namely:

$$G_1(U_n, W) = v_1(n); \ldots ; G_k(U_n, W) = v_k(n).$$

But these k equations are *linearly dependent* because as seen above their sum is always equal to 1. Thus, each training case yields in fact only $(k - 1)$ non-linear independent constraints to be satisfied by the w unknown parameters. Hence a training set of size S provides $(k - 1) \times S$ non-linear equations to be satisfied by w unknown parameters. To avoid overfitting, one should

"ideally" have $w < (k - 1)S$, which yields the following dimensional constraint on "h" and "f":

Parameters parsimony constraint $h(f + k + 1) < (k - 1)S - k$.

Here, we have $k = 5$ patient classes and $S = 45$ because we use *leave-one-out* training. Hence, the parameters parsimony constraint on (h, f) becomes here as $h(f + 6) \leq 175$.

2.4.3 Implementation of automatic learning

After selecting f highly discriminating input features for our MLP, we also fix the size h of its hidden layer, making sure that f and h verify the parsimony constraint $h(f + 6) \leq 175$. During training of this MLP, when a case is ambiguously diagnosed as belonging to 2 classes $CL(i)$ and $CL(j)$, we assign to this case MLP outputs $v_i = v_j = 1/2$. Automatic learning is performed by standard Gradient Descent, to minimize the average "Cross Entropy" between true outputs and MLP-computed outputs. Indeed, for classification tasks, minimizing cross entropy is generally more efficient than minimizing mean squared error.

Since our training set is of moderate size, we implement the classical leave-one-out technique. Namely, one eliminates one patient from the training set before performing automatic learning; then one evaluates by 0 or 1 the correctness of our MLP classification of this left-out patient. Final performance is the average of these evaluations over all possible choices of the left-out patient.

2.4.4 Input selection for minimal size MLP classifiers

We want to select "features" discriminating between 5 classes of patients $CL(1), \ldots CL(5)$ of sizes $s(1) \ldots s(5)$. This requires solving at least 10 basic tasks of the type "discriminate $CL(p)$ versus $CL(q)$". Consider any feature F computable from fMRI recordings, and hence providing for each patient PAT a number $F(PAT)$. For each class $CL(p)$, the $s(p)$ patients $[PAT\ 1, PAT\ 2, \ldots]$ of $CL(p)$ generate a list $V(p)$ of $s(p)$ values $F1 = F(PAT1), F2 = F(PAT2), \ldots$ The probability distribution $\mu_p(F)$ of $[F1, F2, \ldots]$ is unknown, and difficult to estimate if $s(p)$ is not very large. For feature efficiency evaluation only, we roughly approximate $\mu_p(F)$ by a uniform distribution on an unknown interval $J_p(F)$. A classical algorithm to estimate $J_p(F)$ is the "German Tank Estimate" (see Appendix), which computes first the min $m(p)$ and the max $M(p)$ of the list $V(p)$, and then extends adequately the interval $[m(p), M(p)]$.

When $J_p(F)$ and $J_q(F)$ are nearly disjoint intervals, we naturally consider that the classes $CL(p)$ and $CL(q)$ are strongly separated by feature F. To quantify this notion, define the *separability* $sep(J, J')$ of two intervals J and J' with intersection $I = J \cap J'$ by:

$$sep(J, J') = \min\left(\frac{|J - I|}{|J|}, \frac{|J' - I|}{|J'|}\right),$$

where K is the length of interval K. Then $0 \le sep(J, J') \le 1$, and values of $sep(J, J')$ close to 1 mean that J, J' are nearly disjoint. Now define $DIS_{p,q}(F) = sep(J_p(F), J_q(F))$ as the *discriminating power* of feature F to differentiate between classes $CL(p)$ and $CL(q)$.

To select a small but efficient set of "f" input features for our MLP classifier, we start from the $(148)(149)/2$ mutual information coefficients $MIreg(m, n)$ at the cortex regions "scale". For each pair $m \le n$, the "regions based" feature $F_{m,n} = MIreg(m, n)$ provides for each patient an indicator of the neuronal interaction level between cortex regions REG_m and REG_n. Fix any two distinct classes $CL(p), CL(q)$. For all (m, n) with $1 \le m \le n \le 148$, we compute the power $POW(m, n) = DIS_{p,q}(F_{m,n})$ of feature $F_{m,n}$ to discriminate $CL(p)$ vs $CL(q)$. Then we rank the features $F_{m,n}$ by decreasing $POW(m, n)$, and select only the top two $F_{m,n}$ having highest discriminating power $POW(m, n)$.

