Robust 3D Convolutional Neural Network With Boundary Correction for Accurate Brain Tissue Segmentation

BEIBEI HOU[1], (Member, IEEE), **GUIXIA KANG**[1,2], **NINGBO ZHANG**[1], **AND CHUAN HU**[1], (Member, IEEE)

[1]Key Laboratory of Universal Wireless Communications, Ministry of Education, Beijing University of Posts and Telecommunications, Beijing 100876, China
[2]Wuxi BUPT Sensory Technology and Industry Institute CO. LTD, Wuxi 214001, China

Corresponding author: Guixia Kang (gxkang@bupt.edu.cn)

ABSTRACT The morphology, symmetry, and volume of brain tissue are good indicators for measuring the central nervous system disease progression. The objective of this paper is to segment cerebrospinal fluid (CSF), gray matter (GM), and white matter (WM) automatically with multi-modality magnetic resonance scans. A novel coarse-to-fine method is proposed to segment CSF, GM, and WM using two cascade 3D convolutional neural networks. The first densely connected fully convolutional network (DC-FCN) is designed with feature reuse, which can take full advantage of the spatial information and alleviate computer memory limitation. The second 6-CNN is designed to correct boundary voxel, which can further reduce computational cost while improving the segmentation accuracy. As of today, our method ranks the 3rd on the MRBrainS13 challenge, outperforming most of the participant methods when using available input modalities (T1, T1-IR, and T2-FLAIR). In addition, we also verify the proposed framework on the IBSR dataset, which demonstrates the effectiveness of the boundary correction strategy. Through accurate segmentation of brain tissue, neuroimaging physicians can be assisted in assessing disease progression and even localizing lesions.

INDEX TERMS Multi-modality, coarse-to-fine, 3D convolutional neural networks, cascade framework.

I. INTRODUCTION

Neuroimaging is an indispensable examination instrument to measure brain diseases in clinical practice. magnetic resonance imaging (MRI) is employed fashionably owing to its security, ability to represent soft tissues and visualization in three-dimensional. It is the only imaging technique that can provide the most information at a time currently, i.e. multi-modality. Nowadays, numerous toolboxes have been presented for the analysis of multi-modality MR scans [1]–[3].

The anatomic region labeling that is a prerequisite for the quantitative analysis of neuroimaging data attracts much attention. Generally, a human brain can be segmented into three main tissues, involving cerebrospinal fluid (CSF), gray matter (GM) and white matter (WM) [4], which behaves a good indicator for measuring the disease progression [5]. For example, brain tissue atrophy is a common manifestation

of neurological disorders, such as Alzheimer's [6] and Huntington's disease [7], multiple sclerosis [8], bipolar disorder or schizophrenia [9]. Consequently, quantitative characterization of brain tissue in MRI is conventional for the diagnosis of neurological diseases.

However, the segmentation of CSF, GM and WM in MRI is a challenging task. On one hand, MRI-related challenges embody numerous inevitable artifacts in brain images, such as partial volume effect (PVE), intensity nonuniformity (INU) and noises. On the other hand, subject-related challenges subsume personalized cortical thickness and morphology of the cortex, which blurs the boundary among brain tissues.

To relieve the workload on experts while improving the reliability of the radiologist assessment, a great effort has been devoted to develop the automated segmentation methods [10]. Atlas-based method [11], [12] is designed upon

the assumption that the spatial relationships among images complying with their anatomy, which is the most popular approach for brain image segmentation [13]. To improve the accuracy of segmentation, multi-atlas [14] and probabilistic-atlas [15] have been proposed in succession, which virtually aggravates the computational burden on the model. Therefore, a trade-off must be considered between fitting precision and computational cost [16], [17]. In addition, machine learning methods are also potential for medical image analysis [18], [19], including markov random field (MRF) [4], [20], [21], random forest (RF) [22], [23], support vector machine (SVM) [24], [25] and their combinations [26], [27]. However, traditional machine learning methods perform not well in generalization and a set of imaging features which require specific expertise in anatomy is necessary.

It is noticeable that deep learning methods behave best when it comes to static spatial information extraction. Typically, convolutional neural networks (CNNs) have shown its superiority in several computer vision tasks, such as the ImageNet challenge [28]. Recently, CNNs are also prevalent in medical image analysis [29] owing to its flexibility. Moeskops *et al.* [30] presented a multi-scale CNN framework fusing multiple patch size and multiple convolution kernel size. It simplified the tissue segmentation into a classification problem by extracting the patch in voxel and judging which tissue it belongs to. Subsequently, Nie *et al.* [31] employed fully convolutional networks (FCNs) to segment the isointense phase brain MR images. It performed efficiently because it avoids the problem of duplicate storage and computational convolution caused by numerous voxel. Brosch *et al.* [32] proposed a convolutional encoder network (CEN) by constructing two interconnected pathways, i.e. convolutional and de-convolutional pathway. CEN performed pre-training on the input images using restricted boltzmann machines (RBMs). However, a higher requirement is put forward for computer on the efficient partitioning of network architectures. When the depth of the network reaches a certain level, the computer cannot model successfully due to the limited memory-bound.

To accurately and efficiently segment brain tissues from volumetric MRIs, a novel coarse-to-fine segmentation framework is proposed by synthesizing the correlation between context-based and voxel-based learning method. Specifically, our pipeline consists of two cascade stages. The first stage is coarse segmentation, where the densely connected fully convolutional network (DC-FCN) [33] is designed in combination with convolution and deconvolution operations. It performs down-sample before traditional convolutional operations to decrease the size of data, which lowers the computer requirements as well as achieves significant acceleration. The second stage is boundary correction in which a fine boundary definition is performed. It amends the boundaries of the tissues in voxel level and yields the final segmentation results. To the best of our knowledge, this is the first work to segment brain tissue using cascaded CNNs.

The main contributions of this work are summarized as below:

(1) A coarse-to-fine cascade segmenting framework is proposed taking into account not only the overall brain but also the detail of the boundary between brain tissues.

(2) DC-FCN with three-dimensional convolution and deconvolution operations is constructed for the coarse brain tissue segmentation. It achieves better performance with fewer parameters through repeated utilization of features.

(3) The concept of boundary correction is presented. Only a very small number of voxel are selected (∼5% of total brain volume) for the boundary correction model, which decreases the redundancy of patches noticeably.

(4) The generalization capabilities of our cascade framework are illustrated not only on the internet brain segmentation repository (IBSR) datasets, but also with an online assessment in the MR brain image segmentation challenge (MRBrainS).

The rest of this paper is organized as follows. The proposed framework is presented in Section II, where we describe the detailed cascade framework. In Section III, the multiple databases are described, along with the experimental results. The discussion is presented in Section IV and the main conclusion is drawn in Section V.

II. METHOD
An overview of the proposed segmentation framework is illustrated in Figure 1. The brain tissue is automatically segmented in a cascaded scheme, including preprocessing and segmentation framework. This work is detailed as follows:

A. PREPROCESSING
Considering the MRI-related and subject-related challenges described above, specific preprocessing steps are required before brain tissue segmentation. With the extensive development of deep learning, the impact of some preprocessing steps on the final segmentation performance is negligible gradually [34]. Therefore, only the skull stripping and intensity normalization are carried out for structural MR scans.

1) SKULL STRIPPING
From the perspective of images, non-brain tissue in structural MRI will affect the accurate segmentation of brain tissue seriously, such as scalp, skull and so on. In this paper, the skull stripping is achieved by BET [35] from FSL software.

