IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 29, 2019, accepted February 24, 2019, date of publication March 11, 2019, date of current version March 29, 2019.

*Digital Object Identifier 10.1109/ACCESS.2019.2904143*

# Hippocampal Segmentation From Longitudinal Infant Brain MR Images via Classification-Guided Boundary Regression

**YEQIN SHAO[1,2], JAEIL KIM[3], YAOZONG GAO[4], QIAN WANG[5], WEILI LIN[6], AND DINGGANG SHEN[6,7], (Fellow, IEEE)**

[1]School of Transportation and Civil Engineering, Nantong University, Jiangsu 226019, China
[2]Fujian Provincial Key Laboratory of Information Processing and Intelligent Control, Minjiang University, Fuzhou 350121, China
[3]School of Computer Science and Engineering, Kyungpook National University, Seoul 41566, South Korea
[4]Shanghai United Imaging Intelligence, Ltd., Shanghai 201210, China
[5]Institute for Medical Imaging Technology, School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai 200030, China
[6]Department of Radiology and BRIC, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA
[7]Department of Brain and Cognitive Engineering, Korea University, Seoul 02841, South Korea

Corresponding authors: Qian Wang (wang.qian@sjtu.edu.cn) and Dinggang Shen (dinggang_shen@med.unc.edu)

The work of Y. Shao and Q. Wang were supported in part by the National Natural Science Foundation of China (NSFC) under Grant 61671255, in part by the National Key R&D Program of China under Grant 2017YFC0107602, in part by the Science and Technology Commission of Shanghai Municipality under Grant 16410722400, and in part by the Open Fund Project of Fujian Provincial Key Laboratory of Information Processing and Intelligent Control, Minjiang University, under Grant MJUKF-IPIC201808. The work of D. Shen was supported by the National Institutes of Health (NIH) under Grant EB006733, Grant EB008374, Grant EB009634, Grant MH100217, Grant MH108914, Grant AG041721, Grant AG049371, and Grant AG042599.

**ABSTRACT** Hippocampal segmentation from infant brain MR images is indispensable for studying early brain development. However, most of the hippocampal segmentation methods were developed for population-based adult brain images, which are not suitable for longitudinal infant brain images acquired in the first year of life due to the low image contrast and variable development patterns of the hippocampal structure. To address these challenges, we propose a classification-guided boundary regression method to first detect hippocampal boundaries in the longitudinal infant brain images and then use those detected boundaries to guide the deformable model for final segmentation. Specifically, we first employ a classification-guided regression forest to predict the 3D displacements from individual image voxels to the potential hippocampal boundaries. These predicted displacements then determine the boundary maps by a voting strategy. Second, we iteratively enhance the voted hippocampal boundary map by incorporating the spatial context information given the tentative boundary estimation of the current time point. Besides, the longitudinal context information from all time points of the temporal sequence of the same subject (i.e., given their tentative segmentation results) is also utilized to facilitate accurate segmentation. Finally, a deformable model is applied to the enhanced voted boundary maps for achieving the longitudinal hippocampal segmentation. The experiments on infant brain MR images acquired from 2-week-old to 1-year-old show promising hippocampal segmentation results, indicating the applicability of our method in early brain development study.

**INDEX TERMS** Infant brain, longitudinal MR image, hippocampal segmentation, classification-guided boundary regression, deformable segmentation.

## I. INTRODUCTION

Accurate segmentation of brain structures in magnetic resonance (MR) images is a prerequisite to explore brain development [1]–[3]. Among the various brain structures, hippocampus plays an essential role in memory and learning functions [4]–[6]. Morphometric development of hippocampus is an important biomarker for investigating brain structure and function alterations in many clinical studies [7], [8]. Particularly, infant brain undergoes critical and fast postnatal development in the first year of life, which is the foundation of human brain development. Accurate hippocampal

The associate editor coordinating the review of this manuscript and approving it for publication was Wenming Cao.

segmentation from infant brain MR images is highly desired to study early brain development [9] and neurodevelopment disorders (such as autism and attention deficit hyperactivity disorder [10]).

However, it is still a great challenge to accurately segment the hippocampi from infant brain MR images due to the following reasons.

(1) *Ambiguous hippocampal boundary.* Image intensities in the hippocampi are very similar to those in the nearby structures, such as amygdala, caudate nucleus, and thalamus [11]. Therefore the image contrast around the hippocampal boundary is relatively low, which results in the ambiguous boundary (shown in FIGURE 1).

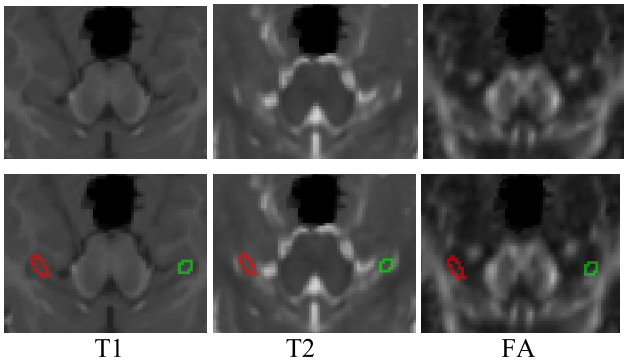



**FIGURE 1.** Illustration of the hippocampi in the T1, T2, and FA images of an infant brain. The red and green contours at the bottom row indicate the left and the right hippocampi, respectively. It can be observed that the image contrast is low near the boundaries of the hippocampi.

(2) *Small hippocampal volume.* The size of the hippocampi is smaller than many nearby structures. Thus, it is difficult to accurately locate and segment the hippocampi in the entire MR image.

To address the above challenges, many hippocampal segmentation methods have been proposed. These segmentation methods mainly fall into 2 categories.

### A. ATLAS-BASED METHODS

#### 1) SINGLE-ATLAS METHODS

In the atlas-based methods [12]–[14], single-atlas methods are first presented. A single atlas image is registered to the testing image. Then, the label map of the atlas is warped to the testing image for final segmentation with the estimated deformation field. For example, Barnes *et al.* [12] performed global affine registration to obtain a hippocampal region-of-interest (ROI), and another local affine registration to refine the obtained ROI for the final segmentation. Carmichael *et al.* [13] evaluated the atlases selection and the registration methods for better hippocampal segmentation. Kwak *et al.* [14] adopted a graph-cut algorithm to optimize the atlas-based hippocampal segmentation result. Among the single-atlas methods, the highly-variable anatomy of target tissue often hinders the segmentation accuracy.

#### 2) MULTI-ATLAS METHODS

To alleviate the high anatomical variability across images, multiple atlases are also adopted for the atlas-based hippocampal segmentation. These methods register multiple atlases to the testing image, and then derive the final segmentation by label fusion. For example, Kim *et al.* [2] used the multi-atlas strategy and the sequences of location-adaptive classifiers to segment the hippocampi from 7T MR images. Hao *et al.* [3] adopted a local label learning strategy to fuse labels using statistical machine learning technique. Dong *et al.* [15] proposed to learn a hypergraph of image voxels to propagate the labels from multiple atlases for hippocampal segmentation in the testing image. Wu *et al.* [16] proposed both multi-scale feature representation and label-specific patch partition to determine the hippocampal label. Wang *et al.* [17] used a multi-atlas method to segment the hippocampi with a joint label fusion strategy. Zhu et al. [18] learned a distance metric to propagate the atlas labels. Guo *et al.* [19] employed hierarchical multi-set kernel canonical correlation analysis to learn feature representation and propagated the labels under a multi-atlas framework. The main drawback of the atlas-based methods is the high computation cost related with the deformable registration (between atlases and the testing image) and also the sophisticated label fusion.

### B. DEFORMABLE-MODEL-BASED METHODS

Deformable model is also widely used for the hippocampal segmentation in MR images. Zarpalas *et al.* [20] employed adaptive gradient distribution boundary map in the Active Contour Model (ACM) to segment the hippocampi. Hajiesmaeili *et al.* [21] presented an initialization technique, which divides the hippocampus into head, middle and tail, under the ACM framework. Gao *et al.* [22] proposed a region-based signed pressure force function to outline the hippocampal boundary. Luo *et al.* [23] employed 3D Active Appearance Model (AAM) to segment the hippocampi. Hu *et al.* [24] used AAM to capture global shape characteristics of the target structure and refined the segmentation result in a patch-wise way. Ettaïeb *et al.* [25] employed Active Shape Model (ASM) based on the spatial distance to segment the hippocampi. Although the deformable-model-based methods are extensively studied, they often suffer the difficulty in designing the initialization and deformation strategy. Besides, the size of infant hippocampi is small, hence the deformable-model-based methods tend to obtain suboptimal segmentation on infant hippocampal images.

Recently, boundary regression [26] achieves promising segmentation performance in medical image analysis. Chen *et al.* [27] outlined the target organ boundary with landmarks, and predicted the boundary by detecting each landmark. Wang *et al.* [28] represented the boundary with key boundary points and used multi-dimensional support vector regressor to jointly estimate those boundary points on multiple imaging modalities. Although these methods can predict

the target boundary, they need to train a large number of point detectors to outline the whole target boundary, which is highly time-consuming. Recently, Shao *et al.* [29], [30] adopted a single regression forest to effectively predict the nearest boundary point from each image voxel and thus achieved the whole target boundary through a large number of votes. Based on the local image patch around each voxel under consideration, a regression forest is trained to predict the displacement from each image voxel to its nearest hippocampal boundary point. In this way, the entire hippocampal boundary can be voted by collecting the nearest boundary point predictions from a large number of image voxels. Although the boundary regression method is robust to many image segmentation tasks, it is still challenged by the extremely low image contrast around the infant hippocampal boundary (c.f. FIGURE 8(a)), especially when the local image appearance is merely considered by the regression forest.

