Infant Brain Extraction in T1-Weighted MR Images Using BET and Refinement Using LCDG and MGRF Models

Amir Alansary, et al.

nsi ™

NeuroSpectrum Insights, Inc.

info@neurospectruminsights.com
www.neurospectruminsights.com

Infant Brain Extraction in T1-Weighted MR Images Using BET and Refinement Using LCDG and MGRF Models

Amir Alansary, *Member, IEEE*, Marwa Ismail, Ahmed Soliman, *Member, IEEE*, Fahmi Khalifa, *Member, IEEE*, Matthew Nitzken, *Member, IEEE*, Ahmed Elnakib, *Member, IEEE*, Mahmoud Mostapha, *Member, IEEE*, Austin Black, Katie Stinebruner, Manuel F. Casanova, Jacek M. Zurada, *Life Fellow, IEEE*, and Ayman El-Baz, *Senior Member, IEEE*

Abstract—In this paper, we propose a novel framework for the automated extraction of the brain from T1-weighted MR images. The proposed approach is primarily based on the integration of a stochastic model [a two-level Markov–Gibbs random field (MGRF)] that serves to learn the visual appearance of the brain texture, and a geometric model (the brain isosurfaces) that preserves the brain geometry during the extraction process. The proposed framework consists of three main steps: 1) Following bias correction of the brain, a new three-dimensional (3-D) MGRF having a 26-pairwise interaction model is applied to enhance the homogeneity of MR images and preserve the 3-D edges between different brain tissues. 2) The nonbrain tissue found in the MR images is initially removed using the brain extraction tool (BET), and then the brain is parceled to nested isosurfaces using a fast marching level set method. 3) Finally, a classification step is applied in order to accurately remove the remaining parts of the skull without distorting the brain geometry. The classification of each voxel found on the isosurfaces is made based on the first- and second-order visual appearance features. The first-order visual appearance is estimated using a linear combination of discrete Gaussians (LCDG) to model the intensity distribution of the brain signals. The second-order visual appearance is constructed using an MGRF model with analytically estimated parameters. The fusion of the LCDG and MGRF, along with their analytical estimation, allows the approach to be fast and accurate for use in clinical applications. The proposed approach was tested on *in vivo* data using 300 infant 3-D MR brain scans, which were qualitatively validated by an MR expert. In addition, it was quantitatively validated using 30 datasets based on three metrics: the Dice coefficient, the 95% modified Hausdorff distance, and absolute brain volume difference. Results showed the capability of the proposed approach, outperforming four widely used BETs: BET, BET2, brain surface extractor, and infant brain extraction and analysis toolbox. Experiments conducted also proved that the proposed framework can be generalized to adult brain extraction as well.

Index Terms—BET, infant brain, isosurfaces, LCDG, MGRF, MRI, skull stripping.

I. INTRODUCTION

BRAIN extraction is the process of removing all the outer tissues (e.g., eyes, dura, scalp, and skull) around the brain, which consists of the gray matter and white matter (WM), while the inclusion of cerebrospinal fluid in the brain depends on the application. Brain extraction is a primary step in neuroimaging analysis as well as a preprocessing step for many brain analysis algorithms like intensity normalization, registration, classification, and segmentation. Accordingly, accurate brain extraction is crucial for these algorithms to work properly. For instance, in cortical thickness estimation, inaccurate skull stripping (e.g., failing to remove the dura or missing brain parts) can result in an overestimation or underestimation of the cortical thickness [1]. Extraction of brains from adult scans is a difficult process, yet is more challenging in the case of infants due to many factors including the difference in size of the corresponding anatomy, and the variability in the properties of the signal acquired. This explains the fact that most of the work in the literature is devoted to brain extraction from adult scans, and that fewer techniques are focused on infant brain extraction.

Many brain extraction approaches have been developed to extract the brain from T1-weighted MR brain images. These methods use different techniques, such as, region growing, deformable models, atlas-based and label fusion, and hybrid algorithms. Park and Lee [2] automatically identified two seed points of the brain and nonbrain regions using morphological operations. Then, two-dimensional (2-D) region growing was employed based on the knowledge of brain anatomy. This was extended by Roura *et al.* [3], where a multispectral adaptive region growing algorithm was proposed. This enabled the use of the approach proposed in [2] to axial views and provided more reliable results for different patient populations and different MRI scanners (1.5 and 3 T). Smith [4] developed an

automated deformable model-based method, which is widely known as the brain extraction tool (BET). In their approach, the deformable contour is guided by a set of locally adaptive forces, which include morphological and image-based terms in addition to a surface smoothness constraint. Liu *et al.* [5] presented another deformable model-based brain extraction method using a set of Wendland's radial basis functions [6]. Their deformable model is directed by an internal force to consider the smoothness constraint and an external force to impose the intensity contrast across the boundaries. Finally, the brain contours are constructed by integrating the 2-D coronal and sagittal slices that were separately computed to obtain a complete 3-D brain volume. Also, Zhuang *et al.* [7] used a level-set-based deformable model that combines two forces: the mean curvature of the curve and the intensity characteristics of the cortex in MR images. Baillard *et al.* [8] developed a deformable model-based approach to find the brain surface. As an alternative solution for initializing the first contour manually, an atlas-based technique is used to make the brain extraction process entirely automatic. After atlas registration and initial segmentation, the brain is finally segmented based on level sets with adaptive parameters depending on the input data. Wang *et al.* [9] used an atlas-based approach for skull stripping, along with a deformable-surface-based approach guided by local intensity information and prior information that is learned from a set of real brain images. Sadananthan *et al.* [10] presented a skull-stripping approach using graph cuts, which consists of two steps. An initial brain mask is generated using intensity thresholding as a first step. Then, a graph-theoretic image segmentation method is applied to position cuts which remove narrow connections. Zhang *et al.* [11] developed a method for brain extraction by estimating image intensity parameters to construct a binary image of the head. Then, an initial contour is estimated. The final brain is extracted using an improved geometric active contour model which extends the solution of the boundary leakage problem to make it insusceptible to the inhomogeneity of intensity. Somasundaram and Kalavathi [12] developed a contour-based method to segment the brain from T1-, T2-, and proton density-weighted MRI of human head scans in two phases. The first phase involves the extraction of brain regions in the middle slice, with a landmark circle drawn at the center of the extracted brain region. The second phase extracts brain regions in the rest of the slices with reference to that landmark circle. Leung *et al.* [13] presented a brain extraction technique using a template library. Multiple best-matched atlases are selected by comparing the target image to all the atlases in that library. In order to segment the target image optimally after image registration, label fusion techniques do the task of combining the labels from all atlases incorporated. Iglesias *et al.* [14] developed a learning-based brain extraction algorithm that has two models: a discriminative model based on a random forest classifier trained to detect the brain boundary, and a generative model based on finding the contour with highest likelihood according to the discriminative model, which is refined later using graph cuts. Segonne *et al.* [15] presented a hybrid approach that combines watershed algorithms and deformable surface models. The watershed, with a preflooding height, construct an initial estimate of the brain volume us-
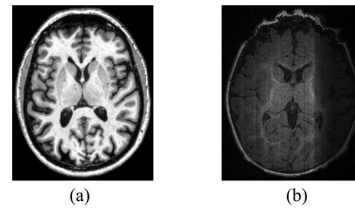

Fig. 1.   T1-weighted MRI axial views for (a) adult and (b) infant brains.

ing a single WM voxel as a global minimum. Then, a surface deformation process is used to correct the initial segmentation inaccuracies. A statistical atlas is used finally to potentially correct the segmentation. Beare *et al.* [16] introduced marker-based watershed scalper for brain extraction in T1-weighted MR images that is built using filtering and segmentation components from the insight toolkit framework. Rex *et al.* [17] developed a metaalgorithm that uses four freely available brain extraction algorithms: brain surface extractor (BSE) [18], BET [4], 3dIntracranial [19], and MRI watershed from FreeSurfer [20]. For extracting the brain, an atlas is used to define which extraction algorithm or combination of extractors works best defining the brain in each anatomic region.

