3D-brain segmentation using deep neural network and Gaussian mixture model

Duy M. H. Nguyen, Huy T. Vu, Huy Q. Ung, Binh T. Nguyen*
Department of Computer Science
University of Science, Vietnam
nhmduy.hcmus@gmail.com
vuthanhhuy.hcmus@gmail.comn
ungquanghuy.hcmus@gmail.com
ngtbinh@hcmus.edu.vn

Abstract

Automatic segmentation of major brain tissues from high-resolution magnetic resonance images (MRIs) plays an important role in clinical diagnostics and neuroscience research. In this paper, we present a novel approach to extract brain tissues including gray matter, white matter and cerebrospinal fluid by using Gaussian mixture models (GMMs), Convolution neural networks (CNNs) and Deep neural networks (DNNs). GMMs are applied to classify voxels which have distinct intensity information and are easy to recognize while DNNs and CNNs are treating voxels which are similar in appearance and usually recognized insufficiently by traditional approaches. The empirical results on IBSR 18 *dataset show that the proposed method outperforms* 13 *state-of-the-art algorithms, surpassing all the other methods by a significant margin.*

1. Introduction

Automatically segmenting important brain tissues including gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF) from high-quality MRIs has played a crucial role in clinical diagnostics and neuroscience research for helping to assess many diseases, such as the Alzheimer's disease, brain tumor, Parkinson's disease, hydrocephalus, and schizophrenia [18][20][37]. Using magnetic resonance images, one can detailedly visualize anatomical structures of human brain and deeply understand significant functions in human brain.

In the last two decades, there exist a lot of brain image segmentation methods e.g. supervoxel based analysis [9][21][26], parametric or non-parametric methods [5][16][32][45], and atlas based approaches [2][38]. Wells and co-workers [44] propose an adaptive segmentation technique by extracting the information of issue intensity properties and intensity inhomogeneities from MRIs and using the expectation maximization (EM) algorithm. Not only giving more accurate segmentation for each tissue class, the approach also gives a better visualization for reconstructing a three-dimensional structure of human brain. Song and colleagues describe a weighted probabilistic neural network for estimating the probability density function of pixels from brain MRIs [25][35]. They assume the voxel intensities in each target region have a Gaussian distribution, and then use a GMM [4] for segmenting brain tissues from MRIs. GMM parameters can be estimated by maximizing the likelihood of observed images via the EM algorithm [25]. Ji et al. [16] present a fuzzy local Gaussian mixture model (FLGMM) for automatic segmentation of brain tissues from MRIs with an assumption that the local image data within the neighborhood of each voxel satisfy GMMs. Although these methods may produce high accuracies, there are still two major disadvantages preventing them from achieving better results [21]. The first challenge is the heterogeneity of voxels in each tissue region, which is caused by the inconstancy of magnetic field and susceptibility during MRI scans. This effect significantly increases the intraclass variability and decreases the interclass distance. The second challenge comes from the difficulty in choosing meaningful features. Using mistakenly features which are redundant and noisy may considerably decrease the performance.

In recent years, CNNs have obtained lots of success in pattern recognition and machine learning [23]. Unlike classical machine learning techniques which require plenty of works to find appropriate features for classification, CNNs do not need to create a set of hand-crafted features. Taking these advantages, there are various studies which have been developed by using CNNs in the field of medical image analysis [6][7][8][28], especially in tasks related to brain image segmentation. Zhang and colleagues[46] propose an efficient approach by way of using CNNs for segmenting three major brain tissues (GM, WM, and CSF) in MR

brain images of very young infants. De Brebisson et al.[13] present a deep neural network for brain segmentation in adult MR brain images from the MICCAI 2012 challenge on multi-atlas labelling. However, due to different research purposes, these studies are not implemented and compared on public datasets such as the Internet Brain Segmentation Repository (IBSR) dataset.

In this paper, we present an efficient approach for brain tissue segmentation and compare the performance with 13 states-of-the-art methods on a public dataset - IBSR18 [33], which has been widely used in the recent studies [43][21]. The proposed method is constructed by applying both GMM and deep neural networks. The contribution of our approach are three folds. First, we present an automatic algorithm under a general framework, which can simultaneously exploit the power of traditional approaches and modern techniques based on deep learning. Instead of treating the segmentation of brain tissues in the whole image, our deep network is designed to investigate at positions which usually are mistakenly recognized. Hence, it helps to increase the capability of feature representation for complex voxels. Second, since our approach does not make any assumption of the structure and the shape information of brain tissues, it can be easily applied to different medical segmentation tasks by replacing the GMM with another traditional method performing well in the corresponding problem. Finally, beside leveraging CNNs, which can learn high-level feature representations, we combine the global spatial information with feature maps learned from CNNs and then use new features in DNNs. Consequently, the proposed method can simultaneously represent both local and global features for each voxel in 3D structure. The extensively experimental results in IBSR18 dataset show that our proposed method outperforms other approaches.

The rest of this paper is organized as follows. In section 2, we present a proposed method for $3D$ brain segmentation by using GMMs and DNNs. In section 3, we describe our implementation details and show experiment results. The last section gives concluding remarks and several directions for future work.

2. Method

In order to segment a $3D$ brain volume into three classes (GM, WM and CSF), we classify each voxel in the $3D$ structure into one of these brain tissues. The flowchart of a proposed method can be illustrated in Figure 1, which consists of a GMM and two classifiers trained by DNNs. The first deep neural network classifier, called uncertain voxels detector, is applied to scan the whole image for detecting certain and uncertain voxels. In our approach, the term *certain voxel* means the voxels having distinct intensity information and easily recognized while the term *uncertain voxel* indicates the voxels having the similarity appearance and usually recognized insufficiently by traditional approaches. The detected certain voxels by the first DNN are fed into a GMM model, which is easily trained and requires less computational resources. For further discrimination of uncertain voxels, we put them into the second deep neural network classifier. Finally, we combine two predicted results calculated from GMM model and the second DNN classifier to create final segmentation results. The details of each model will be described in next sections.

