PLAGIARISM COMPARISON SCAN REPORT

Content Type	TEXT	TEXT
Values	Text content used	Text content used

First Content **1%** matched	Second Content **1%** matched

A Deep Network for Joint Registration and Parcellation of Cortical SurfacesAbstract. Cortical surface registration and parcellation are two essential steps in neuroimaging analysis. Conventionally, they are performed independently as two tasks, ignoring the inherent connections of these two closely-related tasks. Essentially, both tasks rely on meaningful cortical feature representations, so they can be jointly optimized by learning shared useful cortical features. To this end, we propose a deep learning framework for joint cortical surface registration and parcellation. Specifically, our approach leverages the spherical topology of cortical surfaces and uses a spherical network as the shared encoder to rst learn shared features for both tasks. Then we train two task-specic decoders for registration and parcellation, respectively. We further exploit the more explicit connection between them by incorporating the novel parcellation map similarity loss to enforce the region-of-interests boundary consistency, thereby providing extra supervision for the registration task. Conversely, parcellation network training also benets from the registration, which provides a large amount of augmented data by warping one surface with manual parcellation map to another surface, especially when only few manually-labeled surfaces are available. Experiments on a dataset with more than 600 cortical surfaces show that our approach achieves large improvements on both parcellation and registration accuracy (over separately trained networks) and enables training high-quality parcellation and registration models using much fewer labeled data.Keywords Surface Registration Parcellation Deep Neural Network.1IntroductionCortical surface registration and parcellation are two essential tasks in surfacebased neuroimaging analysis. Cortical surface registration estimates a deformation eld to align cortical features from dierent scans and establishes vertexwise cortical correspondences across individuals or time points [4,13,16], thus enabling subsequent analyses, e.g., group comparison or longitudinal studies. Cortical surface parcellation learns the mapping between the cortical features and the anatomically or functionally meaningful regions, thereby parcellating the cortex into dierent regions of interest (ROIs) [2,11,14]. Conventionally, these two tasks are generally performed independently [6], ignoring their inherent close relationship. However, they both rely on learning eective cortical feature representations, one for inferring the deformation eld and the other for predicting 2parcellation labels, which means that they explore cortical features in dierent views and thus conducting them together enables better regularization for more meaningful and robust feature representation, and therefore, in short, they can help each other. For example, for the parcellation task, with the vertex-wise correspondence established by cortical surface registration, the manual parcellation labels from one subject or an atlas can be easily propagated to a new subject, thus helping the parcellation task to better learn the mapping from features to parcellation labels. For the registration task, the parcellated ROI boundaries can be used as an extra guidance (in addition to cortical geometric or functional features) to better learn the mapping from features to deformation eld for registering two surfaces.Therefore, in this paper, for the rst time, we explore the idea of joint registration and parcellation of cortical surfaces through a deep spherical neural network, leveraging the spherical topology of the cerebral cortex. To extract more useful cortical features, we design a shared encoder to rst learn the common features shared by both tasks. Then we train two task-specic decoders for registration and parcellation, respectively. We further exploit the more explicit connection between the two tasks which is formulated by a novel parcellation map similarity loss. This loss forces the warped predicted parcellation map of the moving surface to match the manual parcellation map of the xed surface, e.g., an atlas, thereby providing extra supervision of the ROI boundary consistency for the registration task. Conversely, this loss can be considered as a data augmentation method that generates quasi-ground-truth parcellations for the unlabeled surfaces to help the semi-supervised learning of the parcellation network [12]. Consequently, the two tasks can mutually guide each others training and boost each others performance compared with separately trained networks, especially with limited labeled data, which is of signicant importance for practical use cases where only few manually labeled surfaces are available.2MethodOur goal is to train a joint registration and parcellation network (JRP-Net) that generates the parcellation map of a given cortical surface based on its cortical features and simultaneously yields the individual-to-atlas deformation eld, more accurately with few manually labeled surfaces. To this end, we formulate our approach as in Fig. 1. Our JRP-Net consists of three modules a shared encoder (SE) for extracting mutual high-level features, a registration decoder (RD) for cortical surface registration and a parcellation decoder (PD) for surface parcellation. Let M, F be the moving surface and xed surface dened on the sphere S 2 discretized using icosahedron subdivisions [4]. The SE module learns the mapping f SE between an input surface map and extracted features Z Z F f SE (F), Z M f SE (M). The RD module takes the extracted features from both F and M as input and outputs the spherical velocity eld u u f RD (Z F , Z M) and further derives the dieomorphic deformation eld exp(u) using

6 scaling and squaring layers as in[2]