As mentioned earlier, there has not been much work in the literature addressing infant brain extraction. To the best of our knowledge, we are briefly presenting all the existing techniques. Chiverton *et al.* [21], for example, used parameter estimation to fit a Gaussian model to a predefined histogram. The extraction process then employed 3-D morphological operators. Another approach was proposed by Kobashi *et al.* [22], where connected triangles were used in order to construct an initial surface model, to be deformed by moving the vertices. Fuzzy rules were used to define the positions of the triangles. Pport *et al.* [23] presented an approach called hybrid skull stripping. A binary mask was generated using morphological operators to define the brain tissue outer boundary, which was later followed by region growing and edge detection. In order to improve the accuracy, a unique threshold value was identified for each slice in the volume using *k*-means clustering. Shi *et al.* [24] applied multiple brain extractions using BET and BSE on test subjects, where the parameters of each technique are learned from the training data. Also, the representative subjects are selected as exemplars and used to guide brain extraction of new subjects in different age groups. A level-set-based fusion method was further used to combine the multiple brain extractions in order to get the final result.

In summary, different brain extraction approaches have been developed; however, they have their own drawbacks. Some of them give better results when removing nonbrain tissue while losing some brain parts, and others give better results when extracting the whole brain while keeping some nonbrain tissue parts [25], [26]. Atlas-based approaches are very time consuming and their performance heavily depends on the registration accuracy between the atlas and the test subject, in addition to the difficulty of constructing an infant brain atlas [27]. Moreover,
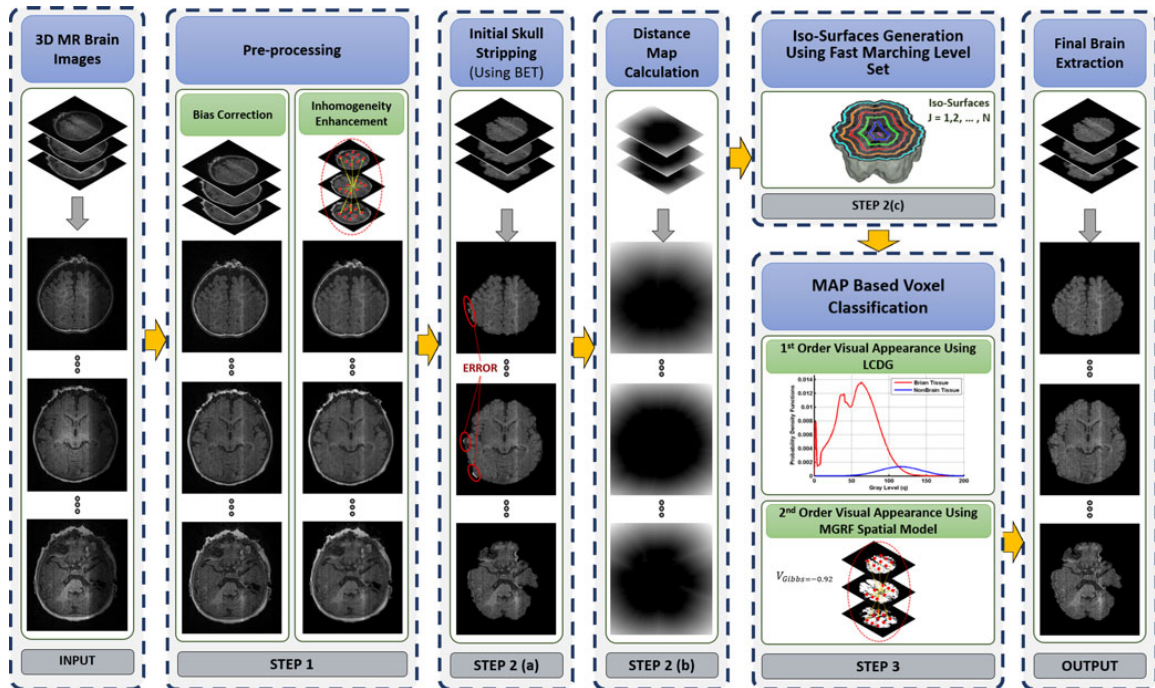
Fig. 2.   Basic steps of the proposed framework for brain extraction from infant MR brain data.

the majority of the existing methods are developed for adult MR brain images and fail to accurately extract the brain from MR infant images due to the reduced contrast and higher noise [28]. Infant brain MRI extraction meets with challenges stemming from image noise, inhomogeneities, artifacts, and discontinuities of boundaries arising from similar visual appearance of adjacent brain structures (see Fig. 1). Furthermore, accurate infant brain extraction contributes much to the analysis, treatment, and early diagnosis of brain injury and disorders resulting from the infant prematurity.

II. METHODS

We present a hybrid framework, as shown in Fig. 2, that possesses the ability to accurately extract brain tissue from infant MR brain images. The proposed framework integrates both stochastic and geometric approaches and consists of three basic steps: 1) bias correction and inhomogeneity enhancement, 2) initial skull stripping and isosurfaces generation, and 3) final brain extraction using the visual appearance features of the MR brain images. The proposed approach avoids many of the shortcomings of the methods presented in the literature. One of the main advantages of it is that it does not require shape priors. Moreover, it is not atlas-based, and thus, its accuracy is not affected by any related processes, such as registration. Also, the integration of geometric features guarantees retaining all brain parts. Inaccurate skull stripping can drastically affect subsequent processes, such as shape analysis for diagnosis of autism [29].

The proposed approach aims at handling all sources of such inaccuracies. Details of the proposed approach are outlined in the following sections.

A. Bias Correction and Inhomogeneity Enhancement

Illumination nonuniformity of infant brain MRIs, which is known as bias field, limits the accuracy of the existing brain extraction approaches. Therefore, to accurately extract the brain, it is important to account for the low-frequency intensity nonuniformity or inhomogeneity. In this paper, we use a 3-D generalized Gauss–Markov random field (GGMRF) model [30] that is applied after bias correction of the brain using the nonparametric approach proposed in [31]. This step reduces noise effects and removes (smooth) inconsistencies of the MRI data by accounting for the 3-D spatially homogeneous pairwise interactions between the gray levels of the MRI data. Namely, the gray-level values $q \in \mathbf{Q} = \{0, \ldots, Q-1\}$ are considered as samples from a 3-D GGMRF model [30] represented for the 26-neighborhood of a voxel. The maximum *a posteriori* estimates [30] and voxelwise stochastic relaxation (iterative conditional mode [32]) of q values of each brain MR scan are employed as follows:

$$\widehat{q}_s = \arg \min_{\widetilde{q}_s} \left[|q_s - \widetilde{q}_s|^{\alpha} + \rho^{\alpha} \lambda^{\beta} \sum_{r \in \nu_s} \eta_{s,r} \left| \widetilde{q}_s - q_r \right|^{\beta} \right] \quad (1)$$