2.1. Gaussian mixture model for certain voxels

GMM is a parametric probability density function represented as a weighted sum of Gaussian component densities [4]. The parameters of the model can be estimated from training data by using an iterative Expectation-Maximization (EM) algorithm [29]. Several studies have been developed basing on GMM to describe the intensity information for brain tissues and achieving the positive performance [25][40]. One of motivations behind these approaches is the histogram of each brain tissue has a normal distribution shape and furthermore, the peaks of these three histograms are obviously separated. Figure 2 illustrates the intensity information of CSF, GM and WM. It is important to emphasize that one of the most difficult challenges among these methods is histograms of different tissue regions usually overlap each other. Additionally, the intensity range of different samples in training data is unstable. These problems make methods based on GMMs not sufficiently to represent the characteristics of the voxels located in overlapping areas, and then obtain a low segmentation accuracy.

In order to overcome these shortcomings, we elaborate a novel approach for GMM inspired from DNNs to effectively segment brain tissues. Instead of treating the segmentation for whole voxels, our method only focuses on voxels having distinct characteristics and easily detected, which can be obtained from an uncertain voxels detector. We normalize the intensity of voxels before employing GMM model by using Eq. (1):

$$v_{i,\,n}^{\text{nor inten}} = v_{i,\,n}^{\text{ori inten}} \Big/ \left(2 \sum_{j=1}^{N_n} v_{i,\,n}^{\text{ori inten}} \Big/ N_n \right) - 0.5, \quad (1)$$

where $v_{i,n}^{\text{nor inten}}$ and $v_{i,n}^{\text{ori inten}}$ are the normalized and original intensities of the i^{th} voxel of the n^{th} brain; N_n is the number of brain voxels of the n^{th} brain.

The proposed GMM contains two GMM sub-models. The first one is used to cluster on the training data and then applied to the testing data. This sub-model is called *generalizing model* since it learns from all $3D$ training volumes . We combine all voxels from all the training into one dataset

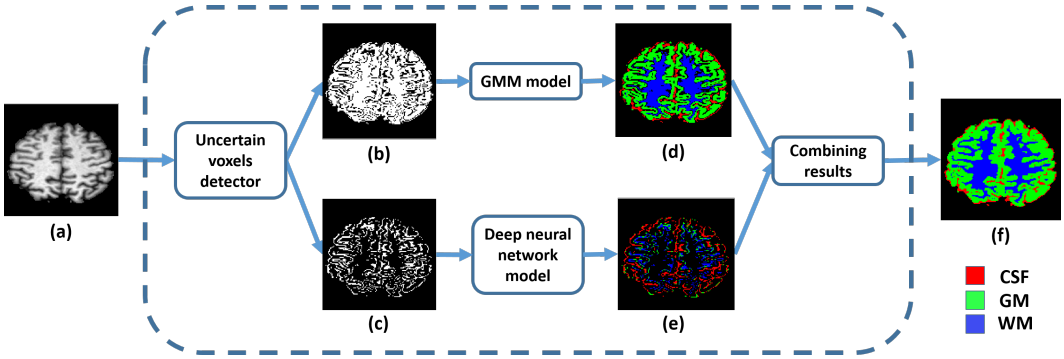


Figure 1. Flow chart of our system. (a) MR 3D scan; (b) Certain voxels region; (c) Uncertain voxels region; (d) Predictions for certain voxels region; (e) Predictions for uncertain voxels region; (f) Combined result.

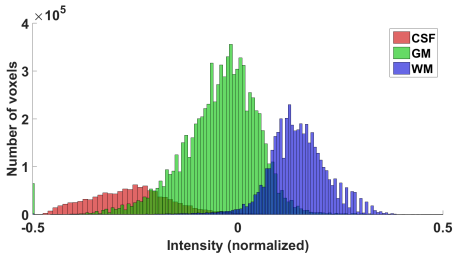


Figure 2. Intensity histograms of each tissue type collected from the training dataset.

and use GMM to model the normalized intensity distribution in this dataset. The likelihood probability of $v_{i,k}^{\text{nor inten}}$, which is normalized intensity feature of the i^{th} voxel from the k^{th} brain, is modeled as a mixture of three Gaussian components:

$$f_{gen}(v_{i,k}^{\text{nor inten}}|\theta) = \sum_{j=1}^{3} \alpha_j f(v_{i,k}^{\text{nor inten}}|\theta_j) \tag{2}$$

for $k = \overline{1, K}, i = \overline{1, N_k}$, where K represents the number of $3D$ training volumes, N_k denotes the number of voxels of the k^{th} $3D$ training volume. Here, $j = 1, 2, 3$ representing three types of brain tissues of interest. Being constrained by the conditions $\alpha_j > 0$, with $j = 1, 2, 3$ and $\sum_{j=1}^{3} \alpha_j = 1$, α_j is the j^{th} mixture coefficient. Parameter $\theta_j = (\mu_j, \sum_j)$ denotes the mean and the variance of the j^{th} Gaussian component function, respectively. We use the expectation maximization (EM) algorithm to estimate the parameters for (2).

The second model is called *the particularizing model* as this GMM is a clustering model applied to each $3D$ testing volume only. In contrast to the generalizing GMM which clusters on all of the $3D$ training volumes, this particularizing model clusters on each $3D$ testing volume and then applies to itself again to compute the posterior probability. The likelihood probability of $v_{i,n}^{\text{nor inten}}$ (the normalized intensity feature of the i^{th} voxel from the n^{th} $3D$ testing volume) can be formulated by:

$$f_{par}(v_{i,n}^{\text{nor inten}}|\eta) = \sum_{j=1}^{3} \beta_j f(v_{i,n}^{\text{nor inten}}|\eta_j) \tag{3}$$

for $i = \overline{1, N_n}$ and $\beta_j, \eta_j (j = 1, 2, 3)$ are parameters needed to be estimated, N_n is the number of voxels of the n^{th} $3D$ testing volume. For each $3D$ testing volume, the EM algorithm is employed to estimate the parameters for (3) to best fit the intensity distribution in the brain.

