A Novel Graph Neural Network to Localize Eloquent Cortex in Brain Tumor Patients from Resting-State fMRI Connectivity

Naresh Nandakumar[1(✉)], Komal Manzoor[2], Jay J. Pillai[2], Sachin K. Gujar[2], Haris I. Sair[2], and Archana Venkataraman[1]

[1] Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, USA
nnandak1@jhu.edu

[2] Department of Neuroradiology, Johns Hopkins School of Medicine, Baltimore, USA

Abstract. We develop a novel method to localize the language and motor areas of the eloquent cortex in brain tumor patients based on resting-state fMRI (rs-fMRI) connectivity. Our method leverages the representation power of convolutional neural networks through specialized filters that act topologically on the rs-fMRI connectivity data. This Graph Neural Network (GNN) classifies each parcel in the brain into eloquent cortex, tumor, or background gray matter, thus accommodating varying tumor characteristics across patients. Our loss function also reflects the large class-imbalance present in our data. We evaluate our GNN on rs-fMRI data from 60 brain tumor patients with different tumor sizes and locations. We use motor and language task fMRI for validation. Our model achieves better localization than linear SVM, random forest, and a multilayer perceptron architecture. Our GNN is able to correctly identify bilateral language areas in the brain even when trained on patients whose language network is lateralized to the left hemisphere.

Keywords: Rs-fMRI · Graph Neural Network · Language localization

1 Introduction

The eloquent cortex consists of sensorimotor and language areas in the brain that are essential for human functioning. Given its importance, localizing and subsequently avoiding the eloquent cortex is a crucial step when planning a neurosurgery. However, this localization is challenging due to the varying anatomical boundaries of these networks and the effects of the tumor. For example, it has been shown that motor and language functionality in brain tumor patients can be displaced due to neural plasticity [1]. The gold standard for eloquent mapping is intraoperative electrical stimulation, which is highly invasive and requires the

M. D. Schirmer et al. (Eds.): CNI 2019, LNCS 11848, pp. 10–20, 2019.
https://doi.org/10.1007/978-3-030-32391-2_2

patient to be awake during surgery. The noninvasive alternative is task-fMRI. However, severely impaired patients, such as those with advanced brain tumors, may not be able to perform these tasks, thus reducing the reliability of the fMRI activation maps. Resting-state fMRI (rs-fMRI) captures spontaneous fluctuations in the brain, which can be used to identify functional systems in the absence of an experimental paradigm. Hence, rs-fMRI may provide an alternative for motor and language localization in critically ill patients [2].

Automatically identifying the eloquent cortex in brain tumor patients is a challenging problem with limited success in the literature. The work of [3] addresses the problem of shifting anatomical boundaries by matching functional brain regions across individuals via a diffusion map representation of task-fMRI. However, this method has yet to generalize to rs-fMRI. With regards to rs-fMRI, the work of [4] describes a method to obtain subject-specific functional parcellations of brain tumor patients using a Markov Random Field prior. However, this method is validated on a coarse functional parcellation which is unsuitible for presurgical mapping. The work of [5] describes a method to compute language laterality from rs-fMRI by comparing connectivity between fixed areas of expected language activation. However, this study stopped short of *localization*, which is the main clinical need. The authors of [2] propose a semi-automated method to determine the language network from group ICA maps of rs-fMRI data. However, this method relies on manual thresholding for each patient. Finally, the work of [6] describes a multi-layer perceptron to classify resting-state networks at the voxel level based on seed correlation maps, which was then extended to identify the language network in three separate tumor cases [7]. However, this method is computationally expensive, requires a large amount of training data, and has only been evaluated on a limited dataset.