In this paper, we propose a classification-guided boundary regression method to accurately determine the hippocampal boundary based on local image patches, and then use the obtained boundary map to drive the deformable model for final hippocampal segmentation in the infant brain MR images. Specifically, a boundary regression forest [30] is first employed to regress the 3D displacement from each image voxel to its nearest hippocampal boundary point. Then, all image voxels cast their votes at the potential nearest boundary points with the predicted displacements, thus generating the hippocampal boundary map (as shown in FIGURE 8(a)). Since the classification map implies the shape information of the hippocampi, we further integrate a voxel-wise classification forest into the boundary regression forest, hence generating both the boundary displacement map (outputted by the boundary regressor) and the hippocampus classification map (outputted by the classification forest). Here, we call the boundary regression forest, integrated with a classification forest, as the classification-guided boundary regression method (c.f. Section II.A). Since the appearance characteristics of the infant brain structures in MR images change dynamically during the first year of life, especially those of the white matter and gray matter, we learn different classification-guided boundary regression forests for different time points (2-week, 3-month, 6-month, 9-month and 12-month) of the first year of life.

Due to the ambiguous hippocampal boundary, we integrate the spatial and the longitudinal contextual information into the classification-guided boundary regression forest to refine the boundary map. The outputted hippocampus classification map and boundary displacement map aforementioned are again inputted into the classification-guided boundary regression forest to extract the *spatial* context information for better learning performance (c.f. Section II.B.1). Additionally, to consider the longitudinal consistency of multiple time points, we also extract the *longitudinal* context information from the tentative segmentation results of other time points of the same subject (c.f. Section II.B.2). The tentative segmentation results are achieved by a deformable model on the boundary map, which is derived from the outputted boundary displacement map (c.f. Section II.C). With these *spatial* and *longitudinal* context information, we can iteratively refine the estimation of the hippocampal boundary map.

In the final, the refined boundary map is used to drive the deformable model to complete the final segmentation of the hippocampi (c.f. Section II.C). Since the infant hippocampal boundary is ambiguous, we also integrate the shape priors into the deformable model for accurate hippocampal segmentation. To achieve better accuracy, the infant hippocampi are segmented on the multi-modal neuroimaging data, including T1-, T2-weighted MR images, and fractional anisotropy (FA) image from diffusion tensor imaging (DTI). Experimental results show that our method achieves promising segmentation results for the infant brain image sequences acquired from 2-week-old to 1-year-old.

The contributions of our proposed method are three-fold. (1) We propose to integrate voxel-wise classification into the boundary regression to exploit the implicit hippocampal shape information for accurate boundary regression. (2) Besides the spatial context information, the longitudinal context information of the same subject is applied to facilitate the consistent segmentation of a temporal sequence of hippocampi from one subject. (3) We employ a deformable model to incorporate shape priors for the challenging hippocampal segmentation.

The rest of this paper is organized as follows. Section II gives an overview of the classification-guided boundary regression method, followed by the details of each component in our method. Section III shows the experimental results. Section IV presents the discussion. Finally, the paper concludes in Section V.

## II. METHODS

Our method aims to accurately segment the hippocampi from the longitudinal infant brain MR images by the proposed classification-guided boundary regression (c.f. FIGURE 2). Specifically, we **first** combine the classification and the regression forests for the classification-guided regression forest to obtain the hippocampal boundary map (c.f. Section II.A). **Second**, to refine the hippocampal boundary map iteratively, the classification-guided regression forests integrate both the *spatial* context information from the tentative classification/regression output of the same time point and the *longitudinal* context information from the tentative segmentation output of other time points (c.f. Section II.B). **Third**, a deformable model is driven by the improved boundary map to complete the final hippocampal segmentation (c.f. Section II.C).

### A. CLASSIFICATION-GUIDED BOUNDARY REGRESSION

Recently, random forest achieves promising performance in image processing [27], [31]–[37] due to its efficiency and robustness. A classical random forest [38] , which is a regression forest with continuous targets or a classification forest

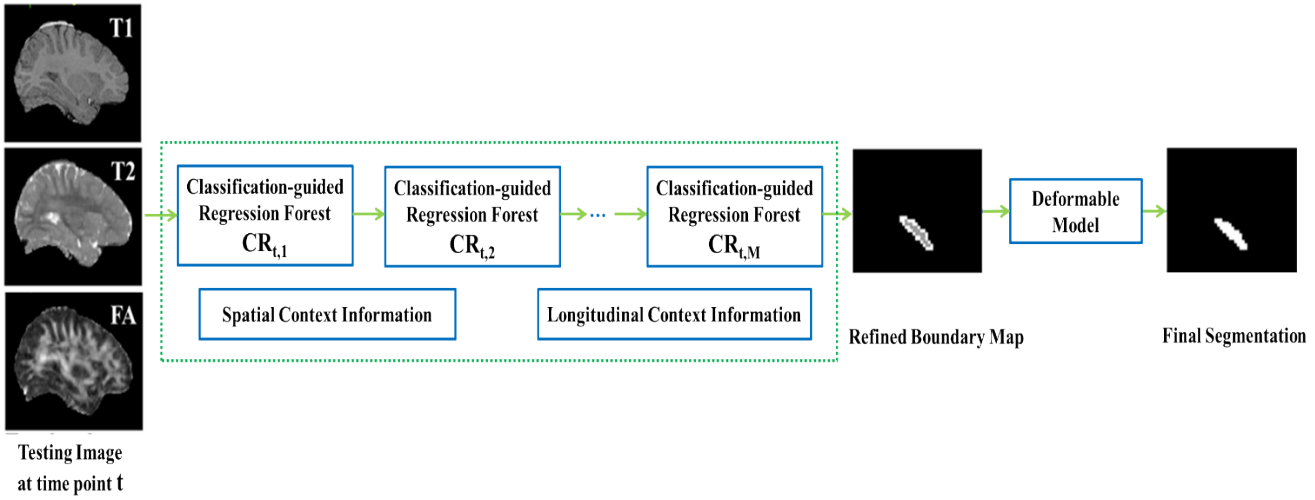



**FIGURE 2.** Flowchart of our method at time point $t$. For simplicity, we only take the left hippocampus as example.

with discrete labels, generally contains an ensemble of binary decision trees. Each tree randomly selects a subset of training samples for the root node. Then each tree grows from the root node by iterative node splitting with a sequence of features selected from a large feature pool and the corresponding optimal thresholds. The node splitting for a decision tree repeats until the tree reaches the maximum depth or the minimum number of training samples in a leaf node. Using a set of trained decision trees, a random forest can predict the classification label or the regression target for each testing sample. The final prediction of a testing sample is obtained by fusing all outcomes of decision trees in the forest.

In this work, to exploit the shape information implied in voxel-wise hippocampal classification, we integrate the classification forest into the boundary regression forest for accurate hippocampal segmentation. Therefore, we need to build (1) a classification forest that estimates the voxel labels (classification labels) within an image patch in a structured way, and (2) a regression forest that predicts the displacement vector (regression target) from the central voxel of a patch to the nearest hippocampal boundary point. The details are described in the following sections.

### 1) STRUCTURED CLASSIFICATION FOREST

Motivated by [39], we use the structured classification forest to exploit the structured information of the target organ for accurate hippocampal identification in the image. Different from the traditional classification forest [40], the structured classification forest predicts the labels for a local patch of voxels simultaneously, instead of the label of a single voxel. With the structured classification forest, a hippocampal classification map can be achieved based on an intensity image. The structured classification forest of this paper is detailed as follows:

**Sampling:** In the training stage, we randomly draw a large number of samples $\{\boldsymbol{p}_i\}$ $(i = 1, 2, \ldots, K)$ near the manually annotated hippocampal boundary using a Gaussian distribution for each training image. Mathematically, $\boldsymbol{p}_i = \boldsymbol{q}_i + \boldsymbol{N}(\boldsymbol{q}_i)\delta$, where $\boldsymbol{q}_i$ is a randomly selected point on the manually annotated hippocampal boundary, $\boldsymbol{N}(\boldsymbol{q}_i)$ is the normal direction of the boundary at $\boldsymbol{q}_i$, and $\delta = \mathcal{N}(0, \sigma)$ indicates a random offset (positive or negative) along the normal direction. In this way, the samples are concentrated around the hippocampal boundary, which makes our method specific to the hippocampal detection and segmentation. In the testing stage, to identify the hippocampi in the testing image, all image voxels are used to predict the hippocampi.

**Feature Extraction:** Given the voxel samples, i.e., $\{\boldsymbol{p}_i\}$ $(i = 1, 2, \ldots, K)$, each sample is represented by a randomly generated, extended Haar-like feature vector $F_i$ (i.e., $F_i = \left[F_{i,T1}^T, F_{i,T2}^T, F_{i,FA}^T\right]^T$) from the respective local image patches on the T1, T2, and FA images. The extended Haar-like feature vector $F_{i,\theta}(\theta \in \{T1, T2, FA\})$ on each image modality includes a set of extended Haar-like features. Each extended Haar-like feature consists of one or more cubic blocks defined as Eq. (1).

$$f_{i,\theta} = \sum_{j=1}^{U} \beta_j \bar{b}_{i,\theta}(\mathbb{L}_j, \mathbb{S}_j), \tag{1}$$

where $\theta \in \{T1, T2, FA\}$, $U$ is the number of the cubic blocks (randomly selected between "1" and "2"), $\bar{b}_{i,\theta}(\mathbb{L}_j, \mathbb{S}_j)$ is the average image intensity on modality $\theta$ of the cubic block, which locates at $\mathbb{L}_j$ with the size of $\mathbb{S}_j$ within the image patch centered at $\boldsymbol{p}_i$, and $\beta_j = (-1)^{(j-1)}$ is the polarity of the $j$-th cubic block. To exploit the abundant appearance information of the local patch centered at $\boldsymbol{p}_i$, the number $U$, location $\mathbb{L}_j$ and size $\mathbb{S}_j$ of the cubic block are randomly determined.

