



Cortical Analysis of Heterogeneous Clinical Brain MRI Scans for Large-Scale Neuroimaging Studies

Karthik Gopinath^(✉), Douglas N. Greve, Sudeshna Das, Steve Arnold,
Colin Magdamo, and Juan Eugenio Iglesias

Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital and Harvard Medical School, Boston, USA

kgopinath@mgh.harvard.edu

Abstract. Surface analysis of the cortex is ubiquitous in human neuroimaging with MRI, e.g., for cortical registration, parcellation, or thickness estimation. The convoluted cortical geometry requires isotropic scans (e.g., 1 mm MPRAGEs) and good gray-white matter contrast for 3D reconstruction. This precludes the analysis of most brain MRI scans acquired for clinical purposes. Analyzing such scans would enable neuroimaging studies with sample sizes that cannot be achieved with current research datasets, particularly for underrepresented populations and rare diseases. Here we present the first method for cortical reconstruction, registration, parcellation, and thickness estimation for clinical brain MRI scans of any resolution and pulse sequence. The methods has a learning component and a classical optimization module. The former uses domain randomization to train a CNN that predicts an implicit representation of the white matter and pial surfaces (a signed distance function) at 1 mm isotropic resolution, independently of the pulse sequence and resolution of the input. The latter uses geometry processing to place the surfaces while accurately satisfying topological and geometric constraints, thus enabling subsequent parcellation and thickness estimation with existing methods. We present results on 5 mm axial FLAIR scans from ADNI and on a highly heterogeneous clinical dataset with 5,000 scans. Code and data are publicly available at <https://surfer.nmr.mgh.harvard.edu/fswiki/recon-all-clinical>.

1 Introduction

Clinical MRI exams account for the overwhelming majority of brain MRI scans acquired worldwide every year [19]. These exams comprise several scans acquired during a session with different orientations (axial, coronal, sagittal), resolutions, and MRI contrasts. The acquisition hardware and pulse sequence parameters differ significantly across (and even within) centers, leading to highly heterogeneous data. Since cortical thickness is a robust biomarker in the study of normal aging [25] and many brain disorders and diseases [21, 22, 24], methods that can

extract parcellations and thickness measurements from clinical scans (while registering to a reference spherical coordinate frame) are highly desirable. However, cortical analysis of clinical scans is complex due to large slice spacing (resulting in incomplete cortex geometry description) and heterogeneous acquisitions (hindering supervised approaches leveraging image intensity distributions).

Existing neuroimaging research studies [12] rely on isotropic scans with good gray-white matter contrast (typically a 1 mm MPRAGE) and utilize prior information on tissue intensities. Classical cortical analysis pipelines like FreeSurfer [5,9] generate two triangle meshes per hemisphere, one for the white matter (WM) surface and one for the pial surface, while preventing self-intersections. The spherical topology of the surfaces enables mapping coordinates to a sphere, thus enabling computation of vertex-wise statistics across subjects in a common space.

Over the last two years, machine learning approaches for cortical reconstruction on 1 mm MPRAGEs have emerged. Methods based on signed distance functions (SDF) like DeepCSR [4] or SegRecon [11] predict voxel-wise SDFs for the WM and pial surfaces. The final meshes are computed as the SDF isosurfaces and do not guarantee topological correctness. PialINN [18] uses an explicit representation to project the pial surface from the WM surface, which is assumed to be topologically correct. Approaches based on surface deformation like TopoFit [14] or Vox2Cortex [3] use image and graph convolutions to predict a deformation that maps a topologically correct template mesh to an input MRI, thus generating WM and pial surfaces. However, these approaches neither prevent self intersections nor guarantee topological correctness.

Contribution: Our proposed method allows cortical analysis of brain MRI scans of any orientation, resolution, and MRI contrast without retraining, making it possible to use it out of the box for straightforward analysis of large datasets “in the wild”. The proposed method combines two modules: a convolutional neural network (CNN) that estimates SDFs of the WM and pial surfaces, and a classical geometry processing module that places the surfaces while satisfying geometric constraints (no self-intersections, spherical topology, regularity). The CNN capitalizes on recent advances in domain randomization to provide robustness against changes in acquisition – in contrast with existing learning approaches that can only process images acquired with the same resolution and MRI contrast as the scans they were trained on. Finally, our method’s classical geometry processing gives us geometric guarantees and grants instant access to an array of existing methods for cortical thickness estimation, registration, and parcellation [9].