In this paper, we propose the first end-to-end model that uses convolutional neural networks (CNNs) to identify eloquent cortex in brain tumor patients. Our problem loosely resembles image segmentation, for which deep learning approaches using CNNs have made great strides [8]. However, rs-fMRI captures correlated patterns of activity rather than local similarities, which cannot be represented by a traditional spatial convolution. Therefore, deep learning for rs-fMRI has focused almost exclusively on perceptron architectures [6] and patient wise classification [9], rather than network analysis. Our approach blends the ideas of image segmentation and functional network extraction. Namely, we construct a similarity graph from rs-fMRI data that summarizes functional connectivity between ROIs. These graphs are then input to a novel graph neural network (GNN) which leverages convolutional filters designed to act topologically upon similarity matrices [10]. The output of our GNN is a vector that classifies each node in the graph as either eloquent cortex, tumor, or background gray matter. We train and evaluate four separate GNN's to perform either language or motor classification. The motor classes are divided into three regions of the motor strip corresponding to finger, tongue, or foot movements. Our loss function reflects the large class-imbalance in our data, as eloquent cortex and tumor represent a

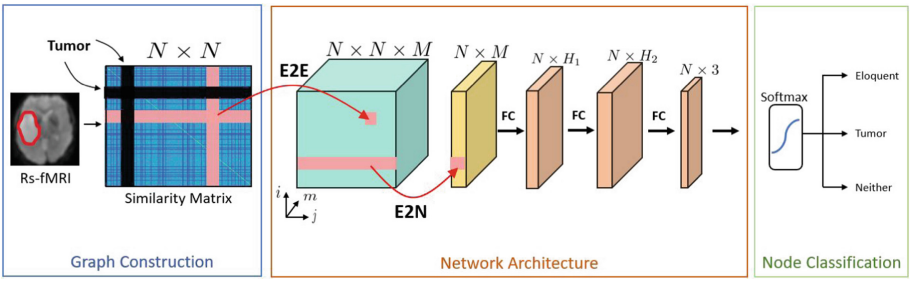

Fig. 1. The overall workflow of our model. Left: Graph construction encodes fMRI and tumor information. Middle: Our GNN architecture employs E2E, E2N, and FC layers for feature extraction. Right: We perform a node (parcel) identification task.

small fraction of the brain. Our model outperforms three baseline approaches in eloquent cortex detection, overall accuracy, and AUC.

2 A Graph Neural Network for Node Identification

The underlying assumption of our framework is that, while the anatomical boundaries of the eloquent cortex, particularly the language network, may shift, its connectivity with the rest of the brain will remain consistent [2]. We construct a weighted graph from the rs-fMRI data. We use a deep learning framework to capture complex interactions in the connectivity data. Our GNN node classifier extracts salient edge-node relationships and node features within the graph using a combination of specialized convolutional filters and fully-connected (FC) layers. An important distinction in our problem is the presence of large anatomical lesions, i.e, the brain tumors. Since the tumors often encroach into the gray matter, we introduce "missing" full rows and columns into our graph. These missing rows and columns are the most salient features of the data, therfore we introduce two baseline class labels, "tumor" and "background gray matter" to avoid biasing the algorithm. Figure 1 outlines our overall pipeline from graph construction to node classification.

Graph Construction. Let N be the number of brain regions in our parcellation and T be the number of time points for a rs-fMRI scan. We define $\mathbf{x}_i \in \mathbb{R}^{T\times 1}$ as the average time series extracted from parcel i. We construct graph $\mathbf{W} \in \mathbb{R}^{N\times N}$

$$\mathbf{W}_{i,j} = \exp\left[\frac{\langle \mathbf{x}_i, \mathbf{x}_j\rangle}{\epsilon} - 1\right] \tag{1}$$

where $\langle\cdot,\cdot\rangle$ represents the Pearson's correlation coefficient between time courses and $\epsilon \geq 1$ controls decay speed. By construction, rows and columns of $\mathbf{W}$ that correspond to tumor are "missing" and computationally set to zero to indicate that they are not functionally similar to any other region in the brain. Our choice of $\epsilon \geq 1$ along with the form of Eq. (1) asserts that $\mathbf{W}_{i,j} > 0$ for all non-tumor

regions. Therefore, even two healthy parcels with a strong negative correlation will still be more functionally similar than tumor regions in our model. Our framework assumes that tumor boundaries have been delineated for each patient.