**Node Splitting:** With the feature vectors and the corresponding labels of those drawn samples, we need to build classification decision trees by node splitting with a set of best features and corresponding thresholds. Here, the best

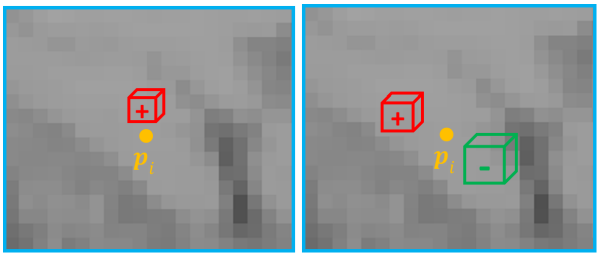



**FIGURE 3.** Illustration of the extended Haar-like feature. The whole cyan rectangle indicates the image patch centered at $p_i$. The red and green cubes indicate the cubic blocks, the location and size of which are randomly generated.

feature and the threshold for each node are determined by the maximum information gain of the classification labels upon node splitting. To avoid expensive computation, in this paper, the information gain is defined in Eq. (2):

$$\mathrm{IG}_i^C = \mathrm{E}(O_i) - \sum_{\mathcal{H}\in\{L,R\}} \frac{|O_{i,\mathcal{H}}|}{|O_i|}\mathrm{E}\left(O_{i,\mathcal{H}}\right), \tag{2}$$

$$\mathrm{E}(O_i) = \frac{1}{|O_i|}\sum_{\boldsymbol{l}\in O_i} \left\|\boldsymbol{l}-\bar{\boldsymbol{l}}\right\|_2^2, \tag{3}$$

where $\mathrm{IG}_i^C$ is the information gain for classification decision tree, $O_i$ is the samples in the $i$-th node prior to the splitting, $O_{i,\mathcal{H}}(\mathcal{H}\epsilon\{L,R\})$ indicates the samples in the left or the right children nodes after the splitting, and $\mathrm{E}(\cdot)$ evaluates the variation of the sample labels. In particular, $\boldsymbol{l}$ is a patch-based label vector, generated by concatenating labels in a target image patch, and $\bar{\boldsymbol{l}}$ is the mean of the label vectors in $O_i$. In this paper, each leaf node in a decision tree stores the probability distribution of labels for each voxel of the local patch. The probability distribution of labels is computed as the classical classification forest [38]. Given a testing image, we extract extended Haar-like feature vector from an image patch for each image voxel. Then we go through the trained decision tree, based on the feature vector, to predict the voxel labels in the local patch centered at the image voxel.

### 2) CLASSIFICATION-GUIDED BOUNDARY REGRESSION

To accurately segment the infant hippocampi, we propose the classification-guided boundary regression by integrating the structured classification into the boundary regression forest. In the training stage, we draw voxel samples $\{\boldsymbol{p}_i'\}$ $(i = 1, 2, \ldots, K')$ and generate their Haar-like features from the multi-modal images as Section II.A.1). Then, we compute the displacement vector $\boldsymbol{d}_i'$ from each sample $\boldsymbol{p}_i'$ to its nearest hippocampal boundary point as the regression target, and the label vector $\boldsymbol{l}_i'$ from the local patch centered at each sample $\boldsymbol{p}_i'$ of the annotation image as the classification label (c.f. FIGURE 4).

With the triples (the Haar-like features, displacement vectors and label vectors) of all training samples, we train a classification-guided regression forest $CR_1$. Different from both classification forest and regression forest, the node

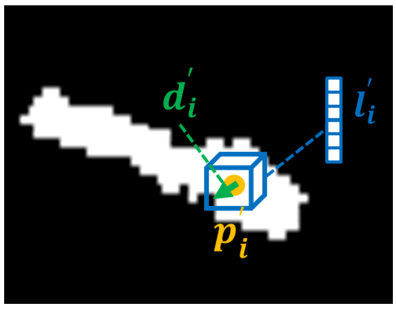



**FIGURE 4.** Illustration of the displacement vector and label vector of our classification-guided boundary regression. It is an annotation image. The yellow point is a sample point, the green arrow indicates the displacement vector from the sample point to the nearest boundary point, and the blue cubic indicates the structured label, which can be concatenated as a label vector.

splitting of the classification-guided regression forest is determined by the auxiliary information gain $\mathrm{IG}_i = (1-\alpha)\mathrm{IG}_i^C + \alpha IG_i^R$, where $\alpha\epsilon[0,1]$ is a scalar, $\mathrm{IG}_i^C$ is computed from the classification part (defined as Eq. (2)), and $\mathrm{IG}_i^R$ is computed from the regression part (defined as Eq. (4)).

$$\mathrm{IG}_i^R = \mathrm{H}(O_i) - \sum_{\mathcal{H}\in\{L,R\}} \frac{|O_{i,\mathcal{H}}|}{|O_i|}\mathrm{H}\left(O_{i,\mathcal{H}}\right), \tag{4}$$

$$\mathrm{H}(O_i) = \frac{1}{|O_i|}\sum_{\boldsymbol{y}\in O_i} \|\boldsymbol{y}-\bar{\boldsymbol{y}}\|_2^2. \tag{5}$$

In particular, $\mathrm{H}(\cdot)$ indicates the variation of regression targets, $\boldsymbol{y}$ is a regression target in the $i$-th node, and $\bar{y}$ is the mean of regression targets in $O_i$. In this way, each leaf node in the classification-guided regression forest stores both boundary displacement vectors and classification label vectors of samples falling in this node. For the purpose of testing, we record the selected features and the respective thresholds of node splitting during the training stage.

In the testing stage, we test the voxels in a bounding box of a testing image to predict the hippocampal boundary. The hippocampal bounding box is a rectangular region, centered at the initially estimated hippocampal location. Here, the initial hippocampal location on the testing image is determined by the detected hippocampal landmarks (as described in Section II.C.1). Within the bounding box, we extract the Haar-like features $\widehat{\boldsymbol{F}'}$, whose pattern are previously recorded by the trained forest, from the local patch of each image voxel $\widehat{\boldsymbol{p}'}$. Then, with the trained classification-guided regression forest $CR_1$, we estimate both the displacement vector $\widehat{\boldsymbol{d}'}$ from $\widehat{\boldsymbol{p}'}$ to the nearest hippocampal boundary point, and the label vector $\widehat{\boldsymbol{l}'}$ of the local patch centered at $\widehat{\boldsymbol{p}'}$. Since the voxel samples closer to the hippocampal boundary can provide more reliable information than those far-away voxels in boundary regression, we adopt a weighted voting strategy to integrate the contribution of each testing voxel based on the magnitude of estimated displacement. Specifically, with $\widehat{\boldsymbol{p}'}$ and the estimated displacement $\widehat{\boldsymbol{d}'}$, we cast a weighted vote at the potential boundary point $\widehat{\boldsymbol{p}'} + \widehat{\boldsymbol{d}'}$. Here, the weight is

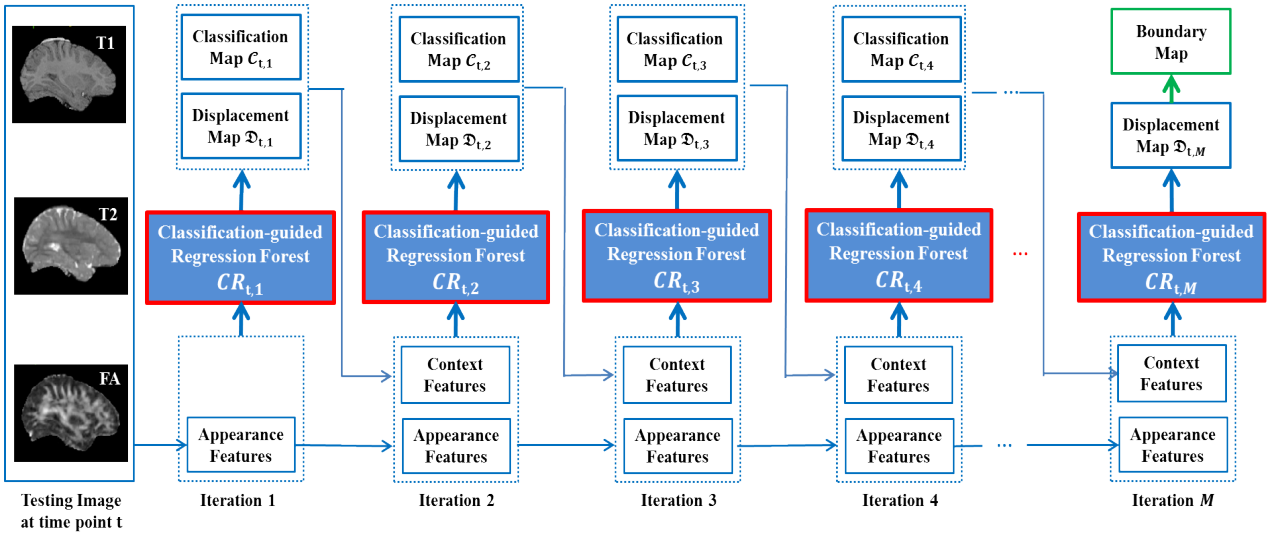



**FIGURE 5.** Flowchart of the classification-guided boundary regression with spatial context information for the time point $t$.

inversely proportional to the magnitude of $\widehat{\boldsymbol{d}}'$. By summing up all weighted votes from all voxels in the bounding box, we can finally get the hippocampal boundary map. Meanwhile, with the structured classification, the label of each voxel can be predicted by the voxel itself and the voxels in the local patch. By averaging the classification labels of each voxel in the bounding box, we can also acquire the hippocampus classification map, which can improve the hippocampal boundary map in an iterative way (c.f. FIGURE 5).

