HHS Public Access

Cortical Surface Parcellation using Spherical Convolutional Neural Networks

Prasanna Parvathaneni[1,★], **Shunxing Bao**[1,★], **Vishwesh Nath**[1], **Neil D. Woodward**[2], **Daniel O. Claassen**[3], **Carissa J. Cascio**[2], **David H. Zald**[4], **Yuankai Huo**[1], **Bennett A. Landman**[1], **Ilwoo Lyu**[1]

[1]Electrical Engineering and Computer Science, Vanderbilt University, TN, USA

[2]Psychiatry & Behavioral Sciences, Vanderbilt University Medical Center, TN, USA

[3]Neurology, Vanderbilt University Medical Center, TN, USA

[4]Psychology, Vanderbilt University, TN, USA

Abstract

We present cortical surface parcellation using spherical deep convolutional neural networks. Traditional multi-atlas cortical surface parcellation requires inter-subject surface registration using geometric features with slow processing speed on a single subject (2–3 hours). Moreover, even optimal surface registration does not necessarily produce optimal cortical parcellation as parcel boundaries are not fully matched to the geometric features. In this context, a choice of training features is important for accurate cortical parcellation. To utilize the networks efficiently, we propose cortical parcellation-specific input data from an irregular and complicated structure of cortical surfaces. To this end, we align ground-truth cortical parcel boundaries and use their resulting deformation fields to generate new pairs of deformed geometric features and parcellation maps. To extend the capability of the networks, we then smoothly morph cortical geometric features and parcellation maps using the intermediate deformation fields. We validate our method on 427 adult brains for 49 labels. The experimental results show that our method outperforms traditional multi-atlas and naive spherical U-Net approaches, while achieving full cortical parcellation in less than a minute.

Keywords

cortical surface parcellation; spherical deformation; spherical U-Net; surface registration

1 Introduction

Regional-based morphological analysis is a widely adapted approach in neurodevelopmental studies. For valid regional analysis, cortical surfaces need to be consistently subdivided into multi-regions based on cortical parcellation protocols in anatomical or functional fashion [5,10,16]. However, consistent labeling of cortical regions is challenging due to the

ilwoo.lyu@vanderbilt.edu.
★P. Parvathaneni and S. Bao contributed equally to this work.

complicated cortical folds and inter-subject variability. Typically, manual labeling is tedious and time-consuming, and there exists labeling inconsistency even across experts. In contrast, a multi-atlas cortical parcellation approach [7] expedites the labeling task with algorithmic consistency. It generally tends to provide better performance as the number of atlases increases. Unfortunately, inter-subject registration is unavoidable in this approach to align multiple atlases to a target subject with significant computational demands proportional to the number of atlases.

With an increasing quantity of imaging data, convolutional neural networks (CNNs) are readily available to handle image segmentation problems on a structured grid. Yet, traditional CNNs architectures are still immature in handling non-uniform data with high complexity. This is mainly because spatial coherence incorporated with existing deep architectures is optimized on standard Euclidean image grids in addition to large memory requirement. In this regard, spherical CNNs recently emerge with efficient operations on a spherical domain. Cohen *et al.* and Esteves *et al.* [3,6] proposed spherical CNNs architectures to achieve computational efficiency as well as numerical accuracy. Although they work effectively on classification or regression tasks, semantic segmentation tasks were not fully addressed. Later, general semantic segmentation in a spherical domain was well discussed in [9].

Cortical surface mesh is of high complexity that still hampers practical use of existing CNNs due to their limited scalability on large size mesh. A few recent pioneering studies led into drawing the attention of CNNs to surface parcellation, unlike well-developed volumetric segmentation. Cucurull *et al.* [4] targeted cortical parcellation on only a few ROIs due to memory capacity. Gopinath *et al.* [8] proposed better capability with their graph CNNs for full cortical parcellation on adult brains with comparable results to a traditional approach [7]. The equal importance of training features is also emphasized in recent studies with the central theme being the specific design of the features for accurate cortical parcellation. For example, Gopinath *et al.* [8] utilized spectral features for better cortical alignments. Wu *et al.* [15] proposed geometry-aware spherical features to use a standard image CNNs architecture.