During testing, for each certain voxel, results from two GMMs (which are posterior probabilities of that voxel belonging to one of the three classes), are averaged to achieve the final probability. After that, the class with the highest posterior probability is assigned to the voxel. Equations (4) - (7) explain this combining step. Figure 1 shows an example of all predictions by GMMs for certain voxels.

$$P_{ge}(v_{i,n}) = \left(P_{ge}^{\text{CSF}}(v_{i,n}), P_{ge}^{\text{GM}}(v_{i,n}), P_{ge}^{\text{WM}}(v_{i,n})\right) \tag{4}$$

$$P_{pa}(v_{i,n}) = \left(P_{pa}^{\text{CSF}}(v_{i,n}), P_{pa}^{\text{GM}}(v_{i,n}), P_{pa}^{\text{WM}}(v_{i,n})\right) \tag{5}$$

$$P(v_{i,n}) = \left(\sum_{k \in \{ge, pa\}} P_k(v_{i,n})\right) \Big/ 2 \tag{6}$$

$$y(v_{i,n}) = \max_{\text{coor}} \left(P(v_{i,n})\right) \tag{7}$$

with $\max_{\text{coor}}(w) = \{i \in \{1, 2, 3\}, w_i = \max\{w_1, w_2, w_3\}\}$

for all $w = (w_1, w_2, w_3) \in \mathbb{R}^3$, $v_{i,n}$ is the i^{th} voxel of the n^{th} $3D$ testing volume, $P_{ge}^{c}(v_{i,n})$ and $P_{pa}^{c}(v_{i,n})$ are the results received from the generalizing, particularizing GMM model respectively, indicating the posterior probability of the i^{th} voxel of the n^{th} the brain belonging to class c, which is including CSF, GM or WM. Finally, $y(v_{i,n})$ is the tissue prediction for the i^{th} voxel of the n^{th} $3D$ testing volume.

This joint process has several advantages. First, because of treating only certain voxels, it can reduce the number of

voxels in overlapping areas; hence, enhancing the capability of the histogram representations for the different tissues. Furthermore, the GMM framework based on two GMM sub-models allows capturing the global characteristics of intensity information in the training data while preserving particular characteristics in each individual brain. From the extensive experiments, we discover that the performance of the GMM when dealing with certain voxels is up to 96%, demonstrating the effective capability of discrimination of the proposed method.

2.2. Uncertain voxels detector

CNNs and DNNs are two important branches of deep learning. Deep learning is currently a new area of machine learning, whose one of main purposes is modeling high-level features by applying a deep graph with multiple processing layers [23]. The major difference between CNNs and DNNs is that the CNNs have layers of convolution and pooling. Based on this specific architecture, layers in CNNs can extract better features (feature maps) from input images. Both CNNs and DNNs have been achieved the state-of-the-art performance on many works related to images, sounds, and texts [14][22][23][27]. In recent studies, a lot of methods have effectively employed CNNs and DNNs for object segmentation in medical imaging such as e.g lung, liver, and brain tumors [10][15][24]. These approaches are formulated in an end-to-end framework, taking input images and directly computing the corresponding probability maps. However, they usually require lots of computational resources. In this paper, we propose a novel framework inspired by CNNs and DNNs to segment the brain tissues from 3D structures in MRIs.

Different from previous studies, the presented method focuses on all voxels that traditional approaches have not performed well. Not only utilizing advantages from current techniques, we use a deep learning approach for improving the performance on brain tissue segmentation from MRIs. Beside extracting the patch images for training CNNs, we also combine them with the global spatial information [25] from $3D$ coordinates of each voxel. All these informations are later represented together through DNNs in order to enhance the capability of the feature representation in both global and local ways. In fact, our proposed framework includes two classifiers. The first classifier called uncertain voxels detector, which will be described in this section. One of its functions is to detect whether a voxel in a brain image is a certain or an uncertain voxel. It is important to emphasize that all certain voxels are fed in to a GMM model while uncertain voxels are used as input data for the second classifier. The second classifier can be considered as a fine discrimination model for further analyzing all characteristics of uncertain voxels and discussed in details in the next section.

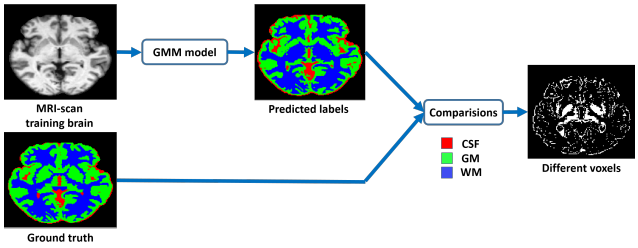


Figure 3. Extracting incorrectly predicted voxel for training the uncertain voxels detector as positive samples.

The inputs of uncertain voxels detector are voxels from a $3D$ MRI and the computed results are the answers: being certain or uncertain voxels. We create the training data for this detector by running the aforementioned trained GMM classifier on the training data and then comparing its predicted values with the ground truth. The incorrect predictions from the GMM classifier are considered as uncertain voxels and all are used as positive samples for training our deep neural network. In principal, incorrectly predicted voxels are usually found at the borders between tissue types as shown in Figure 3. Beside incorrectly predicted voxels, a number of correctly voxels are also chosen randomly and included in the training data as negative samples (certain voxels). The number of positive and negative samples are chosen to be relatively equal. In total, we have approximately 4000000 samples including both negative and positive ones.