where q_s and $\widetilde{q}_s$ are the original gray-level values and their expected estimates, respectively, at the observed 3-D location,
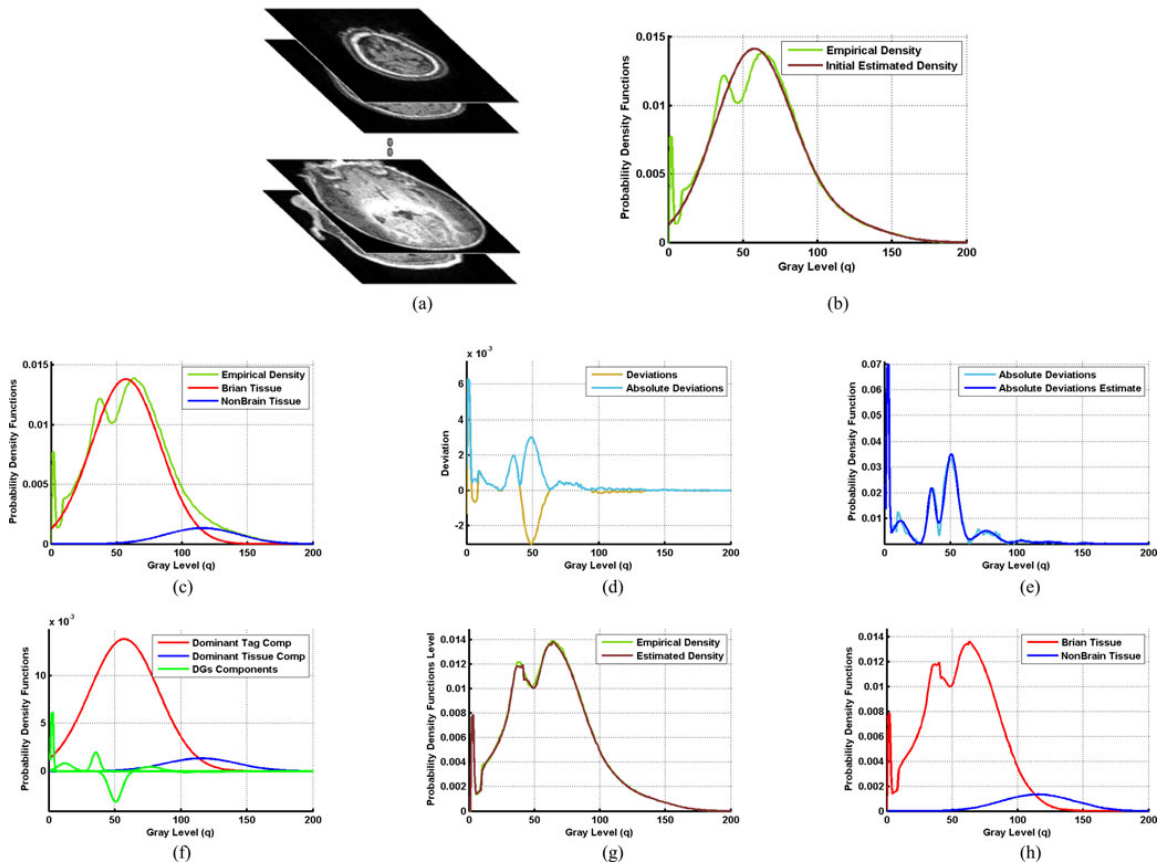
Fig. 3. (a) Typical MR infant brain images, and (b) estimated density (c) using only two dominant Gaussian components, (d) deviation between empirical and estimated densities, (e) estimated density of absolute deviation, (f) LCDG components, (g) final estimated density, and the (h) final estimated marginal density for each class.

$s = (x, y, z)$; ν_s is the 26-neighborhood system; $\eta_{s,r}$ is the GGMRF potential, and ρ and λ are scaling factors. The parameter $\beta \in [1.01, 2.0]$ controls the level of smoothing (e.g., $\beta = 2$ for smooth versus $\beta = 1.01$ for relatively abrupt edges). The parameter $\alpha \in \{1, 2\}$ determines the Gaussian, $\alpha = 2$, or Laplace, $\alpha = 1$, prior distribution of the estimator. In Fig. 2, Step 1 demonstrates the effect of preprocessing of our framework.

B. Initial Skull Stripping and Isosurfaces Generation

The second step of the proposed framework is to initially remove the nonbrain tissue from the MR images. To accomplish this, we used the BET [4] technique, which is a widely used deformable model-based approach for stripping the skull from brain MRIs. The process of skull stripping is very sensitive to the assigned BET factor. Using a high value would result in removing most of the undesired nonbrain tissues, yet losing details of the brain tissues. On the other hand, a lower BET factor

would preserve most of the brain tissues as well as nonbrain tissues. The BET factor used for best results was 0.38. In Fig. 2, Step 2(a) demonstrates the idea. While the BET extracted the brain without losing any of its parts using the assigned BET factor, it failed to remove all nonbrain tissues. For some clinical applications, such as cortical thickness measurement, inaccurate skull stripping results in an over- or underestimation of the thickness. This can drastically affect further analysis, such as feature extraction needed for abnormality detection. Therefore, it is important to account for inaccurate skull-stripping results after the BET step. To achieve this, we propose an additional processing step based on the geometric features of the brain to account for BET's skull-stripping errors. Since the nonbrain tissues are brighter than brain tissue, this step exploits the visual appearance features of the MR brain data. Namely, an evolving isosurface-based approach is proposed to remove the nonbrain tissues, which is guided by the MR data visual appearance features [see Fig. 2, Steps 2(b)–(c)]. First, a set of nested tangent
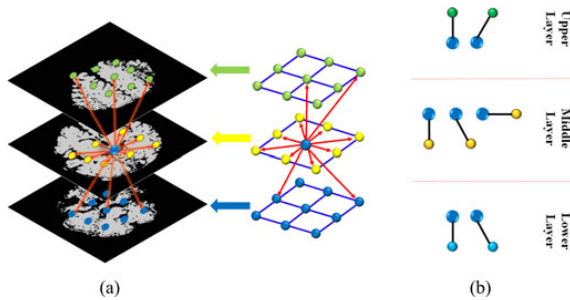
Fig. 4. (a) Graphical illustration for the 3-D neighborhood system and (b) a sample of the different pairwise cliques for the second-order MGRF.
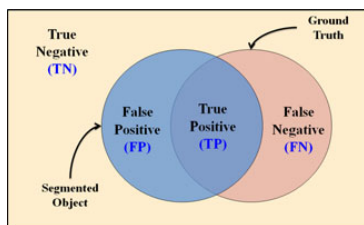

Fig. 5. Segmentation errors calculation between the segmented and ground truth objects for the determination of the DSC.
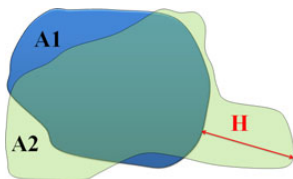

Fig. 6. Schematic illustration for the HD calculation.

surfaces (i.e., isosurfaces) is generated by the fast marching level sets (FMLS) [33], using the calculated distance map of the extracted brain from the BET step. Then, classification of voxels as brain or nonbrain is conducted and is presented in the following section.

C. Modeling Visual Appearance Using 3-D Joint MGRF

In order to classify MRI voxels as brain or nonbrain, we need to accurately model MR data visual appearance. To achieve this goal, we will use a 3-D joint Markov–Gibbs random field (MGRF) model, which is described as follows. Let $\mathbf{Q} = \{0, \ldots, Q-1\}$ and $\mathbf{L} = \{\text{"brain","nonbrain"}\}$ denote the sets of gray levels q and region labels L, respectively. Let $\mathbf{R}$ denote a 3-D arithmetic lattice supporting a given gray-scale image $\mathbf{g} : \mathbf{R} \rightarrow \mathbf{Q}$ and its region map $\mathbf{m} : \mathbf{R} \rightarrow \mathbf{L}$. The 3-D T1-weighted MR images $\mathbf{g}$ and its map $\mathbf{m}$ are described with the following joint probability model:

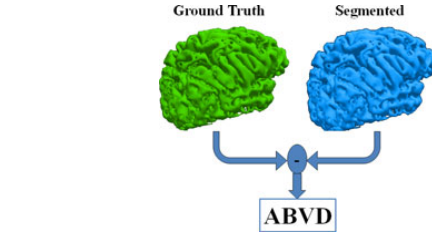$$P(\mathbf{g}, \mathbf{m}) = P(\mathbf{g}|\mathbf{m})P(\mathbf{m}) \qquad (2)$$