Neural Network Architecture. Our GNN architecture employs both convolutional and FC layers to process node information. While traditional convolutional layers assume a grid-like organization to extract spatially local features, our GNN uses one edge-to-egde (E2E) and one edge-to-node (E2N) layer developed in [10], which act topologically on similarity graph data. These convolutional filters span full rows and columns of the graph and were originally designed to perform regression from diffusion MRI connectivity. Mathematically, an E2E filter is composed one row filter, one column filter, and a learned bias, which totals $2N + 1$ parameters. Let $m \in \{1, \cdots, M\}$ be the E2E filter index, $\mathbf{r}^m \in \mathbb{R}^{1\times N}$ be the m-th row filter, $\mathbf{c}^m \in \mathbb{R}^{N\times 1}$ be the m-th column filter and $\mathbf{b} \in \mathbb{R}^{M\times 1}$ be the E2E bias. The feature map $\mathbf{A}^m \in \mathbb{R}^{N\times N}$ output from E2E filter m and activation function $\boldsymbol{\phi}$ is computed as

$$\mathbf{A}^m_{i,j} = \boldsymbol{\phi}\Big(\sum_{n=1}^{N} \mathbf{r}^m_n \mathbf{W}_{i,n} + \mathbf{c}^m_n \mathbf{W}_{n,j} + \mathbf{b}_m\Big) \tag{2}$$

Intuitively, an E2E filter for node pair (i, j) computes a weighted sum of edge strengths over all edges connected to either node i or j. Even with symmetric input $\mathbf{W}$, E2E filters and corresponding feature maps are not necessarily symmetric. This asymmetry is desirable, as functional systems in the brain tend to be lateralized. We use the E2E layer to encode multiple different views (maps) of the edge-to-egde similarities within our connectome data.

The E2N layer condenses our representation from size $N \times N \times M$ after the E2E layer to $N \times M$, analogous to M features for each node. The E2N filter is simply a 1D convolution along the columns of each feature map. Let $\mathbf{g}^m \in \mathbb{R}^{N\times 1}$ be the m-th E2N filter and $\mathbf{d} \in \mathbb{R}^{M\times 1}$ be the E2N bias. The E2N output $\mathbf{a}^m \in \mathbb{R}^{N\times 1}$ from input $\mathbf{A}^m$ is computed as

$$\mathbf{a}^m_i = \boldsymbol{\phi}\Big(\sum_{n=1}^{N} \mathbf{g}^m_n \mathbf{A}^m_{n,i} + \mathbf{d}_m\Big). \tag{3}$$

Mathematically, the E2N filter computes a single value for each node i by taking a weighted combination of edges associated with it. Our motivation for using this layer is to collapse our representation along the second dimension to obtain M features for each node. This step is similar in nature to extracting graph theoretic features, such as node centrality. In particular, we have a representation that encodes the relationship each node has to its connectivity matrix [11].

Our node identification network uses a cascade of three FC layers of sizes $M \times H_1$, $H_1 \times H_2$ and $H_2 \times 3$ respectively. We apply activation functions between each layer. The FC layers find nonlinear combinations of the features to best discriminate class membership for each brain parcel. Overall, our network takes $N \times N$ input and outputs an $N \times 3$ matrix for classification. Notice that the

first input dimension N is maintained throughout our whole network and is not transformed. Therefore, our network maintains node structure to ultimately discriminate class membership for all nodes within one connectome at a time. As shown in Fig. 1, one design choice we make is to set $H_2 > H_1$. Empirically, this relationship robustly captures the structure of our class membership.

