



Improving human cortical sulcal curve labeling in large scale cross-sectional MRI using deep neural networks

Prasanna Parvathaneni^{a,*}, Vishwesh Nath^b, Maureen McHugo^c, Yuankai Huo^a, Susan M. Resnick^d, Neil D. Woodward^c, Bennett A. Landman^{a,b,c}, Ilwoo Lyu^{b,*}

^a Electrical Engineering, Vanderbilt University, Nashville, TN, USA

^b Computer Science, Vanderbilt University, Nashville, TN, USA

^c Department of Psychiatry and Behavioral Science, Vanderbilt University, Nashville, TN, USA

^d National Institutes of Health, Bethesda, MD, USA

ARTICLE INFO

Keywords:

Sulcal curve
Sulcal labeling
Cortical surface
DNN
Shape
Analysis

ABSTRACT

Background: Human cortical primary sulci are relatively stable landmarks and commonly observed across the population. Despite their stability, the primary sulci exhibit phenotypic variability.

New Method: We propose a fully automated pipeline that integrates both sulcal curve extraction and labeling. In this study, we use a large normal control population ($n = 1424$) to train neural networks for accurately labeling the primary sulci. Briefly, we use sulcal curve distance map, surface parcellation, mean curvature and spectral features to delineate their sulcal labels. We evaluate the proposed method with 8 primary sulcal curves in the left and right hemispheres compared to an established multi-atlas curve labeling method.

Results: Sulcal labels by the proposed method reasonably well agree with manual labeling. The proposed method outperforms the existing multi-atlas curve labeling method.

Comparison with Existing Method: Significantly improved sulcal labeling results are achieved with over 12.5 and 20.6 percent improvement on labeling accuracy in the left and right hemispheres, respectively compared to that of a multi-atlas curve labeling method in eight curves ($p < 0.001$, two-sample t-test).

Conclusion: The proposed method offers a computationally efficient and robust labeling of major sulci.

1. Introduction

The human cortex is one of the most complex anatomical structures with substantial variation in shape across individuals. Despite its complexity, the cortical sulci are known as relatively stable regions that embed consistent cortical folds (Gogtay et al., 2004; Kochunov et al., 2005; Northoff et al., 1999) by which the cerebral cortex can be subdivided into functionally and structurally homogeneous regions. From a morphological view, each individual sulcus (or sulcal region) can be well represented as a curve by tracing its sulcal fundus. By taking advantage of such a representation, sulcal curves have played key roles as a distinguishing indicator in cortical surface registration (Kim et al., 2005), brain development and degeneration (Gogtay et al., 2004; Sowell et al., 2003) and morphological variability (Fillard et al., 2007; Amunts et al., 2005). Yet, a concrete representation of cortical sulci is not commonly agreed upon due to an unclear anatomic boundary between sulci and gyri (Rademacher et al., 1992; Caviness et al., 1996; Lyu et al., 2010). With recent advance of 3D cortical surface

reconstruction techniques; (Kim et al., 2005; Ono et al., 1990; Lyu et al., 2017; Thompson et al., 1996) cortical geometric features greatly support sulcal curve extraction (Agrawal et al., 2017; Lohmann et al., 1999; Rettmann et al., 2002; Shattuck et al., 2009).

We focus on major sulcal curves that are relatively stable across subjects. These curves can be used as robust features for shape correspondence for reducing the spatial ambiguity in surface registration (Lyu et al., 2010; Thompson et al., 2004; Joshi et al., 2012a; Tao et al., 2002; Lyu et al., 2015; Agrawal et al., 2017). Moreover, these curves (Ono et al., 1990) could serve as biomarkers that aid in understanding the developmental growth or disease conditions in the human cortex. In this context, improved accuracy in sulcal curve labeling is essential and with the proposed method, we can label sulcal curves at much faster rate with high accuracy.

Despite the success in extracting sulcal curves, labeling of these curves is still an open and challenging problem because of high complexity and variability in cortical folding patterns (Lyu et al., 2017). For example, the inferior temporal sulcus is highly variable with several

* Corresponding authors.

E-mail addresses: Prasanna.Parvathaneni@vanderbilt.edu (P. Parvathaneni), Ilwoo.Lyu@vanderbilt.edu (I. Lyu).

<https://doi.org/10.1016/j.jneumeth.2019.108311>

Received 23 January 2019; Received in revised form 24 May 2019; Accepted 11 June 2019

Available online 12 June 2019

0165-0270/© 2019 Elsevier B.V. All rights reserved.

discontinuous pieces (Ono et al., 1990). Such variability hampers consistent labeling of cortical sulci, which is challenging even for neuro-anatomists. Sulcal labeling could be sensitive to variations in labeling protocols, and even in the same protocol, delineation of sulci can vary across experts (e.g., inconsistent endpoint delineation) (Ono et al., 1990; Thompson et al., 1996; Lohmann et al., 1999). Although one could employ cortical parcellation for sulcal labeling, it could still have challenges, as sulci do not always obey cortical parcellation boundaries. Accurate labeling is essential to derive meaningful inferences from brain-related changes in health and disease.

Manual sulcal labeling is a tedious task and needs a very high level of expertise in achieving high accuracy that agrees with sulcal nomenclature (Ono et al., 1990). This necessitates well-developed neuroanatomical conventions in the existence of high individual variability (Ono et al., 1990; Thompson et al., 1996; Lohmann et al., 1999). Joshi et al. (Joshi et al., 2012a) proposed a multi-atlas labeling method that uses predefined curves. In these approaches, sulcal curves from an atlas brain surface are projected onto the target subject surface and evolve along sulcal fundi via a level-set approach. Statistical shape models (Tao et al., 2002) define a shape variation prior as projections of landmarks onto tangent planes to the sphere. A watershed approach (Rettmann et al., 2002) extracts regions around sulcal fundi that embed some meaningful geometric characteristics like geodesic depth. However, their method is not fully automated, as sulcal regions are manually selected from the extracted sulcal regions. Another semi-automated approach (Shattuck et al., 2009) guides the user through a cortical surface delineation protocol implemented as a tool with an interface incorporated into BrainSuite software (Greenspan et al., 2016) reducing the inter-rater variability. All these manual/semi-automatic methods need manual intervention in addition to very long processing time.

In the last decades, fully automated techniques have been proposed to overcome the need for manual intervention. For example, (Shi et al., 2009) proposed a probabilistic graphical model of sulci, from which major sulcal curves are jointly labeled by solving a maximum a posteriori (MAP) estimation. (Tu et al., 2007) used a discriminative model using a boosting tree to extract major cortical sulci without employing any user-defined rules. On the other hand, (Joshi et al., 2012b) proposed an atlas registration method that transfers the sulcal curves from an atlas surface to the subject surface and then refines the sulcal curves' locations to closely follow the sulcal fundi using geodesic curvature flow. However, this approach uses only a single subject atlas. (Lyu et al., 2010) proposed a multi-atlas sulcal curve labeling using spectral point matching. This method does not require surface registration. (Mangin et al., 2015) noted that "the future of sulcus recognition is in pattern matching methods informed by a very large dataset of manually labeled sulci". In this context, readily accessible large scale neuroimaging dataset is highly likely to capture sulcal variability with the improved accuracy of labeling the sulcal curves on unseen subjects. As the gold standard is not available, learning variability in the cortical folding patterns from big data can better infer new models that encode the complex folding patterns. However, handling such huge data with existing methods could be practically implausible or requires a well-defined prior model that can fully incorporate sulcal variability (Ono et al., 1990; Shattuck and Leahy, 2002; Pantazis et al., 2010).

Recently, deep neural networks (DNN) have become popular in the medical imaging field (Greenspan et al., 2016; Litjens et al., 2017) involving large datasets due to the accuracy, speed and flexibility offered by these models. In neuroimaging, deep learning has shown its success in a wide range of applications including anatomical brain segmentation (de Brebisson and Montana, 2015), brain tumor segmentation (Havaei et al., 2017), biological psychiatry (Mechelli, 2018), deep unsupervised learning in traumatic brain injury (Minaee et al., 2018), epileptic discharge detection for EEG-fMRI (Hao et al., 2018), segmentation of deep brain regions in MRI and ultrasound (Milietari et al., 2017), inter-scanner harmonization in diffusion MRI (Nath et al., 2018), understanding sensory cortex (Yamins and DiCarlo, 2016), et

cetera. Convolution neural nets (CNN) are a specific type of DNN that uses convolution and pooling layers. These are widely used for image recognition tasks as they reflect the translation-invariant nature of most images. U-Net (Ronneberger et al., 2015) is a deep CNN model that is adapted from fully convolutional networks. It can work with very few training images and provide more precise segmentations. CNNs are successfully applied for segmentation and other applications in volume images. However, relatively little work has applied these approaches to 2-manifolds of cortical surfaces. For example, the application of CNNs is extended onto non-structured data with geometric deep learning (Bronstein et al., 2017; Boscaini et al., 2016). (Seong et al., 2018)) implemented a graph CNN that samples the data over a surface and reshapes the data to make it compatible with conventional CNN toolbox. (Cucurull et al., 2018)) used graph based methods for performing parcellation of two regions (44 and 45) in Broca's area using structural and functional features on the cortical surface patches.