3 Results

3.1 Inter-region connectivity input to MLP

For each one of the 10 pairs $CL(p), CL(q)$ with $1 \le p < q \le 5$, the preceding selection algorithm selects two regions Reg_m, Reg_n. This algorithm, then, generates a set of 20 pairs of cortex regions REG_m, REG_n with strongly discriminating coefficients $MIreg(m, n)$ (Fig. 4). To respect the parameter parsimony constraint, we keep only a set of 18 of these Inter-region Connectivity Coefficients, to be used below as a set of 18 inputs features $F_{m1,n1}, \ldots, F_{m18,n18}$ for our first MLP classifier.

3.2 Intra-region connectivity input to MLP

The intra-region connectivity coefficient $MIreg(m, m)$ of cortex region Reg_m is an indicator of the level of neuronal interactions within Reg_m. The selection algorithm outlined above can detect the regions Reg_m for which the feature $F_{m,m} = MIreg(m, m)$ has strong power $POW(m, m)$ to discriminate between at least two classes $CL(p), CL(q)$. We have thus selected a set of 21 such cortex regions, 21 intra-connectivity coefficients (Fig. 5), from which we keep only a set of 18 coefficients to be used below as input features of our second MLP classifier.

3.3 Discriminating power for inter- and intra-region connectivity coefficients:

For each one of the 10 basic discrimination tasks $CL(p)$ vs $CL(q)$ the highest discriminating power $DIS(p, q)$ reached for this task among all the
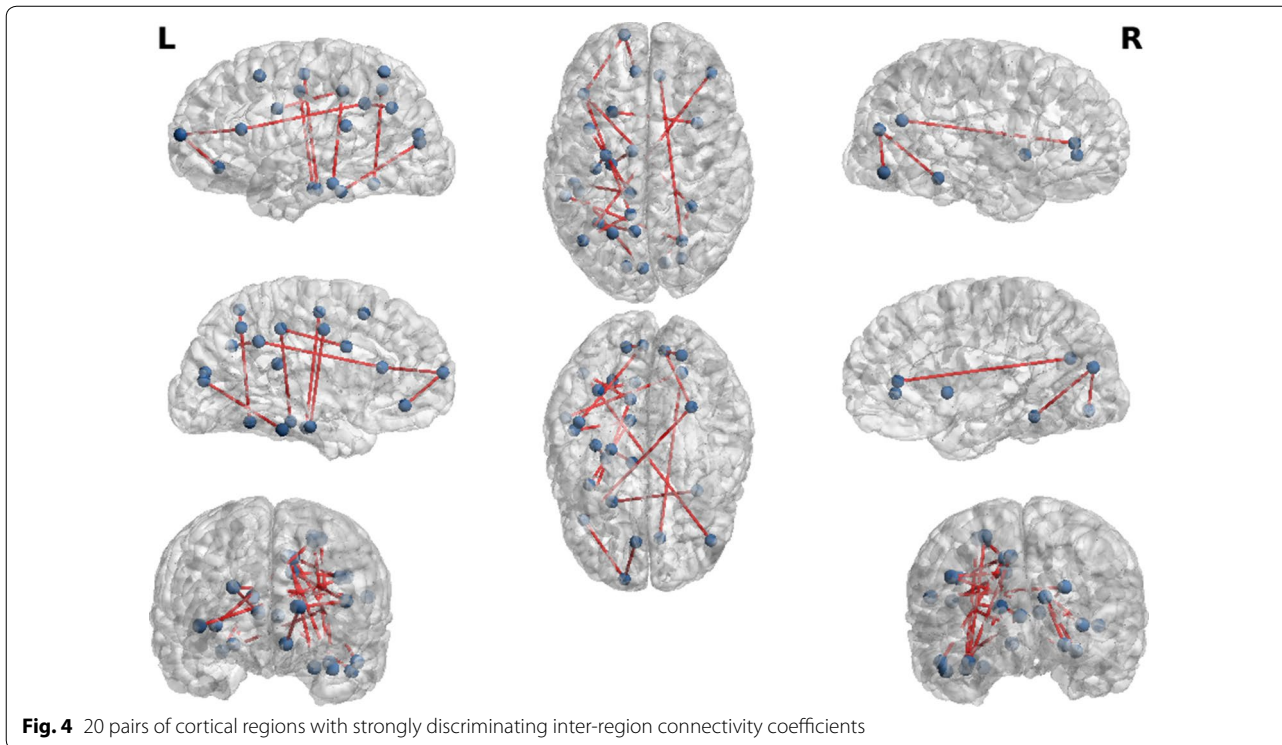