### B. SPATIAL AND LONGITUDINAL CONTEXT INFORMATION

If the image voxels independently predict their displacement vectors toward the hippocampal boundary, the resulted boundary map could be incoherent. To this end, we refine the voted hippocampal boundary map by combining the classification-guided boundary regression with (1) the spatial context information from the tentative classification/regression outputs of the same time point (c.f. Section II.B.1), and (2) the longitudinal context information from other time points (c.f. Section II.B.2) in the temporal sequence of the same infant subject.

#### 1) SPATIAL CONTEXT INFORMATION

To improve the boundary map, we enforce the spatial constraint in the classification-guided boundary regression under the auto-context model [41]. The main idea is to train a set of classification-guided regression forests $CR_m(m = 1, 2, \ldots, M)$, each of which is trained with both the appearance features of the multi-modal images and the spatial context features extracted from the outputs (i.e., the classification map and displacement map) of a previously trained forest $CR_{m-1}$. Note that the first classification-guided regression forest $CR_1$ is trained only by the appearance features of the multi-modal images. The spatial context features here are extracted in the same way as the Haar-like appearance features (c.f. Section II.A.1). The classification-guided regression forests can iteratively generate better classification maps and displacement maps, which contribute to more precise displacement estimation. Then, the final displacement vectors, implied in the displacement map, generate the boundary map by voting at the predicted boundary points. The flowchart of the classification-guided boundary regression with the spatial context information is shown in FIGURE 5.

#### 2) LONGITUDINAL CONTEXT INFORMATION

To obtain more accurate displacement estimation, we integrate the longitudinal context information from different time points of the same subject. Specifically, for each subject, its tentative segmentation results from different time points are mapped to the specific time point under consideration by rigid registration [42] (c.f. FIGURE 6). On the mapped segmentation results, the longitudinal context features are extracted in the same way as the Haar-like appearance features (c.f. Section II.A.1). Then, the *longitudinal* context features from the mapped segmentation results of other time points, and the *spatial* context features (c.f. Section II.B.1) from the classification/regression maps of the same time point, are inputted into the classification-guided regression forest with *weights*. Since the images from later time points usually have higher image contrast, which tends to produce more reliable segmentation results, the *weights* of the longitudinal context features increase gradually for the later time points.

Finally, the flowchart of our proposed boundary regression model can be summarized as FIGURE 7 below. Given a testing image, the learned classification-guided regression forests are iteratively employed to generate the displacement estimation. In each iteration, the context features from the previous iteration, and the appearance feature from the testing image are inputted into the trained classification-guided regression forest to generate a new classification map and a new displacement map. Note that, the context features here include the spatial context features (from the classification

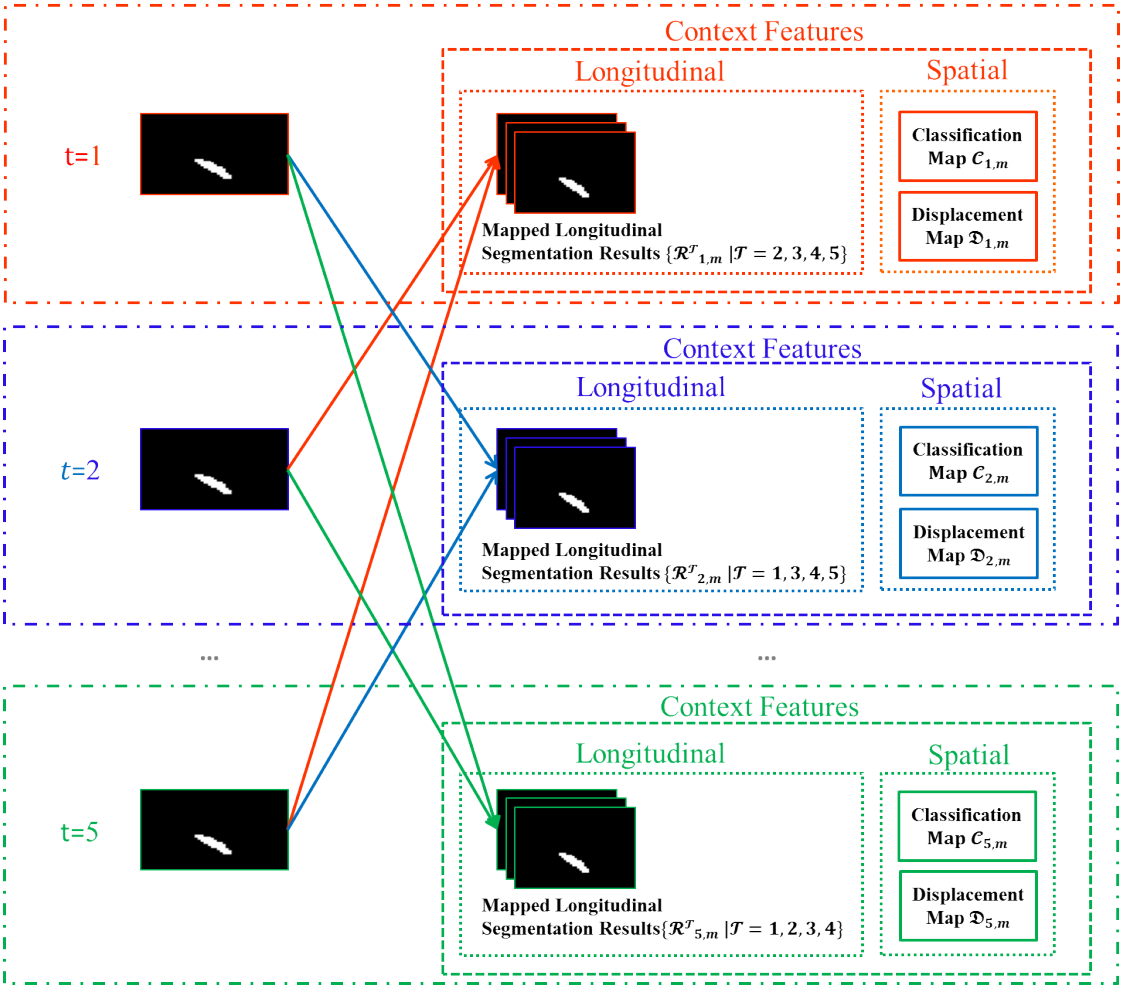



**FIGURE 6.** Longitudinal and spatial context features of a subject at iteration *m*.

map and the displacement map of the same time point), and longitudinal context features (from the segmentation results of other time points). By repeating this step, a sequence of displacement maps can be iteratively generated.

Meanwhile, a sequence of the gradually refined boundary maps can also be acquired with the estimated displacement vectors and the aforementioned weighted voting strategy. The refined boundary maps are illustrated in FIGURE 8. In the experiment, we observe that the boundary map in our method converges by 5 iterations.

### C. DEFORMABLE SEGMENTATION DRIVEN BY BOUNDARY MAP

The obtained boundary map highlights the target hippocampal boundary with high votes, which can be used as an external force to guide the active shape model (ASM) [43]. Besides, ASM can integrate the shape priors into the segmentation procedure to alleviate the ambiguous boundary. Therefore, we apply ASM on the boundary map for accurate hippocampal segmentation.

#### 1) DEFORMABLE MODEL INITIALIZATION WITH POINT REGRESSION

To segment the hippocampi in infant brain MR images with ASM, proper initialization is a prerequisite. Inspired by the point regression method [27], [31], [44], we automatically detect six utmost points for each hippocampus (namely superior, inferior, left, right, anterior, and posterior points of the hippocampus) as landmarks.

Specifically, in the training stage, we randomly draw samples $\{\tilde{\boldsymbol{p}}_i\}$ $(i = 1, 2, \ldots, \tilde{K})$ around a specific annotated landmark. Subsequently we extract the extended Haar-like features from multi-modality images and the obtained boundary map as Eq.(1)), and then compute the displacement vector $\tilde{\boldsymbol{d}}_i$ from each sample $\tilde{\boldsymbol{p}}_i$ to the specific landmark as the regression target. With these Haar-like features and the corresponding 3D displacement vectors, a regression forest $\tilde{R}_{\mathcal{L}}(\mathcal{L} = 1, 2, \ldots, 12)$ can be trained [38] for the $\mathcal{L}$-th annotated landmark.