In this study, our goal is to label sulcal curves by using a U-Net model (Ronneberger et al., 2015). The novelty of our approach is a new application of deep learning on cortical surfaces using a U-Net model. We aim to achieve substantively faster labeling with better accuracy than multi-atlas approaches (Lyu et al., 2010; Joshi et al., 2012a). Briefly, we use a large dataset of healthy controls to train a neural network for performing sulcal curve extraction and labeling. We train samples with pre-labeled sulcal curves and geometric feature maps. These features are mapped onto a 2-D polar plane to fully utilize a 2-D U-Net model. Specifically, we use the following geometric features: sulcal curve distance map, mean curvature, surface parcels, and spectral features (eigenfunctions of 2-manifold). We enhance the accuracy of the model using an independent dataset of manual labels. In the experiments, we show that our approach outperforms the standard curve labeling method (Lyu et al., 2010) that is considered as reference. Also sulcal curve labeling is performed in less than a minute compared to that of 32 minutes for baseline method.

2. Methods

2.1. Data acquisition

We analyzed structural data from two cohorts of participants. We first constructed a sulcal curve atlas dataset using 21 subjects from the publicly available KIRBY21 database (11 males; 10 females; age range = 22–61) (Lebed et al., 2013). Structural images in the atlas dataset were acquired with (3T Philips MPRAGE sequence with a $1 \times 1 \times 1.2 \text{ mm}^3$ resolution and an FOV of $240 \times 204 \times 256 \text{ mm}^3$). Our primary analyses were carried out using data from 784 participants in the Baltimore Longitudinal Study of Aging (BLSA) (349 males; 435 females; mean age = 72 years; range = 25–99) (Shock, 1984). Participants in this study were scanned 1–7 times over a period of 8.5 years on a Philips 3 T scanner using a 3D "magnetization prepared rapid gradient echo" (MPRAGE) sequence. Each image had $170 \times 256 \times 256$ voxels with $1.2 \times 1 \times 1 \text{ mm}^3$ resolution. The local Institutional Review Boards approved the study, and all participants provided written informed consent at each visit. From the BLSA dataset, we created 3 samples of participants. The training set consisted of 1–7 longitudinal scans from 759 individuals for a total of 1374 separate MRI sessions. The validation dataset was constructed from single sessions from 28 individuals that are not included as part of the training set. Single session data from 22 participants not included in the training or validation datasets were used to create a test dataset.

2.1.1. Sulcal curve atlas

For sulcal curve atlases from the KIRBY21 dataset, eight primary sulcal curves were manually labeled on both hemispheres by an expert according to a published sulcal curve labeling protocol (Pantazis et al., 2010): central sulcus (CS), superior temporal sulcus (STS), superior frontal sulcus (SFS), inferior frontal sulcus (IFS), occipitotemporal

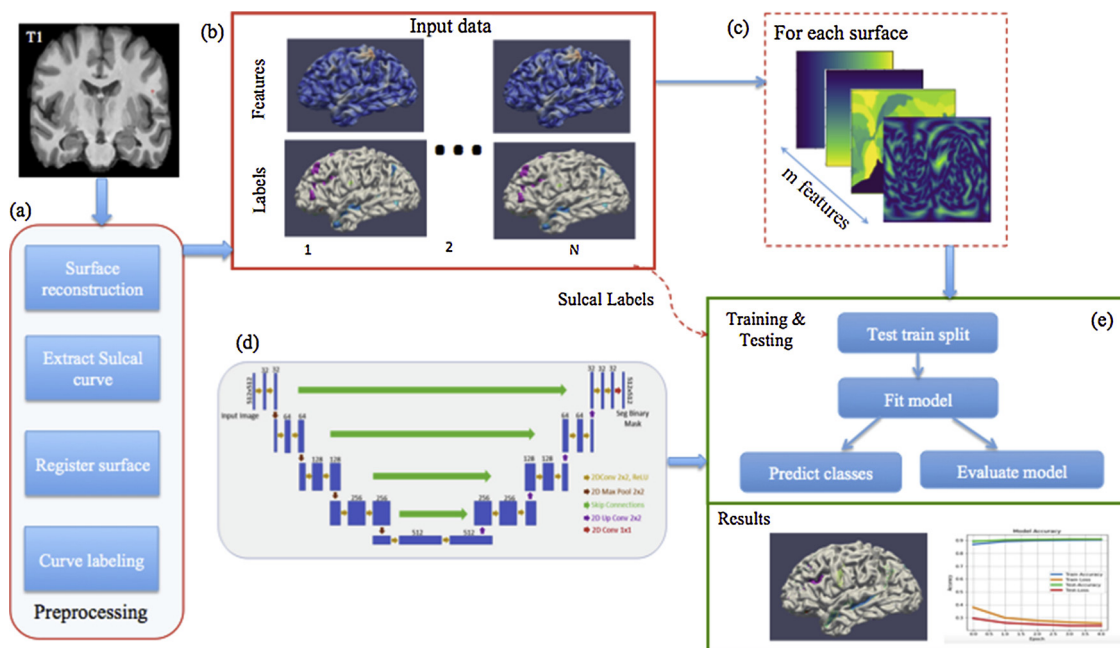


Fig. 1. An overview of the proposed method: (a) Preprocessing steps include surface reconstruction, sulcal curve extraction, surface registration, and curve-labeling. (b) Cortical features are generated including distance maps, spectral features, mean curvature, parcellation, and sulcal labels. (c) Features defined on spherical polar coordinates are mapped onto a uniformly spaced grid. (d) The features are then passed into 2-D U-net model and fitted with batch size of 10 for maximum of 20 epochs. (e) Labels are predicted on the test set for evaluation.

sulcus (OTS), cingulate sulcus (CingS), calcarine sulci (CalcS), and olfactory sulcus (OLF) (see Fig. 5 for examples).

2.2. Preprocessing

Structural data were processed for cortical surface reconstruction using the MaCRUISE pipeline (Huo et al., 2016). Sulcal curves were extracted using the TRACE method (Lyu et al., 2017). Sulcal curves were labeled using a curve labeling protocol (Lyu et al., 2010). All datasets were quality checked for each label and any faulty surfaces or extremely inconsistent curve labeling results are excluded. Note that we used these labels for training the deep learning model using various geometric features including distance map, surface parcels, mean curvature and spectral features. Cortical features were finally projected onto a plane to feed CNN for training with the U-Net model. An overview of the proposed workflow is shown in Fig. 1.

2.2.1. Sulcal curve extraction

The TRACE method (Lyu et al., 2017) was applied on each hemisphere to extract a set of sulcal curves along the sulcal fundic regions. Briefly, candidate sulcal points were selected to form a topological graph of the points, and Dijkstra trajectories over the graph to delineate optimal sulcal curves. This includes primary and secondary sulcal curves as a set of points (a subset of vertices) without curve labels, which needs to be pruned and labeled through the proposed pipeline. In this paper, we used default parameter settings as suggested in (Lyu et al., 2017). The code is available at <https://github.com/ilwoolyu/CurveExtraction>.

2.2.2. Surface registration

Although surface registration is not a mandatory step for the reference method (Lyu et al. (2010)) it could improve the quality of sulcal curve-labeling after surface registration. We first mapped each hemisphere onto the unit sphere while minimizing area distortion (Dale et al., 1999) and then established cortical surface correspondence (Lyu et al., 2018) (<https://github.com/ilwoolyu/HSD>). A custom template was obtained by averaging in a group-wise fashion (Lyu et al., 2018) of

co-registered 21 subjects from the publicly available dataset (Landman et al., 2011) of the Kirby Research Center for Functional Brain Imaging in Baltimore (<http://mri.kennedykrieger.org/databases.html>).

2.2.3. Sulcal curve labeling for model training

We used the standard curve labeling method (Lyu et al., 2010) to assign labels to primary sulcal curves by pruning minor sulcal curves and branches. The reference method employs multi-atlas to determine a label for each individual sulcus. It establishes a point-wise curve correspondence with each atlas and finds the best match across the established correspondence. However, we found in this work that a majority vote shows better performance since the best match can sometimes work poorly if only partial perfect match (few points) with a particular atlas yields the highest score among the other atlases. Therefore, final labels were assigned if at least half of atlases agree.

2.3. Deep neural network for sulcal curve labeling

We used cortical geometric features to capture both cortical folding patterns and individual variability. First, the sulcal curve distance map (geodesic distances between sulcal curves) was generated to represent cortical folding patterns. As complementary features to cortical folds, we used mean curvature and surface parcellation labels (Huo et al., 2018a). In addition, spectral features were computed to capture intrinsic geometric characteristics in the embedding space being spanned by the eigenvectors associated with the first five smallest eigenvalues. These features were fed into the neural network for sulcal label prediction.