In this paper, we propose a novel cortical parcellation approach using a deep spherical U-Net [9] that can naturally encode relatively large surface mesh. In particular, we focus on parcellation-specific inputs and their augmentation for efficient utilization of the architecture and accurate parcellation results. Specifically, we compute deformation fields to generate deformed geometric features that best fit ground-truth parcel boundaries using a spherical surface registration method [12]. Since the networks lack generalization of input features, we further propose data augmentation driven by intermediate deformation fields rather than dipole moment variation that overcomes only rotational invariance. This can thus offer a rich set of plausible training samples by leveraging geometric features and their deformation. The key contributions include (i) novel features optimized over cortical parcel boundaries and (ii) data augmentation driven by their intermediate deformation fields. Figure 1 shows an overview of the proposed method.

2 Methods

2.1 Objective

We denote the ith cortical label by $z_i \in \mathbb{Z}^+$. Given N cortical labels $\mathbf{L} = \{z_1, \ldots, z_N\}$ and a cortical surface $\Omega \in \mathbb{R}^3$, our objective is to estimate a mapping $F: \Omega \to \mathbf{L}$ to determine a label for each cortical location.

2.2 Parcel Boundary Alignment

Deformation field.—Let $M: \mathbb{S}^2 \to \mathbb{S}^2$ denote a continuous spherical deformation field. Given $\mathbf{x} \in \Omega$ and its corresponding location $\hat{\mathbf{x}}$, the deformation field M holds

$$\hat{\mathbf{x}} = M(\mathbf{x}). \tag{1}$$

To estimate M, we use a spherical surface registration method [12][1] that reconstructs M by a linear combination of spherical harmonics coefficients; i.e., M is a function of spherical harmonics degree l. For convenience, let $M_l(\cdot)$ denote a deformation field at degree l in the remainder of the paper. An advantage of this method is to easily generate incremental deformation fields by adding basis functions due to orthonormality of spherical harmonics bases that smoothly morph subjects to a target template (e.g., M_0 = rigid body alignment, M_{10} = more local non-rigid deformation). Hence, once the deformation fields are computed with a high degree, the intermediate fields can be reconstructed without recomputing low degree again. We can then use all intermediate deformation fields for data augmentation by adding basis functions later.

Boundary map.—Optimal geometric alignment does not necessarily provide optimal cortical parcellation despite their high correlation (see precentral gyrus in Fig. 2 for example). Also, it is important in training to have well-shaped features. Therefore, we compute deformation fields that align parcel boundaries for more accurate prediction. To compute such deformation fields, we need two steps: (1) boundary extraction and (2) the extracted boundaries as a continuous function. Given ground-truth parcel labels F, we can obtain boundaries by finding points:

$$\partial \Omega = \{\mathbf{s} \in \Omega | F(\mathbf{s}) \neq F(\mathbf{x}): \forall \mathbf{x} \in \mathbf{N}(\mathbf{s})\}, \tag{2}$$

where $\mathbf{N}(\cdot)$ is a set of neighboring vertices on Ω. Now, we need to represent boundaries as a continuous function on Ω to allow derivatives required for the objective function in [12]. The idea is to compute the geodesic distance between Ω. Let $T(\mathbf{x}): \Omega \to \mathbb{R}^+$ denote the minimum travel-time Ω to $\forall \mathbf{x} \in \Omega$. The travel-time $T(\mathbf{x})$ holds the following eikonal equation with a unit propagation speed:

$$\begin{aligned} \|\nabla T(\mathbf{x})\| &= 1, & \mathbf{x} \in \Omega, \\ T(\mathbf{x}) &= 0, & \mathbf{x} \in \partial \Omega. \end{aligned} \tag{3}$$

The solution is thus equivalent to the geodesic distance from Ω. In this work, we use the ordered upwind method [14]. The distance map T is of different scale for each region across subjects. For better surface registration, we further normalize T with respect to the maximum distance per parcel, similar to the distance map normalization in [13], which provides consistent measurements across parcellation maps.