Fig. 7. 3-D schematic illustration for the ABVD estimation.
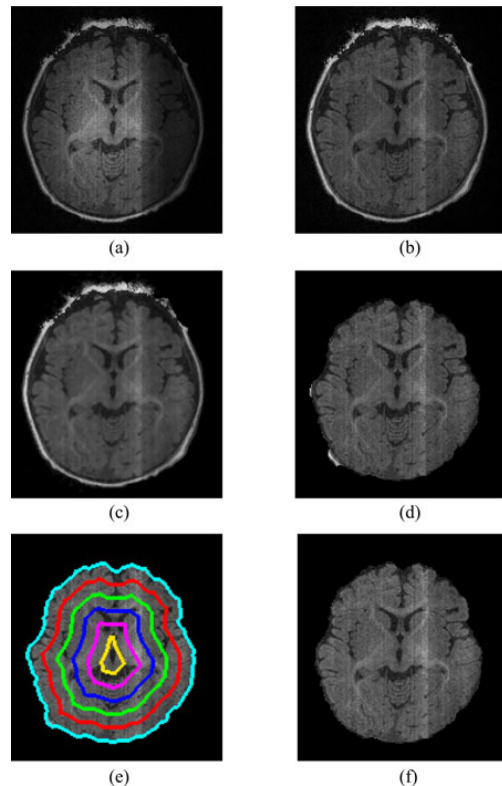
Fig. 8. Stepwise brain extraction using our framework: (a) Original MR image, (b) bias-corrected image, (c) GGMRF-edge preserved image obtained with $\rho = 1$, $\lambda = 5$, $\beta = 1.01$, $\alpha = 2$, and $\eta_{s,r} = \sqrt{2}$, (d) extracted brain using BET [4], (e) isosurfaces used to remove nonbrain tissues, and (f) final extracted brain.

where $P(\mathbf{m})$ is an unconditional probability distribution of maps, and $P(\mathbf{g}|\mathbf{m})$ is a conditional distribution of the images given the map. The ultimate goal is to accurately estimate $P(\mathbf{g}|\mathbf{m})$ and $P(\mathbf{m})$, which are described next.

1) First-Order Visual Appearance $P(\mathbf{g}|\mathbf{m})$: To accurately approximate the marginal probability distributions of the brain and nonbrain tissue, the empirical gray-level distribution of a given brain data is precisely approximated with a linear combination of discrete Gaussians (LCDG) with positive and negative

Input   Output

Fig. 9. More stripping results for different cross sections from different subjects showing reliable brain extraction using the proposed approach.

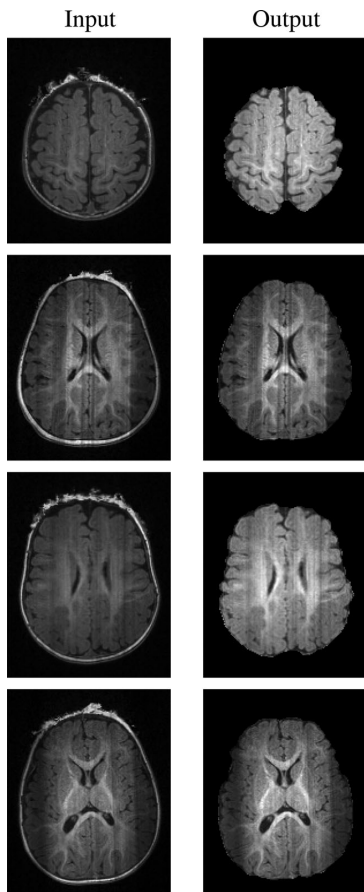Fig. 10. 3-D results of the proposed approach on two independent infant subjects projected into the axial, coronal, and sagittal views.

components [34]. The LCDG restores brain and nonbrain transitions more accurately than a conventional mixture of only positive Gaussians. This results in obtaining a better initial map $\mathbf{m}$ formed by voxelwise classification of the image gray values. Next, the LCDG is explained in more details.

Let $\boldsymbol{\Psi}_\theta = (\psi(q|\theta) : q \in \mathbf{Q})$ defines a discrete Gaussian (DG),[1] where $\theta = (\mu, \sigma)$, integrating a continuous 1-D Gaussian density with mean μ and variance σ^2 over successive gray-level intervals [34]. The LCDG with two dominant positive DGs and $M_{\mathrm{p}} \geq 2$ positive and $M_{\mathrm{n}} \geq 0$ negative subordinate DGs is defined as [34]

$$P_{\mathbf{w},\boldsymbol{\Theta}}(q) = \sum_{i=1}^{M_{\mathrm{p}}} w_{\mathrm{p}:i}\psi(q|\theta_{\mathrm{p}:i}) - \sum_{j=1}^{M_{\mathrm{n}}} w_{\mathrm{n}:j}\psi(q|\theta_{\mathrm{n}:j}) \quad (3)$$

---

[1]A DG $\Psi_\theta = (\psi(q|\theta) : q \in \mathbf{Q})$ with $\theta = (\mu, \sigma^2)$ is defined as $\psi(q|\theta) = \Phi_\theta(q + 0.5) - \Phi_\theta(q - 0.5)$ for $q = 1, \ldots, Q - 2$, $\psi(0|\theta) = \Phi_\theta(0.5)$, and $\psi(Q - 1|\theta) = 1 - \Phi_\theta(Q - 1.5)$, where $\Phi_\theta(q)$ is the cumulative Gaussian function with the mean μ and the variance σ^2.
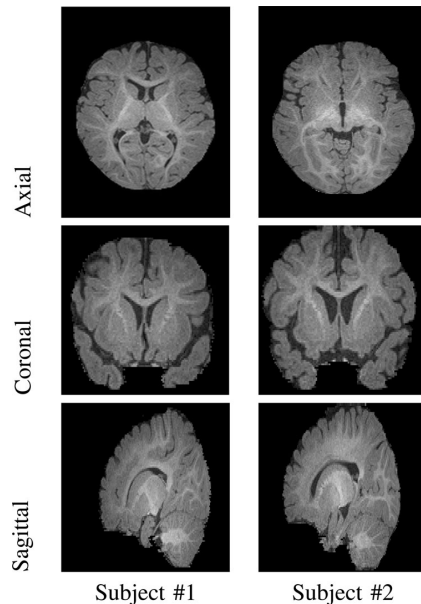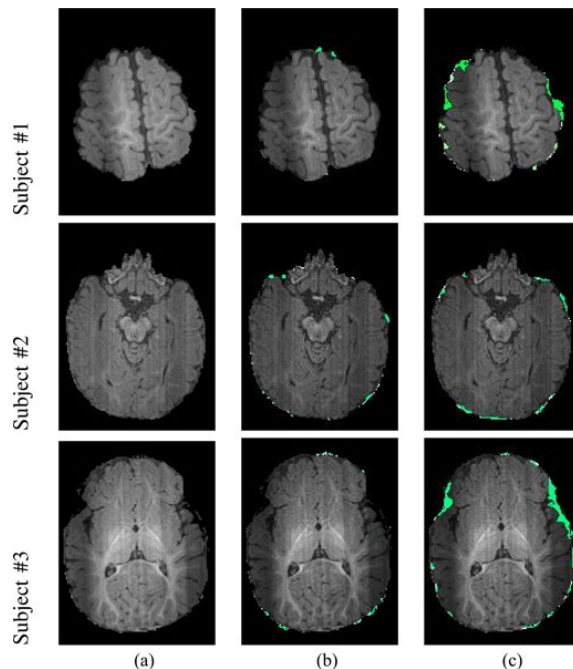