2.3.1. Sulcal curve distance map

To represent cortical folding patterns, we computed a geodesic distance map $u: \mathbb{R}^3 \rightarrow \mathbb{R}$ on the cortical surface Ω between the sulcal curves. We set all the extracted sulcal curves as a source $C \subseteq \Omega$. By letting $c \subseteq C$, the minimum travel time $u(x): \mathbb{R}^3 \rightarrow \mathbb{R}$ from the source to any point $x \in \Omega$ can be obtained from the following propagation equation with some speed function $\exists F \in \mathbb{R}^+$,

$$u(c) = 0 \tag{1}$$

$$\|\nabla u(\mathbf{x})\|F\left(\mathbf{x}, \frac{\nabla u(\mathbf{x})}{\|\nabla u(\mathbf{x})\|}\right) = 1 \tag{2}$$

This wavefront propagation formulation is a static Hamilton-Jacobi partial differential equation. We set a constant speed function in every direction by letting $F = 1$ as follows:

$$\|\nabla u(\mathbf{x})\| = 1 \tag{3}$$

By solving Eq. (3), we have the minimum travel time with unit speed. This is equivalent to the geodesic distance between sulcal curves, which generates the geodesic distance map u of the sulcal curves. Typically, gyral regions have high values whereas sulcal regions have a zero distance. We later feed this map to the network for training.

Given a threshold $\xi \in \mathbb{R}^+$, sulcal regions can be segmented by a binary mask $M: \mathbb{R} \rightarrow [0,1]$ as follows

$$M(u; \xi) = f(\mathbf{x}) = \begin{cases} 1, & \text{if } u \leq \xi \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

This threshold widens the extracted sulcal curves to prevent them to being too narrow to provide sufficient geometric information within sulcal regions. We empirically set $\xi = 10$ to sufficiently cover regions along a single sulcus similar to the previous studies (Lyu et al., 2017; Lebed et al., 2013).

2.3.2. Spectral features of cortical surface

Spectral features were generated for each cortical surface to feed the neural network. We first build a graph $G = \{V, E\}$, where V is a set of vertices and E is a set of edges. We then setup a $|V|$ by $|V|$ weighted adjacency matrix W that stores node affinities. A diagonal node degree matrix D encodes the sum of all the point affinities (vertex degree) at point i .

$$d_i = \sum_{j=1}^{|V|} W_{ij}. \tag{5}$$

A diagonal node-weighting matrix G is given by the exponential of negative mean curvature. The node weighting at point i is defined as,

$$g_i = e^{-h_i} \tag{6}$$

where h_i is mean curvature at point i . Finally a general Laplacian operator is formulated on the connectivity adjacency matrix as the following $|V|$ by $|V|$ matrix:

$$L = G^{-1}(D - W) \tag{7}$$

Since L is symmetric, positive, and semi definite, its eigenvalues Λ

and their associated eigenvectors U hold the following form:

$$L = U \Lambda U^{-1} \tag{8}$$

We then use (Lombaert et al., 2013) to correct sign ambiguity in eigenvectors. Here we refer to spectra of a fixed subject being arbitrarily chosen from the dataset. The spectral features were given by the eigenvectors associated with the first five nonzero smallest eigenvalues, which were ultimately fed as five additional input channels into the neural network.

2.3.3. Planar mapping

In general, the cortical surface has a genus-zero form that is not yet fully compatible with a neural network optimized for a uniform grid representation. To address this issue, we represent a cortical surface with polar coordinates (θ, ϕ) at (x, y, z) as follows.

$$\theta = \arctan\left(\frac{y}{x}\right) \tag{9}$$

$$\phi = \arctan\left(\frac{z}{\sqrt{x^2 + y^2}}\right) \tag{10}$$

where θ is the azimuth angle and ϕ is the elevation angle. However, a polar coordinate system inherently has non-uniform angular representations, which yields substantial length distortion around the poles. To reduce such distortion, we rotate each sphere such that the poles are located around insular and ventricle regions so as to minimize length distortion of sulcal curves around the poles (Joshi et al., 2012c). The inspiration for choosing a pole is drawn from Auzias et al.'s approach (Auzias et al., 2013). Given a set of sulcal curves $C \subseteq \Omega$ for each of N subjects, we approximate global pole location by maximizing the following objective function:

$$\hat{z} = \operatorname{argmax}_{\|z\|=1} \sum_{N}^{i=1} \sum_{c \in C_i} \arccos^2(zc) \tag{11}$$

with underlying assumptions that the cortical anatomy on the spheres is roughly aligned and the initial value for z is chosen such that it is roughly located away from most of the sulcal curves (i.e., insular and ventricle regions). Spherical reparameterization is done by updating the pole as obtained from Eq. (11).

In this way, we reduce the planar projection distortion on the primary sulci of interest. After the planar projection, we resampled the plane with a 512×512 resolution that is optimized for convolutional layers of the neural network. The cortical features and label maps were then used for training. Fig. 2 illustrates the distance map normalized between 0 and 1 on a cortical surface, a sphere and after projection to a

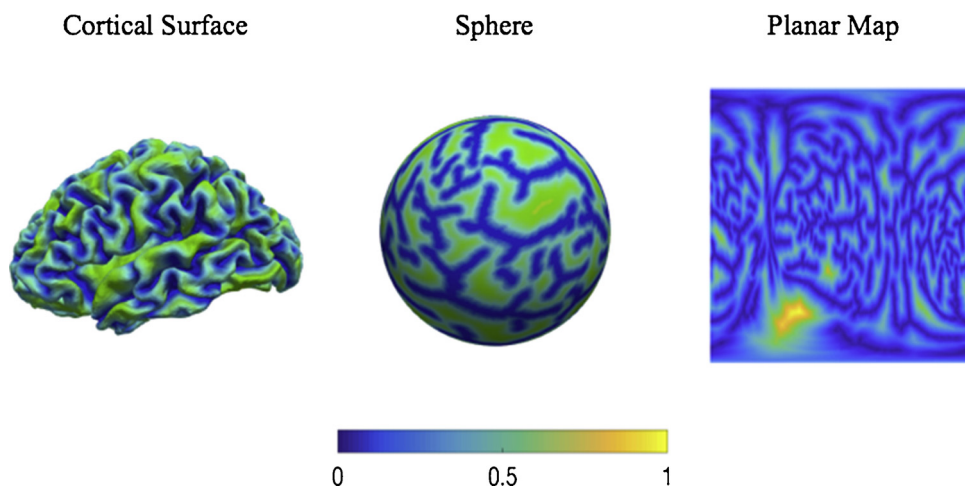


Fig. 2. Qualitative representation showing normalized distance map feature on cortical surface, sphere and planar map.

planar map.

2.3.4. Training

The network architecture is inspired and motivated from the original U-Net design that has been known to perform well for image segmentation tasks (Huo et al., 2018b). A 5-level architecture has been used for each sulcal label with curve overlap as the loss function. A batch size of 10 was used to meet with GPU memory constraints while also achieving better performance. Hence, the network weights are saved using check pointing and batch images are fed into the training model using a fit generator.

The network consists of 2D convolution layers followed by batch normalization and a rectified linear unit (ReLU) activation, which yields better performance during preliminary evaluations. Max pooling is used after each convolution layer. In the last layer, 1×1 convolution produces probability of sulcal curve label from a sigmoid activation function. We can threshold this probability map to derive desirable label information. In this study the threshold was empirically set to 0.7. Each network was trained until convergence for a maximum of 20 epochs. Each epoch took about 6 min to train on an NVIDIA Titan Xp. and Intel(R) Xeon(R) CPU E5-2630 v4 @ 2.20 GHz and 32GB RAM. Finally, as the network was trained on planar maps, the predicted labels were mapped back onto the unit sphere:

$$x = \cos \theta \cos \phi, \quad (12)$$

$$y = \cos \theta \sin \phi, \quad (13)$$

$$z = \sin \theta \quad (14)$$

2.3.5. Post processing

Predicted labels are generated based on the various features supplied to the training model and do not consider cortical shape to delineate labels to only primary sulcal curves. This can yield to issues such as partial labeling on the primary curves or labeling extra minor branches due to the limited ability in cortical shape representation in the U-Net. As our goal is to label primary sulcal curves, the extra minor branches need to be pruned, and missing label information needs to be filled before further analysis (see Fig. 3). Since a sulcal fundus is represented by a single curve, it is reasonable for final sulcal label decision to include or exclude whole branches as major sulci are generally delineated along the entire sulcal fundi. We propose curve length-based pruning to trim the minor branches. Briefly, we measure both the length (i.e., geodesic distance on Ω) of the extracted sulcal curves in C by TRACE and that of the predicted sulcal curves by the proposed method. For each predicted curve, we compute a length ratio r proportional to its corresponding extracted curve in C . If the extracted and predicted curves are perfectly matched, we have $r = 1$, whereas $r < 1$ if the extracted curves are partially labeled. Thus, this quantity is used to prune the predicted curves ($r < 0.2$) or fill partially missing labels along the extracted curves ($r > 0.8$). Although we empirically choose such thresholds in this study, they could be learned from datasets for better decision-making.