For each voxel used in training this deep network, we extract two types of features. The first type includes local features, which are three adjacent image slices of size 11×11 along the x axis, three adjacent image slices of size 11×11 along the y axis, and three adjacent image slices of size 11×11 along the z axis. These features describe the normalized intensity of the predicted voxel itself and its surrounding voxels. The second type of features consists of global features, which are the coordinates of the voxel in the $3D$ volume. These coordinates are normalized by scaling into $[0, 1]$ and then being zero-centered:

$$v_{i,n}^{\text{nor coor}} = \Big(\big(v_{i,n}^{\text{ori coor}} - \alpha\big) \Big/ (\beta - \alpha) \Big) - 0.5 \qquad (8)$$

$$\text{with } \alpha = \min_{j \in \overline{1,N_n}} \{v_{j,n}^{\text{ori coor}}\},\ \beta = \max_{j \in \overline{1,N_n}} \{v_{j,n}^{\text{ori coor}}\}$$

where $v_{i,n}^{\text{nor coor}}$ is the normalized coordinate and $v_{i,n}^{\text{ori coor}}$ is the original coordinate in $3D$of i^{th} voxel of n^{th} brain. The label for each voxel is its ground truth which has value 1 for uncertain voxel and 0 for certain voxel.

Figure 4 illustrates the feature extraction step and the neural network structure used in our deep networks. First, each of the three adjacent slices of MRIs along each axis is propagated to three consecutive convolution and max pooling layers. Depths of these three convolution layers are 13,

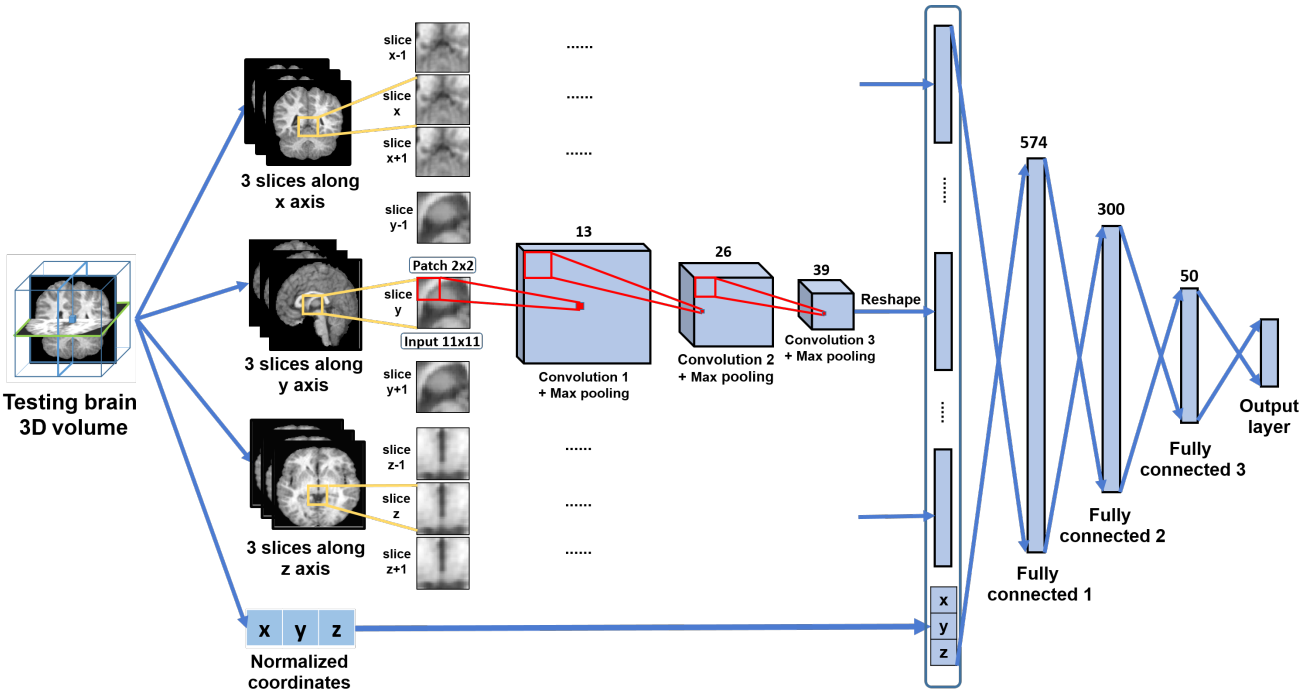


Figure 4. Feature extraction step and the structure of our deep neural networks.

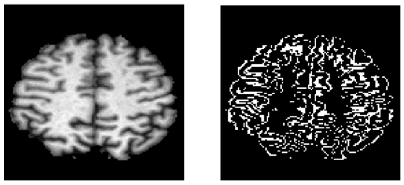
26, and 39 respectively. The receptive field size for each layer is 2×2. Then, the results can be reshaped into a vector. Nine vectors from nine slices of images and the normalized coordinates are concatenated to form an input vector for the DNN. This DNN has three layers with sizes of 574, 300, and 50 respectively. The output layer has the size of two, indicating the probability of the input voxel belonging to the certain or uncertain voxel class.

The training dataset is divided into batches of size 150. Let V denote a batch of training voxels, and W denote the weights of our neural network including both CNNs and DNNs layers. We use the Adam [19], an effective gradient-based optimization algorithm to optimize the objective function:

$$L(V;W) = \left(-\sum_{v_i \in V} \log p(c_i|v_i;W) \right) \Big/ \text{size}(V) + \lambda ||W||^2 \tag{9}$$

The first term is the mean negative-log likelihood loss where c_i is the target class (certain or uncertain voxel) for the voxel v_i in the batch V. The second term is for weight decay regularization where λ controls the importance of the term [11].

This approach has several benefits. First, instead of learning with all voxels in the image, our framework learns at the positions insufficiently recognized by the GMM. Therefore, it can further discriminate for the voxels having complex characteristics and require lower resources for training process than other techniques. Furthermore, since our method does not require any assumption of the shape and the structure information of brain tissues, it can easily be applied to improve the performance of other $3D$ segmentation tasks by replacing the GMM with a traditional approach (which can work well for the corresponding object). In our experiments, about 78% voxels of the testing $3D$ volume are predicted to be certain voxels. The rest 22% uncertain voxels will be classified in the second deep neural.



(a) (b)

Figure 5. (a) Input MRI scan of a testing 3D volume; (b) Uncertain voxels predicted by detector.