Fig. 4 20 pairs of cortical regions with strongly discriminating inter-region connectivity coefficients
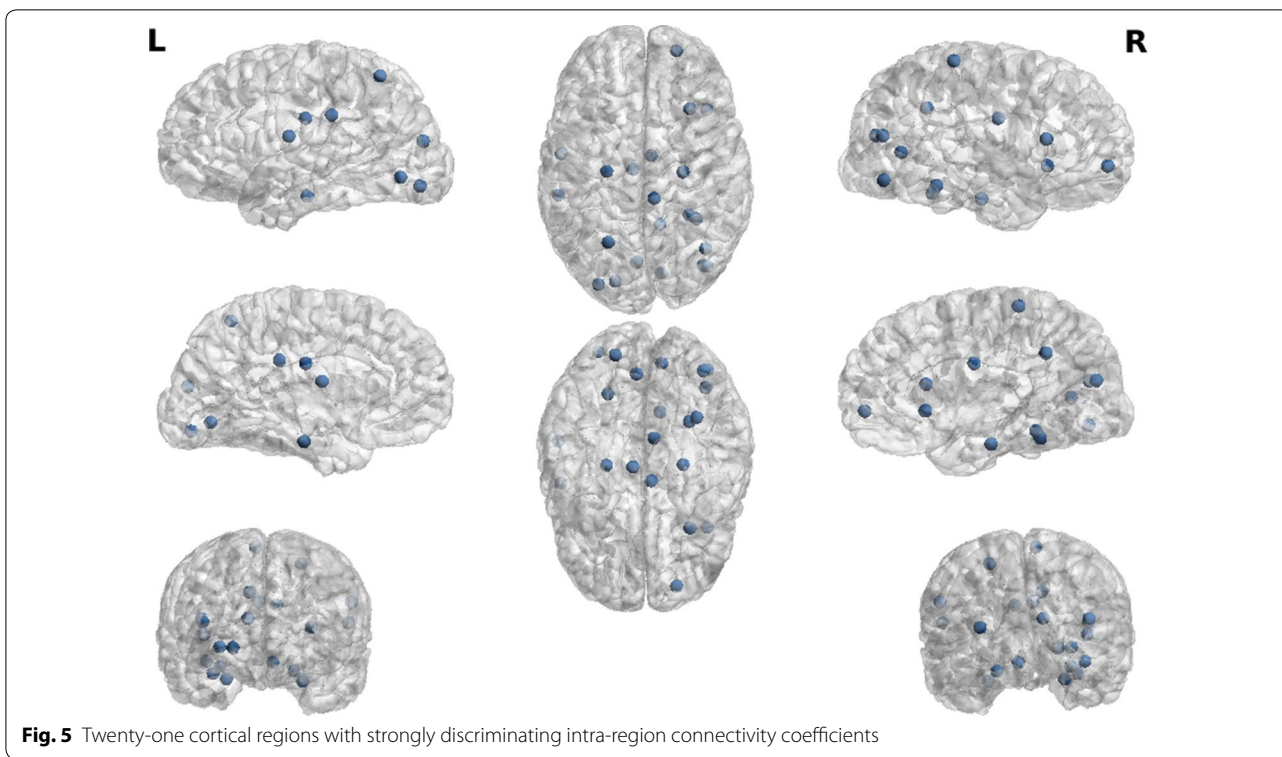
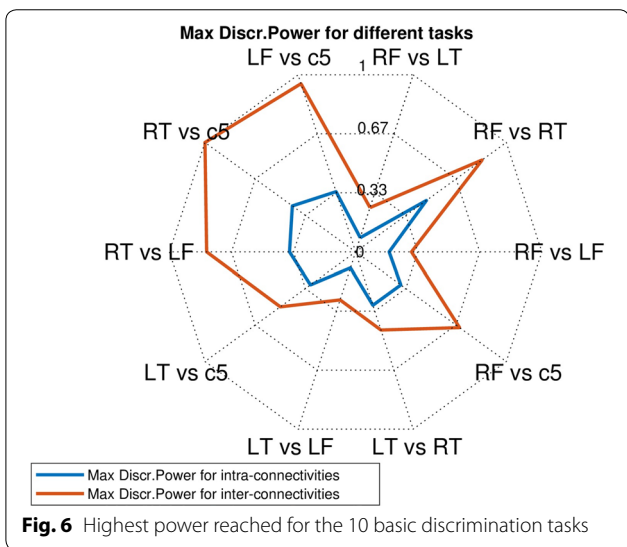**Fig. 5** Twenty-one cortical regions with strongly discriminating intra-region connectivity coefficients