[13]. The PD module takes Abbreviated paper title3MShared Encoder (SE)Registration Decoder (RD)Shared Encoder (SE)Feature extraction path Feature concatenationParcellation Decoder (PD)Average convexityMean curvatureWarpSkip connection Parcellation path Registration path Spherical mapping and resamplingFParcellation Decoder (PD)Feature similarity loss Parcellation map similarity loss 6 scaling and squaring layersTrans. ConvTrans. ConvTrans. ConvConv BlockConv BlockConv Block40962x4840962x1610242x3210242x9610242x322562x1922562x642562x64642x25640962x2Trans. ConvLayerLayerLayerBlockBlockBlockBlockBlockBlockBlock40962x240962x1640962x1610242x1610242x3210242x3210242x3240962x3640962x22562x322562x642562x64642x1282562x64642x64Trans.Trans.Conv BlockTrans.Trans.Conv40962x16Conv BlockConv BlockConv BlockConvConvConvTrans.Pool LayerConvConvConvConvConvConvPool10242x1610242x32Pool LayerPool2562x322562x64Pool LayerPool642x64642x128BlockBlockBlockConvConvConv2562x642562x6410242x3210242x3240962x1640962x36ConvConvFig. 1. Our JRP-Net for joint registration and parcellation of cortical surfaces. The input cortical surfaces are color-coded by two geometric features, i.e., average convexity (up) and mean curvature (down). The Conv Block contains repeated 1-ringConvBNReLU. The Trans. Conv represents the spherical transposed convolution for upsampling the surface. The input and output sizes are denoted before and after each operation. See Sec. 2 for more interpretations of the math symbols.the extracted features as input and outputs the parcellation map P f P D (Z). We will detail the network architectures (Sec. 2.1), the specic losses (Sec. 2.2) and training strategy (Sec. 2.3) to train our JRP-Net eectively in the following.2.1 Network ArchitectureWe construct our network based on the popular Spherical U-Net [14]. It leverages the spherical topology of cortical surfaces and extends convolution, pooling operations to the spherical space using 1-ring lter on regularly resampled spherical surfaces, and then constructs the network using corresponding spherical operations, achieving promising performance in various tasks, e.g., parcellation [14], registration [16] and harmonization [15].Shared Encoder (SE). Our SE module shares a similar architecture with the encoder part of Spherical U-Net. It consists of 4 repeated 1-ring-Convolution Batch Normalization (BN)ReLU layers in four resolutions with three spherical mean pooling layers between them. This basic encoder architecture has demonstrated good representation ability in many tasks [10] and can be trained to extract shared features at various spatial scales for both registration and parcellation tasks. 4Parcellation Decoder (PD). The PD module predicts the parcellation map based on the extracted feature maps Z using a similar decoder architecture as in Spherical U-Net, with modications to the feature channels and resolutions.Registration Decoder (RD). Our RD module is similar to a recent unsupervised learning framework [16]. It is used to fuse the separately extracted high-level features Z F and Z M at multiple scales and predict the corresponding deformation eld. Compared to previous registration methods [16,1] that directly take the concatenation of F and M as input and output the deformation eld using a single network, our novel registration network (SERD) has two main advantages. 1) Extracting features exclusively using SE and computing deformation eld exclusively using RD makes the training objectives of the encoder and the decoder more explicit and distinct and thus easier to accomplish 2) Our deep multi-scale feature concatenation in RD enables more eective learning of the dierences of the high-level boundary features of F and M, which is important for computing the deformation eld aligning F and M [7].2.2 Loss Functions Feature Similarity Loss (L fs). L fs is applied to the registration network to enforce the feature similarity between the warped moving surface (moved surface) and the atlas surfacecov(F, M) , L fs (F, M,) F M 2 cc F M (1)where is the learned deformation eld and M represents the moved surface maps, cov(,) is the covariance, is the standard deviation and cc is the weight for the correlation coecient term.Spherical Deformation Smoothness Loss (L s). We use the same operator s as in [16] on the spherical surface to approximate the spherical deformations gradients on the sphere and accordingly penalize the gradients asN 1 L s ( ) s R v n (u) 2 , N n1(2)where R v n (u) denotes the local 1-ring velocity vectors of vertex v n on a surface with N vertices. Hence, it