Weighted Loss Function. Naturally, there exists a large class imbalance in our setup, as the majority of nodes considered will be background gray matter. We cannot rely on traditional data augmentation techniques to mitigate this imbalance, as our model operates on whole-brain connectivity. To accomodate for the class imbalance, we train our model with a modified version of the Risk-sensitive cross-entropy (RSCE) loss function [12], which is designed to handle membership imbalance in multi-class classification. Let $\hat{y}_c^n$ be the output probability of our network for assigning node n to class c and y_c^n be 1 when node n belongs to class c and 0 otherwise. The loss function per patient is

$$\mathcal{L}(y_c^n, \hat{y}_c^n) = -\frac{1}{N} \sum_{n=1}^{N} \sum_{c=1}^{C} \delta_c \cdot y_c^n \log(\hat{y}_c^n) \tag{4}$$

where δ_c is the risk factor associated with class c. If δ_c is small, then we pay a smaller penalty for misclassifying samples that belong to class c. Our strategy is to penalize misclassifying eloquent nodes (false negatives) larger than misclassifying background (false positives) to encourage our model to learn the language and motor distributions given a small number of language training samples.

Neural Network Implementation Details. We implement our network in PyTorch using the SGD optimizer with weight decay $= 5 \times 10^{-5}$ for parameter stability, and momentum $= 0.9$ to improve convergence. For our model, $\epsilon = 1$ and layer dimensions are $M = 8$, $H_1 = 9$, $H_2 = 27$ and $C = 3$. We train our model with learning rate .005 and 80 epochs, which provides for reliable performance without overfitting. The LeakyReLU$(x) = \max(0, x) + 0.33 \cdot \min(0, x)$ activation function is applied at each hidden layer. Empirically, this activation function is robust to a range of initializations. A softmax activation is applied at the final layer for classification. After cross-validation, we set $\delta = (1.3, .3, .15)$ for the eloquent cortex, tumor, and background gray matter classes respectively. With GPU, total training time is within 3–5 min.

2.1 Baseline Comparisons

We evaluate the performance of our GNN against 3 baseline methods. The first baseline is a linear SVM based on graph theoretic measures: node degree, betweenness, closeness, and eigenvector centrality [11]. The second baseline is a random forest (RF) on the stacked rs-fMRI similarity features of each node. We omit tumor class and nodes for SVM and RF as the algorithms do not exploit the spatial consistency of the similarity matrix. The last baseline is a multi-layer perceptron (MLP) to observe how adding specialized E2E and E2N layers changes performance for this task. The MLP maintains the same input-output relationship, total parameter number, activations, and loss function as the GNN.

3 Experimental Results

Dataset and Preprocessing: We evaluate the GNN on rs-fMRI data from 60 patients who underwent preoperative mapping as part of their presurgical workup. The data was acquired using a 3.0 T Siemens Trio Tim (TR = 2000 ms, TE = 30 ms, FOV = 24 cm, res = $3.59 \times 3.59 \times 5$ mm). The fMRI was processed using SPM8. The steps include slice timing correction, motion correction and registration to the MNI-152 template. The rs-fMRI was bandpass filtered from 0.01 to 0.1 Hz, spatially smoothed with a 6 mm FWHM Gaussian kernel, scrubbed using the ArtRepair toolbox in SPM, linearly detrended, and underwent nuisance regression using CompCor.

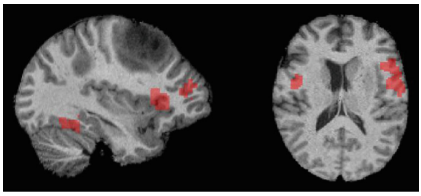

Fig. 2. Left: One left-hemisphere language network (red) subject. Right: One bilateral language network subject. (Color figure online)

Our dataset includes three different motor paradigms that were designed to target distinct parts of the motor homonculus [13]: finger tapping, tongue moving, and foot tapping. Since the task-fMRI data was acquired for clinical purposes, only 38 patients performed the finger task, 41 patients performed the tongue task, and 18 patients performed the foot task. Our ground truth language annotations are derived from task-fMRI activations of the same 60 patients during two language paradigms: sentence completion (SC) and silent word generation (SWG). Our dataset includes 55 patients with left-hemisphere language networks and 5 patients with bilateral networks. The fMRI underwent slice timing correction, motion correction and registration to the MNI-152 template. The General Linear Model (GLM) implemented in SPM8 was used to derive task-fMRI activation maps. The task activation maps were confirmed by an expert neuroradiologist as consistent with the information provided during presurgical planning. The tumor boundaries for each patient were manually delineated by a medical fellow using the MIPAV software. Figure 2 shows language areas (red) for two separate subjects to illustrate the heterogeniety of our cohort.