2) INTENSITY NORMALIZATION
From the perspective of intensity, it should be mapped into a standard scale. For each scan, it first subtracts its mean and then divides its variance, which can speed up the procedure of model training as well.
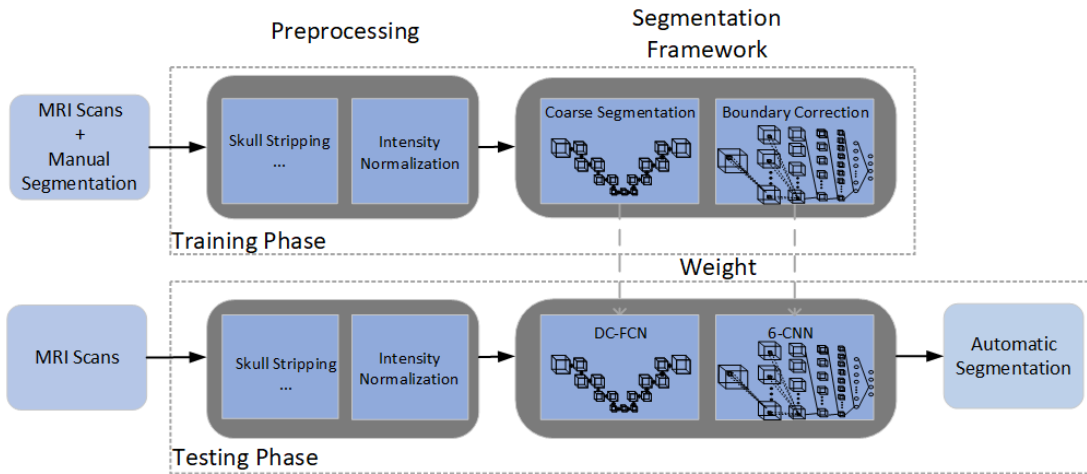
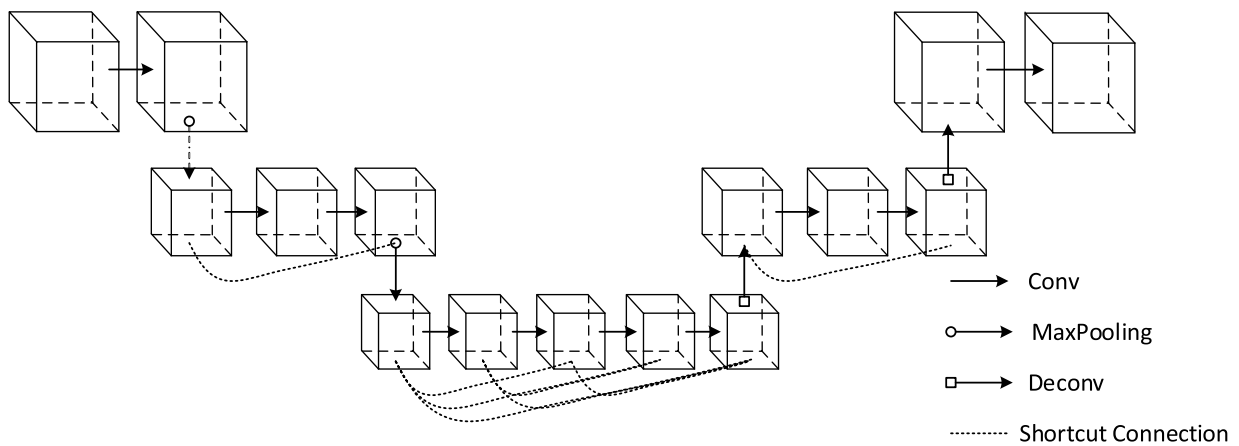**FIGURE 1.** Overview of the proposed segmentation pipeline.

FIGURE 2. A simple schematic of the DC-FCN used in coarse segmentation.

B. SEGMENTATION FRAMEWORK

The main challenge of brain tissue segmentation is the determination of boundaries between tissues. Accordingly, a pipeline with two cascade components is designed, which includes DC-FCNs that produce overall segmentation maps and CNNs that impose boundary correction on the output of DC-FCNs and produce the fine segmentation labels. The detail of segmentation framework is revealed in Algorithm 1.

1) COARSE SEGMENTATION

In computer vision, the extraction and utilization of features affect the performance of models directly. In the field of medical image analysis, the width and depth of the network are usually limited to computer memory, resulting in infinite precision. Densely connected convolutional network (DenseNet) takes advantage of features by short connections to enhance the transmission between features, making the network narrower with fewer parameters. Different from the residual network [36], DenseNet concatenates all the layers on the premise of guaranteeing the maximum

transmission of information among layers. Each layer in DenseNet has direct access to gradients from the loss function and the original input signal, leading to an implicit deep supervision. It improves the delivery of information and gradients throughout the network, avoiding the shortcoming that the residual network is easy to trap in the local optimum.

The DC-FCN is designed referring to the architecture of U-net [37] for rough segmentation in brain MR images. As shown in Figure 2, it consists of two symmetric pathways, that is, contracting-pathway with convolutional operations [38] and expansive-pathway with deconvolution operations [39]. In order to improve the accuracy as well as has fewer parameters, the kernel size in the convolutional layer and deconvolution layer is set to $3*3*3$, batch normalization (BN) is used after each convolutional operation and rectified linear unit (ReLU) is used as the activation function. In the last layer, the feature map is converted to the result with a specific depth, so its kernel size is set to $1*1*1$ and the activation function is soft-max.

Algorithm 1 Hybrid Scheme for Segmentation Framework

Given notations

(N_x, N_y, N_z): size of input volume

L_1: the depth of global network

L_2: the depth of classification network

U(●): 3D up-sampling

D(●): 3D down-sampling

B(●): Boundary voxels selection

P(●): Record the position of the boundary voxel

A(●): Substitute the value of the corresponding position

Input

$I \in \mathbb{R}^{N_x \times N_y \times N_z}$: T1-weighted volume

$O_1^i(●)$: the operation in *i-th* layer of coarse segmentation model

$O_2^i(●)$: the operation in *i-th* layer of classification model

S_i: the output of *i-th* layer of coarse segmentation model

SS_i: the voxel output of *i-th* layer of classification model

$S_F \in \mathbb{R}^{N_x \times N_y \times N_z}$: Final segmentation map

Phase 1 - Coarse Segmentation

1. **for** $i = 1 : L_1$ **do**
2.   **if** $i = 1$ **then**
3. $S_0 = D(I) \in \mathbb{R}^{N_x/2 \times N_y/2 \times N_z/2}$
4.   **else if** $i = L_1$ **then**
5. $S_{L1} = U(S_{i-1})$
6.   **else**
7. $S_i \leftarrow O_1^i(W_1^i, S_{i-1})$
8.   **end if**
9. **end for**
10. $B = B(S_{L1})$
11. $P = P(B, S_{L1})$

Phase 2 – Boundary Correction

12. $SS_0 = B$
13. **for** $i = 1 : L_2$ **do**
14. $SS_i \leftarrow O_2^i(W_2^i, SS_{i-1})$
15. **end for**
16. $S_F = A(S_{L1}, P, SS_{L2})$
17. **return** S_F

The DC-FCN performs an overall analysis of the input image with an output size that is consistent with the input image. Therefore, we design a loss function based on the dice similarity index to achieve the coarse segmentation of brain tissue as shown in Eq.(1).

$$E_1 = -\frac{1}{n}\sum_x \sum_k (1 - 2\frac{y_{1,k}^x \times \hat{y}_{1,k}^x}{\left|y_{1,k}^x\right| + \left|\hat{y}_{1,k}^x\right|}) + \frac{\lambda}{2n}w_1^2, \quad (1)$$

where $y_{1,k}^x$ represents the probability prediction of image x on category $k (k \in \{1, 2, 3, 4\})$ in the first stage, w_1 denotes the weight in the DC-FCN, and λ is regularization coefficient which is set to 0.05 in this study.