can encourage the deformation eld to be smooth.Supervised Parcellation Loss (L sp). We use the weighted cross entropy loss to supervise the training of the parcellation network when manual parcellation maps are availableN 1 exp(P(v n)[c]) Wc log(Lsp (P, P) N exp(P(v n)[j])), n1 j1 C (3)where P and P are the predicted and manual parcellation maps respectively, W c is the inverse proportions of the c-th ROIs area in P , P(v n)[j] represents the probability of vertex v n predicted as label j and c is the manual label of Abbreviated paper title5v n . Considering the real applications of cortical surface analysis, we assume F (atlas) always has manual labels and thus L sp (P, P) L sp (P F , P F), when only F is labeled, otherwise L sp (P, P) L sp (P F , P F) L sp (P M , P M) when both F and M are labeled.Parcellation Map Similarity Loss (L ps). We use the multi-class Dice loss, which addresses imbalanced parcellation labels inherently [12]K K (v) (PM k 1 L ps , P M v P F k , , P M k 2)(v) (P F) 1 Dice(PF k K) K P F k (v) v (P M k)(v), k1 k1 v (4)where k indicates a ROI label (out of K ROIs) and v is the vertex location. The P M in Eq. 4 is manually labeled parcellation map when available or predicted using f P D otherwise. Hence it can provide extra ROI boundary consistency supervision for the registration task and augmented quasi-ground-truth parcellations of M for the parcellation task, when M is not labeled.2.3 Training StrategyTo train the proposed JRP-Net, we design a progressive training strategy to learn the network parameters in an easy-to-hard manner with three steps. 1) We train the parcellation network (SEPD) by minimizing L sp using the strong supervision from all available manually labeled surfaces for 2,000 iterations (i.e., 200 epochs for 10 labeled surfaces when batch size is 1, which empirically pretrains the parcellation network suciently). 2) We jointly train registration and parcellation networks (SEPDRD) by optimizing L fs s L s sp L sp , where represents the weights, using all available surfaces for 10 epochs. 3) We incorporate L ps and optimize the full objective L fs s L s sp L sp ps L ps for 100 epochs (or early stopped when obtaining stable results). Herein, L ps teaches PD to parcellate an unlabeled surface such that the predicted parcellation matches the inversely warped manual parcellation of the atlas via RD. Conversely, Lps enforces RD to correctly warp the predicted parcellation map via PD to match the atlas parcellation map, thus enabling mutual learning between the two tasks.3Experiments and Results3.1 Experimental SettingWe used an infant dataset with 623 cortical surfaces, which were reconstructed via an in-house computational pipeline tailored for infants. We mapped each surface onto the sphere with 2 features at each vertex, i.e., sulc (average convexity) and curv (mean curvature) [3], and labeled it into 36 gyrus-based regions based on the parcellation protocol in [2]. We used the public UNC 4D infant cortical surface atlas [5] as the xed surfaces. We randomly split the data into 60% for training, 10% for validation, and 30% for testing and made sure the longitudinal surfaces from the same subject are in the same set. All the results reported are then solely based on the hold-out test set. 6Baselines. We used our SEPD architecture in Fig. 1 (similar to Spherical U-Net [14]) as the baseline parcellation model (BL-Parc) trained via L sp , and SERD architecture as the baseline registration model (BL-Reg) trained via L fs s L s . Jointly training two independent BL-Parc and BL-Reg models (i.e., task-specic encoders for each task instead of the SE) is called JRP-woSEwL ps . Similarly, JRP-wSE-woL ps means jointly training with SE but without L ps . We also compared with 3 available registration methods, FreeSurfer [4], Spherical Demons [13], and an unsupervised learning approach [16].Implementation Details. We implemented our method using PyTorch. We used Adam optimizer with a xed learning rate 5e-4 for all deep learning experiments. Sulc and curv values were rstly normalized between [-1, 1]. The loss weights are cc 1.2, s 10.0, sp 2.0 and ps 5.0. Note that when computing L fs , we assign dierent weights for the similarities of sulc and curv maps, 0.75 for sulc and 0.25 for curv, because sulc is a more robust cortical folding measure, while curv is more variable and contains more noises for registration. These parameters are empirically determined. We used the ocial codes of Spherical Demons and FreeSurfer (7.1.0) for their experiments. We run all the experiments on a PC with an