2.3.6. Evaluation

We used the sulcal curve editor (<https://github.com/ilwoolyu/SulcalCurveEditor>) to manually label major sulcal curves on 22 subjects via the sulcal delineation protocol (Pantazis et al., 2010). These labels were used for evaluation to validate the proposed DNN predicted labels. These results were compared with a standard curve labeling method (Lyu et al., 2010). Since true curve correspondence was unavailable between major curves, the closest distance measures at sulcal points may not be able to capture the missing or extra curves. In particular, we have computed the sulcal Dice coefficient (SDC) as described below for evaluation. This measure summarizes labeling accuracy as well as false positives in a single metric. From Eq. (4), we have two masks M_A and M_B on a cortical surface Ω , which cover two

corresponding major curves A and B, respectively, as follows.

$$M_A(\mathbf{x}) = M(u_A(\mathbf{x}); \xi) \quad (15)$$

$$M_B(\mathbf{x}) = M(u_B(\mathbf{x}); \xi) \quad (16)$$

where u_A and u_B are geodesic distance from A and B. SDC can be computed by taking surface integral of M_A and M_B .

$$SDC = \frac{\iint M_A(\mathbf{x})M_B(\mathbf{x})\delta\Omega}{\iint M_A(\mathbf{x})\delta\Omega + \iint M_B(\mathbf{x})\delta\Omega} \quad (17)$$

It is challenging to derive analytical solutions to the above surface integrals on an arbitrary cortical surface. In triangular mesh, we can instead approximate the solution by counting the numbers of vertices within the masks, where the numerator is the intersection between the vertices belonging to both M_A and M_B .

In evaluation, we computed SDC for the proposed and reference methods with respect to manual labeling. For each label, we summarized mean SDC values across all the subjects. Two sample t-tests are performed between the reference and proposed methods. Finally, we used false discovery rate (FDR) (Benjamini and Hochberg (1995)) on resulting SDC for multi-comparison correction.

3. Results

We found that the proposed method significantly improved sulcal curve labeling in 9 out of 16 sulcal curves across the left and right hemispheres compared to the reference method (Lyu et al. (2017)), after multi-comparison correction via false discovery rate (Storey, 2011). The average SDC is improved by 12.5 percent for the left hemisphere and 20.6 percent for the right hemisphere. Fig. 4 shows mean SDC values for each of the eight curves on both hemispheres. There is marginal, but not statistically significant, improvement of SDC across both hemispheres in SFS (12.8%), STS (11.3%), and IFS (15.3%) using the proposed method. The highest improvement of SDC is seen in CS with over 28.8% in left hemisphere and 45.3% in right hemisphere with SDC above 0.93 in both hemispheres ($p < 0.05$).

Fig. 5 shows qualitative comparisons of the reference and proposed methods with manual labeling for a single subject on the left hemisphere. The predicted CS from the proposed method is consistent with manual labeling while the reference method only captures about a half of the total curve. Some false positives on minor branches are exhibited in the reference method for STS and SFS that are not present in the proposed method. The proposed method consistently label IFS agreed with manual labeling whereas the reference method includes extra branches (false positives) or is missing a portion of the corresponding curve (false negatives). CingS is consistent with manual labeling for both methods. There is a false positive branch apparent towards the frontal region in both methods (Fig. 5), while the standard curve labeling method also has a false negative in the parietal region in which the curve label is not captured. OTS and OLF are consistent across all the methods except for extra minor branch on OLF.

Although sulcal labels by the proposed method reasonably well agree with manual labeling, the proposed method sometimes misses a portion of the corresponding curve or shows extra branches of the curve, as shown in Fig. 6. CS and OLF are well matched with manual labels (Fig. 6 a). While OTS is also well matched to the manual labels in this example, there are apparent false positives or false negatives in the remaining five curves (Fig. 6. b–g).

Fig. 7 shows an overlay of all eight curves using manual labeling from 22 subjects. They are overlaid on the resampled template surface from 21 subjects in the Kirby21 dataset (Lebed et al., 2013) after surface registration (Huo et al., 2018a). Heterogeneity is observed in most of the sulcal curves. In particular, variability in IFS appears to be higher in the right hemisphere compared to that of left hemisphere. Similarly, higher variability is seen across the frontal region of the SFS curve and the entire region of STS curves across the subjects.

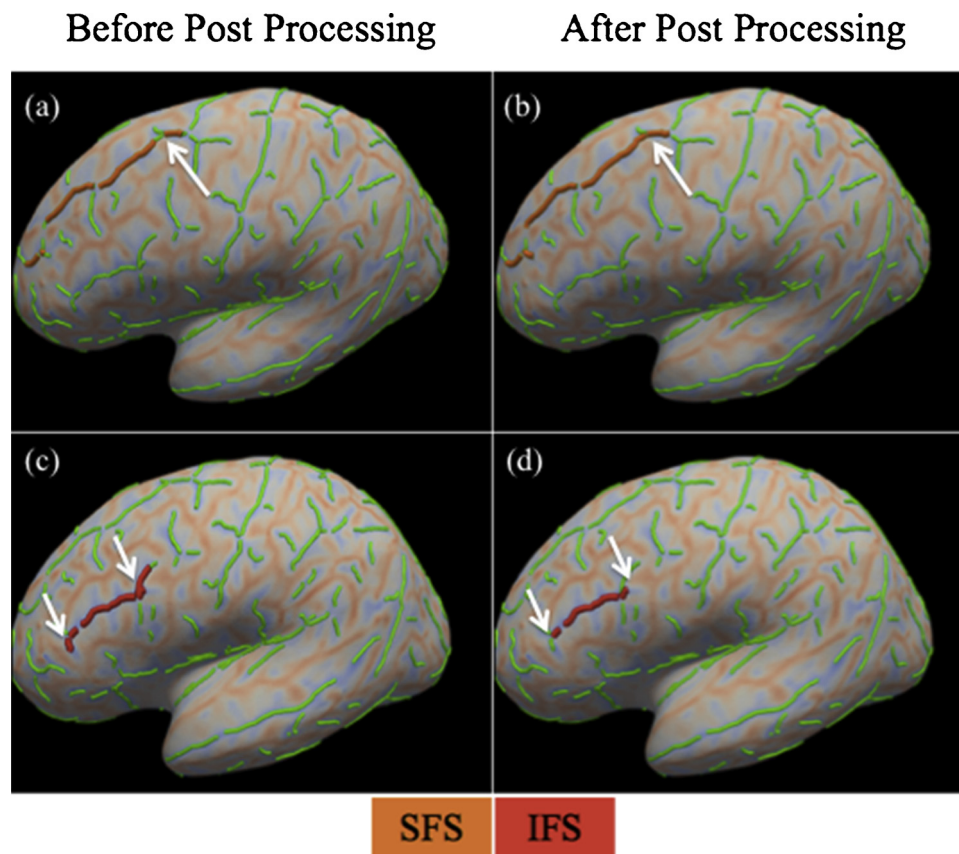


Fig. 3. Example of pruning and filling of sulcal curves in post processing. Inflated surfaces are used for better visualization. Mean curvature is shown as the background on the inflated surface. Sulcal curves in C by TRACE are shown in green lines. Arrows indicate the areas of change before and after post processing. (a) SFS curve with missing branch before post processing. (b) SFS curve after filling. (c) IFS curve with extra branches before post processing. (d) IFS curve after pruning extra branches.

Fig. 8 highlights the disparities in left and right hemisphere between the proposed method and manual labeling. Shorter curves are seen in the frontal region of SFS for labeling from the proposed method compared to that of manual labeling in the left hemisphere. In the right hemisphere, the difference in curve lengths of the two methods is not as pronounced as seen in the left hemisphere (Sowell et al., 2002; Amunts et al., 2000); however, there is higher curve variability in the frontal region. In contrast, STS displays false positives with longer curves in labeling from the proposed method compared to that of manual labeling in the left hemisphere. Moreover, there is also higher curve variability in the left hemisphere compared to that of the right hemisphere. This is in agreement with the higher SDC seen in right hemisphere for STS compared to that of the left hemisphere (Fig. 4). Similarly, CingS has higher variability in the predicted curves in the left hemisphere compared to that of the right hemisphere.