Fig. 11. Comparative infant brain extraction results for three independent subjects: (a) ground truth images and the brain extraction results obtained using (b) the proposed and (c) the iBEAT [24] techniques.
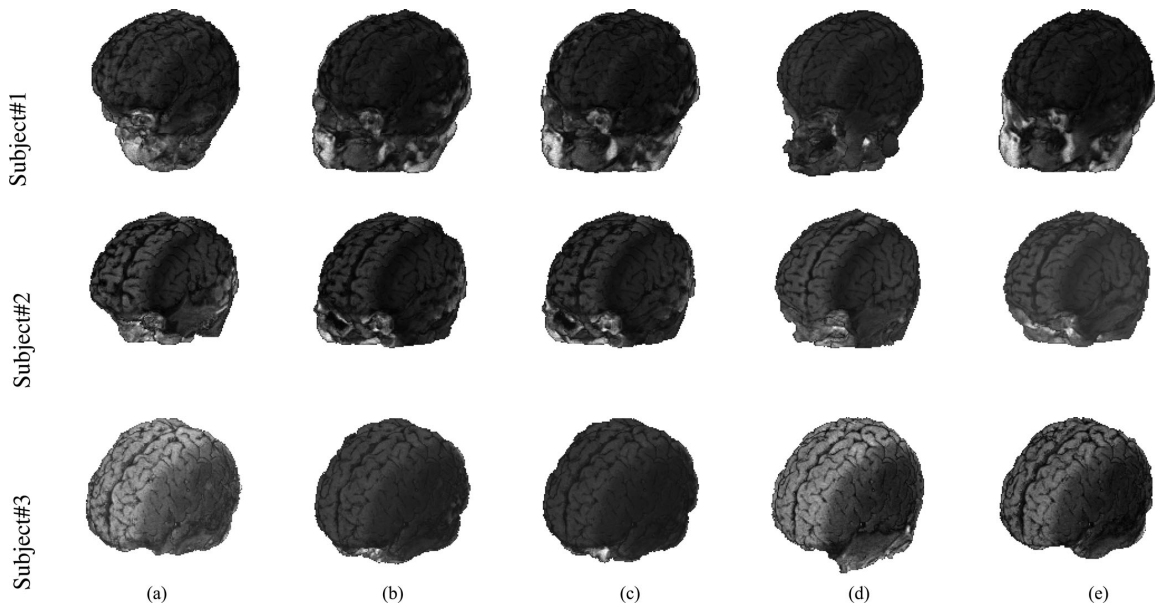
Fig. 12. 3-D visualization of the segmented brain tissue for three different subjects using (a) our approach, (b) BET [4], (c) BET2 [40], (d) BSE [18], and (e) iBEAT [24].

where all the weights $\mathbf{w} = [w_{\mathrm{p}:i}, w_{\mathrm{n}:j}]$ are nonnegative and meet an obvious constraint $\sum_{i=1}^{M_{\mathrm{p}}} w_{\mathrm{p}:i} - \sum_{j=1}^{M_{\mathrm{n}}} w_{\mathrm{n}:j} = 1$. All the LCDG parameters, including the numbers of DGs, are estimated from the mixed empirical distribution to be modeled using the modified expectation-maximization-based algorithm introduced in [35].

The marginal intensity distributions of the MR infant images have two dominant modes: one mode for brain tissues, and a second mode for nonbrain tissues. Fig. 3 shows the steps to build the LCDG models of the two modes. First, the marginal empirical probability distribution of the input gray-level images [see Fig. 3(a) are collected [see Fig. 3(b)]. Then, the obtained empirical distribution is approximated with a mixture of two positive DGs relating each to a dominant mode [see Fig. 3(c)]. Second, the deviations between the empirical and estimated distributions [see Fig. 3(d)] are approximated with the alternating "subordinate" components of the LCDG described in [35]. Finally, the obtained positive and negative subordinate mixtures [see Fig. 3(f)] are added to the dominant mixture yielding the final mixed LCDG model [see Fig. 3(g)], which is partitioned into two LCDG submodels [one per class, see Fig. 3(h)] by incorporating the subordinate DGs with the dominant terms so that the misclassification rate obtained is minimal [34].

2) 3-D Second-Order Visual Appearance ($P(\mathbf{m})$): In order to overcome noise effect and to ensure segmentation homogeneity, the spatial interactions between the region labels of a brain map $\mathbf{m}$ are also considered using Potts MGRF model. This model is identified using the nearest voxels' 26-neighbors [see Fig. 4(a)] and analytical bivalued Gibbs potentials [see

Fig. 4(b)]. The 3-D MGRF model is defined by El-Baz [36] as

$$P(\mathbf{m}) \propto \exp \sum_{(x,y,z) \in \mathbf{R}} \sum_{(\xi, \zeta, \kappa) \in \nu_s} \mathbf{V}(m_{x,y,z}, m_{x+\xi, y+\zeta, z+\kappa}) \ (4)$$

where $\mathbf{V}$ is the bivalue Gibbs potential that depends on the equality of the nearest pair of labels

$$V = \begin{cases} V(\lambda, \ \lambda') = V_{\mathrm{eq}}, \ \text{if} \ \lambda = \lambda' \\ V(\lambda, \ \lambda') = V_{\mathrm{ne}}, \ \text{if} \ \lambda \neq \lambda' \end{cases}. \quad (5)$$

Let $f_{a,\mathrm{eq}}(\mathbf{m})$ denote the relative frequency of the equal label pairs in the equivalent voxel pairs $\{((x,y,z), (x+\xi, y+\zeta, z+\kappa)) : (x,y,z), (x+\xi, y+\zeta, z+\kappa) \in \mathbf{R}; (\xi, \zeta, \kappa) \in \nu_s\}$. The initial $\mathbf{m}$ results in approximate analytical maximum likelihood potentials estimates [36]

$$V_{\mathrm{eq}} = -V_{\mathrm{ne}} \approx 2f_{\mathrm{eq}}(\mathbf{m}) - 1 \quad (6)$$

that allow for computing the voxelwise probabilities $p_{x,y,z}(m_{x,y,z} = \lambda)$ of each label $\lambda \in \mathbf{L}$. In Fig. 2, Step 3 shows the classification process. Algorithm 1 summarizes the basic steps of the proposed framework.

III. PERFORMANCE EVALUATION METRICS

The performance of the proposed framework was evaluated using three performance metrics: 1) the Dice similarity coefficient (DSC) [37], 2) the 95% modified Hausdorff distance (MHD) [38], and 3) the absolute brain volume difference (ABVD). The following sections explain the three metrics in more detail.

TABLE I
COMPARATIVE ACCURACY OF OUR APPROACH VERSUS THE METHODS IN [4],[18], [24], AND [40] BY THE DSC, MHD, AND ABVD ON 30 INFANT DATASETS WITH
AVAILABLE GROUND TRUTH ("SD"—STANDARD DEVIATION)