Deformed data.—For input features for training, we use standard cortical geometric features: mean curvature ($iH(\mathbf{x}) \in \mathbb{R}$) from inflated surface (for global cortical folding agreement), sulcal depth ($SD(\mathbf{x}) \in \mathbb{R}$) and mean curvature ($H(x) \in \mathbb{R}$) from cortical surface (for local cortical folding agreement). To create a template, we co-register training samples in an iterative averaging manner [11]. Here, we compute a distance map of the mode (most frequent) cortical labels across the training set after their registration to the template using the three geometric features. We then register the normalized distance map T to the template distance map at $I = 10$, which produces deformation fields M_{10}. Note that we found no noticeable improvement of the boundary alignment after I becomes greater than 10 in practice. Finally, the deformed data in training are given by a tuple $P_{10}(x) = [iH(M_{10}(\mathbf{x})), SD(M_{10}(\mathbf{x})), H(M_{10}(\mathbf{x})), F(M_{10}(\mathbf{x}))]$.

2.3 Data Augmentation

The proposed feature deformation is latent. It is valid only if an unseen surface has similar geometric patterns to the deformed features. However, it is unlikely to happen unless a fairly large number of training data are given, which suggests data augmentation to predict unseen data better. Thus, our goal is to generate intermediate deformed features between subjects and the target template. In this way, we can include smooth deformation trajectories as additional plausible training samples. Specifically, we create all intermediate samples as follows:

$$\bigcup_{l=0}^{10} \left\{ \left[iH\big(M_l(\mathbf{x})\big), SD\big(M_l(\mathbf{x})\big), H\big(M_l(\mathbf{x})\big), F\big(M_l(\mathbf{x})\big) \right] \right\}. \tag{4}$$

To exploit more samples, we also compute deformation fields that align the three geometric features to the template in a similar manner. Figure 2 illustrates an example of deformed features along their deformation trajectory.

2.4 Deep Learning Architecture

We adapt a state-of-the-art spherical U-net architecture designed for segmentation tasks [9] that can be naturally extended to cortical spherical parametrization. In this method, the convolutional kernels are predefined as differential operators for the 1st and 2nd derivatives, which yields fast convolution as well as superior performance over existing spherical networks in their benchmarks. In our framework, three geometric features with their augmentation are provided to input channels and N labels (after the deformation) to output channels. In training, we incrementally reconstruct deformation fields from 0 to 10, which generates 11×2 times of the original size of the training set (deformation driven by parcel boundary and geometric feature). At the end of the testing stage, we refine predicted

parcellation maps with a standard graph cut technique [2] to remove potential isolated regions and to create smooth parcel boundaries.

3 Experimental Setup

We used T_1-weighted scans on healthy adults ($n = 427$) from 23 to 34 years old, acquired from a Phillips 3T scanner. The cortical surfaces and their spherical mapping were reconstructed via a standard FreeSurfer pipeline with a large number of vertices ($\approx 160k$). We used only left hemispheres. The BrainCOLOR protocol [10] ($N = 49$ ROIs) was used for labeling with manual correction.

We trained the spherical U-Net on NVIDIA Titan Xp with a batch size of 4 at level 5 of the icosahedral subdivision due to memory capacity. We used the cross-entropy loss, and a total of 5,205,008 parameters were optimized by the Adam optimizer. The initial learning rate was set to 0.01 with a step decay of 0.9 per 20 epochs. We randomly divided our data into three sets: training (80%), validation (10%), and testing (10%). Thus, $385 \times 11 \times 2 = 8,470$ training samples were used in our framework after data augmentation. The optimal weights with the lowest validation loss were chosen up to 100 epochs, and each epoch took about 41 minutes for training of the 8,470 data. For a fair comparison, we applied the same graph-cut technique [2] on all the baseline methods. To avoid potential errors introduced by misalignment, we also used the aligned features rigidly to the template, i.e., $P_0(\mathbf{x}) = [iH(M_0(\mathbf{x})), SD(M_0(\mathbf{x})), H(M_0(\mathbf{x})), F(M_0(\mathbf{x}))])$.

4 Results

For proof of concept, we trained a spherical U-Net model [9] with the proposed deformed features driven by only M_{10}. From the testing set, we then provided the deformed geometric features P_{10} driven by their optimal boundary alignments. The Dice overlap was 88.53 ± 1.05%. This indicates that prediction is quite accurate if boundary-driven geometric features P_{10} are provided, which is a strong assumption in practice. We observed low Dice overlap of 78. ± 24 4.48% when we fed the rigid features P_0 from the same testing set to the networks, which is expected as the networks lack generalization. After the proposed data augmentation, we observed Dice overlap of 86.59 ± 1.53% closer to that with the deformed features driven optimal boundary alignment.