Given a testing image, we employ the trained regression forest $\tilde{R}_{\mathcal{L}}$ to predict the displacement vector from each image voxel to the $\mathcal{L}$-th landmark, and then cast a vote at the

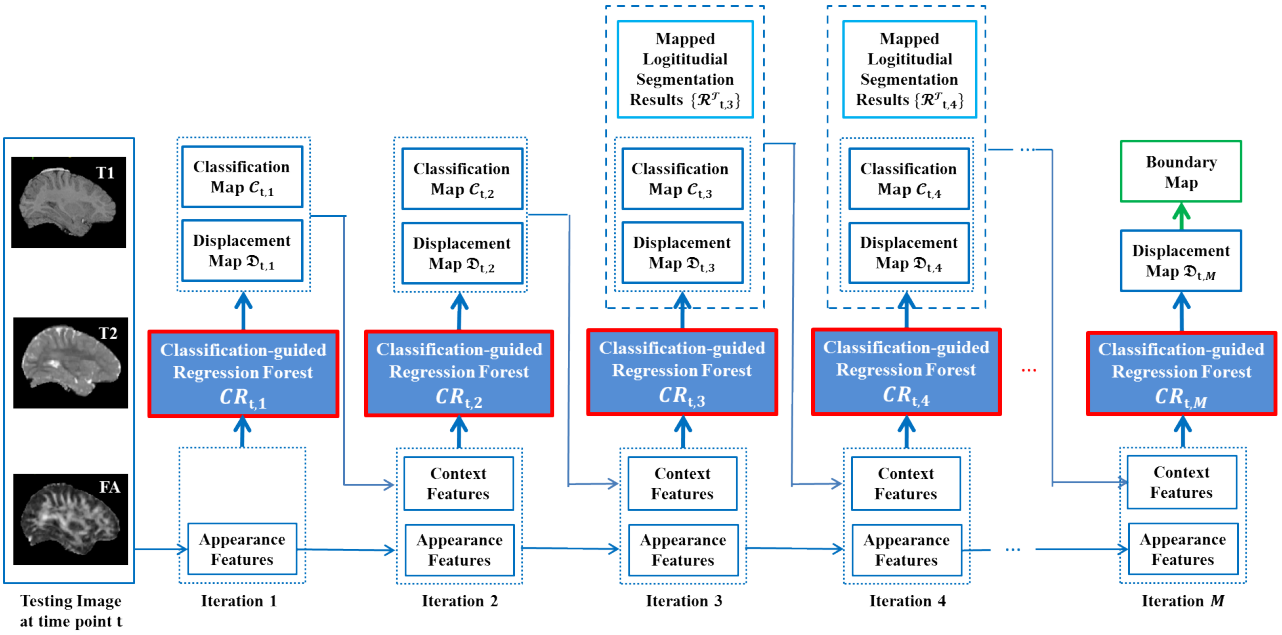



**FIGURE 7.** Flowchart of our proposed method for the time point $t$.

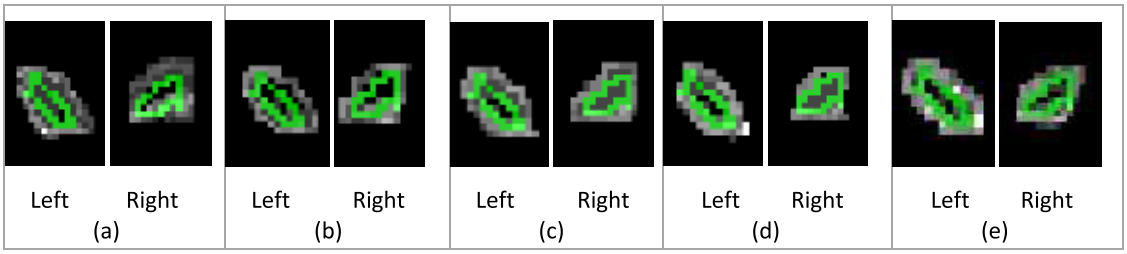



**FIGURE 8.** The boundary map of a certain subject evolves with the use of spatial and the longitudinal context information. Green curves represent the manually-delineated hippocampal boundaries. With the increase of iteration, the accuracy of the boundary map can be gradually improved. To highlight the voted hippocampal boundaries, we only show the left and right hippocampal regions of the obtained boundary map at each iteration. (a) Iteration 1. (b) Iteration 2. (c) Iteration 3. (d) Iteration 4. (e) Iteration 5.

potential landmark position (c.f. FIGURE 9(a)). The position with the maximal vote is considered as the $\mathcal{L}$-th landmark. It is worth noting that, different from the previous boundary regression, the point regression for landmark detection is to predict a specific landmark point, instead of the nearest boundary point (c.f. FIGURE 9).

With all detected landmarks, we compute the affine transformation between the corresponding points of the mean shape $\bar{S}$ (described in Section II.C.2 below) and the detected landmarks. Using the affine transformation, the mean shape is warped onto the testing image as the deformable model initialization.

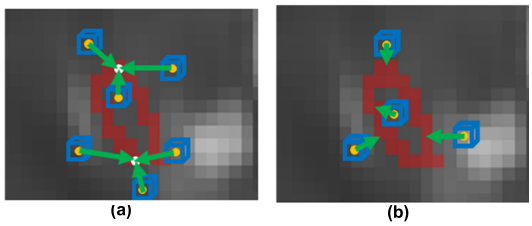



**FIGURE 9.** Illustration of point regression for landmark detection and boundary regression for boundary detection. Red contour indicates the hippocampal boundary, and white points indicate two typical target boundary landmark points. Each blue cube represents the image patch centered at a voxel sample (i.e., the yellow point), and each green arrow represents the displacement vector from the patch center to a target boundary point. (a) Point regression for landmark detection. (b) Boundary regression for boundary detection.

### 2) SHAPE MODEL

Due to the ambiguous hippocampal boundary, we further incorporate the shape priors (shape model) into the deformable model for accurate hippocampal segmentation. To establish the shape model for the infant hippocampi in brain MR images, we first build a shape model by principal component analysis (PCA) based on hippocampal surfaces from the training subjects. Specifically, we extract the target hippocampal surfaces from manually segmented images in the training set by the marching cube algorithm [45]. One typical surface is selected from the dataset as the reference mesh and further smoothed. Here, the number of the vertices

on the reference mesh is 420 for both the left and the right hippocampi. Then, the surface registration method [46] is used to warp the reference mesh onto other hippocampal surface meshes for building surface correspondences. All these surfaces are further affine aligned into a common space, and then used to build the PCA shape space with 95% variation. A shape $\boldsymbol{S}$ in the PCA shape space can be represented as $\boldsymbol{S} = \bar{\boldsymbol{S}} + A\boldsymbol{b}$. Here, $\bar{\boldsymbol{S}}$ is the mean shape, $A$ is a matrix of principal components, and $\boldsymbol{b}$ is the parametric perturbation to generate the shape $\boldsymbol{S}$.

#### 3) DEFORMABLE SEGMENTATION

Given a testing stage, the initial shape is inferred according to the detected landmarks as described in Section II.C.1. Each point of the hippocampal shape model can be independently deformed on the boundary map towards the real target boundary, by searching for the position with the maximal boundary votes along the respective surface normal direction. Then, the intermediate deformed shape $\widehat{\boldsymbol{S}}$ is refined by the learned PCA shape model as follows, $\widehat{\boldsymbol{b}} = A^T(\Lambda^{-1}\widehat{\boldsymbol{S}} - \bar{\boldsymbol{S}})$, $\widehat{\boldsymbol{b}} \in [-3\lambda_i, 3\lambda_i]$. Here $\lambda_i$ is the eigenvalue regarding the $i$-th principal component of the shape. $\Lambda^{-1}$ is a transformation to project the deformed shape into the PCA shape space. During deformable segmentation, the model deformation and the shape refinement alternate, under the guidance of both the boundary map and the PCA shape space. With the iteration increasing, the shape constraint gradually decreases. The deformable segmentation finally converges when the intermediate deformed shape *no longer* changes.

## III. EXPERIMENT RESULTS

In the experiments, we validate our method on the T1, T2 and FA images from 10 infant subjects at each time point of 2 weeks, 3 months, 6 months, 9 months and 12 months of age. The T1- and T2-weighted images were both scanned with a Siemens head-only 3T scanner. The T1-weighted MR images were obtained at the resolution of $1 \times 1 \times 1\text{mm}^3$ with 144 sagittal slices, while the T2-weighted MR images were obtained at the resolution of $1.25 \times 1.25 \times 1.95\text{mm}^3$ with 64 axis slices. Diffusion weighted images were obtained at the resolution of $2 \times 2 \times 2\text{mm}^3$ with 60 axial slices. Additionally, for each subject, the T2-weighted MR images were aligned to the corresponding T1-weighted MR images; the FA images were first aligned to the corresponding T2 images and then transferred to the corresponding T1 image spaces. Since the T1, T2 and FA images from each subject share the same brain anatomical space, they can be accurately aligned with rigid registration by FLIRT [42] using correlation ratio as the similarity measure. As the image preprocessing step, all images in the dataset were isotropically resampled to $1 \times 1 \times 1\text{mm}^3$. Standard preprocessing operations [47] were then applied to each subject, including skull stripping, bias-field correction, histogram matching, and also removal of the cerebellum and brain stem. Besides, for each subject, the manual hippocampi delineation of an experienced radiologist from UNC-hospital is used as the ground truth for quantitative evaluation.

To quantitatively evaluate our proposed method, we employ Dice similarity coefficient (DSC) [48] and average surface distance (ASD) as the metric.