Fig. 6 Highest power reached for the 10 basic discrimination tasks

Table 2 Automatic classification accuracy in a leave-one-out design based on 18 strongly discriminating inter-connectivity coefficients at cortex regions scale

Patient class	RF	LT	RT	LF	Other
Accuracy (%)	82	90.5	100	87	90

3.4 Patient classification by minimal MLP with highly discriminating inputs

3.4.1 MLP classifier based on inter-region connectivity coefficients

The set of 18 inter-connectivity coefficient $MIreg(m, n)$ selected above provides the $f = 18$ inputs of our first MLP classifier, for which we impose a hidden layer size $h = 7$. The 4 layers of this MLP have then sizes (18, 7, 5, 5). Our parameter parsimony equation becomes $h(f + 6) = 168 < 175$, and is hence correctly verified, again assuring a reasonable robustness of the MLP training results.

After training this small-size MLP, we evaluate its performances by leave-one-out technique. The global percentage of successful patient classifications over all 5 patient classes was $89 \pm 2.1\%$. In Table 2, we display the percentage of successful classifications within each patient class. Most errors occur for discrimination between Classes RF and LF.

connectivity coefficients $MIreg(m, n)$, is a crude indicator of "how well" the task may be solved by an efficient classifier. In Fig. 6, we display in red the ten $DIS(p, q)$ values reached by inter-connectivity coefficients and in blue the $DIS(p, q)$ values reached by intra-connectivity coefficients. Clearly the red values dominate the blue values.

Hekmati *et al. Brain Inf.* (2020) 7:13

Page 10 of 13

Table 3 Classification accuracy in leave-one-out design based on 18 highly discriminating intra-connectivity coefficients

Patient class	RF	LT	RT	LF	Other
Accuracy (%)	70	80	100	87	100

Table 4 Classification accuracy for SVM in leave-one-out design based on 18 highly discriminating connectivity coefficients

Patient class	RF	LT	RT	LF	Other
Accuracy (inter-connectivity)	%59	%64	%83	%78	%69
Accuracy (intra-connectivity)	%57	%62	%73	%71	%64

3.4.2 MLP classifier based on 18 intra-region connectivity coefficients

The set *of* 18 intra-connectivity coefficients $MIreg(m,m)$ selected above provides $f = 18$ inputs for our second MLP classifier which has again a 4-layer architecture with dimensions (18, 7, 5, 5) (see Section 2) and hence satisfies our parsimony equation. After training this second MLP, performance evaluated by leave-one-out technique demonstrates a global percentage of successes of $88 \pm 2.1\%$. Table 3 displays classification accuracy within each patient class. Results are slightly weaker than for our first MLP, particularly for Class RF and Class LF. Discrimination between these two classes generate most errors of classification.

3.5 Patient classification by support vector machine (SVM) with highly discriminating inputs:

The set of 18 inter-connectivity/intra-connectivity coefficients $MIreg(m,n)$ selected above can be input to a multiclass SVM classifier that uses a one-vs-rest approach. Table 4 displays classification accuracy within each patient class (leave-one-out technique).

4 Discussion

Our goal was to explore automatic classification of fMRI data in young epileptic patients to identify the origin of epileptic seizures. Given the moderate size of our data set, one challenge was to avoid using machine learning techniques involving large numbers of parameters. We have solved this challenge by introducing a rigorous "parameters parsimony principle".

A second challenge was to identify strongly discriminating input features computable from our fMRI data. To this end, we have systematically used large matrices of mutual information coefficients between all pairs of roughly 780 time series extracted from each patient's fMRI data. We have introduced a computable version of mutual information between any two cortex regions, as an indicator of neuronal interactions between these two regions. This led us to define and compute the discriminating power of all these $148^2/2$ inter-region connectivity coefficients.

We have then extracted one set of 18 strongly discriminating inter-region connectivity coefficients, and used them as input features for a small-size multi-layer perceptron (MLP) with 4 layers (dimensions: 18, 7, 5, 5), including a "SoftMax" terminal layer. After automatic training by a leave-one-out technique, this MLP provided 89% successful patients classification. For comparison, a standard support vector machine applied to our patient classification has yielded less than 71% successes. Our innovative development of mutual information between pairs of cortex regions, and of algorithmic selection of highly discriminating pairs of cortex regions has shown good capacity to extract useful and interpretable brain activity features from fMRI recordings. The small size of the MLP classifier we have thus constructed rigorously avoids overfitting and reaches good performance on our group of epileptic patients. To confirm these exploratory findings, we plan to test our approach on larger groups of patients.