Further Related Work: The parameterization of surfaces as SDFs has been combined with deep neural networks in several domains [20], including cortical reconstruction [4]. Our robustness to MRI contrast and resolution changes is achieved using ideas from the domain randomization literature [26], which involves training supervised CNNs with synthetic images generated from segmentations on the fly at every iteration. These techniques have been successfully applied to MRI analysis [1,13] and use random sampling of simulation param-

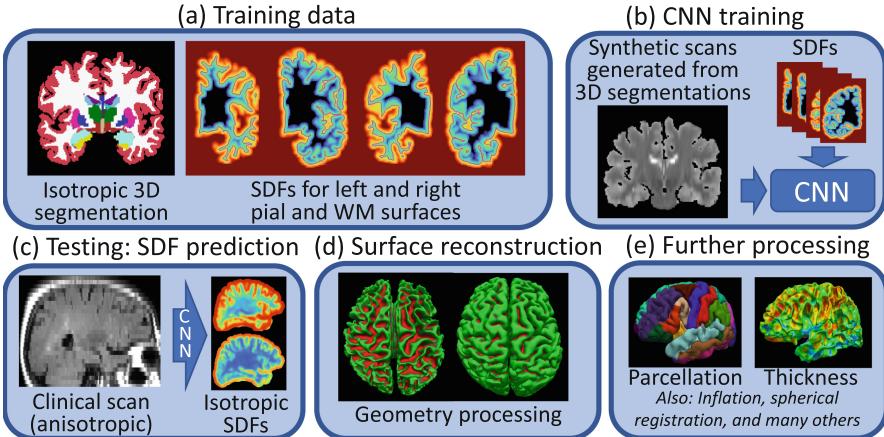


Fig. 1. Overview of our proposed approach for cortical analysis of clinical brain MRI scans of any resolution and MRI contrast, without retraining. The images shown in (c–e) correspond to a real axial FLAIR scan with 5 mm slice spacing and 5 mm thickness.

eters such as orientation, contrast, and resolution from uniform distributions at every mini-batch, which results in unrealistic appearance, making the CNN agnostic to these features.

2 Methods

Our proposed method (Fig. 1) has two distinct components: a learning module to estimate isotropic SDFs from anisotropic scans and a geometry processing module to place the WM and pial surfaces with topological constraints.

2.1 Learning of SDFs

This module estimates isotropic SDFs of the WM and pial surfaces of both hemispheres in a contrast- and resolution-independent fashion. It utilizes a domain randomization approach based on training a voxel-wise SDF regression CNN with synthetic data, which comprises volumetric segmentations and corresponding surfaces (real images are *not* used). Such training data can be obtained “for free” by running FreeSurfer on isotropic T1 scans (we used the HCP dataset [10]).

Given a 3D segmentation and four surface meshes (WM and pial surfaces for each hemisphere; see Fig. 1a), we compute the following input/target pairs at every training iteration. As input, we simulate a synthetic MRI scan of random orientation, resolution, and contrast from the 3D segmentation. For this purpose, we use a Gaussian mixture model conditioned on the (spatially augmented) labels, combined with models of bias field, resolution, and noise similar to [1]. We use random sampling to determine the orientation (coronal, axial, sagittal, or isotropic), slice spacing (between 1 and 9 mm), and slice thickness (between 1 mm

and the slice spacing). The thickness is simulated with a Gaussian kernel across slices. The final synthetic image is upscaled to 1 mm isotropic resolution, such that the CNN operates on input-output pairs of the same size and resolution.

As regression targets, we use voxel-wise SDFs computed from the WM and pial meshes for both hemispheres. The computation of the SDFs would greatly slow down CNN training if performed on the fly. Instead, we precompute them before training and deform them nonlinearly (along with the 3D segmentation) for geometric augmentation during training. While this is only an approximation to the real SDF, it respects the zero-level-set that implicitly defines the surface, and we found it to work well in practice. An example of a synthetic scan and target SDFs used to train the CNN are shown in Fig. 1b.

The regression CNN is trained by feeding the synthetic images to the network and optimizing the weights to minimize the L1 norm of the difference between the ground truth and predicted distance maps. In practice, we clip the SDFs at an absolute value of 5 mm to prevent the CNN from wasting capacity trying to model relatively small variations far away from the surfaces of interest. At test time, the input scan is upcaled to the isotropic resolution of the training data and pushed through the CNN to obtain the predicted SDFs.