Usually, the predicted uncertain region is larger than the actual region. Figure 5b depicts the result predicted by our uncertain voxels detector on a slice of a testing 3D volume.

2.3. Deep neural network classifier for uncertain voxels

With the capability of learning complex patterns, neural network model is expected to give a better prediction

for uncertain voxels than the GMM classifier. Therefore, we employ a deep neural network model to detect uncertain voxels instead of using a GMM.

From that fact that this model is only applied to uncertain voxels, we use uncertain voxels from the training data as training samples. However, as stated above, the predicted uncertain region is usually larger than the actual region, the set of training samples should be enlarged to cover all possible cases. Beside incorrectly predicted voxels by the GMM classifier, we also add into the training dataset: opened boundaries of these voxels (1 pixel opened in four directions: top, left, right, bottom), voxels whose posterior probabilities returned from GMM classifier are lower than 0.7, and all voxels in CSF. By this way, our network will learn more samples to cover all possibly predicted uncertain voxels. Figure 6b shows voxels used as training samples for our deep neural network classifier.

The features of uncertain voxels are extracted by the same way as the training process of the uncertain voxel detector, illustrated in Figure 4. The only one difference is the size of the output layer is three (representing three classes). The label for each voxel is its ground truth tissue type which has value of 1 for CSF, 2 for GM, and 3 for WM. After predicting labels for both certain and uncertain regions using two different classifiers, we combine all results into the whole brain. As these are two individual processes, our proposed method can be run in parallel to increase the performance. Figure 1 illustrates results computed from the two classifiers and their combination.

(a) (b)

Figure 6. (a) Incorrectly predicted voxels by GMM; (b) Extended uncertain region used for training the deep neural network classifier.

3. Experiments and Results

3.1. Dataset and Preprocessing

We evaluate the proposed method on the IBSR 18 dataset, which is popularly used in recent studies, and compare its performance with 13 state-of-the-art methods for brain tissue segmentation task [21][25][43]. This dataset consists of MRIs and $3D$ ground truth volume of 18 brains of size $256 \times 256 \times 128$ with 1.5 mm slide thickness. The ground truth volume is divided into four different structures: non-brain, CSF, GM and WM.

In the experiments, we randomly use 12 brains for training our model and do testing on 6 remaining brains. The skull of each brain image is eliminated by using a binary mask supported from the ground-truth ones in the dataset. We employ the normalization for all the intensity values by using (1) to ensure the similarity of the spatial structure 's information between individual brains.

3.2. Implementation Details

We implement our system by using both MATLAB and Python programs, and especially working with TensorFlow library [1] for deep neural networks.

The network architecture is designed in the same way for two deep neural networks as shown in Figure 4. In the training process, all weights of networks are randomly taken in a normal distribution with mean 0 and standard deviation 0.1. The learning rate is initialized with a value as $1e-04$ and the parameters using for Adam [19] are chosen as $\beta_1 = 0.9$, $\beta_2 = 0.999$ and $\epsilon = 1e-08$. In the experiments, we choose the ReLU function as an activation function. In addition, dropout layers are also handled with a dropout rate 0.5 after each layer in DNN for preventing overfitting problem [36]. Based on the experiment data, the proposed method took about 32 hours for training on a workstation with Intel(R) Core(TM) i7$-$6700K CPU @4.00Ghz and a NVIDIA GeForce GTX 980 GPU. The average time for processing one brain testing with size $256 \times 256 \times 128$ is just 122 seconds, so method is fast enough to be used in clinical practice.

3.3. Qualitative Evaluation

A few distinct typical brain images and the corresponding tissue regions segmented by our method are illustrated in Figure 7. From the ground-truth and input images, we notice that these brain tissue regions have very complex appearances and great variations at different slices, which can make the segmentation task very challenging. Moreover, the intensity values of voxels from separated brain tissues overlaps considerably. For this reason, it may create mistakes on the classification methods based on the intensity information. Leveraging advantages of DNNs and the global spatial information of each voxel, our proposed method can efficiently segment the brain regions even for voxels lying at the borders of tissues. We observe in Figure 7 that the first DNN is capable of accurately detecting uncertain voxels on the whole image. Indeed, most of them are located at the border of brain tissues. Furthermore, the second dDNN also classifies correctly these uncertain voxels into three groups: CSF, GM and WM. However, there still exist a few false positives and false negatives in final results. This is rooted in the similarities in the intensity and the coordinates of adjacent bordering voxels. Despite of a few incorrect predictions, our proposed method still correctly segment most of the brain tissues in the testing set.