In comparison, we performed surface parcellation using multi-atlas and spherical U-Net [9] with P_0. In multi-atlas, we propagated labels from all training samples to a single subject after surface registration [12], and their final labels were determined by majority voting. Such a large number of atlases (= 385) generally results in accurate parcellation due to less bias to atlas selection with computational demands (about a day: registration for 3–5 minutes per atlas). Also, the spherical U-Net was trained with P_0. We note that the spherical U-Net with P_0 is presented in this paper first time for evaluation.

The Dice overlap was 82.73 ± 1.86% and 85.23 ± 1.57% for multi-atlas and spherical U-Net approaches, respectively. Of these approaches, ours achieved the highest Dice overlap with statistical significance in paired t-tests ($p < 0.05$). Note that both spherical U-Net and our approach used exactly the same input features P_0 and no deformed features were provided

(i.e., no registration step involved), which hence yields very fast cortical parcellation ($<$ a minute). Figure 3 shows an example of resulting cortical parcellation maps for the three approaches. We further performed paired t-tests to observe ROI-wise improvement on individual parcels. The test statistics revealed that our approach significantly improved parcellation accuracy after false discovery rate (FDR) [1] for multi-comparison correction (q = 0.05). Our approach outperforms multi-atlas (46 regions) and spherical U-Net (24 regions) as shown in Fig. 4. It is noteworthy that no regions were found with significantly reduced Dice overlap.

5 Conclusion

We presented a cortical parcellation method using spherical U-Net with novel features optimized over cortical parcellation boundaries. To enhance the capability of the spherical U-Net, we also incorporated intermediate deformed features along trajectories of the deformation fields. In the experiments, the proposed method achieved qualitatively and quantitatively better performance. Furthermore, full cortical parcellation was obtained in less than a minute.

Acknowledgments.

References

1. Benjamini Y, Hochberg Y: Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal statistical society: series B (Methodological) 57(1), 289–300 (1995)

2. Boykov Y, Kolmogorov V: An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. IEEE Transactions on Pattern Analysis & Machine Intelligence 26(9), 1124–1137 (2004) [PubMed: 15742889]

3. Cohen TS, Geiger M, Köhler J, Welling M: Spherical cnns. arXiv preprint arXiv:1801.10130 (2018)

4. Cucurull G, Wagstyl K, Casanova A, Veli kovi P, Jakobsen E, Drozdzal M, Romero A, Evans A, Bengio Y: Convolutional neural networks for mesh-based parcellation of the cerebral cortex. In: International Conference on Medical Imaging with Deep Learning (2018)

5. Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL, Dale AM, Maguire RP, Hyman BT, et al.: An automated labeling system for subdividing the human cerebral cortex on mri scans into gyral based regions of interest. Neuroimage 31(3), 968–980 (2006) [PubMed: 16530430]

6. Esteves C, Allen-Blanchette C, Makadia A, Daniilidis K: Learning so(3) equivariant representations with spherical cnns. In: European Conference on Computer Vision (2018)

7. Fischl B, Van Der Kouwe A, Destrieux C, Halgren E, Ségonne F, Salat DH, Busa E, Seidman LJ, Goldstein J, Kennedy D, et al.: Automatically parcellating the human cerebral cortex. Cerebral cortex 14(1), 11–22 (2004) [PubMed: 14654453]

8. Gopinath K, Desrosiers C, Lombaert H: Graph convolutions on spectral embeddings: Learning of cortical surface data. arXiv preprint arXiv:1803.10336 (2018)

9. Jiang C, Huang J, Kashinath K, Marcus P, Niessner M, et al.: Spherical cnns on unstructured grids. arXiv preprint arXiv:1901.02039 (2019)

10. Klein A, Dal Canton T, Ghosh SS, Landman B, Lee J, Worth A: Open labels: online feedback for a public resource of manually labeled brain images. In: 16th Annual Meeting for the Organization of Human Brain Mapping (2010)

11. Lyttelton O, Boucher M, Robbins S, Evans A: An unbiased iterative group registration template for cortical surface analysis. Neuroimage 34(4), 1535–1544 (2007) [PubMed: 17188895]