NVIDIA RTX2080 Ti GPU and an Intel Core i7-9700K[1]

CPU.3.2 ResultsWe evaluate all the methods using the Dice overlap metric computed as in Eq. 4. For parcellation task, it is between predictions and the ground truth parcellations for registration task, it is between moved manual parcellation maps and the xed (atlas) parcellation map. We also provide the within-group correlation coecient (CC) of moved cortical features as in [16,5,9] for additionally evaluating the within-group

spatial normalization accuracy[6]

for registration task.Comparison of Registration-Only Methods.

As shown in Table 1,[4]

using our BL-Reg alone for cortical surface registration achieves comparable accuracy with available registration methods, but is 50 times faster than [16], 500 times faster than [13,4]. Note that conventional methods [13,4] do not have GPU implementation and we report the time on CPU for them in Table 1 (other methods are evaluated on GPU). We also run BL-Reg on the same CPU, it takes 0.4s and thus is still much faster than [13,4]. Compared to [16] that directly concatenates F and M as input in a coarse-to-ne manner, our SERD architecture fuses high-level features in deep feature space, thus avoiding the time-consuming reinterpolation of deformations in original spherical space while obtaining better results. This indicates the eective learning of high-level features in SE and the deformation computation based on boundary dierences in RD. To additionally validate the topology-preserving registrations, we computed the folded triangles [8] in moved surfaces and found there is no folded triangles for all methods.Ablation Study on Shared Encoder (SE). Comparing JRP-wSE-woLps with BL-Parc and BL-Reg in Table 1, we can see that jointly training SE improves Dice over separately training two networks by 0.86% and 0.84% for parcellation and registration tasks respectively, which indicates that SE learns and exploits the shared useful features between the two tasks successfully. Abbreviated paper title7ParcellationDice (%)88.48(3.68)---
-88.48(3.68)89.34(2.47)89.98(2.53)RegistrationModelsParc. only (BL-Parc)Reg. only (FreeSurfer [4])Reg. only (Spherical Demons [13])Reg. only (Zhao et al. [16])Reg. only (BL-Reg) (Ours)JRP-woSE-wL ps (Ours)JRP-wSE-woL ps (Ours)JRP-wSE-wL ps (Ours)Run timeDice (%)-CC sulc-CC curv-0.2792(0.0813)0.2720(0.1237)0.2955(0.0810)0.2968(0.0796)0.3313(0.0794)0.3095(0.0953)0.01s76.69(5.52) 0.7788(0.0549)76.58(5.83) 0.7825(0.0675)77.03(5.66) 0.7859(0.0526)77.34(3.82) 0.7946(0.0374)88.78(3.12) 0.8209(0.0339)78.18(4.39) 0.7996(0.0466)90.23(2.56) 0.8288(0.0317)
0.3359(0.0676)30min1.5min10s0.16s0.01s0.16s0.16s0.16sTable 1. Average (standard deviation) performance of cortical surface parcellation and registration on the hold-out test set using dierent models trained with all manual parcellations in the training set. Parc. means parcellation, Reg. means registration.Ablation Study on Parcellation Map Similarity Loss (L ps). Table 1 shows incorporating L ps substantially improves the performance, with 10% and 1.5% Dice improvement over separately trained networks for registration and parcellation, respectively. For JRP-woSE-wL ps , the L ps is computed between two manual parcellations according to Eq. 4, thus it does not provide any information for training the independent parcellation network but provides maximum extra supervision for the registration network and leads to huge improvement on registration accuracy. Further with SE (JRP-wSE-wL ps), L ps back-propagates gradients from RD to SE and thus also enhance the performance of parcellation task by using more eective features in SE.How Registration and Parcellation Help Each Other. We simulate the common practical use case where only few manually labeled surfaces are available. We randomly used N surfaces manual parcellations to train BL-Parc, BL-Reg and JRP-Net (JRP-wSE-wL ps) on the training set. To achieve more reliable results, we repeated the experiment 10 times for each model (each time with randomly selected manual parcellations). Fig. 2 shows the results on the test set averaged over 10 times experiments. We can see that in the extreme case where only atlases parcellations are available (11 atlases in UNC 4D atlas [5]), our method still achieves good performance, with large improvements over independently trained models. This demonstrates that registration did help parcellation by providing estimated supervision from augmented parcellation maps on unlabeled surfaces. Conversely, the registration performance is also boosted by the parcellation task (when N3) even using predicted parcellations to en-Fig. 2. Test Dice of the models trained with dierent numbers of manual parcellations. 8Moving surfaceFreeSurferSpherical Demons Zhao et al. [16]BL-RegJRP-wSE-wFixed surface12.0-12.0Parc. mapSulc mapParc. mapSulc mapParc. mapSulc mapSulcFig. 3. For each case (left to right) the moving surface sulc map (rst row) and manual parcellation map (second row), corresponding moved maps by dierent methods, and xed surface maps. Note that all the methods fairly take the features (sulc and curv) as input and output the deformation eld. The manual parcellation maps are only for validating the registration performance and were not used to drive the registration.force the anatomical consistency. Also note that the lines of registration Dice and parcellation Dice of our JRP-Net in Fig. 2 are highly coincident. A paired t-test with a signicance level 0.05 on the 10 times experiments shows there is no signicant dierence between them, which could be a strong evidence that the two tasks mutually guide each others training to successfully learn the optimal features using SE and nd the optimal solution for both tasks. Fig. 3 shows some practical registration cases from the test set that available registration methods all fail to align but are correctly aligned using our JRP-Net (N50). The suboptimal results of available methods are an inherent problem because they only use feature similarity as the registration objective, while our JRP-Net incorporating the parcellation map similarity can eectively solve this problem.4ConclusionIn

this paper, we propose[5]

the JRP-Net for joint registration and parcellation of cortical surfaces, which are connected by the shared encoder and parcellation map similarity loss. By leveraging the inherent relation between the two tasks, our shared encoder is eective in extracting more meaningful features shared by both tasks. The parcellation map similarity loss further enables eective mutual training between the two tasks. Both visual and quantitative results show large improvements of our method in both registration and parcellation over separately learned networks. In

future, we will release our model[3]

for enhancing cortical surface registration and parcellation for the research community.