DSC is an overlap measure between the segmented organ and the manual ground truth:

$$DSC = \frac{2TP}{2TP + FP + FN} \times 100\% \tag{6}$$

where $TP$ is the number of the correctly labeled organ voxels, $FP$ is the number of falsely labeled organ voxels, and $FN$ is the number of falsely labeled background voxels.

Average surface distance is the average distance between the surface of segmented organ (SEG) and that of the manual ground truth (GT):

$$ASD = \frac{1}{2}\left(\frac{\sum_{u\in \text{GT}} d(u, \text{SEG})}{|\text{GT}|} + \frac{\sum_{v\in \text{SEG}} d(v, \text{GT})}{|\text{SEG}|}\right) \tag{7}$$

where $d(u, \text{SEG})$ is the minimal distance of voxel $u$ on the ground truth surface to the voxels on the segmented organ surface, $d(v, \text{GT})$ is the minimal distance of voxel $v$ on the segmented organ surface to the voxels on the ground truth surface, and $|\cdot|$ is the cardinality of a set.

### A. PARAMETER SETTING

We use leave-one-out cross-validation to evaluate our method. The parameter setting is as follows: the number of the trees in each regression forest is 10; the maximum depth of each tree is 15; the size of the feature pool for splitting each node is 1000; the minimum number of the samples in a leaf node is 5; the patch size for extracting the Haar-like features is $8\times8 \times 8$mm; the number of the samples (i.e., $K = K' = \tilde{K}$) drawn from each training image is 10000; the number of iterations for auto-context model is 5; the beginning iteration integrated with longitudinal information is 3; and the number of the iterations for deformable model is 20.

### B. BOUNDARY REGRESSION VERSUS CLASSIFICATION-GUIDED BOUNDARY REGRESSION

To compare the conventional boundary regression ([29], [30]) with the proposed classification-guided boundary regression, we apply the ASM model on the boundary maps generated from the boundary regression and the classification-guided boundary regression, respectively. For a fair comparison, both methods are based on the image appearance features with one iteration. During the ASM-based deformable segmentation, the deformable model searches on the boundary map for the maximal vote along the normal direction of the surface. Note that both deformable models adopt the same initial shape and deformation strategy.

To quantitatively compare the boundary regression with the classification-guided boundary regression in hippocampal segmentation, the DSC metric between automatic segmentation and the ground truth is computed. It can be observed from

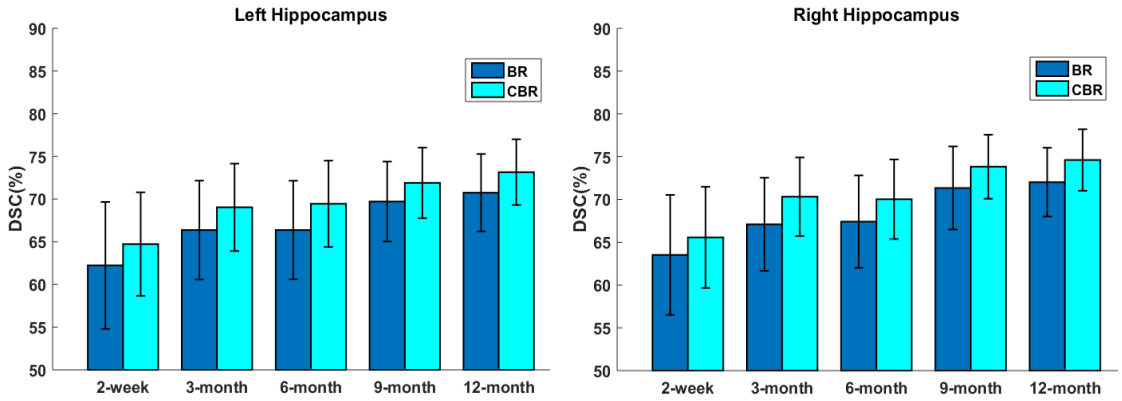



**FIGURE 10.** Comparison of hippocampal segmentation results between the boundary regression (BR) and the classification-guided boundary regression (CBR).

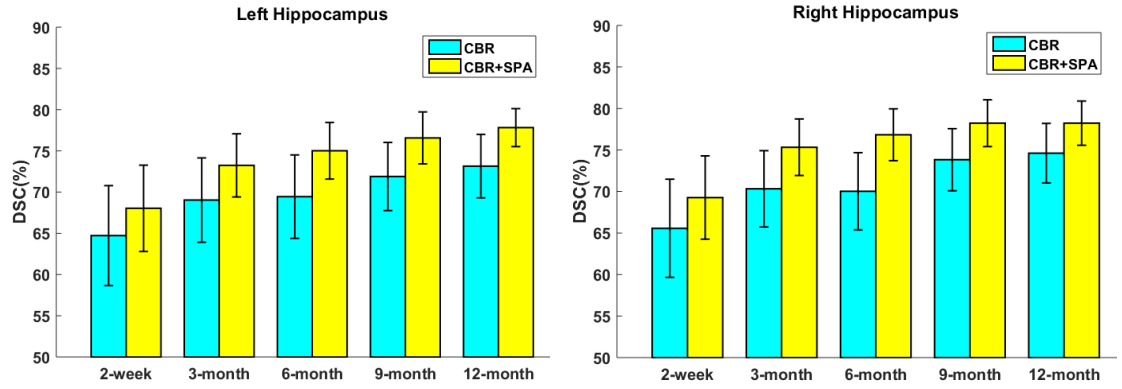



**FIGURE 11.** Comparison of hippocampal segmentation results between the classification-guided boundary regression (CBR) and the classification-guided boundary regression with spatial context information (CBR+SPA).

FIGURE 10 that, on both hippocampi, the classification-guided boundary regression achieves better performance than the boundary regression for the hippocampal segmentation. It proves that the voxel-wise classification forest is effective to achieve a better hippocampal segmentation. Besides, among different time points, the segmentation accuracy of the first time point (the 2-week time point) is lowest for both boundary regression methods. The reason is that the infant hippocampi has too small size, low image contrast and ambiguous boundary at early stage of life.

### C. IMPORTANCE OF SPATIAL CONTEXT INFORMATION

To study the role of the spatial context information in improving the boundary map, we suppress the spatial context information and evaluate the difference of final segmentation performance. In this paper, we employ the spatial context information under the 5-iteration auto-context model (c.f. FIGURE 5). The first classification-guided regression forest is trained without spatial context information, and the others are trained with the spatial context information extracted from the outputs of the previous classification-guided regression forests. Hence, under the auto-context model, we actually compare the classification-guided boundary regression with 1 iteration and the classification-guided boundary regression with 5 iterations. Based on the generated boundary maps by these two classification-guided boundary regression models, we apply two deformable models with the same setting to segment the hippocampi.

To analyze the importance of the spatial context in our method, we quantitatively compare the segmentation result with the ground truth of the hippocampi. From FIGURE 11, we can observe that, the classification-guided boundary regression with the spatial context information outperforms that without spatial context information on both hippocampi. This means that the spatial context information is effective to improve hippocampal boundary map.

### D. CONTRIBUTION OF LONGITUDINAL CONTEXT INFORMATION

To validate the contribution of the longitudinal context information, we compare our method using only spatial context information, with that using both spatial and longitudinal context information under 5-iteration auto-context model. The ASM model is also applied on the boundary map to segment the hippocampi with the same setting.

FIGURE 12 shows the segmentation performance (in DSC metric) of our method with and without longitudinal context information. We can find that, the DSC of our method with the longitudinal context information is higher than that without longitudinal context information, which indicates

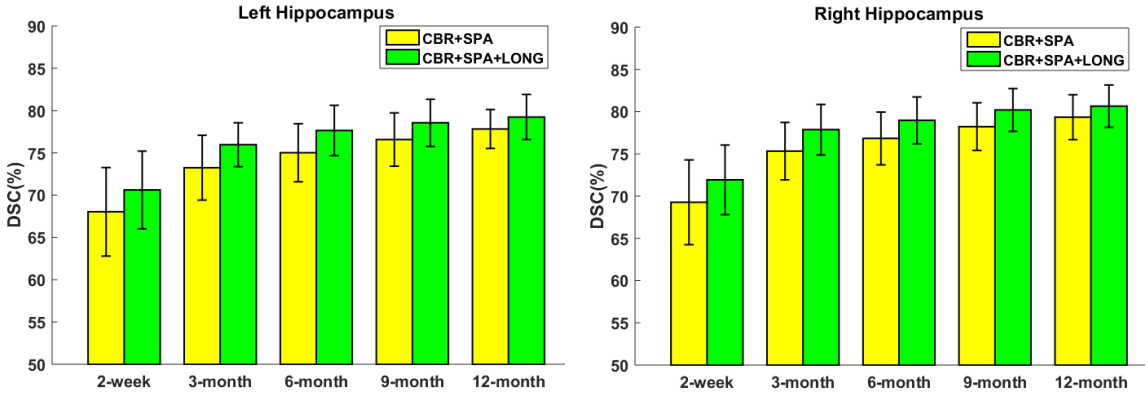



**FIGURE 12.** Comparison of hippocampal segmentation results between our method (CBR+SPA+LONG) and our method without longitudinal context information (CBR+SPA).