| | Evaluation Metric | | | | | |
| | DSC (%) | | MHD (mm) | | ABVD (%) | |
Method	Mean $\pm$ SD	p-value	Mean $\pm$ SD	p-value	Mean $\pm$ SD	p-value
OUR	96.77 ± 1.45	——	4.14 ± 1.76	——	4.95 ± 2.62	——
iBEAT [24]	94.44 ± 2.14	0.0162	9.31 ± 8.96	0.04	7.27 ± 1.92	0.0007
BET2 [40]	89.98 ± 1.53	0.0001	11.03 ± 7.4	0.0033	11.57 ± 7.05	0.03
BET [4]	88.7 ± 2.7	0.0001	13.6 ± 5.6	0.0005	15.7 ± 6.5	0.0001
BSE [18]	88.42 ± 3.03	0.0001	18.4 ± 4.02	0.0005	19.07 ± 3.722	0.015
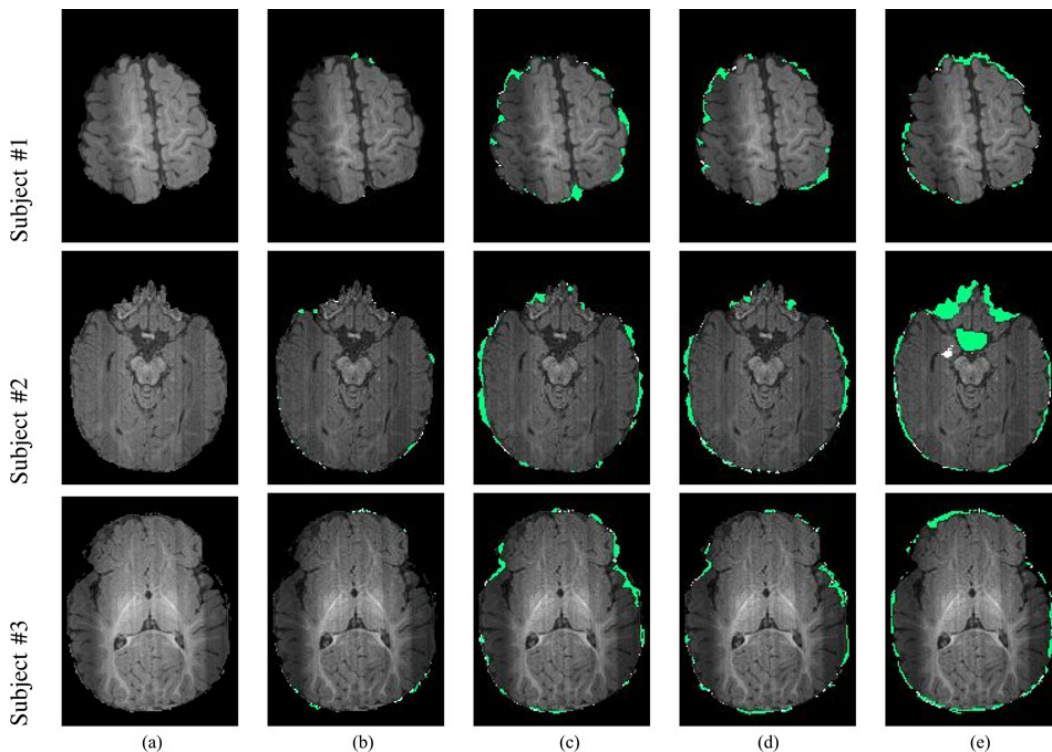
Fig. 13. Comparative infant brain extraction results for three independent subjects: (a) ground truth images and the brain extraction results obtained using (b) our, (c) BET [4], (d) BET2 [40], and (e) BSE [18] approaches.

A. DSC

The DSC characterizes the agreement between the segmented and ground truth objects (as seen in Fig. 5). The DSC measure is given from [37] as

$$DSC = \frac{2TP}{2TP + FP + FN} \qquad (7)$$

where TP, FP, and FN denote the true positive, false positive, and false negative, respectively. Higher DSC values indicate better segmentation, which means that the results match the ground truth better than results with lower DSC values. A DSC value of 0 indicates no overlap, and a DSC value of 1 indicates ideal segmentation (or agreement).

B. MHD

In order to measure the error distance between the segmented and ground truth objects, we used the MHD. The Hausdorff distance (HD) [38] from a set A_1 to a set A_2 is defined as the maximum distance of the set A_1 to the nearest point in the set A_2 (as shown in Fig. 6)

$$H(A_1, A_1) = \max_{c \in A_1} \{ \min_{e \in A_2} \{ d(c, e) \} \} \qquad (8)$$

Algorithm 1 Proposed Brain Extraction Approach

---

1) Perform preprocessing on the MR brain data:
   a) Bias correction of the brain [31].
   b) GGMRF edge preservation, Fig. 2, Step 1.
2) Perform initial skull stripping using BET [4], Fig. 2, Step 2 (a).
3) Calculate the distance map inside the binary mask obtained from BET using FMLS [33], then generate a set of N isosurfaces from this map, Fig. 2, Step 2 (b),(c).
4) Estimate the LCDG models for brain and nonbrain tissues and form an initial map **m** by voxelwise classification, Fig. 2, Step 3.
5) Estimate analytically the Gibbs potentials for the pairwise MGRF model of **m** to identify the MGRF probability, Fig. 2, Step 3.
6) while $j \leq N$
   a) Select the jth isosurfacer and classify its voxels using a Bayes classifier combining the first- and second-order visual appearance features.
   b) Are all the voxels on the selected isosurfaces classified only as brain tissue?
      • No $\longrightarrow$ **Go to Step 6 (a)**.
      • Yes $\longrightarrow$ **Break**
7) Apply connected component analysis to get rid of any scattered nonbrain tissues, Fig. 2, Step 4.

---

where c and e denote points of set A_1 and A_2, respectively, and $d(c, e)$ is the Euclidean distance between c and e.

The bidirectional HD between the segmented region SR and its ground truth GT is defined as

$$H_{Bi}(\text{GT}, \text{SR}) = \max\{H(\text{GT}, \text{SR}), H(\text{SR}, \text{GT})\}. \qquad (9)$$

In this paper, we use the 95th-perecntile bidirectional HD as a metric that measures the segmentation accuracy, which is also known as the MHD.

C. ABVD

In addition to the DSC and the MHD, we have used the ABVD as a third metric for measuring the segmentation accuracy. The ABVD is the percentage volume difference between the segmentation and the ground truth as seen in Fig. 7.

IV. EXPERIMENTAL RESULTS

In order to assess the robustness and performance of the proposed framework, we applied it to 300 T1-weighted MR infant brain datasets which were obtained from the infant brain imaging study [39]. An MR expert assessed the results qualitatively. In addition, the accuracy of the proposed approach was quantitatively validated using 30 datasets with known manually segmented ground truth that were obtained by an MR expert. The MR data were acquired at 3 T and consists of T1- and T2-weighted MR images of infants scanned at approximately 5–9 months old with voxel size of $1 \times 1 \times 1$ mm^3.

A stepwise brain extraction using the proposed approach for a selected axial cross section of one subject is demonstrated in Fig. 8. The input MR image [see Fig. 8(a)] is first bias corrected [see Fig. 8(b)] and applied to the 3-D GGMRF [30] edge preservation [see Fig. 8(c)]. This is followed by an initial brain extraction using BET [4] [see Fig. 8(d)]. As mentioned in Section II, a relatively small BET factor was chosen in order to minimize the loss of brain tissues. The average value concluded from the experiments on the subjects used is 0.38. The proposed isosurfaces-based approach [see Fig. 8(e)] is then employed to achieve the final skull stripping as shown in Fig. 8(f). It is clear from the results in Fig. 8 that the proposed framework provides more accurate infant brain extraction than BET. More stripping results for different cross sections from different subjects are shown in Fig. 9. Moreover, Fig. 10 shows the 3-D extraction results of the proposed two infant subjects projected into the axial, coronal, and sagittal views for visualization.

To highlight the advantage of the proposed framework, we compared its performance to one of the state-of-the-art tools for infant brain extraction, called the infant brain extraction and analysis toolbox (iBEAT) [24]. The version used is iBEAT 1.1 which is a Linux-based software package, and is publicly available at http://www.nitrc.org/projects/ibeat. The iBEAT performs multiple complementary brain extractions using a metaalgorithm, as mentioned in Section I, yet, Figs. 11 (2-D images results) and 12(a) and (e) (3-D volume results) show that the proposed approach provides more accurate stripping results. Also, Table I shows that our approach performs notably better, according to the higher DSC and lower MHD values and ABVD values .

In addition to the comparison with iBEAT, we also compared our approach to three widely used BETs for adult brain extraction: the BSE [18], the BET [4], and the BET2 [40]. This comparison was conducted to show that the existing approaches for adult brain extraction would have low accuracy on infant brains. The comparative accuracy of the proposed approach versus the BSE, BET, and BET2 techniques on representative images for three subjects is shown in Fig. 13. As demonstrated in Fig. 13 and the 3-D extracted brains shown in Fig. 12, our approach extracted the brain tissue more accurately compared with the other approaches. The lower performance of the BET [4] could be caused by its sensitivity to image noise and inhomogeneity because this method relies only on voxels' intensity changes and does not account for spatial voxel interactions. On the other hand, the BET2 approach [40] slightly improves the brain extraction accuracy compared with the BET one. However, unlike the BET and our approach, the BET2 technique requires both T1-and T2-weighted MR images. The BSE [18] succeeds in accurately removing the skull but it removes parts from the brain tissues as well, which may lead to inaccurate results for some clinical application (e.g., cortical thickness underestimation).