12. Lyu I, Kang H, Woodward ND, Styner MA, Landman BA: Hierarchical spherical deformation for cortical surface registration. Medical image analysis 57, 72–88 (2019) [PubMed: 31280090]

13. Lyu I, Kim SH, Girault JB, Gilmore JH, Styner MA: A cortical shape-adaptive approach to local gyrification index. Medical image analysis 48, 244–258 (2018) [PubMed: 29990689]

14. Sethian JA, Vladimirsky A: Ordered upwind methods for static hamilton–jacobi equations: Theory and algorithms. SIAM Journal on Numerical Analysis 41(1), 325–363 (2003)

15. Wu Z, Li G, Wang L, Shi F, Lin W, Gilmore JH, Shen D: Registration-free infant cortical surface parcellation using deep convolutional neural networks In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 672–680. Springer (2018)

16. Yeo BT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, Roffman JL, Smoller JW, Zöllei L, Polimeni JR, et al.: The organization of the human cerebral cortex estimated by intrinsic functional connectivity. Journal of neurophysiology 106(3), 1125 (2011) [PubMed: 21653723]
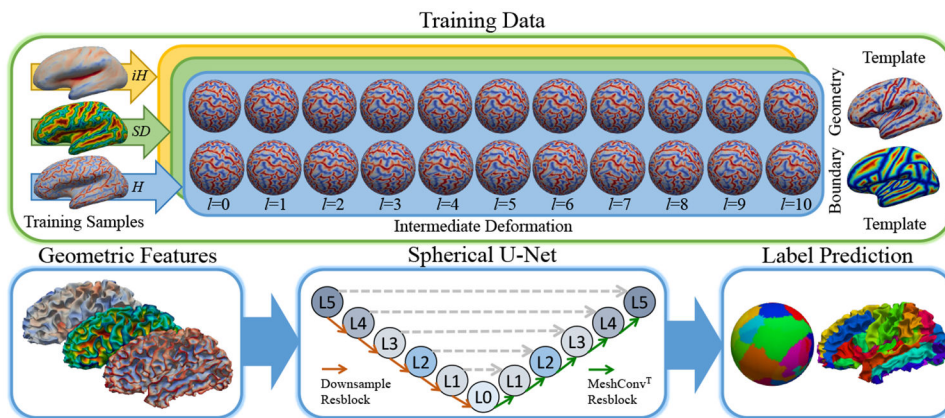
Fig. 1.
An overview of the proposed method. Three geometric features (*iH*, *SD*, *H*) are used for training the spherical U-Net to predict 49 cortical parcellation labels. For each geometric property, intermediate deformation fields draw a total of 11+11 respective samples after boundary and geometric alignment for data augmentation. The cortical parcellation is then performed using the original geometric features of testing subjects.
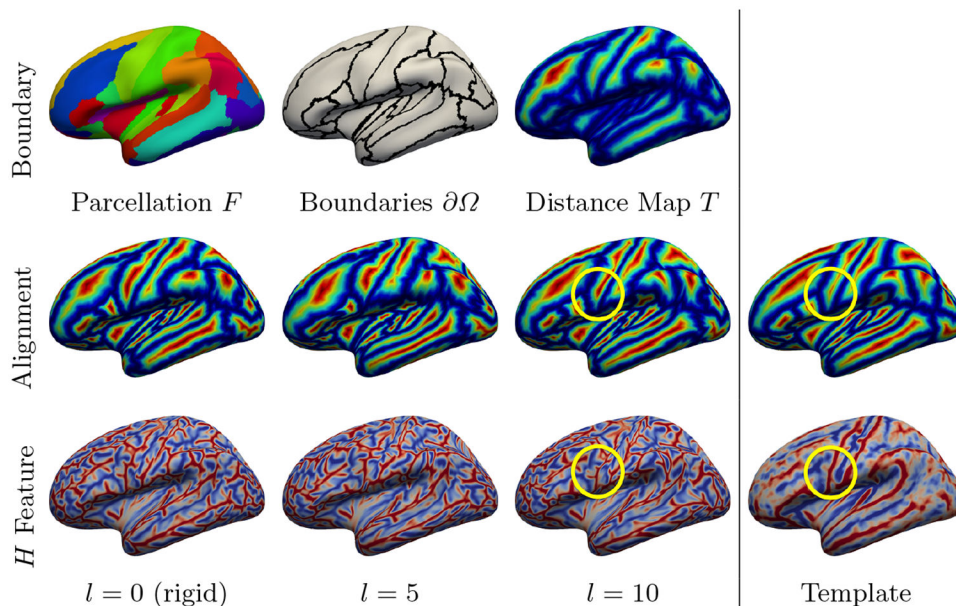
Fig. 2.
Boundary extraction and alignment. (*1st row*) For inputs for training, parcel boundaries are obtained from ground-truth labels (Eq. (2)). The boundaries are used to generate distance map *T* by solving an eikonal equation, and (*2nd row*) smooth trajectory of its deformation to a template is represented by increasing spherical harmonics degree *l*. (*3rd row*) The features for training are accordingly deformed by the deformation fields obtained by the boundary alignment. Note that these boundaries are quite well matched to those of the template, whereas their corresponding deformation on mean curvature *H* does not fully agree with that of the template (*yellow* circles).