**TABLE 1.** Comparison between our method and the state-of-the-art methods in DSC (%).

| Method | 2-week | 3-month | 6-month | 9-month | 12-month |
|---|---|---|---|---|---|
| Guo *et al.* [19] | 65.20±10.10 | 70.30±5.30 | 72.60±7.00 | *NA* | *NA* |
| Zhang *et al.* [49] | 67.40±14.18 | 73.31±7.64 | 74.01±8.12 | 73.50±5.84 | *NA* |
| **Our method** | **71.12±4.53** | **76.88±3.10** | **78.31±2.81** | **78.95±2.75** | **79.87±2.60** |

**TABLE 2.** Comparison between our method and the state-of-the-art methods in ASD (mm).

| Method | 2-week | 3-month | 6-month | 9-month | 12-month |
|---|---|---|---|---|---|
| Guo *et al.* [19] | 0.88±0.33 | 0.76±0.17 | 0.72±0.22 | *NA* | *NA* |
| **Our method** | **0.74±0.25** | **0.69±0.16** | **0.66±0.15** | **0.65±0.14** | **0.63±0.12** |

the contribution of the longitudinal context information in infant hippocampal segmentation. Besides, for longitudinal context information, since the images from later time points have better image contrast, the tentative segmentation results from later time points are better than the early time points. Therefore, the improvement of hippocampal segmentation performance at early time points tends to more than that at the last time point.

### E. COMPARISON WITH STATE-OF-THE-ART METHODS

To justify the effectiveness of our method in infant hippocampal segmentation, we also compare our proposed method with the state-of-the-art methods. The performances of all methods on both hippocampi are summarized in TABLE 1. Note that, these methods are evaluated on the same dataset.

As can been seen from TABLE 1, our method achieves competitive performance on infant hippocampal segmentation to other segmentation methods under comparison. Specifically, on all reported time points, our method achieves better performance than Guo *et al*. [19] and Zhang *et al*. [49].

## IV. DISCUSSION

We have proposed a boundary detection method based on the classification-guided boundary regression to segment the infant hippocampi. Different from the traditional boundary regression, we integrate the implicit shape clue from classification results into the traditional boundary regression for better segmentation. We also employed the spatial and longitudinal context information to achieve a coherent segmentation result. Due to the low image contrast near the hippocampus head and tail, the segmented organ boundary is likely to leak. That is the main reason why we adopted the shape-constrained deformable model after the boundary regression.

Currently, the infant hippocampus dataset consists of only 10 subjects at 5 time points, which is not suitable for deep learning methods to obtain discriminative features. Next, we will collect more subjects from hospitals to further improve the segmentation results of our method, and try popular deep learning methods for better hippocampus segmentation. Besides, the average segmentation time of our method is 591 seconds (with Intel Core i7-4770K 3.50 GHz, 16 GB RAM) for a testing image under the auto-context model, which is similar to that of the one-layer random forest of [49].

## V. CONCLUSION

In this paper, we present a classification-guided boundary regression method with spatial and longitudinal context

information for accurate hippocampal segmentation in MR infant brain images. Specifically, we combine classification forest and regression forest to learn the classification-guided boundary regression method for predicting the entire hippocampal boundary. Then, we integrate the spatial and longitudinal information into the classification-guided regression forest to iteratively refine the voted boundary map. Finally, a deformable model is employed on the boundary map to achieve the final hippocampal segmentation. Validated on 10 subjects from 2-week-old to 1-year-old, our classification-guided boundary regression method achieves better segmentation accuracy than the state-of-the-art methods under comparison.

## REFERENCES


[1] V. Dill, A. R. Franco, and M. S. Pinho, "Automated methods for hippocampus segmentation: The evolution and a review of the state of the art," *Neuroinformatics*, vol. 13, no. 2, pp. 133–150, Apr. 2015.

[2] M. Kim *et al.*, "Automatic hippocampus segmentation of 7.0 Tesla MR images by combining multiple atlases and auto-context models," *NeuroImage*, vol. 83, pp. 335–345, Dec. 2013.

[3] Y. Hao *et al.*, "Local label learning (LLL) for subcortical structure segmentation: Application to hippocampus segmentation," *Hum. Brain Mapping*, vol. 35, no. 6, pp. 2674–2697, Oct. 2014.

[4] T. den Heijer *et al.*, "Structural and diffusion MRI measures of the hippocampus and memory performance," *NeuroImage*, vol. 63, no. 4, pp. 1782–1789, Dec. 2012.

[5] W. B. Scoville and B. Milner, "Loss of recent memory after bilateral hippocampal lesions," *J. Neurol., Neurosurg., Psychiatry*, vol. 20, no. 1, pp. 11–21, Feb. 1957.

[6] J. T. Wixted and L. R. Squire, "The medial temporal lobe and the attributes of memory," *Trends Cogn. sci.*, vol. 15, no. 5, pp. 210–217, May 2011.

[7] K. Van Leemput *et al.*, "Automated segmentation of hippocampal subfields from ultra-high resolution in vivo MRI," *Hippocampus*, vol. 19, no. 6, pp. 549–557, Jun. 2009.

[8] A. Convit *et al.*, "Specific hippocampal volume reductions in individuals at risk for Alzheimer's disease," *Neurobiol. Aging*, vol. 18, no. 2, pp. 131–138, Mar./Apr. 1997.

[9] D. K. Thompson *et al.*, "Optimizing hippocampal segmentation in infants utilizing MRI post-acquisition processing," *Neuroinformatics*, vol. 10, no. 2, pp. 173–180, Apr. 2012.

[10] T. Bartsch, *The Clinical Neurobiology of the Hippocampus: An Integrative View*. London, U.K.: Oxford Univ. Press, 2012.

[11] B. Fischl *et al.*, "Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain," *Neuron*, vol. 33, no. 3, pp. 341–355, Jan. 2002.

[12] J. Barnes *et al.*, "Automatic calculation of hippocampal atrophy rates using a hippocampal template and the boundary shift integral," *Neurobiol. Aging*, vol. 28, no. 11, pp. 1657–1663, Nov. 2007.

[13] O. T. Carmichael *et al.*, "Atlas-based hippocampus segmentation in Alzheimer's disease and mild cognitive impairment," *NeuroImage*, vol. 27, no. 4, pp. 979–990, Oct. 2005.

[14] K. Kwak *et al.*, "Fully-automated approach to hippocampus segmentation using a graph-cuts algorithm combined with atlas-based segmentation and morphological opening," *Magn. Reson. Imag.*, vol. 31, no. 7, pp. 1190–1196, Sep. 2013.

[15] P. Dong, Y. Guo, D. Shen, and G. Wu, "Multi-atlas and multi-modal hippocampus segmentation for infant MR brain images by propagating anatomical labels on hypergraph," presented at the Int. Workshop Patch-based Techn. Med. Imag., Munich, Germany, Oct. 2015.

[16] G. Wu, M. Kim, G. Sanroma, Q. Wang, B. C. Munsell, and D. Shen, "Hierarchical multi-atlas label fusion with multi-scale feature representation and label-specific patch partition," *NeuroImage*, vol. 106, pp. 34–46, Feb. 2015.

[17] H. Wang, J. W. Suh, S. R. Das, J. B. Pluta, C. Craige, and P. A. Yushkevich, "Multi-atlas segmentation with joint label fusion," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 3, pp. 611–623, Mar. 2013.

[18] H. Zhu, H. Cheng, X. Yang, Y. Fan, and Alzheimer's Disease Neuroimaging Initiative, "Metric learning for multi-atlas based segmentation of hippocampus," *Neuroinformatics*, vol. 15, no. 1, pp. 41–50, Jan. 2017.

[19] Y. Guo, G. Wu, P.-T. Yap, V. Jewells, W. Lin, and D. Shen, "Segmentation of infant hippocampus using common feature representations learned for multimodal longitudinal data," presented at the Med. Image Comput. Comput. Assist. Interv., Munich, Germany, Oct. 2015.

[20] D. Zarpalas, P. Gkontra, P. Daras, and N. Maglaveras, "Hippocampus segmentation through gradient based reliability maps for local blending of ACM energy terms," presented at the IEEE 10th Int. Symp. Biomed. Imag., San Francisco, CA, USA, Apr. 2013.

[21] M. Hajiesmaeili, J. Dehmeshki, B. B. Nakhjavanlo, and T. Ellis, "Initialisation of 3D level set for hippocampus segmentation from volumetric brain MR images," presented at the 6th Int. Conf. Digit. Image Process., Athens, Greece, Apr. 2014.

[22] S. Gao, H. Zhou, Y. Lin, and Y. Zhang, "Hippocampus segmentation via active contour model," *J. Inf. Comput. Sci.*, vol. 9, no. 4, pp. 262–266, Oct. 2014.

[23] Z. Luo, X. Zhuang, R. Zhang, J. Wang, C. Yue, and X. Huang, "Automated 3D segmentation of hippocampus based on active appearance model of brain MR images for the early diagnosis of Alzheimer's disease," *Minerva Med.*, vol. 105, no. 2, pp. 157–165, Apr. 2014.

[24] S. Hu, P. Coupé, J. C. Pruessner, and D. L. Collins, "Nonlocal regularization for active appearance model: Application to medial temporal lobe segmentation," *Hum. Brain Mapping*, vol. 35, no. 2, pp. 377–395, Feb. 2014.

[25] S. Ettaïeb, K. Hamrouni, and S. Ruan, "Statistical models of shape and spatial relation-application to hippocampus segmentation," presented at the Int. Conf. Comput. Vision Theory Appl., Lisbon, Portugal, Jan. 2014.

[26] Y. Shao, Y. Gao, Q. Wang, X. Yang, and D. Shen, "Locally-constrained boundary regression for segmentation of prostate and rectum in the planning CT images," *Med. image Anal.*, vol. 26, no. 1, pp. 345–356, Dec. 2015.