2.2 Geometry Processing for Surface Placement

To process real clinical scans, we first feed them to the trained CNN to predict the SDFs for the pial and WM surfaces for both hemispheres (Fig. 1c). To avoid generating topologically incorrect surfaces from these SDFs, we capitalize on the extensive literature on the geometry processing of cortical meshes with classical techniques. For reconstructing WM surfaces, we run SynthSeg [1] on the input scan to obtain two binary masks corresponding to the left and right WM labels. From this point on, processing happens independently for each hemisphere. First, we fill in the holes in the hemisphere’s mask and tessellate it to obtain the initial WM mesh. Then, we smooth the mesh and use automated manifold surgery [8] to guarantee spherical topology. Next, we iteratively deform the WM mesh by minimizing an objective function consisting of a fidelity term and a regularizer.

Specifically: let $\mathcal{M} = (\mathbf{X}, \mathcal{K})$ denote a triangle mesh, where $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_V]$ represents the coordinates of its V vertices ($\mathbf{x}_v \in \mathbb{R}^3$), and \mathcal{K} represents the connectivity. Let $D_w(\mathbf{r})$ be the SDF for the WM surface estimated by our CNN, where \mathbf{r} is the spatial location. The objective function (“energy”) is the following:

$$\begin{aligned} E[\mathbf{X}; D_w(\mathbf{r}), \mathcal{K}] = & \sum_{v=1}^V [\tanh D_w(\mathbf{x}_v)]^2 + \lambda_1 \sum_{v=1}^V \sum_{u \in \mathcal{N}_v} [\mathbf{n}_v^t (\mathbf{x}_v - \mathbf{x}_u)]^2 \\ & + \lambda_2 \sum_{v=1}^V \sum_{u \in \mathcal{N}_v} \{ [\mathbf{e}_{1v}^t (\mathbf{x}_v - \mathbf{x}_u)]^2 + [\mathbf{e}_{2v}^t (\mathbf{x}_v - \mathbf{x}_u)]^2 \}. \end{aligned} \quad (1)$$

The first term in Eq. 1 is the fidelity term, which encourages the SDF to be zero on the mesh vertices; we squash the SDF through a *tanh* function to prevent huge

gradients far away from zero. The second and third terms are regularizers that endow the mesh with a spring-like behavior [5]: \mathbf{n}_v is the surface normal at vertex v ; \mathbf{e}_{1v} and \mathbf{e}_{2v} define an orthonormal basis for the tangent plane at vertex v ; \mathcal{N}_v is the neighborhood of v according to \mathcal{K} ; and λ_1 and λ_2 are relative weights, which we define according to [5] ($\lambda_1 = 0.0006$, $\lambda_2 = 0.0002$). Optimization is performed with gradient descent. At every iteration, self-intersections are monitored and eliminated by reducing the step size as needed [5].

The pial surface is fitted with a very similar procedure, but using the predicted SDF of the pial surface. Figure 1d shows examples of reconstructed surfaces for the axial FLAIR scan from Fig. 1c. Given the fitted WM and pial surfaces, we use FreeSurfer to compute cortical thickness, parcellation, and registration to a common coordinate frame in spherical coordinates (Fig. 1e).

2.3 Implementation Details

Our voxel-wise regression CNN is a 3D U-net [23] trained with synthetic pairs generated on the fly as explained in Sect. 2.1 above. The U-net has 5 levels with 2 layers each, uses $3 \times 3 \times 3$ convolutions and exponential linear activations. The layers have 24^l features, where l is the level number. The last layer uses linear activation functions to model the SDFs. The CNN weights were optimized with stochastic gradient descent using a fixed step size of 0.0001 and 300,000 iterations (enough to converge). At test time, the run time is dominated by the geometry processing (2–3 h, depending on the complexity of the manifold surgery).