Learning 4D Infant Cortical Surface Atlas with Unsupervised Spherical NetworksAbstract. Spatiotemporal (4D) cortical surface atlas during infancy plays an important role for surface-based visualization, normalization and analysis of the dynamic early brain development. Conventional atlas construction methods typically rely on classical group-wise registration on sub-populations and ignore longitudinal constraints, thus having three main issues 1) constructing templates at discrete time points 2) resulting in longitudinal inconsistency among dierent ages atlases and 3) taking extremely long runtime. To address these issues, in

this paper, we propose[5]

a fast unsupervised learning-based surface atlas construction framework incorporating longitudinal constraints to enforce the withinsubject temporal correspondence in the atlas space. To well handle the diculties of learning large deformations, we propose a multi-level multimodal spherical registration network to perform cortical surface registration in a coarse-to-ne manner. Thus, only small deformations need to be estimated at each resolution level using the registration network, which further improves registration accuracy and atlas quality. Our constructed 4D infant cortical surface atlas based on 625 longitudinal scans from 291 infants is temporally continuous, in contrast to the state-of-the-art UNC 4D Infant Surface Atlas, which only provides the atlases at a few discrete sparse time points. By evaluating the intra- and inter-subject

spatial normalization accuracy[6]

after alignment onto the atlas, our atlas demonstrates more detailed and ne-grained cortical patterns, thus leading to higher accuracy in surface registration.Keywords Surface registration Infant cortical surface atlas1IntroductionCortical surface atlases play an important role in neuroimaging studies by providing a common space for normalizing, comparing, and analyzing brain structure and function across dierent individuals and studies [15]. Considering the rapid growth of the cerebral cortex during early brain development, spatiotemporal (4D) surface atlases are essential to characterize and model such dynamic development of infant brains [17] and are typically constructed on the spherical space by taking advantage of the intrinsic spherical typology of the cortex [8]. Conventional 4D atlas construction methods [17,11,8,5,9,1,14] generally perform several rounds of template estimation and individual-to-template registration, and nally average (based on Euclidean distance [8], Wasserstein distance [2], or 2sparse representation [17]) the cortical features in each age group to obtain the atlas for each age. Although these methods and their constructed infant cortical surface atlases are popular and widely used to investigate early brain development in the past decade, they have three main issues. First, these methods require an extremely long runtime due to the iterative group-wise registration process. Second, they create the atlas at each age based on a subset of the data without exploiting the rich information from the whole dataset, leading to suboptimal atlases. Accordingly, the resulting 4D atlases are constructed at discrete, sparse time points based on subpopulations scanned by predened concentrated ages, e.g., 1, 3, 6, 9, 12, 18 and 24 months of age. These atlases are thus temporally discontinuous and cannot cover many important ages during infancy. Although the kernel regression approach [11,1,14] could be applied to interpolate the uncovered ages atlases with specic kernel designs, it was not fully investigated and developed for infants, due to the diculties in modeling the complicated, regionally-heterogeneous and nonlinear early brain development. Third, the methods [1,14] constructing atlases for each age independently based on the subset of the data may cause longitudinal inconsistency among dierent ages atlases when longitudinal subjects are involved. The reason is that, as revealed by previous works [17,11,8], without considering the longitudinal information, the longitudinally-corresponding anatomical points at dierent ages of the same subject are independently deformed to dierent locations in the atlas space, consequently leading to a longitudinally-inconsistent 4D atlas.To address these issues, we propose an unsupervised learning framework