We also designed and implemented an analysis to interpret why interactivity between two specific regions (say A, B) can separate two specific lobes (say L_1, L_2). We found out that for almost all the patients with seizure focus in L_1, we can find *within* L_1 one or more cortex regions having strong interaction with A and with B, while *all* subregions of L_2 have much weaker simultaneous interactions with A and B. This means that A, B are simultaneously excited whenever specific subregions of L_1 become strongly active, while hyperactivity in any subregion of L_2 has a weak impact on the simultaneous activity of A and B. For example, the inter-connectivity between the regions $= A$ *"Transverse frontopolar gyri and sulci"* and $= B$ *"Vertical ramus of the anterior segment of the lateral sulcus"* has strong discriminating power between lobes $L_1 = LF$ and $L_2 = RF$ and we can actually find two subregions U and V of LF with high interactivity with A:

$MI(U,A) = 0.22$ and $MI(V,B) = 0.25$,
while for all subregions K of RF one has:
$MI(K,A) < 0.12$ and $MI(K,B) < 0.11$.

These data suggest the potential for a similar approach to depict specific nodes that constitute important drivers of the seizure network.

Our methods are generally consistent with previous applications of network connectivity to invasive EEG. Wilke et al. observed a correlation of betweenness centrality, a network graph metric, with the location of

Hekmati *et al. Brain Inf.*   (2020) 7:13

Page 11 of 13

network nodes whose resection was likely to result in seizure freedom [51]. Sinha et al. also validated a model of epileptogenicity based on connectivity against surgical outcomes in a cohort consisting largely of adults [38]. They observed a correlation between their index of epileptogenicity and the putative seizure onset zone; they also found areas of high epileptogenicity outside of the resection zone in several cases of unsuccessful surgery. Tomlinson et al. applied a network approach to a pediatric cohort with focal epilepsy [44]. They observed a significant increase in connectivity outside of the seizure onset zone; further, global synchrony (a measure of the overall strength of connectivity) within the field of invasive EEG could be used to classify patients with regard to seizure-free outcome after surgery. These studies not only reinforce the relevance of network features to seizure onset, they highlight the importance of broader areas of dysconnectivity (beyond the traditional zone of seizure onset) in achieving seizure freedom. Invasive monitoring in the form of electrocorticography (ECOG) or stereo EEG (SEEG) is gold standard for localization of the seizure onset zone. These modalities, however, are limited in their spatial sampling; large areas of the brain are left unexplored, leading to the potential for erroneous and biased conclusions [13]. Optimal use of invasive monitoring therefore requires an accurate pre-test hypothesis regarding the location and extent of the epileptogenic network. It is likely, therefore, that optimal patient outcomes would result from the addition of global network approaches to standard invasive monitoring.

5 Conclusion

We developed machine learning algorithms to evaluate the connectivity obtained from resting-state fMRI in terms of differentiating the lobe of seizure origin in a pediatric cohort with focal epilepsy. These findings support the potential for neuroimaging-based network constructs to depict pathophysiologically relevant features of seizure genesis. If these approaches can be tailored to identify individual elements within the seizure network, biomarkers based on functional networks may ultimately contribute to personalized management strategies in children undergoing epilepsy surgery.

Author details
[1] Department of Mathematics, University of Houston, Houston, TX, USA. [2] Department of Radiology, Texas Children's Hospital, Houston, TX, USA. [3] Outcomes & Impact Service, Texas Children's Hospital, Houston, TX, USA. [4] Department of Radiology, Baylor College of Medicine, Houston, TX, USA.

Appendix: "German Tanks" estimate for uniform distributions

Let μ be a uniform probability distribution on an unknown interval $= [a, b]$. Let $X_1, \ldots, X_N$ be N independent random observations having the same probability distribution μ. From these N observations, one can compute the classical "German Tank" (GT) estimators (A_N, B_N) of (a, b), which are unbiased and have minimal variance.