2) BOUNDARY CORRECTION

It is particularly difficult to determine the boundary between tissues in view of the mentioned subject-related challenges.

Accordingly, the cascade pipeline is proposed to overcome the difficulties of tissue segmentation by adjusting boundary voxel iteratively.

The detail of boundary voxel selection are summarized in Algorithm 2.

Algorithm 2 Boundary Voxels Selection

Given notations

(V_x, V_y, V_z): size of volume of interests

Input

$I \in \mathbb{R}^{N_x \times N_y \times N_z}$: T1-weighted volume

$Sc \in \mathbb{R}^{N_x \times N_y \times N_z \times 4}$: Rough segmentation results

Output

B: Filtered border dataset

P: the location dataset corresponding to B
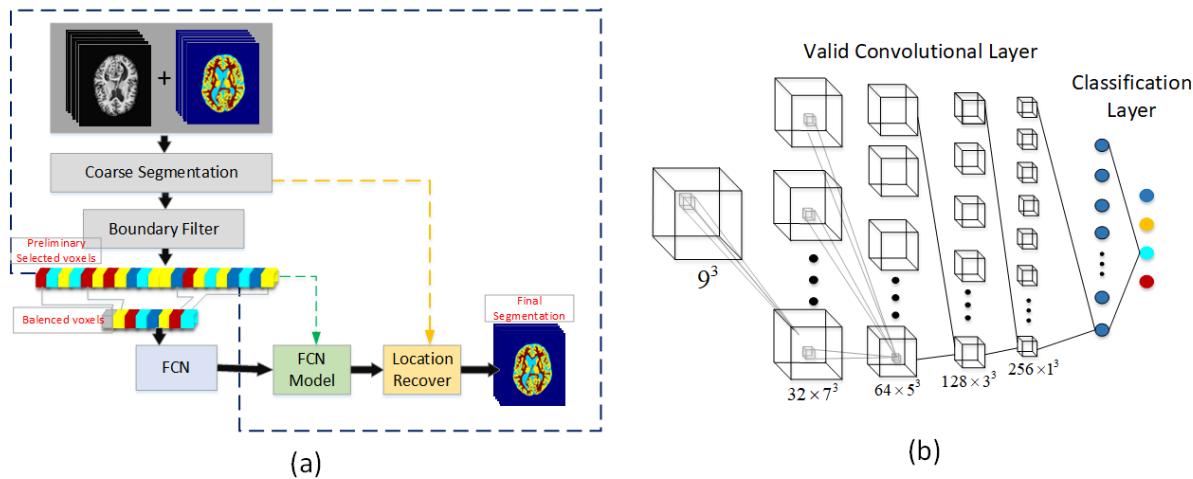
Boundary Voxels Selection

1. $B = [\,]$
2. **for** $c = \{i, j, k\}$ in I **do**
3. $n \leftarrow$ the tissue category of the voxel from S_c at location c
4. $v \leftarrow$ Extracting volume centered at c and size with (V_x, V_y, V_z) from I
5. $p \leftarrow$ Extracting rough segmentation centered at c and size with $(3, 3, 3)$ from S_c in channel n
6.   **if** there is 0 in p **then**
7.   Append volume v to border dataset B
8.   Append location c to location dataset P
9.   **end if**
10. **end for**
11. **return** B, P

The set of boundary patches B is stacked as $B = [n \times v \times v \times v]$, where n denotes the number of the selected training patches and v denotes the size of the selected patches. Meanwhile, the flow of boundary correction is depicted in Figure 3a vividly. The small network is considered to prevent overfitting because of lower variation in tissues in MRI-based patches. Here, a 6-layer convolutional neural network (6-CNN) is designed for boundary correction (see Figure 3b). 6-CNN contains four stacks of convolution operation with a valid mode, followed by a fully-connected, and the soft-max classification layer in the last layer. Among them, the soft-max layer returns four categories which indicate the probability of each voxel belongs to the brain tissue.

In this issue, we simplify the boundary correction to a classification problem, recognizing the categories of tissues which the center voxel belongs to. Then, the typical cross-entropy is used as the loss function to optimize the classification model, as shown in Eq.(2).

$$E_2 = -\frac{1}{n}\sum_x \sum_k C_k y_{2,k}^{x'} \ln \hat{y}_{2,k}^{x'} + \frac{\lambda}{2n}w_2^2, \quad (2)$$

Where C_k indicates the equilibrium coefficient on $k\text{-}th$ tissue to compensate for the imbalance in the class gap. $y_{2,k}^{x'}$ represents the probability prediction of voxel x' on

(a)

(b)

FIGURE 3. (a) The pipeline for boundary correction and the flow of testing is represented in blue box. (b) The proposed CNN model is trained using the multi-modality 3D voxels selected from boundary filter.

category $k(k \in \{1, 2, 3, 4\})$ in the second stage, w_2 is the weight in the 6-CNN.

C. IMPLEMENTATION

Considering the limited training data, we adopt leave-one-out method to optimize as well as validate the proposed cascade architecture. The first DC-FCN is optimized using the stochastic gradient descent (SGD) with learning rate of 0.03, decay of 1E-4, and momentum of 0.6. And the second 6-CNN is optimized by adaptive moment estimation (Adam) [40] with learning rate of 1E-4.

The proposed method is implemented in the Python language, using Keras based on tensorflow libraries. All experiments are performed on a single GTX 1080 GPU with 8 GB RAM memory.

III. EXPERIMENTS AND RESULTS

A. MICCAI 2013 MR BRAIN IMAGE SEGMENTATION

DATA

The MRBrainS challenge is composed of 20 subjects acquired on a 3.0T Philips Achieva MR scans at university medical center Utrecht (Netherlands) [41]. Multi-sequence MRI brain scans, including T1 (TR:7.9ms, TE:4.5ms), T1-IR (TR:4416ms, TE:15ms, TI:400ms), and T2-FLAIR (TR:11000ms, TE:125ms, TI:2800ms), were acquired and used for the challenge. The rigid registration using Elastix [42] and bias correction using SPM8 [43] were performed for all scans. The voxel spacing of all provided sequences was $0.96 \times 0.96 \times 3.00$mm3 after such preprocessing.

- Training dataset: 5 typical samples (2 male + 3 female) withvarying degrees of atrophy and white matter lesions, and the corresponding manual annotations are provided.
- Testing dataset: 15 samples (8 male + 7 female) with labelsheld out by the organizing committee for fair and objective evaluation.