to learn a continuous 4D infant cortical surface atlas, inspired by the recent spherical networks developed on cortical surfaces [19] and an unsupervised atlas construction method [3]. Our approach jointly learns an age-conditional atlas synthesis network and an unsupervised registration network from the entire dataset without manually partitioning the dataset into subgroups, thus is computationally more ecient and powerful in learning rich, useful information from the whole dataset than conventional methods. Moreover, since learning large deformations is dicult [12,22,23], we propose a deep multi-level multi-modal spherical registration network (MM-SRegNet) to decompose the large deformations into small ones at each resolution step and perform cortical surface registration in a coarse-to-ne manner. Further considering the subject-specic longitudinal information of longitudinal subjects frequently seen in large-scale studies, we design a novel longitudinal constraint (LC) loss to enforce the within-subject temporal correspondence after registration to the atlas space, which is important for establishing longitudinally consistent 4D atlases [8]. To sum up, our contributions in this paper are1. We propose an unsupervised learning-based 4D cortical surface atlas construction framework to jointly learn an atlas synthesis network and MMSRegNet for generating the age-conditional atlas and the deformation eld registering the cortical surfaces to the age-matched atlases simultaneously. Besides, our MM-SRegNet enables exible and accurate cortical surface registration using multiple cortical features in a coarse-to-ne manner.) Longitudinalconstraintloss.......................24months3months1monthFig. 1. Overview of our unsupervised learning framework for 4D infant cortical surface atlas construction. The atlas synthesis network G t takes age attributes as input and outputs the conditional atlases, which are then registered to each surface in the input minibatch by MM-SRegNet (Fig. 2). Cortical surfaces of the same subject are grouped together to build the longitudinal constraint loss highlighted in orange color. The blank spheres indicate unavailable data at those time points.2. We efficiently and effectively incorporate longitudinal constraints into our atlas learning framework, which is essential for establishing within-subject longitudinally-consistent correspondences, thus obtaining longitudinally more consistent atlas with higher quality.3. Leveraging our proposed method, we construct the first longitudinally-consistent temporally-continuous 4D infant cortical surface atlas based on 625 longitudinal scans from 291 infants and show that our generated 4D atlas preserves more details of cortical folding patterns compared to the state-of-the-art [17].2MethodAs shown in Fig. 1, our goal is to jointly train a generative model G t that can synthesize atlas given on-demand ages, and a registration model that registers the synthesized atlas to each input surface, thus avoiding the expensive time cost of conventional multi-round registration and iterative template refinement process. Let S S 1 1 , . . . , S 1 n 1 , . . . , S i j , . . . , S N n N denote a longitudinal cortical surface dataset with N subjects and each subject i (i 1, . . . , N) has n i longitudinal scans, S i j denotes the jth time point cortical surface of subject i and t i j is the age of subject i at jth time point. Note that the number of longitudinal scans and the time points of each subject are not necessarily to be the same. Then a baseline framework aims to minimize the loss L BL on the whole datasetL BL L sim (S i j , A t j i i j) c u 2 d i,j u i j 2 s i,j s ui j 2 (1 ,) (Movedatlas ,)MMSRegNet Movedsurfacesofsubject6scalingand squaringlayers......... IndividualsurfacespaceVelocity field( )Deformation field( ) Minibatchofsurfaces( )MinibatchMinibatch1MinibatchFeaturesimilarityloss (AtlasspaceSubjectSubject1SubjectAtlas Title Suppressed Due to Excessive Length3Age 4cov(S i j ,A t j i j)whereLsim (S i j , A t j i j) 1iis the correlation coecient loss for S i j A t j i j i enforcing the similarity between the moved atlas and individual cortical surface maps, A t i j G t (t i j) is the synthesized atlas at age t i j , i j is the deformation eld aligning atlas to jth surface of subject i and exp(v) is computed using