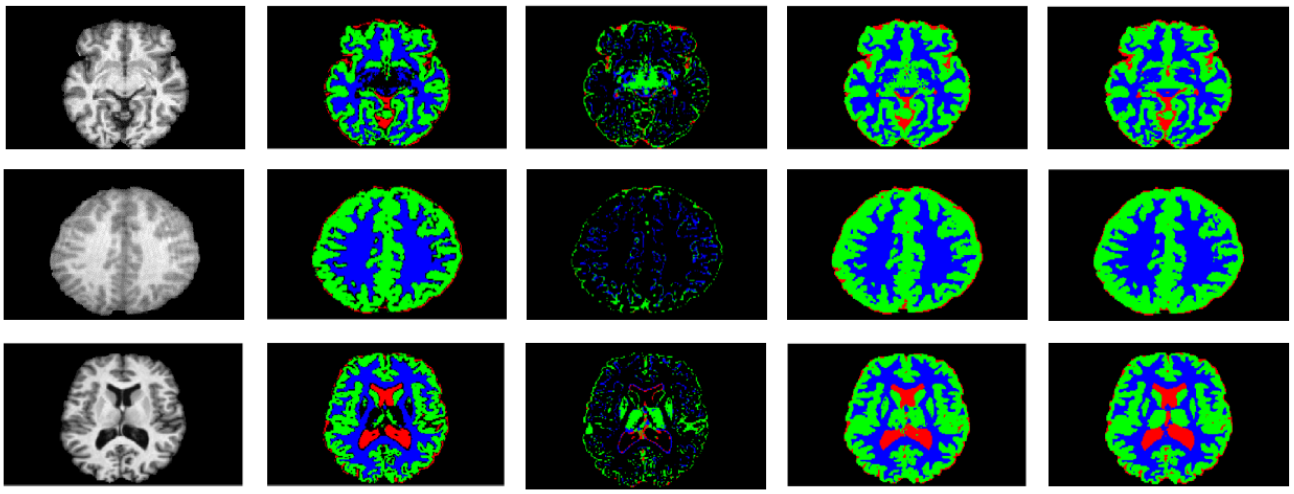


Figure 7. Segmentation results by our proposed method on the IBSR 18 dataset. Three typical brain images are shown in three rows. In each row, from the left to right is the input image, image predicted by GMM model for certain voxel, image predicted by deep neural network classifier for uncertain voxel, image combining between GMM model and deep neural networks classifier and ground truth image from the dataset, respectively.

Method	CSF	GM	WM	Average
Proposed	$\mathbf{0.79 \pm 0.03}$	$\mathbf{0.91 \pm 0.03}$	$\mathbf{0.90 \pm 0.01}$	**0.86**
SITDS	0.67 ± 0.03	0.86 ± 0.01	0.89 ± 0.02	0.80
ITDS	0.60 ± 0.05	0.81 ± 0.03	0.86 ± 0.02	0.75
3L-GMM	0.57 ± 0.19	0.92 ± 0.02	0.87 ± 0.03	0.78
MI	0.52 ± 0.08	0.79 ± 0.04	0.80 ± 0.03	0.70
KNN	0.46 ± 0.16	0.87 ± 0.03	0.86 ± 0.03	0.73
SVPASEG	0.57 ± 0.13	0.90 ± 0.01	0.87 ± 0.02	0.78
SPM8	0.77 ± 0.08	0.91 ± 0.01	0.88 ± 0.01	0.85
WPNN	0.63 ± 0.03	0.83 ± 0.02	0.87 ± 0.03	0.77
GAMIXTURE	0.52 ± 0.15	0.89 ± 0.03	0.87 ± 0.02	0.76
ANN	0.52 ± 0.15	0.87 ± 0.03	0.88 ± 0.03	0.75
SPM5	0.79 ± 0.08	0.89 ± 0.02	0.87 ± 0.02	0.85
MRF	0.53 ± 0.06	0.76 ± 0.03	0.87 ± 0.03	0.72
FCM	0.52 ± 0.15	0.88 ± 0.02	0.88 ± 0.03	0.76
FANTASM	0.53 ± 0.15	0.88 ± 0.02	0.88 ± 0.03	0.76
PVC	0.52 ± 0.15	0.83 ± 0.08	0.84 ± 0.07	0.73
FAST	0.47 ± 0.18	0.88 ± 0.01	0.89 ± 0.02	0.74
KMEANS	0.51 ± 0.06	0.75 ± 0.06	0.78 ± 0.04	0.68

Table 1. Results of various methods in IBSR 18 dataset.

3.4. Comparison

We use the dice coefficient [14] to evaluate our algorithm. The dice coefficient of the ground truth and the prediction for each tissue type in the combined results can be calculated by:

$$Dice\,(V_{gt},\, V_{pd}) = \frac{2|V_{gt} \cap V_{pd}|}{|V_{gt}| + |V_{pd}|} \qquad (10)$$

where V_{gt} represents the ground-truth labels and V_{pd} represents the predicted labels from our model for each tissue type.

We have compared the proposed method with other 13 state-of-the art algorithms in recent studies including the SITDS algorithm, ITDS algorithm [21], 3L-GMM algorithm [25], KNN algorithm [12], SVPASEG algorithm [41], SPM5 algorithm [3], SPM8 algorithm [17], GAMIXTURE algorithm [42], ANN algorithm [39], FCM algorithm [31],

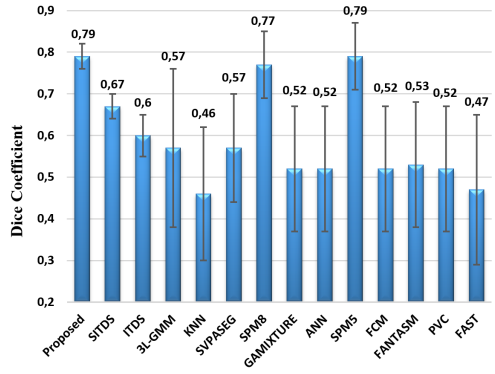


Figure 8. The dice coefficient in CSF of the proposed method compared with the different methods.

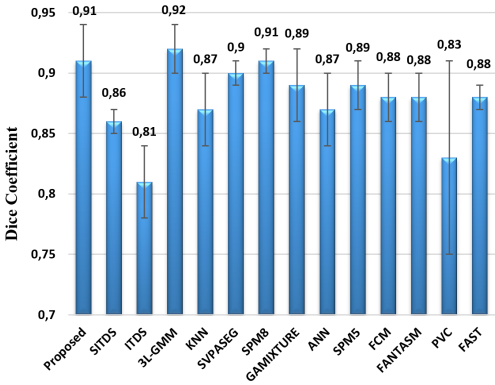


Figure 9. The dice coefficient in GM of the proposed method compared with the different methods.

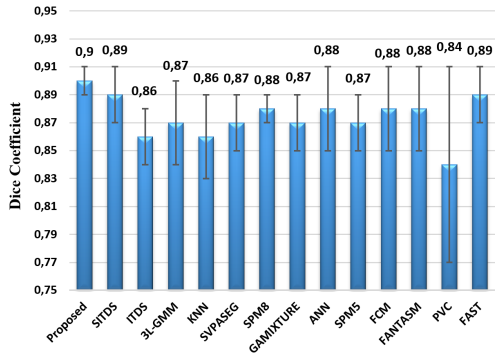


Figure 10. The dice coefficients in WM of our proposed method in comparison with other methods.