EVALUATION

To evaluate the segmentation results, the committee employs three type of measures: voxel-, distance-, and volumetric-based metrics. The dice similarity coefficient (DSC) [44] quantifies the spatial overlap between the manual (M) and the automatic (A) segmentation result. A larger value denotes a better performance. It is defined as:

$$DSC(M, A) = \frac{2|M \cap A|}{|M| + |A|} \times 100\%. \tag{3}$$

The boundary distance between the ground truth and automatic segmentation is depicted by the 95th-percentile of the Hausdorff distance (HD). On this basis, the *K-th* ranked distance [45] is used to suppress the outlier, which avoids the conventional Hausdorff distance is too sensitive to outliers. It is defined as:

$$h_{95}(M, A) = {}^{95}K_{\mathfrak{m} \in \mathbb{M}}^{th} \min_{\mathfrak{a} \in \mathbb{A}} ||a - m||, \tag{4}$$

where ${}^{95}K_{\mathfrak{m} \in \mathbb{M}}^{th}$ is the *K-th* ranked minimum euclidean distance with $K/Nm = 95\%$, $\mathbb{M}$ and $\mathbb{A}$ represent the boundaries set of manual annotation M and automatic output A, respectively. The lower the HD, the higher the proximity between ground truth and automatic segmentation. The HD is defined as:

$$HD(M, A) = \max(h_{95}(M, A), h_{95}(A, M)). \tag{5}$$

The third pattern is the absolute volume difference (AVD), it is defined as:

$$AVD(M, A) = \frac{|V_M - V_A|}{|V_M|} \times 100\%, \tag{6}$$

where V_M and V_A represent the volume of the segmentation result and the reference standard, respectively. Similar to $HD(M, A)$, a smaller value of $AVD(M, A)$ indicates a better segmentation performance.

TABLE 1. Segmentation results on the MRBrainS13 test dataset. Mean DC, HD and AVD split by CSF, GM and WM tissues as well as rank and the final scores are shown. The best value of each score is depicted in bold.

Rank	Methods	CSF			GM			WM			Scores
		DSC	HD	AVD	DSC	HD	AVD	DSC	HD	AVD	
3	Proposed	82.77	2.41	**6.20**	86.00	1.45	5.42	**89.74**	1.82	**5.20**	46
2	MMAN	**84.86**	**2.03**	6.75	**86.40**	1.38	5.72	89.70	1.88	6.28	**38**
6	HDN[50]	83.42	2.26	7.31	86.32	**1.34**	6.19	89.46	**1.78**	6.03	63
7	VoxNet[51]	84.25	2.19	7.69	86.15	1.45	6.60	89.46	1.94	6.05	73
45	SJCE	67.67	4.78	14.99	80.76	3.71	**5.14**	86.83	4.29	6.43	299

Overall, the committee calculates a final score for each component (CSF, GM and WM) and each evaluation measure (DSC, HD, and AVD). Specific details of the assessment can be found in the official challenge website[1]. (http://mrbrains13.isi.uu.nl/details.php)

EXPERIMENTAL DETAILS

Similar to [46] and [47], we rotate 180 degrees along the axial plane first, and flip the rotated and original volumes horizontally. The 3-channel convolutional neural networks are constructed from the provided T1, T1-IR and T2_FLAIR modality. In this experiment, the cerebellum and brainstem are often misclassified to GM and WM due to the limited training samples. Hence, two different pipelines are trained for this particular experiment:

(i) One trained with four-categories labels by attributing the cerebellum and brainstem to the background.
(ii) One trained with five-categories labels by combining cerebellum and brainstem into an additional category.

In the second stage, the boundary voxel of GM is down-sampled to balance the training data, generating a boundary dataset of 2,076,098 patches. And the equilibrium coefficient C_k in eq.(2) is set to 1. The number of epoch is set to 500 in coarse segmentation and 10 in boundary correction.

RESULTS

Table 1 presents the mean DSC, HD, AVD in CSF, GM, WM and overall scores of the proposed pipelines. In this paper, we compare the proposed method with other 4 out of 50 other participants who upload their segmentation results. As of today, our strategy ranks in the third position of the challenge, outperforming the most of participant methods.

So far, the proposed pipeline ranks second place in terms of WM segmentation while ranking third in GM and eleventh in CSF. Table 1 lists the performance of some representative team and its ranking when we submit the results. As shown in Table 1, our proposed pipeline depicts a lower AVD score whether in CSF, GM or WM. However, other approaches obtain better performance in HD such as the similar work proposed by MMAN and HDN. For method SJCE, it performs poor in DC and HD although obtained lowest AVD in GM.

A complete ranking of all the participant can be referred to the challenge website[2]. (http://mrbrains13.isi.uu.nl/results.php)

B. THE INTERNET BRAIN SEGMENTATION REPOSITORY

DATA

The center for morphometric analysis at Massachusetts General Hospital provides the 1.5T MR brain data sets which manual segmentation were provided in the website[3] (http://www.cma.mgh.harvard.edu/ibsr/). A limitation of the IBSR database is that sulcal cerebrospinal fluid voxel are considered as GM. For convenience, we distinguish them as SCFS and SGM.

- IBSR-18 dataset: It is composed of 18 T1-weighted MR images (Age: 38.5±23.3, 14 males and 4 females) from healthy crowds with manual-guided expert segmentation. The voxel size of each scan was $0.94 \times 1.5 \times 0.94$mm^3. The preprocessing steps of MR brain images, such as skull-stripping, image registration, have been performed [48].
- IBSR-20 dataset: It contains 20 subjects (Age: 29.1±4.81, 10 males and 10 females) with T1-weighted MR Image data. All data were normalized and then resliced into 1.0 mm axial, and 1.0 mm sagittal, 3.0 mm coronal scans.

EVALUATION

For comparison with other methods, we just apply the DSC as the evaluation criteria since it is the most common metric.

EXPERIMENT DETAILS

For each database, only T1 image is delivered to the cascade model. For IBSR-18, the CSF with and without sulcal are available. Accordingly, we construct one coarse segmentation and two specific boundary correction models, achieving the transfer learning from SCSF to CSF. For IBSR-20, only labels with SCSF are provided, which negatively affects some methods.