3 Experiments and Results

3.1 Datasets

- **HCP:** we used 150 randomly selected subjects (71 males, ages: 29.9 ± 3.4 years) from HCP [10] to train the U-net. We ran FreeSurfer to obtain the segmentations and SDFs (images are discarded as they are not used in training).
- **ADNI:** in our first experiment, we used 1 mm MPRAGE and corresponding 5 mm axial FLAIR scans of 200 randomly selected subjects from the ADNI dataset [17] (95 males, ages 74.5 ± 7.4 years). This setup enables us to directly compare the results from research- and clinical-grade scans.
- **Clinical:** this dataset comprises 9,735 scans with a plethora of pulse sequence combinations and resolutions from 1,367 MRI sessions of distinct subjects with memory complaints (749 males, ages 18–90) from hospital [2]. Surfaces were successfully generated for 5,064 scans; the rest failed due to insufficient field of view. This dataset includes a wide range of MR contrasts and resolutions. We note that this dataset also includes 581 1 mm MPRAGE scans.

The availability of 1 mm MPRAGEs for some of the subjects enables us to process them with FreeSurfer and use the result as ground truth [16] (Fig. 2a).

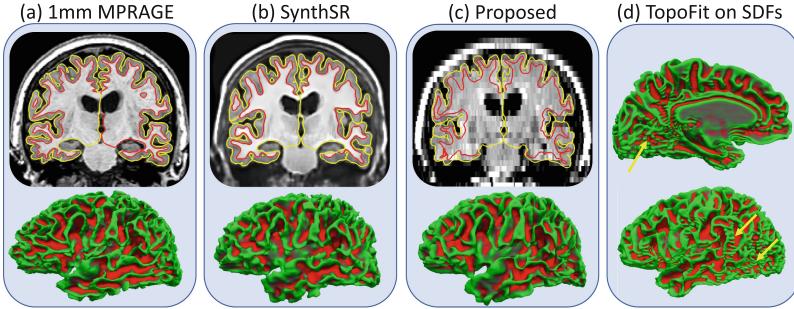


Fig. 2. Qualitative comparison on a 5 mm axial FLAIR scan from ADNI. (a) Ground truth T1 with WM (red) and pial (yellow) surfaces estimated by FreeSurfer (top); and 3D reconstruction of the left WM surface (bottom); (b) Synthetic T1 produced by SynthSR with FreeSurfer surfaces. (c) Our proposed method. (d) Examples of WM surfaces produced by TopoFit, trained to predict on the output of our SDF predictor, show small blobs in different areas (e.g., indicated by yellow arrows) (Color figure online)

3.2 Competing Methods

To the best of our knowledge, the only existing competing method for our proposed algorithm is SynthSR [15], which utilizes a synthetic data generator like ours to turn scans of any resolution and contrast into synthetic 1 mm MPRAGES – which can be subsequently processed with FreeSurfer to obtain surfaces (Fig. 2b). Compared with our proposed approach (Fig. 2c), this pipeline inherits the smoothness of the synthetic MPRAGE, leading to smoother surfaces that may miss larger folds. We also tried training TopoFit [14] on the synthetic images and predicted SDFs, but failed to produce neural networks with good generalization ability, as they led to small blobs on the surfaces at test time (Fig. 2d).

3.3 Results on the ADNI Dataset

Figure 3 summarizes the results on the ADNI dataset. While previous machine learning approaches focus evaluation on distance errors, these can be difficult to interpret. Instead, we evaluate our method using the performance on the downstream tasks that one is ultimately interested in. First, we computed the accuracy of the Desikan-Killiany parcellation [6] produced by SynthSR and our proposed method. Figure 3a shows the results on the inflated surface of the *fsaverage* template. Since the parcellation is computed from the curvature of the WM surface, it is a relatively easy problem. The overlap between the ground truth parcellation and the two competing methods is very high. Dice scores over 0.90 are obtained for almost every region in both methods, and the average across regions is almost identical for both methods (0.95).

We then used the obtained parcellations to study the effect of Alzheimer’s disease (AD) on cortical thickness, using a group study between AD subjects