4. Discussion

Deep learning has shown high efficiency and scalability on large datasets (Greenspan et al., 2016; LeCun et al., 2015). In supervised learning approaches, initial labeling is essential in training for segmentation or classification tasks in medical imaging (Huo et al., 2018b; Chen et al., 2018). This is time-consuming for manual sulcal labeling or semi-automated approaches. In the proposed method, we collected large datasets and employed automated processing pipelines (Lyu et al., 2010, 2017) for curve extraction and labeling to produce training sets with reasonably plausible initial labels. We further performed liberal manual QA of all the curves to ensure filtering out extreme outliers in the data. This step helps prevent the model in learning from outlier examples while retaining the variability needed for training. As initial

labels were generated automatically, the liberal QA was not as time-consuming as the manual or semi-automated approaches. However, it could further be improved by incorporating labeling uncertainty as proposed in (Gros et al., 2019; Wang et al., 2018) that can guide in filtering out the data used for training. This will be promising in future work.

A higher SDC is achieved in all the sulcal curves in both hemispheres except for OLF in the right hemisphere (Fig. 4). This might be because the reference method already has good performance on OLF in the right hemisphere (SDC = 0.85). The false positives or negatives shown in Fig. 6 could have been partially caused because of the heterogeneity across the population among these major sulcal curves, as illustrated in Fig. 7. Such variability has been acknowledged and measured in many previous studies (Thompson et al., 1996; Zilles et al., 1997; Juch et al., 2005). Although the focus in our study is not to measure such variability, a qualitative analysis would be helpful in understanding the major curves. For example, different variability in IFS between hemispheres may explain lower SDC in the right hemisphere than that of the left hemisphere (Fig. 3).

Low performance in certain sulcal curves could also be attributed to labeling inconsistency obtained in the proposed method as highlighted in Fig. 8. Possibly, the inconsistency across the predicted labels in certain anatomical regions has risen from the training data acquired by the reference method even after liberal manual QA. The high variability in sulcal curves with several discontinuous pieces (Ono et al. (1990)) could be another contributing factor. Despite the existence of such high variability, the proposed method achieved better SDC values compared to that of reference method. The contributing factors for improved performance could be the use of multiple geometric features and parcellation labels for training the large dataset used in the proposed

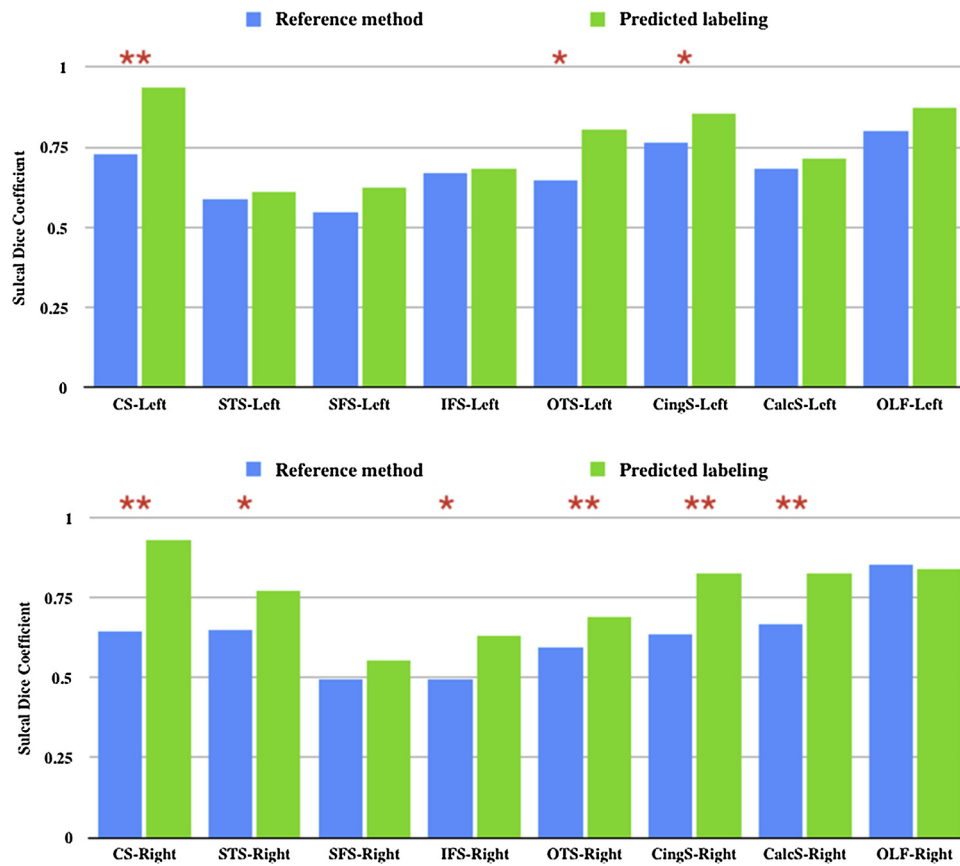


Fig. 4. Sulcal Dice coefficients are shown across eight curves in both left and right hemispheres. * $p < 0.05$. ** $p < 0.005$. Reference method shown in blue bar and Predicted labeling from DNN model is shown in green.

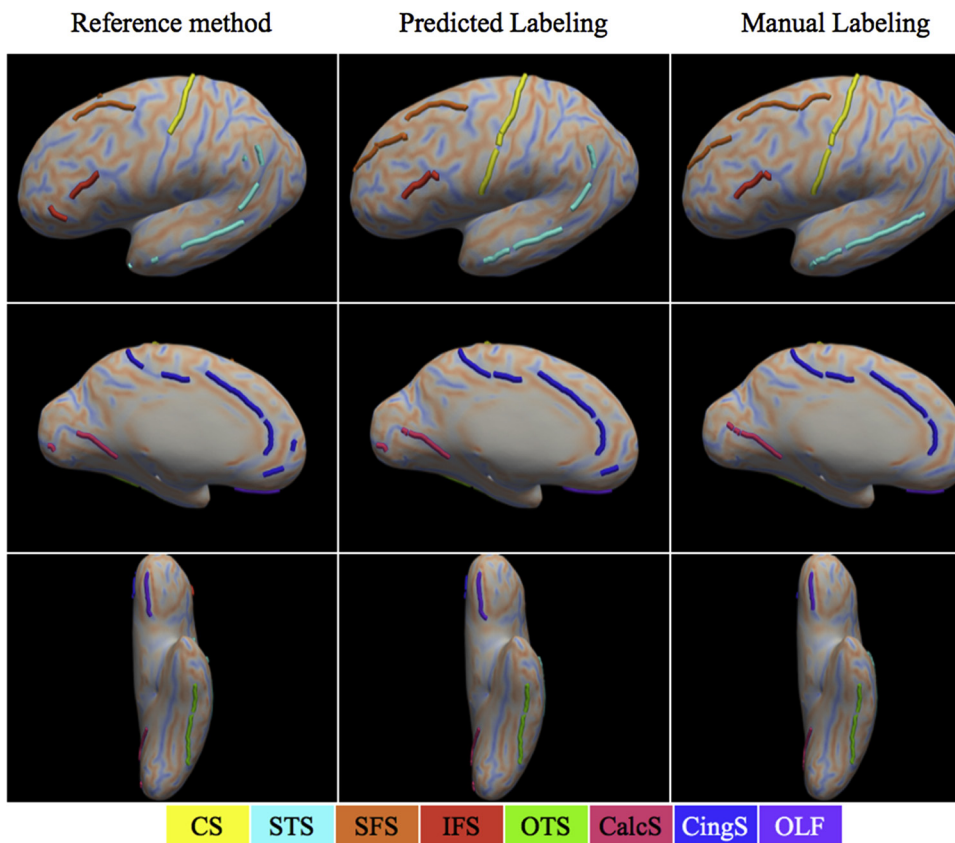


Fig. 5. Qualitative results of eight curves on the left hemisphere for a single subject. Inflated surfaces are used for better visualization. Results are shown in 3 different views for reference, predicted labeling and manual labeling methods for comparison. The color code at the bottom indicates the color associated with each curve. Mean curvature is shown as the background on the inflated surface.