In the IBSR dataset, the category imbalance is still serious although the up- and down-sampling is used. Therefore, the corresponding equilibrium coefficients C_k in eq.(2) is set for each model, as detailed in Table 2. The number of the epoch is set to 100 in coarse segmentation and 10 in boundary correction.
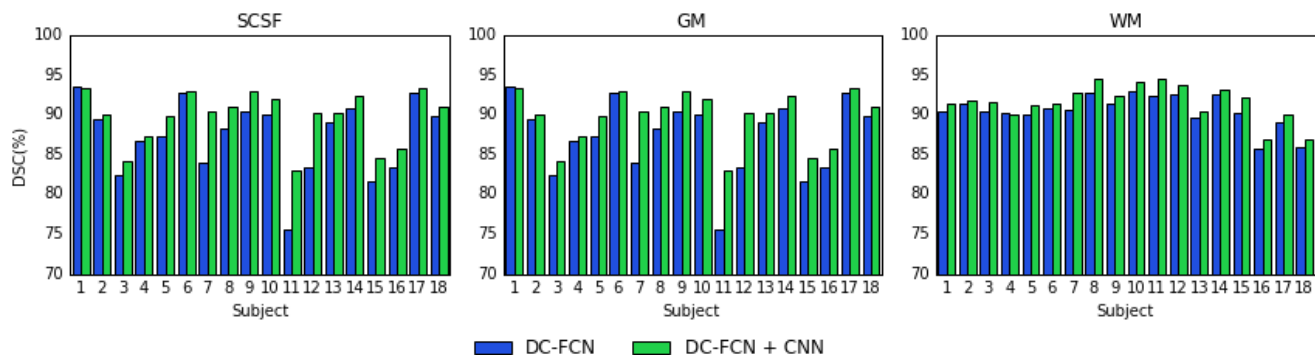
FIGURE 4. DSC of the SCSF, GM and WM categories, measured in each subject in IBSR-18, without (blue) and with (green) BoC.
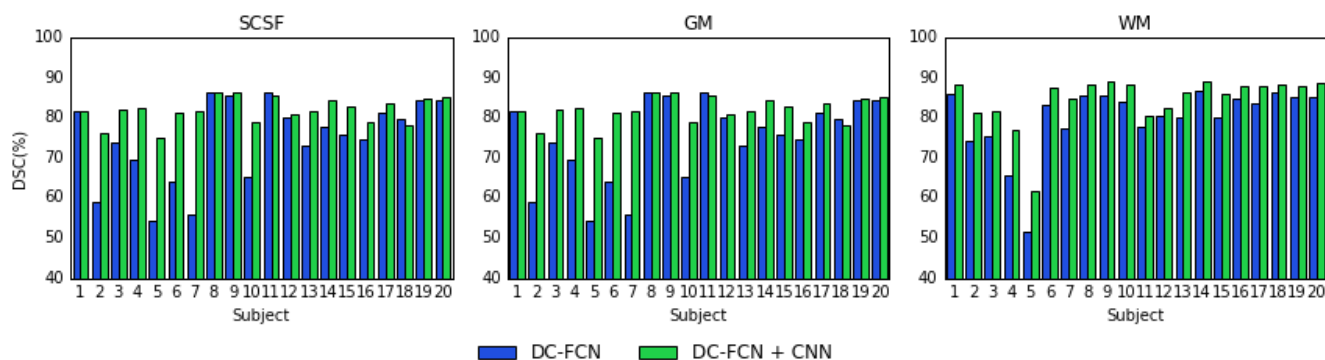

FIGURE 5. DSC of the SCSF, GM and WM categories, measured in each subject in IBSR-20, without (blue) and with (green) BoC.

TABLE 2. Equilibrium coefficients for each boundary correction model on the IBSR dataset.

Datasets	C_0	C_1	C_2	C_3
IBSR-18(CSF)	0.3	0.3	0.1	0.3
IBSR-18(SCSF)	0.1	0.4	0.25	0.25
IBSR-20(SCSF)	0.2	0.4	0.2	0.2

TABLE 3. DSC results (mean±standard deviation) obtained on the IBSR-18 dataset with SCSF for each method.

Methods	SCSF	GM	WM
DC-FCN	87.2±4.6	91.8±1.1	90.5±2.0
DC-FCN+6-CNN	**89.7±3.2**	**93.1±1.0**	**91.6±2.2**
FGFMM[54]	——	88.3±2.4	90.9±2.6
AoC[55]	75.6±4.2	92.4±2.1	90.9±2.0

TABLE 4. Results obtained on the IBSR-20 dataset with SCSF for the DSC (meanstandard deviation).

Methods	SCSF	GM	WM
DC-FCN	74.6±9.9	86.3±5.9	79.9±8.3
DC-FCN+6-CNN	**81.9±3.1**	90.1±4.1	**84.6±6.2**
SVPASEG[56]	76.0±9.0	88.0±4.0	81.0±7.0
AoC[55]	79.5±4.2	**90.9±2.7**	84.4±4.7

RESULTS

Figure 4 depicts a quantitative evaluation of our pipeline for each subject on the IBSR-18 database. Relying on obtained results, it can be observed that the performance of tissue segmentation can be improved by 1∼2% approximately after boundary correction. Meanwhile, the mean DSC scores for all evaluated methods are presented in Table 3. As seen in Table 3, our proposed approach clearly outperforms the rest of available tissue segmentation methods.

Table 4 illustrates the mean DSC value obtained on the IBSR-20 database. It is noticeable that 5∼6% improvement approximately in DSC after boundary correction. Meanwhile, Figure 5 reveals the effect on the boundary correction by characterizing the performance of each evaluated sample in the first and second stage. This phenomenon also illustrates the effectiveness in improving the robustness of tissue segmentation of the boundary correction because it compensates

for the uneven performance in the coarse segmentation stage. Overall, the cascade framework performs more stable after boundary correction.

Comparing with the state-of-the-art approaches, we present the results in Table 3 and Table 4. Atlas of classifiers (AoC) is an informative approach by allowing a statistical summary of the annotated datasets. AoC achieved

higher accuracy than coarse segmentation model DC-FCN and even surpass DC-FCN+6-CNN on GM and WM of IBSR-20 dataset. However, the proposed convolutional neural cascade model realizes a smaller variance and is far superior to other methods in SCSF segmentation. From the aforementioned analysis, the proposed framework performs better than the other evaluated methods.

IV. DISCUSSION

Clinically, the segmentation of brain tissue is of great significance in the diagnosis of nervous system disease. In this paper, a deep learning computational architecture is proposed to segment brain tissue with application to 1.5T and 3T MR images. The proposed pipeline is made up of two CNNs in a cascade manner, where the first network (DC-FCN) segment the brain tissue coarsely while the second network (6-CNN) iteratively correct the boundary voxel which is produced by the previous stage. Although the methods based on CNNs have been applied in brain MRI widely [52]–[54], the idea of boundary correction has not been mentioned yet. From our perspectives, the proposed coarse-to-fine segmentation framework is an interesting contribution to the present study. It is consistent with the habit of manual segmentation and increases the segmentation accuracy significantly. Experimental results show that the proposed framework achieves high-precision segmentation of brain tissue, ranking in the third place in the MRBrainS13 challenge currently. Besides, additional experiments with the IBSR database confirm the effectiveness of the proposed framework as well. Overall, the accuracy of brain tissue segmentation is not only analyzed in patients but also verified in the normal human.

In the proposed coarse segmentation network, DC-FCN, each dense block is an iterative cascade of previous features. It can be viewed as an extension of residual network, performing the iterative concatenate of the previous feature map. There are some appealing changes in application: (1) parameter efficiency, a smaller growth rate makes the network narrower and parameters reduced significantly, which effectively decreases its compute and suppresses over-fitting. (2) feature reuse, all layers can access their previous layers easily so that previously calculated feature map information can be reused easily. (3) feature transfer, enhanced feature transfer facilitates the deep-supervised training. The cross-layer connectivity and multi-scale monitoring of DC-FCN make it ideal for integral tissue segmentation.

Although DC-FCN has achieved brain tissue segmentation, the processing of boundary is not elaborate enough, resulting in blur and smooth in tissue boundaries and failing to meet the clinical application requirements. The introduction of cascade boundary correction can solve the mentioned issue well. It starts from details and realizes the segmentation correction, taking the voxel as focus and extracting characteristics around the voxel. In this article, a 6-layers convolutional neural network is designed to avoid over-fitting caused by lower variation in structural MRI scans. Adequate experiments [58] have proved that the smaller network can meet our demands
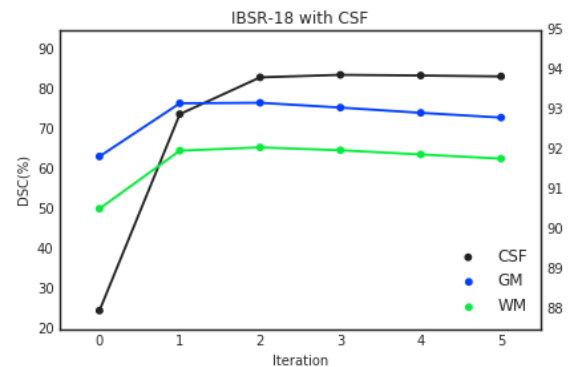
FIGURE 6. Mean DSC with respect to the number of iteration used for migrating from SCSF to CSF on IBSR-18 dataset. For observation, the value of CSF corresponds to the left scale while GM and WM correspond to the right.

and decrease the time it takes to organize the segmentation while ensuring accuracy.