Implementation Details and Evaluation Criteria: We parcellate our rs-fMRI data using the Craddocks atlas [14] with the cerebellar regions removed due to inconsistent acquisition ($N = 384$). Due to different patients performing different tasks, we train and test four separate GNNs, one for language identification and the rest for each motor task. We assign a parcel to the eloquent cortex if a majority of its voxels coincided with the ground truth task activations. We

employ a ten-fold cross validation for the language experiment, and a five-fold cross validation for the motor experiments, as we observed the motor GNNs overfit more easily. For language, we stratify our folds by ensuring at most one bilateral language subject is in each fold. We report eloquent class accuracy as well as overall accuracy for each method that reflect a viable trade-off between true positive rate (TPR) and true negative rate (TNR). We compute and report area under the curve (AUC) by varying hyperparameter settings to approximate ROC. We consider eloquent vs. not eloquent for each ROC statistic reported. We maintain the same hyperparameter values across each of the four experiments. Tumor class accuracy is not reported, as both the MLP and GNN achieved near perfect (≈.995) accuracy due to the assumptions of our setup.

Table 1. Node identification statistics for motor tasks.

Task	Method	Motor	Overall	Sensitivity	Specificity	AUC
Foot	Linear SVM	0.48	0.52	0.46	0.49	0.52
	RF	0.36	0.77	0.34	**0.86**	0.59
	MLP	0.73	0.76	0.63	0.75	0.74
	GNN	**0.84**	**0.81**	**0.78**	0.79	**0.81**
Tongue	Linear SVM	0.58	0.59	0.57	0.55	0.56
	RF	0.42	0.77	0.34	**0.92**	0.65
	MLP	0.75	0.78	0.68	0.77	0.75
	GNN	**0.87**	**0.84**	**0.82**	0.80	**0.83**
Finger	Linear SVM	0.60	0.62	0.56	0.58	0.57
	RF	0.49	0.80	0.43	**0.92**	0.69
	MLP	0.82	0.76	0.78	0.73	0.80
	GNN	**0.88**	**0.87**	**0.84**	0.83	**0.86**

3.1 Motor Class Identification

The motor identification results are reported in Table 1. As seen, our GNN overall outperforms all baselines in nearly all metrics. This notable performance is especially highlighted in the AUC column, as our method has the best trade-off between TPR and FPR. Our results suggest that approaching this problem with a deep learning framework is favorable, as both neural networks outperform the traditional machine learning baselines. Furthermore, we show a marked increase in performance employing the specialized convolutional filters. This suggests that our network learns a more discriminative representation of eloquent cortex at rest than the MLP. Due to different patient subcohorts performing different tasks, we train and test on each motor task separately. Figure 3 shows a coronal view of ground truth (red) and predicted (blue) motor regions for the

foot (left) and tongue (right) tasks in one patient. As seen, our network is able to pinpoint both the midline and the peripheral areas of the motor strip. Our performance suggests that our method is able to localize specific parts of the motor homonculus, which is important in a preoperative setting.

Table 2. Node identification statistics for language class ($N = 60$).

Method	Language	Overall	Sensitivity	Specificity	AUC
Linear SVM	0.56	0.52	0.55	0.49	0.53
RF	0.37	0.77	0.33	**0.89**	0.63
MLP	0.66	0.73	0.61	0.76	0.70
GNN	**0.74**	**0.86**	**0.70**	0.77	**0.76**