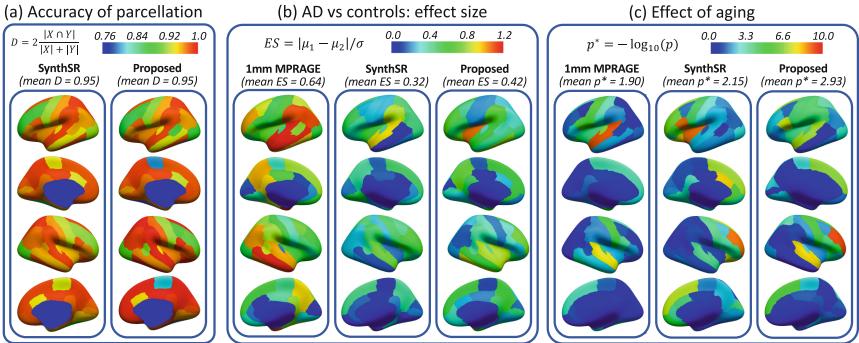


Fig. 3. Summary of results on the ADNI dataset, displayed on the inflated surface of FreeSurfer’s average subject (*fsaverage*). (a) Accuracy of parcellation for SynthSR and our proposed method using Dice scores. (b) Ability to discriminate AD vs controls, measured with effect sizes. (c) Effect of aging, measured as the strength of the (negative) correlation between age and thickness. The strength of the correlation is represented by p-values of a Student’s *t* test assessing whether the correlation significantly differs from zero; note that we log-transform the p-values for easier visualization.

and elderly controls. For this purpose, we first fit a general linear model (GLM) to the cortical thickness at every parcel, using age, gender, and AD status as covariates. We then used the model coefficients to correct the thickness estimates for age and gender, and compared the thicknesses of the two groups.

Figure 3b shows the effect sizes (ES) for the reference 1 mm MPRAGEs and the competing methods. The 1 mm scans yield the expected AD cortical thinning pattern [7], with strong atrophy in the temporal lobe ($ES > 1.0$) and, to a lesser extent, in parietal and middle frontal areas ($ES \sim 1.0$). The average ES across all regions is 0.64. As expected, the thickness estimates based on the FLAIR scans are less able to detect the differences between the two groups. SynthSR loses, on average, half of the ES (0.32 vs 0.64). Most worrying, it cannot detect the effect on the temporal areas (particularly middle temporal). Our method can detect these differences with $ES > 0.6$ in all temporal regions. On average, our method recovers one third of the ES lost by SynthSR (0.42 vs 0.32).

Finally, we studied the effect of aging on cortical thickness using the same GLM as above. Figure 3c shows maps of p -values computed with Student’s *t* distribution, where we have transformed $p^* = \log_{10}(p)$ for easier visualization. Once more, the 1 mm MPRAGEs display the expected pattern [25], with strongest atrophy in superior-temporal and, to less extent, the central and medial frontal gyri. SynthSR fails to detect the superior-temporal effect in the left hemisphere and barely discerns it in the right hemisphere. Our approach, on the other hand, successfully detects these effects. We also note that SynthSR and our method display false positives in frontal areas of the right hemisphere; further analysis (possibly with manual quality control) will be needed to elucidate this result.

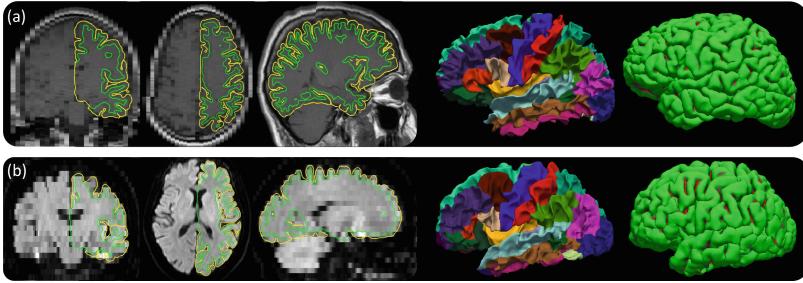


Fig. 4. Sample outputs for heterogeneous scans from the clinical dataset: (a) Sagittal TSE-T1 scan (.4×.4×6 mm). (b) Axial FLAIR (1.7×1.7×6 mm). The WM and pial surfaces are shown on the right. The cortical parcellation is overlaid on the WM surface.

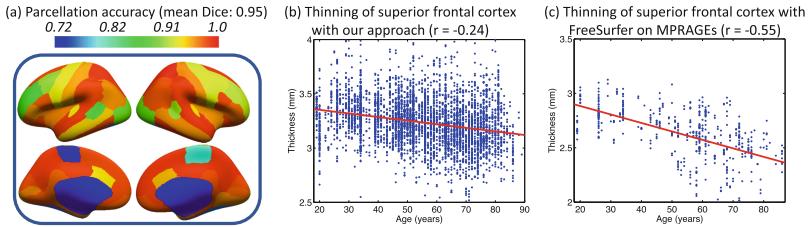


Fig. 5. Results on clinical dataset. (a) Dice scores for parcellation, using FreeSurfer on MPRAGEs as ground truth. (b–c) Thinning of superior frontal cortex in aging, as measured with our method on anisotropic scans (b) and FreeSurfer on MPRAGEs (c).