Nevertheless, different diseases have various criteria for organizational segmentation, so the migration among models is particularly important. In the cascade algorithm proposed in this paper, this idea can be achieved by constructing a major coarse segmentation network and two boundary correction networks. For IBSR-18, the CSF with and without sulcal are available. Accordingly, one coarse segmentation and two specific boundary correction models are constructed, which realizes a transfer learning from SCSF to CSF. In the migration study, Figure 6 depicts the variation of the DSC with the number of iteration for each tissue. Judging from the transition from SCSF to CSF, CSF is the most affected tissue by the boundary model. And the accuracy of GM and WM tissue show a slight downward trend after the first iteration. Overall, the migration model has the comprehensive performance when the number of iteration is 3.

With the design of coarse-to-fine cascade pipeline, we attribute to keep two goals in mind: accuracy and efficiency. For automated tissue segmentation systems for clinical practice, both of them have equal significant impact on clinical application. Figure 7 and Figure 8 illustrate the output in each segmentation stage of samples from the MRBrainS13 and IBSR dataset, respectively. Comparing the results before and after the boundary correction, there is a significant improvement on the segmentation accuracy. However, the trade-off between segmentation and correction is an important factor influencing the performance. In Figure 8, red circles represent the example of under-correction, and black circles indicate the case of over-correction. After many trials, the comprehensive performance is the best when the boundary is corrected three times.

The proposed automatic tissue segmentation architecture has great significance in clinical practice. It has been proven that the morphology of GM is associated with many central nervous system diseases. For example, epilepsy. The blurring boundary between GM and WM is an important biomarker of focal cortical dysplasia (FCD) II [57], which contributes
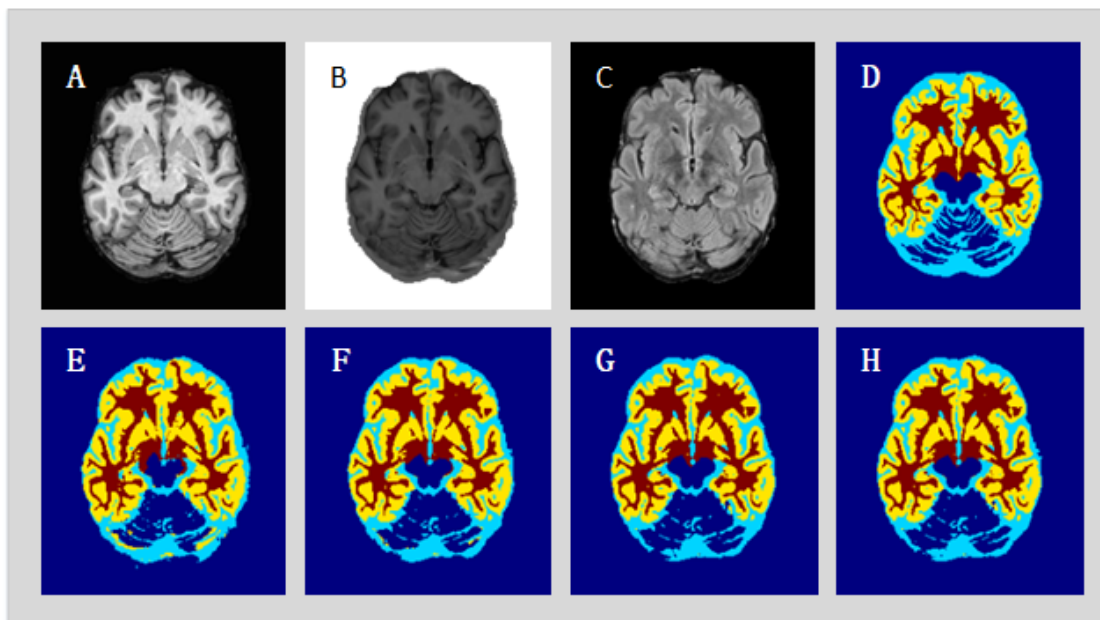
FIGURE 7. Tissue segmentation results on the third training subject of MRBrainS13 dataset (viewed in axial planes). (A) T1 image. (B) T1_IR image. (C) T2_FLAIR image. The tissue segmentation results of manual annotation(D), coarse segmentation(E) and the 1-3th iteration boundary correction(F-H) are visualized with green(CSF), yellow(GM), red(WM).
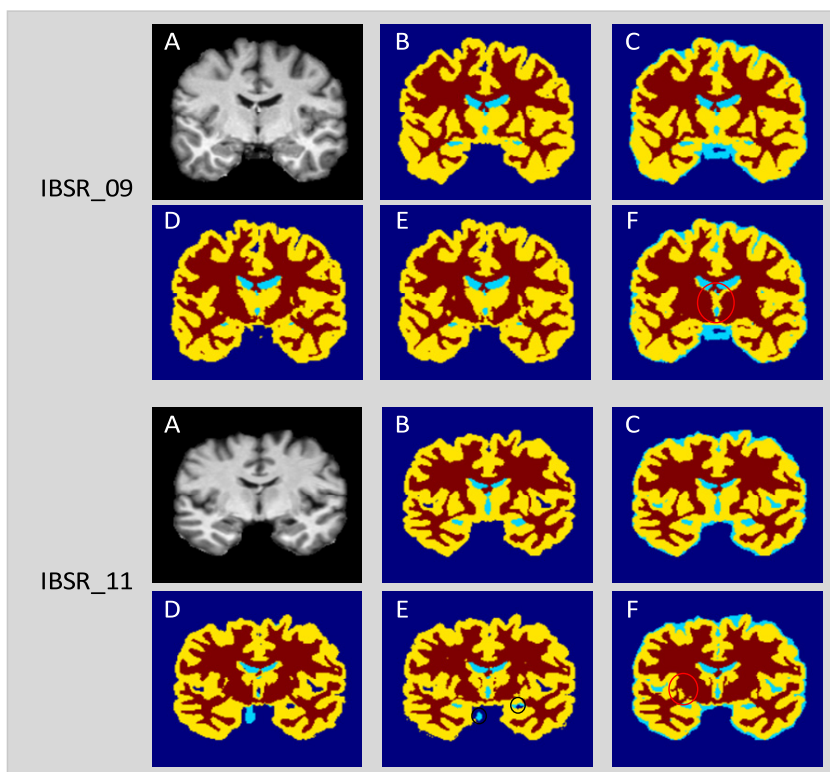
FIGURE 8. Segmentation examples on the best(upper) and worst(bottom) of IBSR-18 dataset. (A) T1 image. Manual annotation of SCSF(B) and CSF(C), coarse segmentation(D), corrected results corresponding to SCSF(E) and CSF(F) are displayed in coronal view. Under-correction area is circled in black and over-correction in red.

to the pathological diagnosis of the disease. In addition, the symmetry and atrophy of tissue are the references for locating the epileptogenic focus. In the future, we will delve into the specific clinical relationship between brain tissue and nervous system diseases, and assist in disease diagnosis and preoperative evaluation.