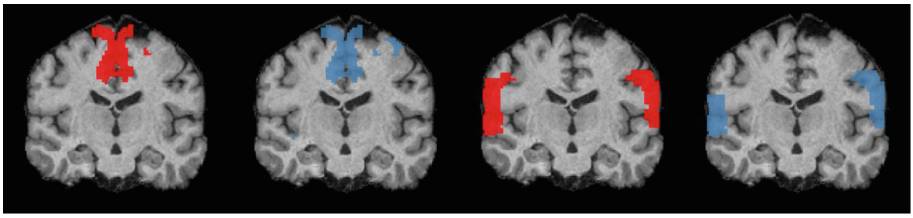

Fig. 3. Ground truth (red) and predicted (blue) motor regions for foot (left) and tongue (right) tasks in one patient. (Color figure online)

3.2 Language Class Identification

Table 2 reports the language identification performance across all methods. Once again, our GNN outperforms the baselines in nearly all methods, with the most notable gains in language accuracy and AUC. The specificity of our GNN is lower than expected due to the hemispheric symmetry of rs-fMRI data. We saw that the most frequent misclassification from our model was assigning contralateral parcels to the language class. We ran two experiments to probe whether our GNN is learning connectivity patterns associated with language rather than memorizing node locations. In Fig. 4 (left), we plot the histogram of true (pink) vs. predicted (blue) language parcels. The x-axis shows how frequently a certain group of parcels was assigned to language in the ground truth and predicted labels. Each bin represents a different group of parcels. Highlighted by the red box, we see that our model tends to overpredict the language class. We assert that this overprediction is viable due to the demands of the clinical application of our work. We compared the GNN output with seed based correlation analysis (SBA), where the "seed" for each patient is selected based on the ground truth task-fMRI activations. The average rs-fMRI time course within the seed location

is correlated with each of the average time courses defined by our parcellation. The correlation maps are thresholded at $\rho > 0.6$ to retain only the strong associations. Figure 4 (right) shows a representative example, where the color of each parcel represents the strength of the connection; red is closer to 0.6 and yellow is closer to 1. Highlighted by the white arrows, there is right-hemisphere over prediction (blue) from our model. However, as shown by the SBA map, these right-hemisphere parcels have high resting-state connectivity with the seed average time course. Our GNN achieves a median of **0.81** dice overlap between the predicted language areas and the seed based correlation maps.

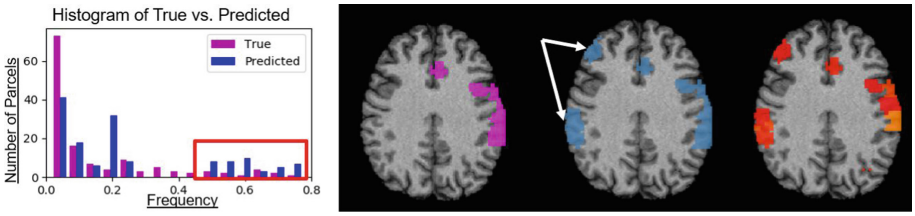

Fig. 4. Left: Histogram of true (pink) vs predicted (blue) language parcels for frequency > 0. Right: Arrows show overprediction overlaps with seed based maps. (Color figure online)

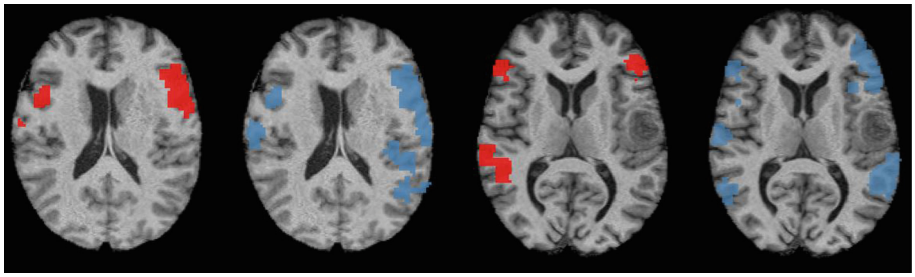

Fig. 5. Ground truth (red) and predicted (blue) for two separate subjects. All bilateral subjects were held out of training. (Color figure online)

Bilateral Language Identification: Our final experiment evaluates whether the GNN can recover a bilateral language network, even when this case is not present in the training data. Here, we trained the model on 55 left-hemisphere language network patients and tested on the remaining 5 bilateral subjects. Our model correctly predicted bilateral parcels in all five subjects. Figure 5 shows ground truth (red) and predicted language maps (blue) for two bilateral subjects. The median language class accuracy for these five cases was **0.62**. Empirically, this is slightly lower than reported in Table 2 due to the lack of training information. This experiment shows that our network learns connectivity patterns, rather than just spatial locations, of the language network.