To compute the GT estimators, denote

$$U_N = \min(X_1, \ldots, X_N), V_N = \max(X_1, \ldots, X_N).$$

One then has [11, 49]:

$$A_N = U_N - (V_N - U_N)/(N - 1);$$
$$B_N = V_N + (V_N - U_N)/(N - 1).$$

References
1. Bartolomei F, Chauvel P, Wendling F (2008) Epileptogenicity of brain structures in human temporal lobe epilepsy: a quantified study from intracerebral EEG. Brain 131(Pt 7):1818–1830. https://doi.org/10.1093/brain/awn111
2. Bartolomei F, Lagarde S, Wendling F, McGonigal A, Jirsa V, Guye M et al (2017) Defining epileptogenic networks: contribution of SEEG and signal analysis. Epilepsia 58(7):1131–1147. https://doi.org/10.1111/epi.13791
3. Biswal BB, Van Kylen J, Hyde JS (1997) Simultaneous assessment of flow and BOLD signals in resting-state functional connectivity maps. NMR Biomed 10(4–5):165–170
4. Bullmore E, Sporns O (2009) Complex brain networks: graph theoretical analysis of structural and functional systems. Nat Rev Neurosci 10(3):186–198. https://doi.org/10.1038/nrn2575

Hekmati *et al. Brain Inf.*   (2020) 7:13

Page 12 of 13

5. Centeno M, Carmichael DW (2014) Network connectivity in epilepsy: resting state fMRI and EEG-fMRI contributions. Front Neurol 5:93. https://doi.org/10.3389/fneur.2014.00093

6. Cohen-Gadol AA, Ozduman K, Bronen RA, Kim JH, Spencer DD (2004) Long-term outcome after epilepsy surgery for focal cortical dysplasia. J Neurosurg 101(1):55–65. https://doi.org/10.3171/jns.2004.101.1.0055

7. Cohen-Gadol AA, Wilhelmi BG, Collignon F, White JB, Britton JW, Cambier DM et al (2006) Long-term outcome of epilepsy surgery among 399 patients with nonlesional seizure foci including mesial temporal lobe sclerosis. J Neurosurg 104(4):513–524. https://doi.org/10.3171/jns.2006.104.4.513

8. Fischl B, Liu A, Dale AM (2001) Automated manifold surgery: constructing geometrically accurate and topologically correct models of the human cerebral cortex. IEEE Trans Med Imaging 20(1):70–80. https://doi.org/10.1109/42.906426

9. Fischl B, Salat DH, van der Kouwe AJ, Makris N, Segonne F, Quinn BT et al (2004) Sequence-independent segmentation of magnetic resonance images. Neuroimage 23(Suppl 1):S69-84. https://doi.org/10.1016/j.neuroimage.2004.07.016

10. Fonov V, Evans AC, Botteron K, Almli CR, McKinstry RC, Collins DL et al (2011) Unbiased average age-appropriate atlases for pediatric studies. Neuroimage 54(1):313–327. https://doi.org/10.1016/j.neuroimage.2010.07.033

11. Goodman L (1954) Someparactical techniques in serial number analysis. J Am Stat Assoc 49(265):97–112

12. Grayson DS, Fair DA (2017) Development of large-scale functional networks from birth to adulthood: a guide to the neuroimaging literature. Neuroimage. https://doi.org/10.1016/j.neuroimage.2017.01.079

13. Hader WJ, Tellez-Zenteno J, Metcalfe A, Hernandez-Ronquillo L, Wiebe S, Kwon CS et al (2013) Complications of epilepsy surgery: a systematic review of focal surgical resections and invasive EEG monitoring. Epilepsia 54(5):840–847. https://doi.org/10.1111/epi.12161

14. Hagmann P, Cammoun L, Gigandet X, Meuli R, Honey CJ, Wedeen VJ et al (2008) Mapping the structural core of human cerebral cortex. PLoSBiol 6(7):e159. https://doi.org/10.1371/journal.pbio.0060159

15. Hekmati R, et al (2018) Automatic classification of large sets of brain activity fMRI time series. Paper presented at the 4th International Conference on Big Data and Information Analytics, Houston.

16. Hekmati R, et al (2018) Machine learning to evaluate fMRI recordings of brain activity in epileptic patients. Paper presented at the Q-Bio Conference, Houston.