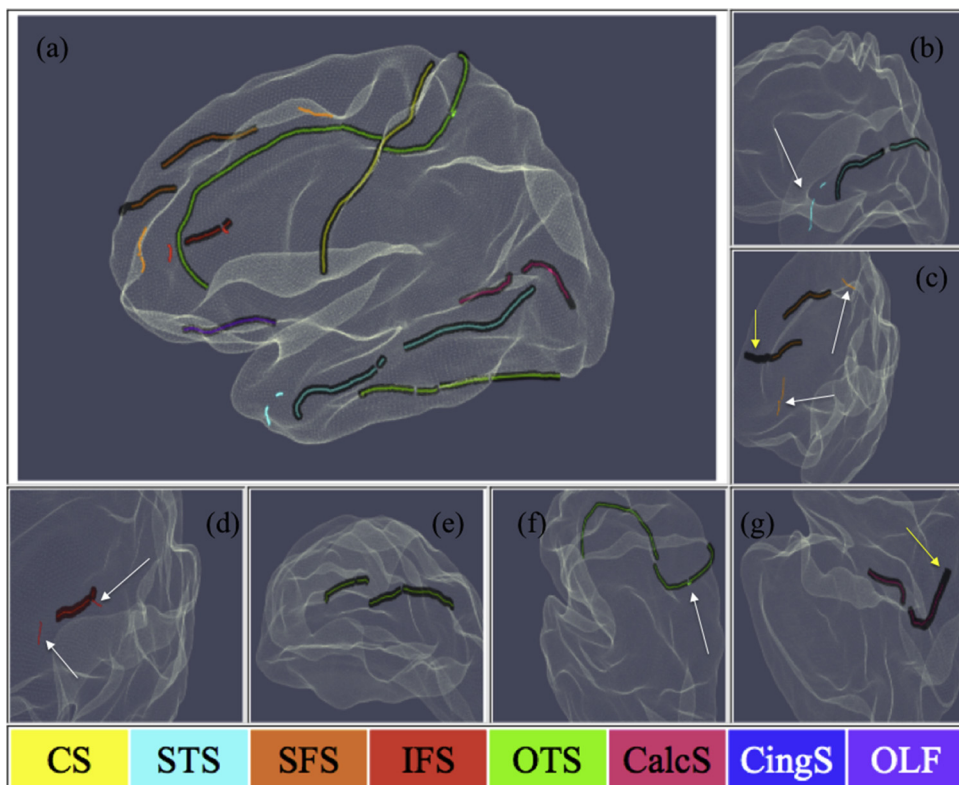


Fig. 6. Illustrative example of inconsistencies seen in predicted labeling. Yellow arrows indicate false negatives and white arrows indicate false positives of predicted labeling compared to that of manual labeling. (a) Overlay of eight manual curves (black tube representation) and predicted curves (line representation) on a single subject. (b) STS with false positives. (c) SFS with a false negative and two false positive branches. (d) IFS with false positives. (e) OTS with no false positives or false negatives. (f) CingS with false positive. (g) CalcS with false negative. All the curves in (b-g) are rotated in the best view for the corresponding curves.

method. We have included multiple (1–7) longitudinal sessions for individual subjects in training data to create a large dataset needed for deep learning to perform well allowing it to generalize to new examples without over fitting. Also having the regularization step though layers like drop out is another factor that is known to help generalizing to new examples. Even though multiple sessions per subject are used in training, we have only included one session per subject in testing for fair comparison. In this work, we have shown the possibility of using a deep learning approach that integrates both sulcal curve extraction and labeling to achieve improved curve-labeling results.

The following observations are made based on the sulcal dice coefficient results from Fig. 4. The quality of sulcal labeling relies on the cortical folding patterns. For example, it is well known that CS and CalcS have a consistent form of continuous curves and CingS is usually either continuous or have 2 branches. Accordingly, higher accuracy is seen in these labels because of the consistency in the cortical folding

pattern. In contrast, STS has either 2 or 3 branches and SFS are not consistently continuous (Lyu et al., 2010; Ono et al., 1990). In such sulci, relatively low accuracy is observed due to the inconsistent folding patterns.

Although the proposed method outperformed the reference method in most curves, there is still a room for improvement. First, in order to use existing architecture for U-Net models designed for uniform grid data, we have mapped cortical surface features onto a plane, which causes mapping distortions. We alleviated distortions to some extent by choosing pole locations and placing the labels away from the poles, yet they are not fully addressed but attenuated. Second, while results from (Lyu et al. (2010)) are based on a young healthy population, the majority of our sample is middle to older age adults. This difference in training and test population demographics could have led to failures in labeling certain curves due to high variation in sulcal shapes between the two age groups. Although we removed outlier cortical surfaces and

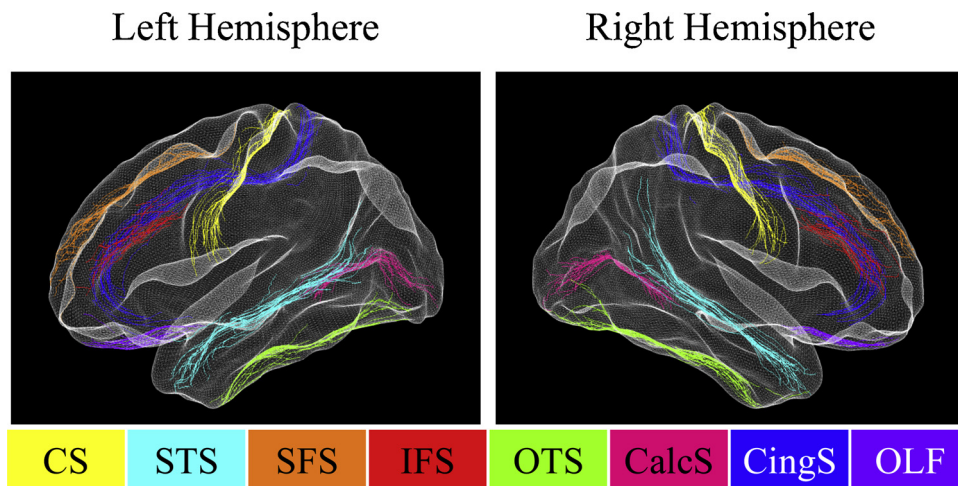


Fig. 7. Eight major sulcal curves manually labeled across 22 subjects in left and right hemispheres.

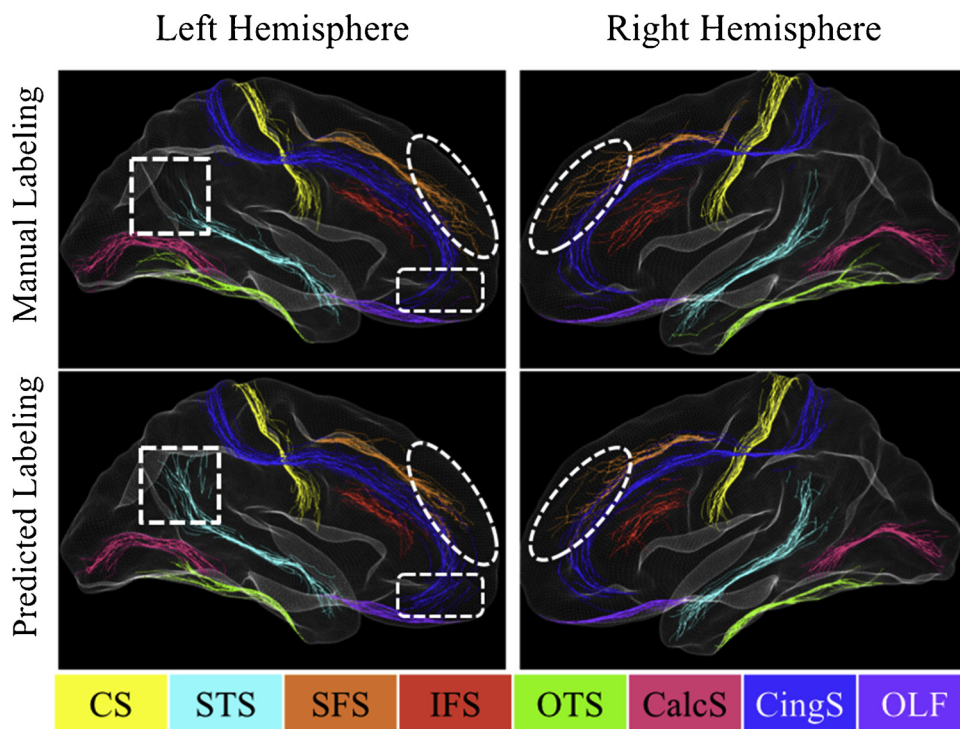


Fig. 8. Illustration highlighting Sulcal curves variability differences between manual and proposed predicted labeling in left and right hemispheres. Dotted squares show differences in STS curves in the left hemisphere. Dotted ovals highlight the differences in SFS curves in both hemispheres. Dotted rectangles indicate differences in CingS curves towards the frontal region.

labeling results from training data, there is still variability and inconsistency within the data used for training this model.

In future work, we will investigate recent spherical CNN methods (Esteves et al., 2018; Coors et al., 2018; Cohen et al., 2018) or graph CNN methods (Edwards and Xie, 2016) in training these datasets to overcome the distortion issues caused with planar mapping of such data. Our approach is not limited to a curve-labeling problem, but it could be extended to tackle other problems like cortical parcellation.

5. Conclusion

In this method, we proposed a DNN predicted labeling method using a U-Net model for labeling eight major sulcal curves on each hemisphere. The proposed method has shown improvement in most of the curves in both left and right hemispheres. We observed significant improvement in CS, OTS, and CingS in the left hemisphere and six curves in the right hemisphere including CS, STS, IFS, OTS, CingS, and CalcS ($p < 0.05$). The overall SDC is improved by 16.6 percent for both hemispheres. CS showed the highest improvement with over 28.8 and 45.3 percent in the left and right hemispheres, respectively, compared to those of the reference method. This provides an opportunity to explore the application of deep learning techniques to methods involving cortical surfaces.