Table I compares our approach with the BSE, BET, and BET2 techniques based on the DSC, MHD, and ABVD metrics. All metrics were obtained by comparing brain extraction results of the 30 datasets against the available ground truth
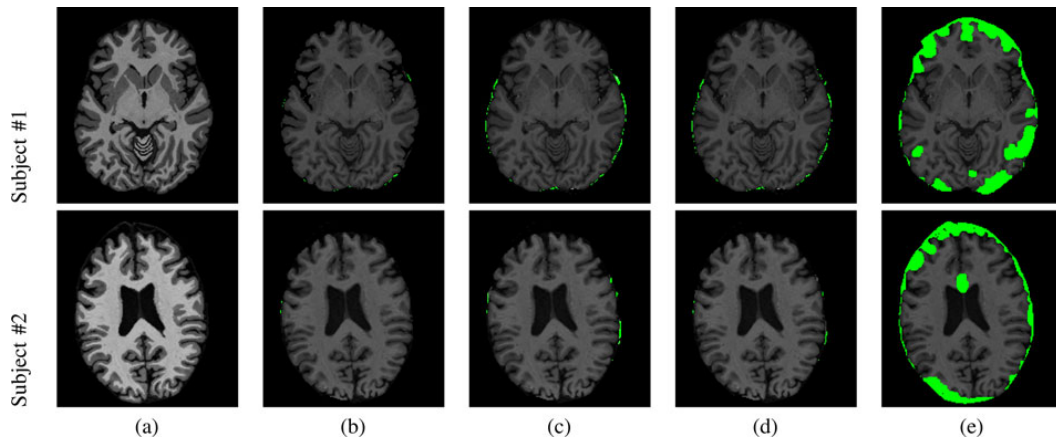
Fig. 14.　Comparative brain extraction results for two independent adult subjects using (b) our approach, (c) BET [4], (d) BET2 [40], and (e) BSE [18]. Ground truth images are shown in (a). The BET factor used for BET and BET2 methods is the default value (0.5). Segmentation errors are highlighted in green. BET and BET2 did not discard all nonbrain tissue, whereas BSE discarded nonbrain, yet some parts of brain tissue were discarded as well.

TABLE II
COMPARATIVE ACCURACY OF OUR APPROACH VERSUS THE METHODS IN [4], [18], AND [40] BY THE DSC, MHD, AND ABVD ON 18 ADULT DATASETS WITH
AVAILABLE GROUND TRUTH ("SD"—STANDARD DEVIATION)

| | Evaluation Metric | | | | | |
| | DSC (%) | | MHD (mm) | | ABVD (%) | |
Method	Mean $\pm$ SD	p-value	Mean $\pm$ SD	p-value	Mean $\pm$ SD	p-value
OUR	93.11 ± 2.4	——	8.43 ± 1.88	——	6.7 ± 1.13	——
BET2 [40]	92.60 ± 2.86	0.0001	10.56 ± 6.19	0.0001	7.42 ± 7.67	0.2035
BET [4]	91.14 ± 2.62	0.0001	11.26 ± 5.83	0.0001	12.69 ± 6.51	0.0001
BSE [18]	89.92 ± 7.11	0.0026	15.12 ± 10.09	0.0001	18.84 ± 22.87	0.0001

segmentation. As demonstrated in Table I, the mean DSC, MHD, and ABVD values for our framework are $96.77 \pm 1.45\%$, 4.14 ± 1.76 mm, and $4.95 \pm 2.62\%$, respectively, which confirm the high accuracy of our approach. Statistical significance of the better performance of our approach with respect to other methods is confirmed by the paired t-tests (p-values are less than 0.05).

In order to show that our approach is not limited to infants and that it could be generalized to adult brains, we applied it on 18 adult subjects and compared its performance to that of BSE, BET, and BET2. The scans are T1-weighted, with voxel size of 0.958 mm $\times$ 0.958 mm $\times$ 3.0 mm. They have been acquired at the UMC Utrecht (the Netherlands) of patients with diabetes and matched controls with varying degrees of atrophy and WM lesions, and ages are 50 and higher. The comparative accuracy of the proposed approach versus BSE, BET, and BET2 on representative images for two subjects are shown in Fig. 14. As demonstrated, our approach [see Fig. 14(b)] extracted the brain tissue more accurately compared with the other approaches. Table II compares our approach on adults with the BSE, BET, and BET2 based on the DSC, MHD, and ABVD metrics. All metrics were obtained by comparing results against the 18 datasets with available ground truth segmentation. As demonstrated in Table II, the mean DSC, MHD, and ABVD

values for our framework are $93.11 \pm 2.4\%$, 8.43 ± 1.88 mm, and $6.7 \pm 1.13\%$, respectively.

V. CONCLUSIONS AND FUTURE WORK

In conclusion, this paper has introduced a novel framework for automated extraction of the brain from 3-D infant MR images. Our experiments show that the fusion of stochastic and geometric models of the brain MRI data leads to more accurate brain extraction, when compared with widely used BETs: iBEAT, BSE, BET, and BET2. The results were evaluated using the DSC, 95% MHD, and the ABVD metrics on 30 infant MR brain datasets. The proposed approach also proved efficient when applied on adult brains, and compared versus state-of-the-art methods. A future extension of this paper would be integrating the proposed skull-stripping approach into a computer-aided diagnostic system for early detection of autism, which is the ultimate goal of our research group. Integrating higher-order cliques of the joint MGRF model could be another extension to the proposed framework, where, currently, only pairwise cliques are exploited. Higher-order cliques (e.g., third- and fourth-order) will better account for noise and large inhomogeneities of the MRI scans of infants.

REFERENCES

[1] A. J. van der Kouwe, T. Benner, D. H. Salat, and B. Fischl, "Brain morphometry with multiecho mprage," *NeuroImage*, vol. 40, no. 2, pp. 559–569, 2008.

[2] J. G. Park and C. Lee, "Skull stripping based on region growing for magnetic resonance brain images," *NeuroImage*, vol. 47, no. 4, pp. 1394–1407, 2009.

[3] E. Roura, A. Oliver, M. Cabezas, J. C. Vilanova, A. Rovira, L. Rami-Torrent, and X. Llad, "Marga: Multispectral adaptive region growing algorithm for brain extraction on axial MRI," *Comp. Methods Prog. Biomed.*, vol. 113, no. 2, pp. 655–673, 2014.

[4] S. M. Smith, "Fast robust automated brain extraction," *Human Brain Mapping*, vol. 17, no. 3, pp. 143–155, 2002.

[5] J.-X. Liu, Y.-S. Chen, and L.-F. Chen, "Accurate and robust extraction of brain regions using a deformable model based on radial basis functions," *J. Neurosci. Methods*, vol. 183, no. 2, pp. 255–266, 2009.

[6] H. Wendland, "Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree," *Adv. Comput. Math.*, vol. 4, no. 1, pp. 389–396, 1995.

[7] A. H. Zhuang, D. J. Valentino, and A. W. Toga, "Skull-stripping magnetic resonance brain images using a model-based level set," *NeuroImage*, vol. 32, no. 1, pp. 79–92, 2006.

[8] C. Baillard, P. Hellier, and C. Barillot, "Segmentation of brain 3D MR images using level sets and dense registration," *Med. Image Anal.*, vol. 5, no. 3, pp. 185–194, 2001.

[9] Y. Wang, J. Nie, P. T. Yap, G. Li, F. Shi, X. Geng, L. Guo, and D. Shen, "Knowledge-guided robust MRI brain extraction for diverse large-scale neuroimaging studies on humans and non-human primates," *PLoS One*, vol. 9, no. 1, pp. 1–23, 2014.