6 scaling and squaring layers as in[2]

[18], which is an eective extension of dieomorphic deformation from Euclidean space to spherical space. u x (x) (x) x is the tangent displacement vector on a unit sphere from x to (x). Hence, the rest terms in Eq. 1 regularize the unbiasedness, extent and smoothness of u, respectively. Note that we use the same s operator as in [20] on the spherical surface to approximate the tangent displacement elds gradients and accordingly penalize it to encourage a smooth deformation eld on the spherical surface.2.1 Longitudinal ConstraintSince the above baseline framework ignores the within-subject longitudinal consistency constraint when involving longitudinal subjects, we are motivated to incorporate it into our framework. To achieve this, we rst group the longitudinal surfaces from the same subject in one minibatch, as shown in Fig. 1. Then we can explicitly enforce the longitudinal correspondence among the surfaces of the same subject after moving them to the atlas space, which is formulated as the longitudinal constraint (LC) lossL LC n i 1 L sim (S i j (i j) 1 , S i j1 (i j1) 1) i j0(2)where ( ) 1 denotes the inverse deformation of , which can be conveniently computed using the method in [18]. This novel L LC is very ecient to integrate, as it adds 0 parameters and only a small overhead over computation to the baseline framework.2.2 Multi-level Multi-modal Spherical Registration NetworkIn unsupervised registration framework, it is known that large deformations are dicult to learn [12,22,23]. To address this issue, a typical method is stacking multiple networks in a pyramid fashion and perform the registration in a coarse-to-ne manner [16,12,22,23]. Inspired by these networks, we propose the MM-SRegNet to perform the coarse-to-ne cortical surface registration. A recent unsupervised cortical surface registration framework [20] employs three Spherical U-Nets [21] to predict spherical deformation elds for three equatorial regions at three orthogonal orientations, thus addressing the polar-distortion issue in spherical surface registration. We build on this method and incorporate the novel spherical scaling and squaring layers [18] to ensure dieomorphic (invertible and topology-preserving) deformations to construct SRegNet and further stack it to construct MM-SRegNet, as shown in Fig. 2. Note that multimodal cortical SRegNetDownsampleScalingandsquaringUpsampleFig. 2. Illustration of our MM-SRegNet. The superscript (6c) indicates the level 6 and the feature (s for sulc, c for curv). A denotes the atlas surface and S is the individual surface. Note that this only provides a brief overview of our method, which can be flexibly extended to other levels and cortical features.features (such as geometry, myelin or functional gradient) are imperfectly correlated [13], therefore our MM-SRegNet also offers much flexibility in the choice of cortical features and thus can provide on-demand atlas flexibly.In this paper, to obtain 4D atlas with fine cortical folding patterns, we follow the popular registration configuration [18,13] to align geometric features, i.e., sulc (average convexity) and curv (mean curvature), measuring cortical folding in a coarse and fine view, respectively. We stack 4 SRegNets at 4 levels (4, 5, 6, 6, with 2,562, 10,242, 40,962, 40,962 vertices, respectively) for aligning sulc, sulc, sulc, curv, respectively. Finally, we deform the generated sulc atlas to the final curv atlas space using the learned deformation field at the last level.2.3 Network and Implementation DetailsThe generative model G t for synthesizing the atlas consists of a fully connected layer with 102423 neurons, a reshape layer and an upsampling layer followed by 3 batch normalization [7] spherical 1-ring convolution [19] ReLU blocks each with [16, 32, 1] channels. Herein, G t takes age t i j as a scalar value input and outputs the age-conditional atlas (409621). For the registration network, since we stack multiple SRegNets to perform the coarse-to-fine registration, we create a smaller version of Spherical U-Net [19] used in SRegNet, with 3 resolution steps and [8, 16, 32] channels, resulting in only 0.5M parameters in each SRegNet.We implemented our method using PyTorch. We trained the atlas synthesis network G t and registration network MM-SRegNet jointly by minimizing the full objective L BL LC L LC in a minibatch style. We trained the network using Adam optimizer with learning rate 5e-4 for 50 epochs. c 1.0, d 0.2, s 1.5, 2.5 and 4.0 at 4th, 5th and 6th level, respectively, and LC 1.0.Note that although looking bigger, our final model is compact with only a small number of parameters, G t with 66K, and MM-SRegNet with 6M parameters in total. Taking about 10 hours to fully train the networks on a PC with) SRegNetMovedsulc map ( )) map( ) SRegNetMovedcurv map ((Sulcatlas Inputsurfacesulc)) map((AttributeCurvatlas InputsurfacecurvTitle Suppressed Due to Excessive Length5 6an

NVIDIA RTX2080 Ti GPU and an Intel Core i7-9700K[1]

CPU, it only needs 5 seconds for one forward inference, including atlas generation and individualto-atlas registration, which is much faster than conventional atlas construction and registration methods, where the whole process can take several days.3Experiments3.1 Dataset and Preprocessing We used a longitudinal infant dataset with 625 scans from 291 subjects, with ages ranging from 0 to 5 years. The distribution of the number of longitudinal scans is, 44% infants with only 1 scan, 24% with 2 scans, 14% with 3, 10% with 4 and 8% with 5 or more scans. Cortical surfaces were reconstructed via iBEAT V2.0 Cloud (httpwww.ibeat.cloud) and then mapped onto the sphere using FreeSurfer [6]. Each surface is then represented as a spherical map with 2 features at each vertex, i.e., sulc and curv, and a parcellation map is also generated based on [4]. We randomly split the dataset into a training set with 418 scans from 185 subjects, and a test set with 207 scans from 106 subjects. In testing stage, we generate the age-matched atlas for each test surface and the corresponding deformation eld aligning it to the age-specic atlas simultaneously.To quantitatively validate the atlas, we calculate the widely-used withingroup

spatial normalization accuracy[6]

, which is measured by the Pearson correlation coecient (PCC) of feature maps and Dice overlap metric of parcellation maps as in [17,8,13,20]. Generally, sharper atlases lead to higher registration accuracy and thus better spatial normalization performance. Therefore, a higher PCC and Dice indicate better within-group alignment and thus a better atlas.3.2 Ablation Studies MM-SRegNet. To demonstrate the eectiveness of our MM-SRegNet, we trained two baseline models on the 6th level, named B-Sulc, B-Curv for generating two baseline sulc and curv atlas with single level SRegNet. Since they use single-modal feature, the most critical problem of these two baselines is that the constructed atlases may not be in the same space, which causes trouble in neuroimaging analysis.