The trained model and code are available on the Github project: (<https://github.com/MASILab/slabelDNN>).

Declaration of Competing Interest

Dr. Yuankai Huo receives a research grant from 12 Sigma Technology. Other co-authors do not have any conflicts of interest to declare.

Acknowledgements

This work was conducted in part using the resources of the Advanced Computing Center for Research and Education at Vanderbilt University, Nashville, TN. This work was supported in part by the Intramural Research Program of the National Institute on Aging, NIH,

and by NIH grants R01EB017230, R01MH102266, UL1 RR024975-01 and NSF CAREER IIS 1452485. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. We thank the participants of the studies. This work was supported in part by NVIDIA Corporation under GPU Grant Program.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.jneumeth.2019.108311>.

References

- Agrawal, P., Whitaker, R.T., Elhabian, S.Y., 2017. Learning deep features for automated placement of correspondence points on ensembles of complex shapes. in Editor (Ed.) (Eds.). Book Learning Deep Features for Automated Placement of Correspondence Points on Ensembles of Complex Shapes. Springer, pp. 185–193.
- Amunts, K., Jancke, L., Mohlberg, H., Steinmetz, H., Zilles, K., 2000. Interhemispheric asymmetry of the human motor cortex related to handedness and gender. *Neuropsychologia* 38 (3), 304–312.
- Amunts, K., Kedo, O., Kindler, M., Pieperhoff, P., Mohlberg, H., Shah, N., Habel, U., Schneider, F., Zilles, K., 2005. Cytoarchitectonic mapping of the human amygdala, hippocampal region and entorhinal cortex: intersubject variability and probability maps. *Anat. Embryol.* 210 (5–6), 343–352.
- Auzias, G., Lefevre, J., Le Troter, A., Fischer, C., Perrot, M., Regis, J., Coulon, O., 2013. Model-driven harmonic parameterization of the cortical surface: HIP-HOP. *IEEE Trans. Med. Imaging* 32 (5), 873–887.
- Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* 289–300.
- Boscaini, D., Masci, J., Rodolà, E., Bronstein, M., 2016. Learning shape correspondence with anisotropic convolutional neural networks. in Editor (Ed.) (Eds.). Book Learning Shape Correspondence With Anisotropic Convolutional Neural Networks. pp. 3189–3197.
- Bronstein, M.M., Bruna, J., LeCun, Y., Szlam, A., Vandergheynst, P., 2017. Geometric deep learning: going beyond euclidean data. *IEEE Signal Process. Mag.* 34 (4), 18–42.
- Caviness Jr, V.S., Meyer, J., Makris, N., Kennedy, D.N., 1996. MRI-Based Topographic Parcellation of Human Neocortex: An Anatomically Specified Method with Estimate of Reliability. *J. Cogn. Neurosci.* 8 (6), 566–587.
- Chen, H., Dou, Q., Yu, L., Qin, J., Heng, P.A., 2018. VoxResNet: Deep voxelwise residual networks for brain segmentation from 3D MR images. *Neuroimage* 170, 446–455.
- Cohen, T.S., Geiger, M., Köhler, J., Welling, M., 2018. Spherical CNNs. arXiv preprint arXiv:1801.10130.
- Coors, B., Condurache, A.P., Geiger, A., 2018. Spherenet: learning spherical representations for detection and classification in omnidirectional images. in Editor (Ed.)