[10] S. A. Sadananthan, W. Zheng, M. W. Chee, and V. Zagorodnov, "Skull stripping using graph cuts," *NeuroImage*, vol. 49, no. 1, pp. 225–239, 2010.

[11] H. Zhang, J. Liu, Z. Zhu, and H. Li, "An automated and simple method for brain MR image extraction," *Biomed. Eng. Online*, vol. 10, no. 1, pp. 1–12, 2011.

[12] K. Somasundaram and P. Kalavathi, "Contour-based brain segmentation method for magnetic resonance imaging human head scans," *J. Comput. Assisted Tomography*, vol. 37, no. 3, pp. 353–368, 2013.

[13] K. K. Leung, J. Barnes, M. Modat, G. R. Ridgway, J. W. Bartlett, N. C. Fox, and S. Ourselin, "Brain MAPS: An automated, accurate and robust brain extraction technique using a template library," *NeuroImage*, vol. 55, no. 3, pp. 1091–1108, 2011.

[14] J. E. Iglesias, C.-Y. Liu, P. M. Thompson, and Z. Tu, "Robust brain extraction across datasets and comparison with publicly available methods," *IEEE Trans. Med. Imaging*, vol. 30, no. 9, pp. 1617–1634, Sep. 2011.

[15] F. Segonne, A. Dale, E. Busa, M. Glessner, D. Salat, H. K. Hahn, and B. Fischl, "A hybrid approach to the skull stripping problem in MRI," *NeuroImage*, vol. 22, no. 3, pp. 1060–1075, 2004.

[16] R. Beare, J. Chen, C. L. Adamson, T. Silk, D. K. Thompson, J. Y. Yang, V. A. Anderson, M. L. Seal, and A. G. Wood, "Brain extraction using the watershed transform from markers," *Front. Neuroinformat.*, vol. 7, no. 32, pp. 1–15, 2013.

[17] D. E. Rex, D. W. Shattuck, R. P. Woods, K. L. Narr, E. Luders, K. Rehm, S. E. Stolzner, D. A. Rottenberg, and A. W. Toga, "A meta-algorithm for brain extraction in MRI," *NeuroImage*, vol. 23, no. 2, pp. 625–637, 2004.

[18] D. W. Shattuck and R. M. Leahy, "Brainsuite: An automated cortical surface identification tool," *Med. Image Anal.*, vol. 6, no. 2, pp. 129–142, 2002.

[19] B. D. Ward, "Intracranial segmentation," Biophysics Research Institute, Medical College of Wisconsin, Milwaukee, WI, USA, 1999.

[20] A. M. Dale, B. Fischl, and M. I. Sereno, "Cortical surface-based analysis: I. Segmentation and surface reconstruction," *NeuroImage*, vol. 9, no. 2, pp. 179–194, 1999.

[21] J. Chiverton, K. Wells, E. Lewis, C. Chen, B. Podda, and D. Johnson, "Statistical morphological skull stripping of adult and infant MRI data," *Comput. Biol. Med.*, vol. 37, no. 3, pp. 342–357, 2007.

[22] S. Kobashi, Y. Fujimoto, M. Ogawa, K. Ando, R. Ishikura, K. Kondo, S. Hirota, and Y. Hata, "Fuzzy-ASM based automated skull stripping method from infantile brain MR images," in *Proc. IEEE Int. Conf. Granular Comput.*, 2007, pp. 632–632.

[23] M. Pport, D. E. I. Ghita, E. Twomey, and P. F. Whelan, "A hybrid approach to brain extraction from premature infant MRI," in *Proc. Scandinavian Conf. Image Anal.*, 2011, pp. 719–730.

[24] F. Shi, L. Wang, Y. Dai, J. H. Gilmore, W. Lin, and D. Shen, "LABEL: Pediatric brain extraction using learning-based meta-algorithm," *NeuroImage*, vol. 62, no. 3, pp. 1975–1986, 2012.

[25] D. W. Shattuck, G. Prasad, M. Mirza, K. L. Narr, and A. W. Toga, "Online resource for validation of brain segmentation methods," *NeuroImage*, vol. 45, no. 2, pp. 431–439, 2009.

[26] C. Fennema-Notestine, I. B. Ozyurt, C. P. Clark, S. Morris, A. BischoffGrethe, M. W. Bondi, T. L. Jernigan, and G. G. Brown, "Quantitative evaluation of automated skull-stripping methods applied to contemporary and legacy images: Effects of diagnosis, bias correction, and slice location," *Human Brain Mapping*, vol. 27, no. 2, pp. 99–113, 2006.

[27] F. Shi, Y. Fan, S. Tang, J. H. Gilmore, W. Lin, and D. Shen, "Neonatal brain image segmentation in longitudinal MRI studies," *NeuroImage*, vol. 49, no. 1, pp. 391–400, 2010.

[28] A. U. Mewes, P. S. Hueppi, H. Als, F. J. Rybicki, T. E. Inder, G. B. McAnulty, R. V. Mulkern, R. L. Robertson, M. J. Rivkin, and S. K. Warfield, "Regional brain development in serial magnetic resonance imaging of low-risk preterm infants," *Pediatrics*, vol. 118, no. 1, pp. 23–33, 2006.

[29] M. Casanova, A. El-Baz, and J. Suri, *Imaging the Brain in Autism*. New York, NY, USA: Springer-Verlag, 2013.

[30] C. Bouman and K. Sauer, "A generalized gaussian image model for edge-preserving MAP estimation," *IEEE Trans. Image Process.*, vol. 2, no. 3, pp. 296–310, Jul. 1993.

[31] N. J. Tustison, B. B. Avants, P. A. Cook, Y. Zheng, A. Egan, P. A. Yushkevich, and J. C. Gee, "N4ITK: Improved N3 bias correction," *IEEE Trans. Med. Imag.*, vol. 29, no. 6, pp. 1310–1320, Jun. 2010.

[32] J. Besag, "On the statistical analysis of dirty pictures," *J. Roy. Statist. Soc. Ser. B Methodological*, vol. 38, pp. 259–302, 1986.

[33] D. Adalsteinsson, "A fast level set method for propagating interfaces," Ph.D. dissertation, Dept. Math. Univ. California, Berkley, 1994.

[34] A. El-Baz, A. Elnakib, F. Khalifa, M. A. El-Ghar, P. McClure, A. Soliman, and G. Gimelrfarb, "Precise segmentation of 3-D magnetic resonance angiography," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 7, pp. 2019–2029, Jul. 2012.

[35] A. Farag, A. El-Baz, and G. Gimel'farb, "Precise segmentation of multimodal images," *IEEE Trans. Image Process.*, vol. 15, no. 4, pp. 952–968, Apr. 2006.

[36] A. El-Baz, "Novel stochastic models for medical image analysis," Ph.D. dissertation, Dept. Electr. Comput. Eng., Univ. Louisville, Louisville, KY, USA, 2006.

[37] K. H. Zou, S. K. Warfield, A. Bharatha, C. M. C. Tempany, M. R. Kaus, S. J. Haker, W. M. Wells III, F. A. Jolesz, and R. Kikinis, "Statistical validation of image segmentation quality based on a spatial overlap index," *Acad. Radiol.*, vol. 11, no. 2, pp. 178–189, 2004.

[38] G. Gerig, M. Jomier, and M. Chakos, "Valmet: A new validation tool for assessing and improving 3D object segmentation," in *Proc. Med. Image Comput. Comput. Assisted Intervention*, 2001, pp. 516–523.

[39] IBIS. Infant brain imaging study. (2007). [Online]. Available: http://www.ibisnetwork.org/

[40] M. Jenkinson, M. Pechaud, and S. Smith, "BET2: MR-based estimation of brain, skull and scalp surfaces," presented at the 11th Annu. Meeting Organization Human Brain Mapping, Toronto, ON, Canada, 2005, vol. 17.

Authors' photographs and biographies not available at the time of publication.