FANTASM algorithm [30], PVC algorithm [34] and FAST algorithm [45]. The results of all methods using KNN algorithm, VPASEG algorithm, SPM5 algorithm, SPM8 algorithm, GAMIXTURE algorithm, ANN algorithm, FCM algorithm, FANTASM algorithm, PVC algorithm and FAST algorithm are obtained from the paper [43] while all methods using SITDS algorithm, ITDS algorithm and 3L-GMM algorithm are reported from their corresponding references.

The segmentation results of the CSF, GM, WM and the average results are depicted in Figure 8, 9 and 10 and Table 1 respectively. The detailed results of all methods are also presented in Table 1. Based on the results of WM in Figure 10, our method achieved the best segmentation performance with average of 90% and standard deviation of 1%. With the results of CSF, the SPM5 algorithm has accuracy equal to our proposed method with 79% but has standard deviation 5% higher than our method. The 3L-GMM algorithm achieves the average 92% at the GM result 1% higher than our proposed method; however, this method does not perform well when classifying on the CSF with the performance lower than 60%. For the overall results in Table 1, it is not difficult to see that our proposed method achieves the best performance in term of overall results and outperforms 13 state-of-the-art algorithms with a stable performance for all cases. As a result, it demonstrates the effectiveness of our approach.

4. Conclusion

In this paper, we have described a novel approach for segmentation effectively three kinds of brain tissues (GM, WM, and CSF) in MRIs. The proposed method integrates both features maps learned from CNNs and the global spatial information extracted from DNNs. It increases the capability of feature representation in both global and local way for each voxel. In addition, our method can be constructed by using a general framework, which allows us to take all advantages of traditional approaches as well as modern approaches based on deep learning simultaneously. The experimental results on the IBSR 18 dataset show that the proposed model outperforms 13 recent state-of-the-art techniques. The execution time of our approach for each 3D voxel brains is about 122 seconds so that it can be fast enough to be used in the clinical practice. In future work, we can apply our ideas into other datasets to obtain more interesting results.

Acknowledgement

We would like to thank The National Foundation for Science and Technology Development (NAFOSTED), University of Science, and Business Intelligence LAB at University of Economics and Law, Viet Nam for supporting us throughout this paper.

References

[1] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, et al. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *arXiv preprint arXiv:1603.04467*, 2016.

[2] X. Artaechevarria, A. Munoz-Barrutia, and C. O. de Solorzano. Combination strategies in multi-atlas

image segmentation: Application to brain mr data. *IEEE Transactions on Medical Imaging*, 28(8):1266–1277, Aug 2009.

[3] J. Ashburner and K. J. Friston. Unified segmentation. *Neuroimage*, 26(3):839–851, 2005.

[4] J. D. Banfield and A. E. Raftery. Model-based gaussian and non-gaussian clustering. *Biometrics*, pages 803–821, 1993.

[5] F. A. Cappabianco, A. X. Falcão, C. L. Yasuda, and J. K. Udupa. Brain tissue mr-image segmentation via optimum-path forest clustering. *Computer Vision and Image Understanding*, 116(10):1047–1059, 2012.

[6] H. Chen, Q. Dou, X. Wang, J. Qin, J. C. Y. Cheng, and P.-A. Heng. *3D Fully Convolutional Networks for Intervertebral Disc Localization and Segmentation*, pages 375–382. Springer International Publishing, Cham, 2016.

[7] H. Chen, D. Ni, J. Qin, S. Li, X. Yang, T. Wang, and P. A. Heng. Standard plane localization in fetal ultrasound via domain transferred deep neural networks. *IEEE Journal of Biomedical and Health Informatics*, 19(5):1627–1636, Sept 2015.

[8] H. Chen, X. Qi, L. Yu, and P.-A. Heng. Dcan: Deep contour-aware networks for accurate gland segmentation. *arXiv preprint arXiv:1604.02677*, 2016.

[9] J. Cheng, J. Liu, Y. Xu, F. Yin, D. W. K. Wong, N. M. Tan, D. Tao, C. Y. Cheng, T. Aung, and T. Y. Wong. Superpixel classification based optic disc and optic cup segmentation for glaucoma screening. *IEEE Transactions on Medical Imaging*, 32(6):1019–1032, June 2013.

[10] J.-Z. Cheng, D. Ni, Y.-H. Chou, J. Qin, C.-M. Tiu, Y.-C. Chang, C.-S. Huang, D. Shen, and C.-M. Chen. Computer-aided diagnosis with deep learning architecture: Applications to breast lesions in us images and pulmonary nodules in ct scans. *Scientific reports*, 6, 2016.

[11] J. T. Chien and T. W. Lu. Tikhonov regularization for deep neural network acoustic modeling. In *Spoken Language Technology Workshop (SLT), 2014 IEEE*, pages 147–152, Dec 2014.

[12] R. de Boer, H. A. Vrooman, F. van der Lijn, M. W. Vernooij, M. A. Ikram, A. van der Lugt, M. M. Breteler, and W. J. Niessen. White matter lesion extension to automatic brain tissue segmentation on {MRI}. *NeuroImage*, 45(4):1151 – 1161, 2009.

[13] A. de Brebisson and G. Montana. Deep neural networks for anatomical brain segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 20–28, 2015.

[14] L. R. Dice. Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302, 1945.

[15] M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P.-M. Jodoin, and H. Larochelle. Brain tumor segmentation with deep neural networks. *Medical Image Analysis*, 35:18 – 31, 2017.

[16] Z. Ji, Y. Xia, Q. Sun, Q. Chen, D. Xia, and D. D. Feng. Fuzzy local gaussian mixture model for brain MR image segmentation. *IEEE Transactions on Information Technology in Biomedicine*, 16(3):339–347, 2012.