- (Eds.). Book Spherenet: Learning Spherical Representations for Detection and Classification in Omnidirectional Images. pp. 518–533.
- Cucurull, G., Wagstyl, K., Casanova, A., Veličković, P., Jakobsen, E., Drozdal, M., Romero, A., Evans, A., Bengio, Y., 2018. Convolutional Neural Networks for Mesh-based Parcellation of the Cerebral Cortex.
- Dale, A.M., Fischl, B., Sereno, M.I., 1999. Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage* 9 (2), 179–194.
- de Brebisson, A., Montana, G., 2015. Deep neural networks for anatomical brain segmentation. in Editor (Ed.) (Eds.). Book Deep Neural Networks for Anatomical Brain Segmentation. pp. 20–28.
- Edwards, M., Xie, X., 2016. Graph based convolutional neural network. arXiv preprint arXiv:1609.08965.
- Esteves, C., Allen-Blanchette, C., Makadia, A., Daniilidis, K., 2018. Learning so (3) equivariant representations with spherical cnns. in Editor (Ed.) (Eds.). Book Learning so (3) Equivariant Representations With Spherical Cnns. pp. 52–68.
- Fillard, P., Arsigny, V., Pennec, X., Hayashi, K.M., Thompson, P.M., Ayache, N., 2007. Measuring brain variability by extrapolating sparse tensor fields measured on sulcal lines. *Neuroimage* 34 (2), 639–650.
- Gogtay, N., Giedd, J.N., Lusk, L., Hayashi, K.M., Greenstein, D., Vaituzis, A.C., Nugent, T.F., Herman, D.H., Clasen, L.S., Toga, A.W., 2004. Dynamic mapping of human cortical development during childhood through early adulthood. *Proc. Natl. Acad. Sci. U.S.A.* 101 (21), 8174–8179.
- Greenspan, H., Van Ginneken, B., Summers, R.M., 2016. Guest editorial deep learning in medical imaging: overview and future promise of an exciting new technique. *IEEE Trans. Med. Imaging* 35 (5), 1153–1159.
- Gros, C., De Leener, B., Badji, A., Maranzano, J., Eden, D., Dupont, S.M., Talbot, J., Zhuoquiong, R., Liu, Y., Granberg, T., 2019. Automatic segmentation of the spinal cord and intramedullary multiple sclerosis lesions with convolutional neural networks. *NeuroImage* 184, 901–915.
- Hao, Y., Khoo, H.M., von Ellenrieder, N., Zazubovits, N., Gotman, J., 2018. DeepIED: An epileptic discharge detector for EEG-fMRI based on deep learning. *Neuroimage Clin.* 17, 962–975.
- Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., Pal, C., Jodoin, P.M., Larochelle, H., 2017. Brain tumor segmentation with deep neural networks. *Med. Image Anal.* 35, 18–31.
- Huo, Y., Plassard, A.J., Carass, A., Resnick, S.M., Pham, D.L., Prince, J.L., Landman, B.A., 2016. Consistent cortical reconstruction and multi-atlas brain segmentation. *Neuroimage* 138, 197–210.
- Huo, Y., Bao, S., Parvathaneni, P., Landman, B.A., 2018a. Improved stability of whole brain surface parcellation with multi-atlas segmentation. in Editor (Ed.) (Eds.). Book Improved Stability of Whole Brain Surface Parcellation With Multi-atlas Segmentation. International Society for Optics and Photonics.
- Huo, Y., Xu, Z., Aboud, K., Parvathaneni, P., Bao, S., Bermudez, C., Resnick, S.M., Cutting, L.E., Landman, B.A., 2018b. Spatially localized atlas network tiles enables 3D whole brain segmentation from limited data. arXiv preprint arXiv:1806.00546.
- Joshi, A.A., Shattuck, D.W., Damasio, H., Leahy, R.M., 2012a. Geodesic curvature flow on surfaces for automatic sulcal delineation. in Editor (Ed.) (Eds.). Book Geodesic Curvature Flow on Surfaces for Automatic Sulcal Delineation. IEEE, pp. 430–433.
- Joshi, A.A., Shattuck, D.W., Leahy, R.M., 2012b. A method for automated cortical surface registration and labeling. *Biomed. Image Regist. Proc.* 7359, 180–189.
- Joshi, A.A., Shattuck, D.W., Leahy, R.M., 2012c. A method for automated cortical surface registration and labeling. in Editor (Ed.) (Eds.). Book A Method for Automated Cortical Surface Registration and Labeling. Springer, pp. 180–189.
- Juch, H., Zimine, I., Seghier, M.L., Lazeyras, F., Fasel, J.H., 2005. Anatomical variability of the lateral frontal lobe surface: implication for intersubject variability in language neuroimaging. *Neuroimage* 24 (2), 504–514.
- Kim, J.S., Singh, V., Lee, J.K., Lerch, J., Ad-Dab'bagh, Y., MacDonald, D., Lee, J.M., Kim, S.I., Evans, A.C., 2005. Automated 3-D extraction and evaluation of the inner and outer cortical surfaces using a Laplacian map and partial volume effect classification. *Neuroimage* 27 (1), 210–221.
- Kochunov, P., Mangin, J.F., Coyle, T., Lancaster, J., Thompson, P., Rivière, D., Cointepas, Y., Régis, J., Schlosser, A., Royall, D.R., 2005. Age-related morphology trends of cortical sulci. *Hum. Brain Mapp.* 26 (3), 210–220.
- Landman, B.A., Huang, A.J., Gifford, A., Vikram, D.S., Lim, I.A., Farrell, J.A., Bogovic, J.A., Hua, J., Chen, M., Jarso, S., Smith, S.A., Joel, S., Mori, S., Pekar, J.J., Barker, P.B., Prince, J.L., van Zijl, P.C., 2011. Multi-parametric neuroimaging reproducibility: a 3-T resource study. *Neuroimage* 54 (4), 2854–2866.
- Lebed, E., Jacova, C., Wang, L., Beg, M.F., 2013. Novel surface-smoothing based local gyrfication index. *IEEE Trans. Med. Imaging* 32 (4), 660–669.
- LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. *Nature* 521 (7553), 436–444.
- Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., van der Laak, J., van Ginneken, B., Sanchez, C.I., 2017. A survey on deep learning in medical image analysis. *Med. Image Anal.* 42, 60–88.
- Lohmann, G., von Cramon, D.Y., Steinmetz, H., 1999. Sulcal variability of twins. *Cereb. Cortex* 9 (7), 754–763.
- Lombaert, H., Sporring, J., Siddiqi, K., 2013. Diffeomorphic spectral matching of cortical surfaces. *Inf. Process. Med. Imaging* 23, 376–389.
- Lyu, I., Seong, J.-K., Shin, S.Y., Im, K., Roh, J.H., Kim, M.-J., Kim, G.H., Kim, J.H., Evans, A.C., Na, D.L., 2010. Spectral-based automatic labeling and refining of human cortical sulcal curves using expert-provided examples. *Neuroimage* 52 (1), 142–157.
- Lyu, I., Kim, S.H., Seong, J.K., Yoo, S.W., Evans, A., Shi, Y., Sanchez, M., Niethammer, M., Styner, M.A., 2015. Robust estimation of group-wise cortical correspondence with an application to macaque and human neuroimaging studies. *Front. Neurosci.* 9, 210.
- Lyu, I., Kim, S.H., Woodward, N.D., Styner, M.A., Landman, B.A., 2017. TRACE: A Topological Graph Representation for Automatic Sulcal Curve Extraction. *IEEE Trans. Med. Imaging*.
- Lyu, I., Styner, M.A., Landman, B.A., 2018. Hierarchical spherical deformation for shape correspondence. in Editor (Ed.) (Eds.). Book Hierarchical Spherical Deformation for Shape Correspondence. Springer, pp. 853–861.
- Mangin, J.-F., Perrot, M., Operto, G., Cachia, A., Fischer, C., Lefèvre, J., Rivière, D., Neurospin, C., 2015. Sulcus Identification and Labeling, Brain Mapping: an Encyclopedic Reference. Academic Press: Elsevier, New York, pp. 365–371.
- Mechelli, A., 2018. 202. Deep learning technology: concepts and applications in biological psychiatry. *Biol. Psychiatry* 83 (9), S81–S82.
- Milletari, F., Ahmadi, S.-A., Kröll, C., Plate, A., Rozanski, V., Maiostre, J., Levin, J., Dietrich, O., Ertl-Wagner, B., Bötzel, K., 2017. Hough-CNN: deep learning for segmentation of deep brain regions in MRI and ultrasound. *Comput. Vis. Image Underst.* 164, 92–102.
- Minaee, S., Wang, Y., Choromanska, A., Chung, S., Wang, X., Fieremans, E., Flanagan, S., Rath, J., Lui, Y.W., 2018. A deep unsupervised learning approach toward MTBI identification using diffusion MRI. arXiv preprint arXiv:1802.02925.
- Nath, V., Parvathaneni, P., Hansen, C.B., Hainline, A.E., Bermudez, C., Remedios, S., Blaber, J.A., Schilling, K.G., Lyu, I., Janve, V., 2018. Inter-scanner harmonization of high angular resolution DW-MRI using null space deep learning. arXiv preprint arXiv:1810.04260.
- Northoff, G., Waters, H., Mooren, I., Schlüter, U., Diekmann, S., Falkai, P., Bogerts, B., 1999. Cortical sulcal enlargement in catatonic schizophrenia: a planimetric CT study. *Psychiatry Res. Neuroimaging* 91 (1), 45–54.
- Ono, M., Kubik, S., Abernathy, C.D., 1990. Atlas of the Cerebral Sulci. Tps.
- Pantazis, D., Joshi, A., Jiang, J., Shattuck, D.W., Bernstein, L.E., Damasio, H., Leahy, R.M., 2010. Comparison of landmark-based and automatic methods for cortical surface registration. *Neuroimage* 49 (3), 2479–2493.
- Rademacher, J., Galaburda, A.M., Kennedy, D.N., Filipek, P.A., Caviness Jr, V.S., 1992. Human cerebral cortex: localization, parcellation, and morphometry with magnetic resonance imaging. *J. Cogn. Neurosci.* 4 (4), 352–374.
- Rettmann, M.E., Han, X., Xu, C., Prince, J.L., 2002. Automated sulcal segmentation using watersheds on the cortical surface. *NeuroImage* 15 (2), 329–344.
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: convolutional networks for biomedical image segmentation. in Editor (Ed.) (Eds.). Book U-net: Convolutional Networks for Biomedical Image Segmentation. Springer, pp. 234–241.
- Seong, S.-B., Pae, C., Park, H.-J., 2018. Geometric convolutional neural network for analyzing surface-based neuroimaging data. *Front. Neuroinform.* 12, 42.
- Shattuck, D.W., Leahy, R.M., 2002. BrainSuite: an automated cortical surface identification tool. *Med. Image Anal.* 6 (2), 129–142.
- Shattuck, D.W., Joshi, A.A., Pantazis, D., Kan, E., Dutton, R.A., Sowell, E.R., Thompson, P.M., Toga, A.W., Leahy, R.M., 2009. Semi-automated method for delineation of landmarks on models of the cerebral cortex. *J. Neurosci. Methods* 178 (2), 385–392.
- Shi, Y., Tu, Z., Reiss, A.L., Dutton, R.A., Lee, A.D., Galaburda, A.M., Dinov, I., Thompson, P.M., Toga, A.W., 2009. Joint sulcal detection on cortical surfaces with graphical models and boosted priors. *IEEE Trans. Med. Imaging* 28 (3), 361–373.
- Shock, N.W., 1984. Normal Human Aging: the Baltimore Longitudinal Study of Aging. Sowell, E.R., Thompson, P.M., Rex, D., Kornsand, D., Tessner, K.D., Jernigan, T.L., Toga, A.W., 2002. Mapping sulcal pattern asymmetry and local cortical surface gray matter distribution in vivo: maturation in perisylvian cortices. *Cereb. Cortex* 12 (1), 17–26.
- Sowell, E.R., Peterson, B.S., Thompson, P.M., Welcome, S.E., Henkenius, A.L., Toga, A.W., 2003. Mapping cortical change across the human life span. *Nat. Neurosci.* 6 (3), 309–315.
- Storey, J.D., 2011. False Discovery Rate: International Encyclopedia of Statistical Science. Springer, pp. 504–508.
- Tao, X., Prince, J.L., Davatzikos, C., 2002. Using a statistical shape model to extract sulcal curves on the outer cortex of the human brain. *IEEE Trans. Med. Imaging* 21 (5), 513–524.
- Thompson, P.M., Schwartz, C., Lin, R.T., Khan, A.A., Toga, A.W., 1996. Three-dimensional statistical analysis of sulcal variability in the human brain. *J. Neurosci.* 16 (13), 4261–4274.
- Thompson, P.M., Hayashi, K.M., Sowell, E.R., Gogtay, N., Giedd, J.N., Rapoport, J.L., de Zubicaray, G.I., Janke, A.L., Rose, S.E., Semple, J., Doddrell, D.M., Wang, Y., van Erp, T.G., Cannon, T.D., Toga, A.W., 2004. Mapping cortical change in Alzheimer's disease, brain development, and schizophrenia. *Neuroimage* 23 (Suppl 1), S2–18.
- Tu, Z., Zheng, S., Yuille, A.L., Reiss, A.L., Dutton, R.A., Lee, A.D., Galaburda, A.M., Dinov, I., Thompson, P.M., Toga, A.W., 2007. Automated extraction of the cortical sulci based on a supervised learning approach. *IEEE Trans. Med. Imaging* 26 (4), 541–552.
- Wang, G., Li, W., Aertsen, M., Deprest, J., Ourselin, S., Vercauteren, T., 2018. Test-time Augmentation With Uncertainty Estimation for Deep Learning-based Medical Image Segmentation.
- Yamins, D.L., DiCarlo, J.J., 2016. Using goal-driven deep learning models to understand sensory cortex. *Nat. Neurosci.* 19 (3), 356–365.
- Zilles, K., Schleicher, A., Langemann, C., Amunts, K., Morosan, P., Palomero-Gallagher, N., Schormann, T., Mohlberg, H., Burgel, U., Steinmetz, H., Schlaug, G., Roland, P.E., 1997. Quantitative analysis of sulci in the human cerebral cortex: development, regional heterogeneity, gender difference, asymmetry, intersubject variability and cortical architecture. *Hum. Brain Mapp.* 5 (4), 218–221.