17. Hosoyama H, Matsuda K, Mihara T, Usui N, Baba K, Inoue Y et al (2017) Long-term outcomes of epilepsy surgery in 85 pediatric patients followed up for over 10 years: a retrospective survey. J NeurosurgPediatr 19(5):606–615. https://doi.org/10.3171/2016.12.PEDS16197

18. Jayakar P, Gaillard WD, Tripathi M, Libenson MH, Mathern GW, Cross JH et al (2014) Diagnostic test utilization in evaluation for resective epilepsy surgery in children. Epilepsia 55(4):507–518. https://doi.org/10.1111/epi.12544

19. Jenkinson M, Smith S (2001) A global optimisation method for robust affine registration of brain images. Med Image Anal 5(2):143–156

20. Kim D-J, Davis EP, Sandman CA, Sporns O, O'Donnell BF, Buss C et al (2016) Children's intellectual ability is associated with structural network integrity. Neuroimage 124:550–556

21. Li Y, Liu Y, Li J, Qin W, Li K, Yu C et al (2009) Brain anatomical network and intelligence. PLoSComputBiol 5(5):e1000395. https://doi.org/10.1371/journal.pcbi.1000395

22. Liao W, Zhang Z, Pan Z, Mantini D, Ding J, Duan X et al (2010) Altered functional connectivity and small-world in mesial temporal lobe epilepsy. PLoS ONE 5(1):e8525. https://doi.org/10.1371/journal.pone.0008525

23. Magalhaes R, Marques P, Soares J, Alves V, Sousa N (2015) The impact of normalization and segmentation on resting-state brain networks. Brain Connect 5(3):166–176. https://doi.org/10.1089/brain.2014.0292

24. Najm I, Jehi L, Palmini A, Gonzalez-Martinez J, Paglioli E, Bingaman W (2013) Temporal patterns and mechanisms of epilepsy surgery failure. Epilepsia 54(5):772–782. https://doi.org/10.1111/epi.12152

25. Ogawa S, Lee TM, Kay AR, Tank DW (1990) Brain magnetic resonance imaging with contrast dependent on blood oxygenation. Proc NatlAcadSci U S A 87(24):9868–9872

26. Paldino MJ, Chu ZD, Chapieski ML, Golriz F, Zhang W (2017) Repeatability of graph theoretical metrics derived from resting-state functional networks in paediatric epilepsy patients. Br J Radiol 90(1074):20160656. https://doi.org/10.1259/bjr.20160656

27. Paldino MJ, Golriz F, Chapieski ML, Zhang W, Chu ZD (2016) Brain network architecture and global intelligence in children with focal epilepsy. AJNR Am J Neuroradiol. https://doi.org/10.3174/ajnr.A4975

28. Paldino MJ, Golriz F, Chapieski ML, Zhang W, Chu ZD (2017) Brain network architecture and global intelligence in children with focal epilepsy. AJNR Am J Neuroradiol 38(2):349–356. https://doi.org/10.3174/ajnr.A4975

29. Paldino MJ, Golriz F, Zhang W, Chu ZD (2019) Normalization enhances brain network features that predict individual intelligence in children with epilepsy. PLoS ONE 14(3):e0212901. https://doi.org/10.1371/journal.pone.0212901

30. Paldino MJ, Zhang W, Chu ZD, Golriz F (2017) Metrics of brain network architecture capture the impact of disease in children with epilepsy. NeuroimageClin 13:201–208. https://doi.org/10.1016/j.nicl.2016.12.005

31. Azencott R, Muravina V, Hekmati R, Zhang W, Paldino M (2019) Automatic clustering in large sets of time series. In: Contributions to partial differential equations and applications. Springer, Berlin, pp 65–75.

32. Rajpoot K, Riaz A, Majeed W, Rajpoot N (2015) Functional connectivity alterations in epilepsy from resting-state functional MRI. PLoS ONE 10(8):e0134944. https://doi.org/10.1371/journal.pone.0134944

33. Reshef DN, Reshef YA, Finucane HK, Grossman SR, McVean G, Turnbaugh PJ et al (2011) Detecting novel associations in large data sets. Science 334(6062):1518–1524. https://doi.org/10.1126/science.1205438

34. Rubinov M, Sporns O (2010) Complex network measures of brain connectivity: uses and interpretations. Neuroimage 52(3):1059–1069. https://doi.org/10.1016/j.neuroimage.2009.10.003

35. Rummel C, Abela E, Muller M, Hauf M, Scheidegger O, Wiest R et al (2011) Uniform approach to linear and nonlinear interrelation patterns in multivariate time series. Phys Rev E Stat Nonlin Soft Matter Phys 83(6 Pt 2):066215. https://doi.org/10.1103/PhysRevE.83.066215

36. Russ SA, Larson K, Halfon N (2012) A national profile of childhood epilepsy and seizure disorder. Pediatrics 129(2):256–264. https://doi.org/10.1542/peds.2010-1371