As shown in Table 1,[4]

the B-M model using our MMSRegNet as the registration network leads to a higher correlation and Dice for both within-subject and across-subject alignment. This is important for building atlases with other cortical features accurately and exibly in the future study. Longitudinal Constraint (LC).

As shown in Table 1,[4]

adding LC loss to the baseline model (named B-MLC) signicantly improve the performance, which is reasonable for correlation metrics, since we directly add the within-subject correlation as an additional loss to train the model. As a more unbiased metric, the Dice improvement is a compelling evidence that incorporating LC strengthens the within-subject temporal correspondence and thus enhances the longitudinal consistency and the quality of the synthesized 4D atlas. Fig. 3 provides a typical Title Suppressed Due to Excessive Length 7 Table 1. Within-subject and inter-subject alignment performance of dierent models.Within-subject alignmentPCCsulc PCCcurv Dice (%)0.8790.051-83.455.76-0.5570.08883.315.790.9070.0440.6770.10085.025.320.9270.0400.7430.06786.525.10Inter-subject alignmentPCCsulc 0.6720.062-0.7280.0600.7510.052PCCcurv -0.3370.0530.4840.0460.5280.057Dice (%)B-SulcB-CurvB-MB-MLC70.924.4069.396.1674.915.4676.124.8513691218240.60.6Fig. 3. Comparison of the learned 4D surface curv atlases using B-M (upper row) and B-MLC (lower row) at discrete ages (months) in the rst two postnatal years.example of the curv atlases learned without LC and with LC. As we can see, the LC encourages longitudinally more consistent 4D cortical surface atlases along with much shaper cortical folding patterns, than the model without LC.3.3 Comparison with State-of-the-art Visual Inspection. We used the B-MLC model to generate the infant cortical surface atlases at corresponding discrete ages with the state-of-the-art UNC 4D infant cortical surface atlas [17] for comparison. Fig. 5 shows the two atlases only in the rst two postnatal years because of the most dynamic postnatal development during this period [10]. We can see that our atlas preserves more detailed cortical folding patterns than UNC atlas in most regions, e.g., insula, inferior parietal and lateral temporal cortex. More importantly, as indicated by arrows, minor cortical folds, which are still developing rapidly during infancy, are well captured by our atlas, but absent in UNC 4D atlas.Quantitative Evaluation. We used a third-party registration tool, Spherical Demons [18], a conventional method solving registration problem using GaussNewton method, to fairly register the hold-out test surfaces to our atlas and UNC atlas, respectively. Note that we choose the age-matched UNC atlas thatFig. 4. Spatial normalization accuracy at each age group based on dierent atlases. 81month3months6months9months12months18months24monthsgyral crownsulcalfundusMeancurvature Ours UNCAverageconvexity Ours UNCFig. 5. Comparison of UNC 4D surface atlas and ours. Note that though limited time points are shown, our method oers continuous 4D atlas at arbitrary time points from 0 to 5 years of age.are closest to the individual surfaces age for each individual surface (11 instances available in UNC atlas, i.e., 1, 3, 6, 9, 12, 18, 24, 36, 48, 60, and 72 months), while our method generates on-demand age-specic atlas at arbitrary time points and thus can provide more accurate and matched atlas to guide the registration. The average PCC of sulc, PCC of curv, and Dice is 0.7060.039, 0.3730.063, 0.8120.049 for UNC atlas, and 0.7200.043, 0.3820.058, 0.8210.030 for our atlas. Detailed results in the rst two years are shown in Fig. 4. We can see that our atlas outperforms UNC atlas in most age groups and shows no signicantly dierences if not better, except PCC of curv and Dice at 1 month likely because of the limited samples in this group, demonstrating the practical usefulness of our higher-quality atlas in improving surface registration and parcellation accuracy.4ConclusionIn this paper, we present three methodological contributions to eciently and accurately build the rst longitudinally-consistent temporally-continuous 4D infant cortical surface atlas. First, we propose an unsupervised learning framework for simultaneously atlas synthesis and individual-to-atlas registration. Taking advantage of the deep learning ability of neural networks, our method can eciently learn the atlas from the whole dataset and rapidly warp the new surfaces to the generated atlas. Second, our novel MM-SRegNet can accurately perform cortical surface registration in a coarse-to-ne manner exibly using multiple cortical features. Third, incorporating the longitudinal constraint further improves within-subject longitudinal consistency and thus the quality of 4D atlas. Compared with the state-of-the-art UNC 4D atlas both visually and quantitatively, our atlas exhibits sharper and much more detailed patterns for more accurately mapping infant brain development. In

future, we will release our model[3]

and atlas to the community to advance the neuroimaging studies of infants.