V. CONCLUSION

Automated tissue segmentation is still an issue of great concern because it is important for assessing the progress of diseases. To relieve the pressure of computation under the premise of accuracy, a two-stage cascade architecture is proposed to segment tissues automatically from volumetric medical data with application to 1.5T and 3T MR scans. The first stage DC-FCN classifies the imaged tissue roughly and the second stage performs detailed correction according to the boundaries produced in the previous phase. Competitive performance is obtained on MRBrainS13 online evaluation as well as IBSR self-evaluation. This is an implication that the proposed framework is applicable to both normal and diseased subjects, that is, it generalizes well.

REFERENCES

[1] M. Behroozi and M. R. Daliri, "Software tools for the analysis of functional magnetic resonance imaging," *Basic Clin. Neurosci.*, vol. 3, no. 5, pp. 71–83, 2012.

[2] M. Behroozi, M. R. Daliri, and H. Boyaci, "Statistical analysis methods for the fMRI data," *Basic Clin. Neurosci.*, vol. 2, no. 4, pp. 67–74, 2011.

[3] W. D. Penny, K. J. Friston, J. T. Ashburner, S. J. Kiebel, and T. E. Nichols, Eds., *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Amsterdam, The Netherlands: Elsevier, 2011.

[4] S. Yousefi, R. Azmi, and M. Zahedi, "Brain tissue segmentation in MR images based on a hybrid of MRF and social algorithms," *Med. Image Anal.*, vol. 16, no. 4, pp. 840–848, 2012.

[5] G.-C. Lin, W.-J. Wang, C.-C. Kang, and C.-M. Wang, "Multispectral MR images segmentation based on fuzzy knowledge and modified seeded region growing," *Magn. Reson. Imag.*, vol. 30, no. 2, pp. 230–246, 2012.

[6] M. R. Daliri, "Automated diagnosis of Alzheimer disease using the scale-invariant feature transforms in magnetic resonance images," *J. Med. Syst.*, vol. 36, no. 2, pp. 995–1000, 2012.

[7] A. Ciarmiello *et al.*, "Brain white-matter volume loss and glucose hypometabolism precede the clinical symptoms of Huntington's disease," *J. Nucl. Med.*, vol. 47, no. 2, pp. 215–222, 2006.

[8] E. Fisher, J.-C. Lee, K. Nakamura, and R. A. Rudick, "Gray matter atrophy in multiple sclerosis: A longitudinal study," *Ann. Neurol., Off. J. Amer. Neurol. Assoc. Child Neurol. Soc.*, vol. 64, no. 3, pp. 255–265, 2008.

[9] H. G. Schnack *et al.*, "Can structural MRI aid in clinical classification? A machine learning study in two independent samples of patients with schizophrenia, bipolar disorder and healthy subjects," *Neuroimage*, vol. 84, pp. 299–306, Jan. 2014.

[10] S. González-Villà, A. Oliver, S. Valverde, L. Wang, R. Zwiggelaar, and X. Lladó, "A review on brain structures segmentation in magnetic resonance imaging," *Artif. Intell. Med.*, vol. 73, pp. 45–69, Oct. 2016.

[11] M. Cabezas, A. Oliver, X. Lladó, J. Freixenet, and M. B. Cuadra, "A review of atlas-based segmentation for magnetic resonance brain images," *Comput. Methods Programs Biomed.*, vol. 104, no. 3, pp. e158–e177, 2011.

[12] N. Shiee, P.-L. Bazin, A. Ozturk, D. S. Reich, P. A. Calabresi, and D. L. Pham, "A topology-preserving approach to the segmentation of brain images with multiple sclerosis lesions," *NeuroImage*, vol. 49, no. 2, pp. 1524–1535, 2010.

[13] A. Ahmadvand and M. Daliri, "Brain MR image segmentation methods and applications," *OMICS J Radiol*, vol. 3, no. 4, p. e130, 2014.

[14] H. Wang and P. Yushkevich, "Multi-atlas segmentation with joint label fusion and corrective learning—An open source implementation," *Frontiers Neuroinform.*, vol. 7, p. 27, Nov. 2013.

[15] F. Rousseau, P. A. Habas, and C. Studholme, "A supervised patch-based approach for human brain labeling," *IEEE Trans. Med. Imag.*, vol. 30, no. 10, pp. 1852–1862, Oct. 2011.

[16] Z. Shahvaran, K. Kazemi, M. S. Helfroush, N. Jafarian, and N. Noorizadeh, "Variational level set combined with Markov random field modeling for simultaneous intensity non-uniformity correction and segmentation of MR images," *J. Neurosci. Methods*, vol. 209, no. 2, pp. 280–289, 2012.

[17] W. Bai *et al.*, "A probabilistic patch-based label fusion model for multi-atlas segmentation with registration refinement: Application to cardiac MR images," *IEEE Trans. Med. Imag.*, vol. 32, no. 7, pp. 1302–1315, Jul. 2013.

[18] P. Moeskops *et al.*, "Automatic segmentation of MR brain images of preterm infants using supervised classification," *NeuroImage*, vol. 118, pp. 628–641, Sep. 2015.

[19] A. W. C. Liew and H. Yan, "Current methods in the automatic tissue segmentation of 3D magnetic resonance brain images," *Current Med. Imag. Rev.*, vol. 2, no. 1, pp. 91–103, 2006.

[20] Y. Zhang, M. Brady, and S. Smith, "Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm," *IEEE Trans. Med. Imag.*, vol. 20, no. 1, pp. 45–57, Jan. 2001.

[21] S. Roy, A. Carass, P.-L. Bazin, S. Resnick, and J. L. Prince, "Consistent segmentation using a Rician classifier," *Med. Image Anal.*, vol. 16, no. 2, pp. 524–535, 2012.

[22] M. Yaqub, M. K. Javaid, C. Cooper, and J. A. Noble, "Investigation of the role of feature selection and weighted voting in random forests for 3-D, volumetric segmentation," *IEEE Trans. Med. Imag.*, vol. 33, no. 2, pp. 258–271, Feb. 2014.

[23] 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.

[24] Y.-T. Liu, H.-X. Zhang, and P.-H. Li, "Research on SVM-based MRI image segmentation," *J. China Universities Posts Telecommun.*, vol. 18, pp. 129–132, Dec. 2011.

[25] S. Damangir *et al.*, "Multispectral MRI segmentation of age related white matter changes using a cascade of support vector machines," *J. Neurol. Sci.*, vol. 322, nos. 1–2, pp. 211–216, 2012.

[26] T. Wu, M. H. Bae, M. Zhang, R. Pan, and A. Badea, "A prior feature SVM-MRF based method for mouse brain segmentation," *Neuroimage*, vol. 59, no. 3, pp. 2298–2306, Feb. 2012.

[27] S. Pereira, A. Pinto, J. Oliveira, A. M. Mendrik, J. H. Correia, and C. A. Silva, "Automatic brain tissue segmentation in MR images using random forests and conditional random fields," *J. Neurosci. Methods*, vol. 270, pp. 111–123, Sep. 2016.

[28] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, Dec. 2012, pp. 1097–1105.

[29] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017.

[30] P. Moeskops, M. A. Viergever, A. M. Mendrik, L. S. de Vries, M. J. N. L. Benders, and I. Išgum, "Automatic segmentation of MR brain images with a convolutional neural network," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1252–1261, May 2016.