4 Conclusion

We have demonstrated a GNN approach to identify the language and motor areas of eloquent cortex in brain tumor patients using rs-fMRI connectivity. Our model learns the resting-state functional signature of both the language and motor network within this tumor cohort by leveraging specialized convolutional filters that encode edge-node relationships within similarity matrices. With higher AUC for eloquent cortex detection, we prove that the features extracted from our GNN are more informative for this task than standard graph theoretic features and features extracted from a MLP. For language, we show that our model can correctly identify bilateral language networks even when trained on only unilateral network cases. Future work will decouple the lateralization problems in detecting language. We aim to add a separate network to our model that will determine which hemisphere(s) the language network is present in. We also aim to extend this work to simultaneously classify language and motor areas in one neural network, rather than training and testing these tasks separately.

Acknowledgements. This work was supported by the National Science Foundation CAREER award 1845430 (PI: Venkataraman) and the Research & Education Foundation Carestream Health RSNA Research Scholar Grant RSCH1420.

References

1. Duffau, H.: Lessons from brain mapping in surgery for low-grade glioma: insights into associations between tumour and brain plasticity. Lancet Neurol. **4**(8), 476–486 (2005)
2. Sair, H.I., et al.: Presurgical brain mapping of the language network in patients with brain tumors using resting-state fMRI: comparison with task fMRI. Hum. Brain Mapp. **37**(3), 913–923 (2016)
3. Langs, G., et al.: Functional geometry alignment and localization of brain areas. In: Advances in Neural Information Processing Systems, pp. 1225–1233 (2010)
4. Nandakumar, N., et al.: Defining patient specific functional parcellations in lesional cohorts via Markov random fields. In: Wu, G., Rekik, I., Schirmer, M.D., Chung, A.W., Munsell, B. (eds.) CNI 2018. LNCS, vol. 11083, pp. 88–98. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00755-3_10
5. Gohel, S., et al.: Resting-state functional connectivity of the middle frontal gyrus can predict language lateralization in patients with brain tumors. Am. J. Neuroradiol. (2019)
6. Hacker, C.D., et al.: Resting state network estimation in individual subjects. Neuroimage **82**, 616–633 (2013)
7. Lee, M.H., et al.: Clinical resting-state fMRI in the preoperative setting: are we ready for prime time? Top. Magn. Reson. Imaging: TMRI **25**(1), 11 (2016)
8. Kamnitsas, K., et al.: Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. MedIA **36**, 61–78 (2017)
9. Khosla, M., Jamison, K., Kuceyeski, A., Sabuncu, M.R.: 3D convolutional neural networks for classification of functional connectomes. In: Stoyanov, D., et al. (eds.) DLMIA/ML-CDS -2018. LNCS, vol. 11045, pp. 137–145. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00889-5_16

10. Kawahara, J., et al.: BrainNetCNN: convolutional neural networks for brain networks; towards predicting neurodevelopment. NeuroImage **146**, 1038–1049 (2017)
11. Opsahl, T., et al.: Node centrality in weighted networks: generalizing degree and shortest paths. Soc. Netw. **32**(3), 245–251 (2010)
12. Suresh, S., et al.: Risk-sensitive loss functions for sparse multi-category classification problems. Inf. Sci. **178**(12), 2621–2638 (2008)
13. Jack Jr., C.R., et al.: Sensory motor cortex: correlation of presurgical mapping with functional mr imaging and invasive cortical mapping. Radiology **190**(1), 85–92 (1994)
14. Craddock, R.C., et al.: A whole brain fmri atlas generated via spatially constrained spectral clustering. Hum. Brain Mapp. **33**(8), 1914–1928 (2012)