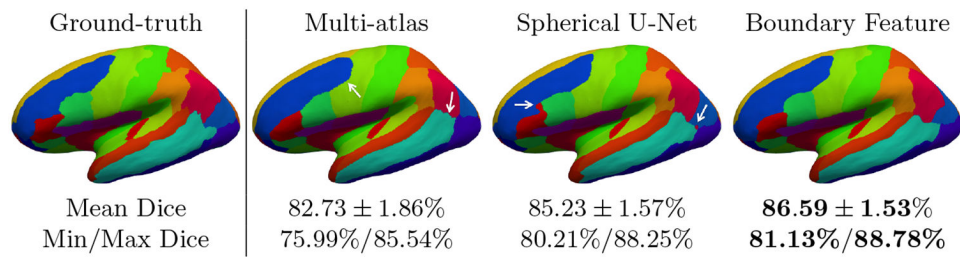|	Ground-truth	Multi-atlas	Spherical U-Net	Boundary Feature
Mean Dice		$82.73 \pm 1.86\%$	$85.23 \pm 1.57\%$	$\mathbf{86.59 \pm 1.53\%}$
Min/Max Dice		$75.99\%/85.54\%$	$80.21\%/88.25\%$	$\mathbf{81.13\%/88.78\%}$

Fig. 3.
Qualitative comparison: ground-truth, multi-atlas, spherical U-Net, and spherical U-Net with the proposed features. Our approach shows better performance than the other methods. The arrows highlight the mismatching regions to the ground-truth.
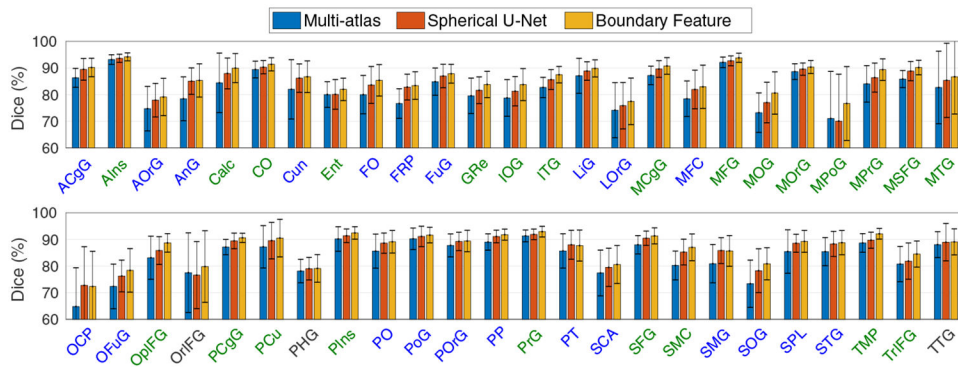
Fig. 4.

Dice overlap of 49 regions on the left hemisphere. Paired *t*-tests reveal improved regions with statistical significance after the FDR correction ($q = 0.05$). 46 and 24 out of 49 regions are improved against multi-atlas and spherical U-Net approaches, respectively. The color in the *x*-axis labels indicates the improved regions: multi-atlas (*blue*), both approaches (*green*), and no improvement (*black*).