[31] D. Nie, L. Wang, Y. Gao, and D. Shen, "Fully convolutional networks for multi-modality isointense infant brain image segmentation," in *Proc. IEEE 13th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2016, pp. 1342–1345.

[32] T. Brosch, L. Y. W. Tang, Y. Yoo, D. K. B. Li, A. Traboulsee, and R. Tam, "Deep 3D convolutional encoder networks with shortcuts for multiscale feature integration applied to multiple sclerosis lesion segmentation," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1229–1239, May 2016.

[33] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. CVPR*, vol. 1, no. 2, Jul. 2017, p. 3.

[34] Z. Akkus, A. Galimzianova, A. Hoogi, D. L. Rubin, and B. J. Erickson, "Deep learning for brain MRI segmentation: State of the art and future directions," *J. Digit. Imag.*, vol. 30, no. 4, pp. 449–459, 2017.

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

[36] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 770–778.

[37] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, Oct. 2015, pp. 234–241.

[38] S. Ji, W. Xu, M. Yang, and K. Yu, "3D convolutional neural networks for human action recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 221–231, Jan. 2013.

[39] A. Dosovitskiy, P. Fischer, J. T. Springenberg, M. Riedmiller, and T. Brox, "Discriminative unsupervised feature learning with exemplar convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 766–772.

[40] D. P. Kingma and J. Ba. (2014). "Adam: A method for stochastic optimization." [Online]. Available: https://arxiv.org/abs/1412.6980

[41] A. M. Mendrik *et al.*, "MRBrainS challenge: Online evaluation framework for brain image segmentation in 3T MRI scans," *Comput. Intell. Neurosci.*, vol. 2015, Aug. 2015, Art. no. 813696.

[42] S. Klein, M. Staring, K. Murphy, M. A. Viergever, and J. P. Pluim, "Elastix: A toolbox for intensity-based medical image registration," *IEEE Trans. Med. Imag.*, vol. 29, no. 1, pp. 196–205, Jan. 2010.

[43] W. Penny, K. Friston, J. Ashburner, S. Kiebel, and T. Nichols, Eds., *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Amsterdam, The Netherlands: Elsevier, 2011.

[44] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.

[45] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge, "Comparing images using the Hausdorff distance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 9, pp. 850–863, Sep. 1993.

[46] M. Havaei *et al.*, "Brain tumor segmentation with deep neural networks," *Med. Image Anal.*, vol. 35, pp. 18–31, Jan. 2017.

[47] K. Kamnitsas *et al.*, "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation," *Med. Image Anal.*, vol. 36, pp. 61–78, Feb. 2017.

[48] T. Rohlfing, "Image similarity and tissue overlaps as surrogates for image registration accuracy: Widely used but unreliable," *IEEE Trans. Med. Imag.*, vol. 31, no. 2, pp. 153–163, Feb. 2012.

[49] P. A. Filipek, C. Richelme, D. N. Kennedy, and V. S. Caviness, Jr., "The young adult human brain: An MRI-based morphometric analysis," *Cerebral Cortex*, vol. 4, no. 4, pp. 344–360, 1994.

[50] J. Dolz, K. Gopinath, J. Yuan, H. Lombaert, C. Desrosiers, and I. B. Ayed. (2018). "HyperDense-Net: A hyper-densely connected CNN for multi-modal image segmentation." [Online]. Available: https://arxiv.org/abs/1804.02967

[51] H. Chen, Q. Dou, L. Yu, J. Qin, and P.-A. Heng, "VoxResNet: Deep voxelwise residual networks for brain segmentation from 3D MR images," *NeuroImage*, vol. 170, pp. 446–455, Apr. 2018.

[52] P. Moeskops, M. A. Viergever, A. M. Mendrik, L. S. de Vries, M. J. N. L. Benders, and I. Išgum, "Automatic segmentation of MR brain images with a convolutional neural network," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1252–1261, May 2016.

[53] S. Bao and A. C. S. Chung, "Multi-scale structured CNN with label consistency for brain MR image segmentation," *Comput. Methods Biomech. Biomed. Eng., Imag. Vis.*, vol. 6, no. 1, pp. 113–117, 2018.

[54] M. Baghdadi, N. Benamrane, and L. Sais, "Fuzzy generalized fast marching method for 3D segmentation of brain structures," *Int. J. Imag. Syst. Technol.*, vol. 27, no. 3, pp. 281–306, 2017.

[55] B. Kodner, S. Gordon, J. Goldberger, and T. R. Raviv, "Atlas of classifiers for brain MRI segmentation," in *Proc. Int. Workshop Mach. Learn. Med. Imag.* Cham, Switzerland: Springer, Sep. 2017, pp. 36–44.

[56] J. Tohka, I. D. Dinov, D. W. Shattuck, and A. W. Toga, "Brain MRI tissue classification based on local Markov random fields," *Magn. Reson. Imag.*, vol. 28, no. 4, pp. 557–573, 2010.

[57] S. J. Hong *et al.*, "Multimodal MRI profiling of focal cortical dysplasia type II," *Neurology*, vol. 88, pp. 734–742, 2017.

[58] S. Valverde *et al.*, "Improving automated multiple sclerosis lesion segmentation with a cascaded 3D convolutional neural network approach," *NeuroImage*, vol. 155, pp. 159–168, Jul. 2017.

GUIXIA KANG received the M.S. degree from Tianjin University, Tianjin, China, and the Ph.D. degree in electrical engineering from the Beijing University of Posts and Telecommunications (BUPT), Beijing. She is currently a Professor at BUPT and the Director of Beijing International S&T Cooperation Base of smart medicine. She has expertise in the physical layer of 5G wireless systems and in the wireless e-Health systems.

From 2002 to 2004, she was a Research Scientist with the Future Radio Concept Department of Siemens, Munich, Germany. She is/was the Project Manager of several national projects, such as the Important National Science and Technology Specific Project, the National 863 project, the National Natural Science Foundation of China, and several international cooperation projects. She has authored one English book (Shaker Verlag, Germany), three Chinese books, and authored or co-authored over 100 journal and conference papers.

Dr. Kang was a recipient of "New Century Talent of Ministry of Education" and "Beijing New Star of Science and Technology." She was also a recipient of the First Prize of Science and Technology Award of the China Communications Association and the Second Prize of the Beijing Science and Technology Progress Award.

NINGBO ZHANG received the Ph.D. degree from the Beijing University of Posts and Telecommunications (BUPT) in 2010. He currently is an Associate Professor with BUPT. He has expertise in the physical layer of 5G wireless systems and the machine to machine communications in IoT.

From 2010 to 2014, he was a Senior Engineer with the Research and Development Wireless Department, Huawei Technologies. Since 2014, he has been the Project Manager of several national projects, such as the National Natural Science Foundation of China, the National Science and Technology Major Project of China, and the National 863 project. He has authored or co-authored over 40 journal and conference papers.

BEIBEI HOU was born in Henan, China, in 1992. She received the B.S. degree from Henan Normal University, Henan, in 2015. She is currently pursuing the M.S. and Ph.D. degrees with the Information and Communication Engineering College, Beijing University of Posts and Telecommunications.

Her research interests include medical image analysis and pattern recognition in brain science.

CHUAN HU was born in Hunan, China. He received the B.Eng. degree from the College of Information Engineering, Xiangtan University, Xiangtan, China, in 2018. He is currently pursuing the M.S. degree with the Beijing University of Posts and Telecommunications.

His research interests include non-linear circuit, biomedical signal processing, image analysis, and deep learning.

• • •