37. Simpson IJ, Cardoso MJ, Modat M, Cash DM, Woolrich MW, Andersson JL et al (2015) Probabilistic non-linear registration with spatially adaptive regularisation. Med Image Anal 26(1):203–216. https://doi.org/10.1016/j.media.2015.08.006

38. Sinha N, Dauwels J, Kaiser M, Cash SS, Brandon Westover M, Wang Y et al (2017) Predicting neurosurgical outcomes in focal epilepsy patients using computational modelling. Brain 140(2):319–332. https://doi.org/10.1093/brain/aww299

39. Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H et al (2004) Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage 23(Suppl 1):S208-219. https://doi.org/10.1016/j.neuroimage.2004.07.051

40. Spencer SS (2002) Neural networks in human epilepsy: evidence of and implications for treatment. Epilepsia 43(3):219–227

41. Supekar K, Uddin LQ, Khouzam A, Phillips J, Gaillard WD, Kenworthy LE et al (2013) Brain hyperconnectivity in children with autism and its links to social deficits. Cell Rep 5(3):738–747. https://doi.org/10.1016/j.celrep.2013.10.001

42. Tang Y, Long J, Wang W, Liao J, Xie H, Zhao G et al (2016) Aberrant functional brain connectome in people with antisocial personality disorder. Sci Rep 6:26209. https://doi.org/10.1038/srep26209

43. Terry JR, Benjamin O, Richardson MP (2012) Seizure generation: the role of nodes and networks. Epilepsia 53(9):e166-169. https://doi.org/10.1111/j.1528-1167.2012.03560.x

44. Tomlinson SB, Porter BE, Marsh ED (2017) Interictal network synchrony and local heterogeneity predict epilepsy surgery outcome among pediatric patients. Epilepsia 58(3):402–411. https://doi.org/10.1111/epi.13657

45. van den Heuvel MP, Mandl RC, Stam CJ, Kahn RS, HulshoffPol HE (2010) Aberrant frontal and temporal complex network structure in schizophrenia: a graph theoretical analysis. J Neurosci 30(47):15915–15926. https://doi.org/10.1523/JNEUROSCI.2874-10.2010

46. van den Heuvel MP, Sporns O (2013) Network hubs in the human brain. Trends CognSci 17(12):683–696. https://doi.org/10.1016/j.tics.2013.09.012

47. van den Heuvel MP, Stam CJ, Kahn RS, HulshoffPol HE (2009) Efficiency of functional brain networks and intellectual performance. J Neurosci 29(23):7619–7624. https://doi.org/10.1523/JNEUROSCI.1443-09.2009

48. Vapnik V (1998) Statistical learning theory, vol 3. Wiley, New York

49. Volz A (2008) A soviet estimate of german tank production. J Slavic Mil Stud 21(3):588–590

50. Wendling F, Bartolomei F, Senhadji L (2009) Spatial analysis of intracerebral electroencephalographic signals in the time and frequency domain: identification of epileptogenic networks in partial epilepsy. Philos Trans A Math Phys Eng Sci 367(1887):297–316. https://doi.org/10.1098/rsta.2008.0220

51. Wilke C, Worrell G, He B (2011) Graph analysis of epileptogenic networks in human partial epilepsy. Epilepsia 52(1):84–93. https://doi.org/10.1111/j.1528-1167.2010.02785.x

52. Zack MM, Kobau R (2017) National and State estimates of the numbers of adults and children with active epilepsy—United States, 2015. MMWR-Morb Mortal Wkly Rep 66(31):821–825. https://doi.org/10.15585/mmwr.mm6631a1

53. Zalesky A, Fornito A, Harding IH, Cocchi L, Yucel M, Pantelis C et al (2010) Whole-brain anatomical networks: does the choice of nodes matter? Neuroimage 50(3):970–983. https://doi.org/10.1016/j.neuroimage.2009.12.027

54. Zhang W, Muravina V, Azencott R, Chu ZD, Paldino MJ (2018) Mutual information better quantifies brain network architecture in children with epilepsy. Comput Math Methods Med 2018:6142898. https://doi.org/10.1155/2018/6142898

55. Zhang X, Tokoglu F, Negishi M, Arora J, Winstanley S, Spencer DD et al (2011) Social network theory applied to resting-state fMRI connectivity data in the identification of epilepsy networks with iterative feature selection. J Neurosci Methods 199(1):129–139. https://doi.org/10.1016/j.jneumeth.2011.04.020

56. Zinkus TP (2018) Pre-surgical planning: multimodality imaging to optimize outcomes in pediatric epilepsy surgery. Mo Med 115(4):365–367

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.