[27] C. Chen, W. Xie, J. Franke, P. A. Grutzner, L. P. Nolte, and G. Zheng, "Automatic X-ray landmark detection and shape segmentation via data-driven joint estimation of image displacements," *Med. image Anal.*, vol. 18, no. 3, pp. 487–499, Apr. 2014.

[28] Z. Wang, X. Zhen, K. Tay, S. Osman, W. Romano, and S. Li, "Regression segmentation for $M^3$ spinal images," *IEEE Trans. Med. Imag.*, vol. 34, no. 8, pp. 1640–1648, Aug. 2015.

[29] Y. Shao, Y. Gao, X. Yang, and D. Shen, "CT prostate deformable segmentation by boundary regression," presented at the Int. MICCAI Workshop Med. Comput. Vis., Cambridge, MA, USA, Sep. 2014, Sep. 2014.

[30] Y. Shao, Y. Guo, Y. Gao, X. Yang, and D. Shen, "Hippocampus segmentation from MR infant brain images via boundary regression," presented at the Int. MICCAI Workshop Med. Comput. Vis., Munich, Germany, Oct. 2015.

[31] T. F. Cootes, M. C. Ionita, C. Lindner, and P. Sauer, "Robust and accurate shape model fitting using random forest regression voting," presented at the 12th Eur. Conf. Comput. Vis. (ECCV), Florence, Italy, Oct. 2012.

[32] A. Criminisi, J. Shotton, D. Robertson, and E. Konukoglu, "Regression forests for efficient anatomy detection and localization in CT studies," presented at the Med. Comput. Vis., Recognit. Techn. Appl. Med. Imag., Beijing, China, Sep. 2011.

[33] G. Fanelli, M. Dantone, J. Gall, A. Fossati, and L. Van Gool, "Random forests for real time 3D face analysis," *Int. J. Comput. Vis.*, vol. 101, no. 3, pp. 437–458, Feb. 2013.

[34] S. K. Zhou, "Shape regression machine and efficient segmentation of left ventricle endocardium from 2D B-mode echocardiogram," *Med. Image Anal.*, vol. 14, no. 4, pp. 563–581, Aug. 2010.

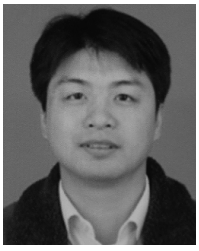
[35] C. Lindner, S. Thiagarajah, J. M. Wilkinson, G. A. Wallis, and T. F. Cootes, "Fully automatic segmentation of the proximal femur using random forest regression voting," *IEEE Trans. Med. Imag.*, vol. 32, no. 8, pp. 1462–1472, Aug. 2013.

[36] C. Chu *et al.*, "Fully automatic cephalometric X-ray landmark detection using random forest regression and sparse shape composition," in *Automatic Cephalometric X-Ray Landmark Detection Challenge*. Beijing, China: ISBI, 2014.

[37] J. Gall, V. Lempitsky, "Class-specific hough forests for object detection," presented at the IEEE Conf. Comput. Vis. Pattern Recognit., Miami Beach, FL, USA, Jun. 2009.

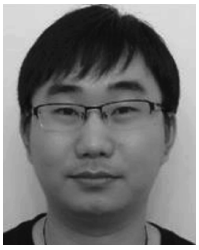
[38] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001.

[39] P. Kontschieder, S. R. Bulo, M. Pelillo, and H. Bischof, "Structured labels in random forests for semantic labelling and object detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 10, pp. 2104–2116, Oct. 2014.
[40] A. Liaw and M. Wiener, "Classification and regression by randomForest," *R News*, vol. 2, no. 3, pp. 18–22, 2002.
[41] Z. Tu and X. Bai, "Auto-context and its application to high-level vision tasks and 3D brain image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 10, pp. 1744–1757, Oct. 2010.
[42] M. Jenkinson, P. Bannister, M. Brady, and S. Smith, "Improved optimization for the robust and accurate linear registration and motion correction of brain images," *NeuroImage*, vol. 17, no. 2, pp. 825–841, Oct. 2002.
[43] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models - their training and application," *Comput. Vis. Image Understand.*, vol. 61, no. 1, pp. 38–59, Jan. 1995.
[44] T. Kohlberger *et al.*, "Automatic multi-organ segmentation using learning-based segmentation and level set optimization," presented at the 14th Int. Conf. Med. Image Comput. Comput. Assist. Intervent., Toronto, ON, Canada, Sep. 2011.
[45] W. E. Lorensen and H. E. Cline, "Marching cubes: A high resolution 3D surface construction algorithm," presented at the 14th Annu. Conf. Comput. Graph. Interact. Techn., Anaheim, CA, USA, Jul. 1987.
[46] A. Myronenko and X. Song, "Point set registration: Coherent point drifts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 12, pp. 2262–2275, Dec. 2010.
[47] L. Wang *et al.*, "LINKS: Learning-based multi-source integration framework for segmentation of infant brain images," *NeuroImage*, vol. 108, pp. 160–172, Mar. 2015.
[48] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, no. 26, pp. 297–302, Jul. 1945.
[49] L. Zhang, Q. Wang, Y. Gao, H. Li, G. Wu, and D. Shen, "Concatenated spatially-localized random forests for hippocampus labeling in adult and infant MR brain images," *Neurocomputing*, vol. 229, pp. 3–12, Mar. 2017.



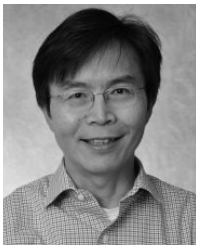
**YEQIN SHAO** received the B.S. degree in computer communications, and the M.S. degree in pattern recognition from the Nanjing University of Science and Technology, Jiangsu, China, in 2002 and 2004, respectively, and the Ph.D. degree in medical image analysis from Shanghai Jiao Tong University, Shanghai, China, in 2015. From 2013 to 2014, he was a Visiting Scholar with the Department of Radiology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. Since 2015, he has been an Associate Professor with the School of Transportation, Nantong University, Jiangsu, China. His research interests include medical image segmentation and machine learning. He is currently a Reviewer of the IEEE TRANSACTION OF MEDICAL IMAGING, *Pattern Recognition*, and MICCAI.

**JAEIL KIM** received the B.S. degree in digital media and computer science from Ajou University, Suwon, South Korea, in 2007, and the M.S. and Ph.D. degrees in computer science from the Korea Advanced Institute of Science and Technology, Daejeon, South Korea, in 2015. From 2014 to 2016, he was a Senior Engineer of the Medical Imaging R&D Group, Samsung Electronics, Suwon. Subsequently, he has been a Postdoctoral Research Associate with the Department of Radiology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. Since 2018, he has been an Assistant Professor with the School of Computer Science and Engineering, Kyungpook National University, Daegu, South Korea. His research interests include medical image analysis, image processing, machine learning, and artificial intelligence.



**YAOZONG GAO** received the B.S. degree in software engineering and the M.S. degree in medical image analysis from Zhejiang University, in 2008 and 2011, respectively, and the Ph.D. degree from the Department of Computer Science, University of North Carolina at Chapel Hill. He was a Computer Vision Researcher with Apple Inc. He is currently directing the Deep Learning Group, United Imaging Intelligence, China. He has published over 90 papers in the international journals and conferences, such as MICCAI, TMI, and MIA. His research interests include machine learning, computer vision, and medical image analysis.

**QIAN WANG** received the B.S. and M.S. degrees in electronic engineering from Shanghai Jiao Tong University, in 2006 and 2009, respectively, and the Ph.D. degree in computer science from The University of North Carolina at Chapel Hill, in 2013. He is currently the Director of the Medical Image Computing Laboratory, Institute of Medical Imaging Technology, School of Biomedical Engineering, Shanghai Jiao Tong University. His research interests include medical image analysis, computer vision, machine learning, artificial intelligence, and translational medical studies.



**WEILI LIN** received the M.S. and Ph.D. degrees in biomedical engineering from Case Western Reserve University, Cleveland, OH, USA. He is currently the Dixie Lee Boney Soo Distinguished Professor of neurological medicine, and the Director of the Biomedical Research Imaging Center, University of North Carolina at Chapel Hill. In addition, he is a Professor of radiology, neurology, biomedical engineering, and the School of Pharmacy. He is the Vice Chair of Basic Research in the Department of Radiology. He has published more than 250 peer reviewed papers. His research interests can be divided into three major areas, including early brain functional and structural development, quantitative brain oxygen metabolism, and hybrid PET/MR.

**DINGGANG SHEN** (F'18) was a tenure-track Assistant Professor with the University of Pennsylvanian, Philadelphia, PA, USA, and a Faculty Member with the Johns Hopkins University, Baltimore, MD, USA. He is currently a Jeffrey Houpt Distinguished Investigator, and a Professor of Radiology, Computer Science, and Biomedical Engineering with the Biomedical Research Imaging Center (BRIC), The University of North Carolina at Chapel Hill. He is currently directing the Image Display, Enhancement, and Analysis Lab, Center for Image Analysis and Informatics, Department of Radiology, and also the Medical Image Analysis Core, BRIC. He has authored over 800 papers in the international journals and conference proceedings. His research interests include medical image analysis, computer vision, and pattern recognition. He is a Fellow of The American Institute for Medical and Biological Engineering, and also a Fellow of The International Association for Pattern Recognition. He serves as an Editorial Board Member for eight international journals. He has also served on the Board of Directors, the Medical Image Computing and Computer Assisted Intervention Society, from 2012 to 2015.