3.4 Results on the Clinical Dataset

The clinical dataset, despite not being clustered into well defined groups as ADNI, enables us to evaluate our method with the type of data that it is conceived for: a heterogeneous set of brain MRI scans acquired “in the wild”. Samples of such scans and outputs produced by our method are shown in Fig. 4. In this experiment, we first used the 581 1 mm MPRAGEs to compute the Dice scores of the Desikan-Killiany parcellation on clinical acquisitions. The results are displayed in Fig. 5a, and show that our proposed method is able to sustain high accuracy in this task (the mean Dice is the same as for ADNI), despite the huge variability in the acquisition protocol of the input scans. As in the previous experiment, we also computed aging curves using all non-1 mm-MPRAGE scans (4,483 in total), while correcting for gender and slice spacing. The fitted curve for a representative region (the superior frontal area, which shows consistent effects in Fig. 3c) is shown in Fig. 5b. While the thinning trend exists, the data are rather noisy and the linear fit ($\rho = -0.24$) underestimates the effect of aging, i.e., the magnitude of the slope. This is apparent when comparing with the fit produced by the 581 MPRAGEs (Fig. 5c, $\rho = -0.55$).

3.5 Discussion and Conclusion

We have presented a novel method for cortical analysis of clinical brain scans of any MRI contrast and resolution that does not require retraining. To the best of our knowledge, this is the first method seeking to solve this difficult problem. The method runs in 2–3 h but could be sped up by replacing some modules (e.g., the spherical registration) with faster learning methods.

Our method provides accurate parcellation across the board, which is helpful in applications like diffusion MRI (e.g., for seeding or constraining tractography with surfaces and parcellations when a T1 scan is unavailable or is difficult to register due to geometric distortion of the diffusion-weighted images). However, we observed increased variability in cortical thickness when processing the highly heterogeneous clinical dataset. Future work will focus on improving the reliability of thickness measurements and provide a confidence for the quality of reconstruction and cortical thickness prediction for lower resolution scans. In such scenarios assessing and modeling geometric covariates (e.g., vertex-wise distance to the nearest slice or angle between surface and acquisition orientation) may help reduce such variability.

Our method and the clinical dataset are publicly available, which enables researchers worldwide to capitalize on millions of retrospective clinical scans to perform cortical analysis currently unattainable in research studies, particularly for rare diseases and underrepresented populations.

Acknowledgment. This work is primarily funded by the National Institute of Aging (1R01AG070988). Further support is provided by, BRAIN Initiative (1RF1MH123195, 1UM1MH130981), National Institute of Biomedical Imaging and Bioengineering (1R01EB031114), Alzheimer’s Research UK (ARUK-IRG2019A-003), National Institute of Aging (P30AG062421)

References

1. Billot, B., et al.: SynthSeg: segmentation of brain MRI scans of any contrast and resolution without retraining. *Med. Image Anal.* **86**, 102789 (2023)
2. Billot, B., Magdamo, C., Cheng, Y., Arnold, S.E., Das, S., Iglesias, J.E.: Robust machine learning segmentation for large-scale analysis of heterogeneous clinical brain MRI datasets. *Proc. Natl. Acad. Sci.* **120**(9), e2216399120 (2023)
3. Bongratz, F., Rickmann, A.M., Pölsterl, S., Wachinger, C.: Vox2Cortex: fast explicit reconstruction of cortical surfaces from 3D MRI scans with geometric deep neural networks. In: CVPR, pp. 20773–20783 (2022)
4. Cruz, R.S., Lebrat, L., Bourgeat, P., Fookes, C., Fripp, J., Salvado, O.: DeepCSR: a 3D deep learning approach for cortical surface reconstruction. In: WACV. pp. 806–815 (2021)
5. Dale, A.M., Fischl, B., Sereno, M.I.: Cortical surface-based analysis: I. segmentation and surface reconstruction. *Neuroimage* **9**(2), 179–194 (1999)
6. Desikan, R.S., 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)