[17] A. John, B. Gareth, C. C., D. Jean, F. Guillaume, F. Karl, D. Gitelman, S. Kiebel, J. Kilner, V. Litvak, R. Moran, W. Penny, M. Rosa, K. Stephan, D. Gitelman, R. Henson, C. Hutton, V. Glauche, J. Mattout, and C. Phillips. Spm8 manual. *Functional Imaging Laboratory, Institute of Neurology*, page 41, 2008.

[18] K. M. Kennedy, E. D. Reese, M. M. Horn, A. N. Sizemore, A. K. Unni, M. E. Meerbrey, A. G. Kalich, and K. M. Rodrigue. Bdnf val66met polymorphism affects aging of multiple types of memory. *Brain research*, 1612:104–117, 2015.

[19] D. Kingma and J. Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[20] B. A. Kirchhoff, B. A. Gordon, and D. Head. Prefrontal gray matter volume mediates age effects on memory strategies. *NeuroImage*, 90:326–334, 2014.

[21] Y. Kong, Y. Deng, and Q. Dai. Discriminative clustering and feature selection for brain mri segmentation. *IEEE Signal Processing Letters*, 22(5):573–577, 2015.

[22] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[23] Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015.

[24] W. Li, F. Jia, and Q. Hu. Automatic segmentation of liver tumor in ct images with deep convolutional neural networks. *Journal of Computer and Communications*, 3(11):146, 2015.

[25] X. Liu and F. Chen. Automatic segmentation of 3-d brain mr images by using global tissue spatial structure information. *IEEE Transactions on Applied Superconductivity*, 24(5):1–5, 2014.

[26] A. Lucchi, K. Smith, R. Achanta, G. Knott, and P. Fua. Supervoxel-based segmentation of mitochondria in em image stacks with learned shape features. *IEEE transactions on medical imaging*, 31(2):474–486, 2012.

[27] V. Mnih, J. M. Susskind, G. E. Hinton, et al. Modeling natural images using gated mrfs. *IEEE transactions on pattern analysis and machine intelligence*, 35(9):2206–2222, 2013.

[28] P. Moeskops, M. A. Viergever, A. M. Mendrik, L. S. de Vries, M. J. N. L. Benders, and I. Igum. Automatic segmentation of mr brain images with a convolutional neural network. *IEEE Transactions on Medical Imaging*, 35(5):1252–1261, May 2016.

[29] S. K. Ng, T. Krishnan, and G. J. McLachlan. The em algorithm. In *Handbook of computational statistics*, pages 139–172. Springer, 2012.

[30] D. L. Pham. Robust fuzzy segmentation of magnetic resonance images. In *Computer-Based Medical Systems, 2001. CBMS 2001. Proceedings. 14th IEEE Symposium on*, pages 127–131, 2001.

[31] D. L. Pham. Spatial models for fuzzy clustering. *Comput. Vis. Image Underst.*, 84(2):285–297, Nov. 2001.

[32] D. L. Pham and J. L. Prince. Adaptive fuzzy segmentation of magnetic resonance images. *IEEE transactions on medical imaging*, 18(9):737–752, 1999.

[33] T. Rohlfing. Image similarity and tissue overlaps as surrogates for image registration accuracy: widely used but unre-

liable. *IEEE transactions on medical imaging*, 31(2):153–163, 2012.

[34] D. W. Shattuck, S. R. Sandor-Leahy, K. A. Schaper, D. A. Rottenberg, and R. M. Leahy. Magnetic resonance image tissue classification using a partial volume model. *NeuroImage*, 13(5):856–876, 2001.

[35] T. Song, M. M. Jamshidi, R. R. Lee, and M. Huang. A modified probabilistic neural network for partial volume segmentation in brain mr image. *IEEE Transactions on Neural Networks*, 18(5):1424–1432, 2007.

[36] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.*, 15(1):1929–1958, Jan. 2014.

[37] J. Steffener, C. Habeck, D. O'Shea, Q. Razlighi, L. Bherer, and Y. Stern. Differences between chronological and brain age are related to education and self-reported physical activity. *Neurobiology of aging*, 40:138–144, 2016.

[38] S. Susstrunk, P. Fua, A. Shaji, A. Lucchi, K. Smith, and R. Achanta. Slic superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(undefined):2274–2282, 2012.

[39] D. Tian and L. Fan. A brain mr images segmentation method based on some neural network. In *2007 1st International Conference on Bioinformatics and Biomedical Engineering*, pages 686–689, July 2007.

[40] G. Tian, Y. Xia, Y. Zhang, and D. Feng. Hybrid genetic and variational expectation-maximization algorithm for gaussian-mixture-model-based brain mr image segmentation. *IEEE transactions on information technology in biomedicine*, 15(3):373–380, 2011.

[41] J. Tohka, I. D. Dinov, D. W. Shattuck, and A. W. Toga. Brain mri tissue classification based on local markov random fields. *Magnetic resonance imaging*, 28(4):557–573, 2010.

[42] J. Tohka, E. Krestyannikov, I. D. Dinov, A. M. Graham, D. W. Shattuck, U. Ruotsalainen, and A. W. Toga. Genetic algorithms for finite mixture model based voxel classification in neuroimaging. *IEEE transactions on medical imaging*, 26(5):696–711, 2007.

[43] S. Valverde, A. Oliver, M. Cabezas, E. Roura, and X. Lladó. Comparison of 10 brain tissue segmentation methods using revisited ibsr annotations. *Journal of Magnetic Resonance Imaging*, 41(1):93–101, 2015.

[44] W. M. Wells, W. E. L. Grimson, R. Kikinis, and F. A. Jolesz. *Adaptive Segmentation of MRI Data*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1995.

[45] S. S. Y. Zhang, M. Brady. Segmentation of brain mr images through a hidden markov random field model and the expectation-maximization algorithm. *IEEE transactions on medical imaging*, 20(1):45–57, 2001.

[46] W. Zhang, R. Li, H. Deng, L. Wang, W. Lin, S. Ji, and D. Shen. Deep convolutional neural networks for multimodality isointense infant brain image segmentation. *NeuroImage*, 108:214–224, 2015.