7. Dickerson, B.C., Bakkour, A., Salat, D.H., Feczko, E., et al.: The cortical signature of Alzheimer's disease: regionally specific cortical thinning relates to symptom severity in very mild to mild AD dementia and is detectable in asymptomatic amyloid-positive individuals. *Cereb. Cortex* **19**(3), 497–510 (2009)
8. Fischl, B., Liu, A., Dale, A.M.: Automated manifold surgery: constructing geometrically accurate and topologically correct models of the human cerebral cortex. *IEEE Trans. Med. Imaging* **20**(1), 70–80 (2001)
9. Fischl, B., Sereno, M., Dale, A.M.: Cortical surface-based analysis: II: inflation, flattening, and a surface-based coordinate system. *Neuroimage* **9**(2), 195–207 (1999)
10. Glasser, M., Sotiropoulos, S., Wilson, J.A., Coalson, T.S.: The minimal preprocessing pipelines for the human connectome project. *Neuroimage* **80**, 105–24 (2013)
11. Gopinath, K., Desrosiers, C., Lombaert, H.: SEGRECON: learning joint brain surface reconstruction and segmentation from images. In: de Bruijne, M., et al. (eds.) MICCAI 2021. LNCS, vol. 12907, pp. 650–659. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87234-2_61
12. Hibar, D., Westlye, L.T., Doan, N.T., Jahanshad, N., et al.: Cortical abnormalities in bipolar disorder: an MRI analysis of 6503 individuals from the ENIGMA bipolar disorder working group. *Mol. Psychiatry* **23**(4), 932–942 (2018)
13. Hoffmann, M., Billot, B., Greve, D.N., Iglesias, J.E., Fischl, B., Dalca, A.V.: SynthMorph: learning contrast-invariant registration without acquired images. *IEEE Trans. Med. Imaging* **41**(3), 543–558 (2021)
14. Hoopes, A., Iglesias, J.E., Fischl, B., Greve, D., Dalca, A.V.: TopoFit: rapid reconstruction of topologically-correct cortical surfaces. In: MIDL (2021)
15. Iglesias, J., Billot, B., Balbastre, Y., Tabari, A., et al.: Joint super-resolution and synthesis of 1 mm isotropic MP-RAGE volumes from clinical MRI exams with scans of different orientation, resolution & contrast. *Neuroimage* **237**, 118206 (2021)
16. Iscan, Z., Jin, T.B., Kendrick, A., Szeglin, B., Lu, H., Trivedi, M., et al.: Test-retest reliability of Freesurfer measurements within and between sites: Effects of visual approval process. *Hum. Brain Mapp.* **36**(9), 3472–3485 (2015)
17. Jack, C.R., Jr., Bernstein, M.A., Fox, N.C., Thompson, P., Alexander, G., et al.: The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. *J. Magn. Reson. Imaging* **27**(4), 685–691 (2008)
18. Ma, Q., Robinson, E.C., Kainz, B., Rueckert, D., Alansary, A.: PialNN: A fast deep learning framework for cortical pial surface reconstruction. In: International Workshop on Machine Learning in Clinical Neuroimaging. pp. 73–81 (2021)
19. Oren, O., Kebebew, E., Ioannidis, J.P.: Curbing unnecessary and wasted diagnostic imaging. *JAMA* **321**(3), 245–246 (2019)
20. Park, J.J., Florence, P., Straub, J., Newcombe, R., Lovegrove, S.: DeepSDF: learning continuous signed distance functions for shape representation. In: CVPR, pp. 165–174 (2019)
21. Pereira, J.B., Ibarretxe, N., Marti, M.J., Compta, Y., et al.: Assessment of cortical degeneration in patients with Parkinson's disease by voxel-based morphometry, cortical folding, and cortical thickness. *Hum. Brain Mapp.* **33**, 2521–34 (2012)
22. Querbes, O., Aubry, F., Pariente, J., Lotterie, J.A., Démonet, J.F., et al.: Early diagnosis of Alzheimer's disease using cortical thickness: impact of cognitive reserve. *Brain* **132**(8), 2036–2047 (2009)
23. Ronneberger, O., Fischer, P., Brox, T.: U-net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28

24. Rosas, H., et al.: Regional and progressive thinning of the cortical ribbon in Huntington's disease. *Neurology* **58**(5), 695–701 (2002)
25. Salat, D.H., et al.: Thinning of the cerebral cortex in aging. *Cereb. Cortex* **14**(7), 721–730 (2004)
26. Tobin, J., Fong, R., Ray, A., Schneider, J., Zaremba, W., Abbeel, P.: Domain randomization for transferring deep neural networks from simulation to the real world. In: IROS, pp. 23–30 (2017)