

Research Article

Segmentation of Brain MR Images by Using Fully Convolutional Network and Gaussian Mixture Model with Spatial Constraints

Jiawei Lai, Hongqing Zhu , and Xiaofeng Ling

School of Information Science & Engineering, East China University of Science and Technology, Shanghai 200237, China

Correspondence should be addressed to Hongqing Zhu; hqzhu@ecust.edu.cn

Received 29 January 2019; Accepted 28 April 2019; Published 13 May 2019

Academic Editor: Erik Cuevas

Copyright © 2019 Jiawei Lai et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Accurate segmentation of brain tissue from magnetic resonance images (MRIs) is a critical task for diagnosis, treatment, and clinical research. In this paper, a novel algorithm (GMMD-U) that incorporates the modified full convolutional neural network U-net and Gaussian-Dirichlet mixture model (GMMD) with spatial constraints is presented. The proposed GMMD-U considers the local spatial relationships by assuming that the prior probability obeys the Dirichlet distribution. Specifically, GMMD is applied for extracting brain tissue that has a distinct intensity region and modified U-net is exploited to correct the wrong-classification areas caused by GMMD or other conventional approaches. The proposed GMMD-U is designed to take advantage of the statistical model-based segmentation techniques and deep neural network. We evaluate the performance of GMMD-U on a publicly available brain MRI dataset by comparing it with several existing algorithms, and the results reported reveal that the proposed framework can accurately detect the brain tissue from MRIs. The proposed learning-based integrated framework could be effective for brain tissue segmentation, which will be helpful for surgeons in brain disease diagnosis.

1. Introduction

Precise segmentation of human brain tissue from magnetic resonance images (MRIs) can aid in identification and diagnosis of neurological diseases, such as Parkinson's disease and Alzheimer's disease. This is a challenging task because brain MRIs are severely affected by intensity nonuniformity, complex structure, and low contrast during acquisition. During the last decades, a large majority of MRI segmentation algorithms were based on machine learning techniques [1]. In general, machine learning can be divided into supervised approaches, semisupervised or unsupervised approaches. For example, both random forests [2–4] and support vector machines [5] are typical supervised techniques commonly used in MRI segmentation. Recently, some new unsupervised algorithms have been proposed for extracting the required brain tissue, including grey matter (GM), white matter (WM), and cerebrospinal fluid (CSF) [6, 7], but clustering analysis is among the most studied [8, 9]. It has been found that clustering performance depends heavily on the selection of the initial cluster centres, and the segmentation results are sensitive to noise. The finite-mixture model (FMM)

has become another important branch in unsupervised segmentation algorithms. It applies an unsupervised learning method to label the observation data and supposes that the intensity of each pixel obeys the chosen probability distribution. Typical FMMs include the Gaussian mixture model (GMM) [10], Student's-*t* mixture model (SMM) [11], circular mixture model [12], Rayleigh mixture model [13], etc. GMM is widely applied to model uncertainty of the data by assuming the conditional probability as a Gaussian distribution. It is the main model because the classical GMM is simple to carry out, and there are very few parameters that need to be estimated. However, GMM suffers from a problem of modelling data with different shapes and sensitivity to outliers. To overcome these shortcomings, segmentation techniques adopting mixture models with longer-tailed probability distributions have widely received consideration. For example, SMM with the longer tail can be regarded as an alternative to GMM [14]. However, one limitation of SMM is that its freedom parameter does not exist in closed form [15]. Another is that it cannot consider the relationships between the local spatial information between neighbouring pixels. An improvement can be obtained by combining the

mixture models with the Markov random field (MRF) [16]. Additionally, the FMM was incorporated with level set-based methods in some studies [17]. The unsupervised methods mentioned above are often subject to various uncertainties, such as the tissue intensity overlaps caused by limitations of the MRI acquisition process. These factors decrease the brain MRI segmentation performance. Instead of employing FMM for modelling the MRI tissue intensity, deep-learning-based algorithms (supervised or unsupervised) have received considerable attention of recently published studies [18]. As one of the parts of broader family of machine learning, deep learning has achieved state-of-the-art results in MRI segmentation [18] or other computer vision tasks including image retrieval [19], image classification [20], object detection [21], etc. The convolutional neural network (CNN) is a deep learning architecture inspired by the biological networks akin to the multilayer perceptron. Classical CNN framework consists of three main building blocks: fully connected layers, pooling layers, and convolutional layers. Some recent studies have claimed that it is also suitable for MRI segmentation and brain tissue detection [22, 23]. However, for the classical CNN, it is a tedious task to recognize image patches. For example, the normal algorithm introduced by Ciresan [24] needs huge amount of time and redundant images for training because the recognition on an image patch acquires a sliding-window operation. Therefore, an end-to-end segmentation method called fully convolutional network (FCN) was recently introduced [25, 26] that can make pixel-wise prediction based on the images' ground truth and can output the label map directly [27]. The purpose of the network FCN is to extract critical feature maps and to restore these maps into the image labels. This procedure focusing on features could be more suitable for precise segmentation, especially for medical images [28]. Actually, the capability of feature representation is the primary reason that FCN achieves great success on object detection, classification, and segmentation, but often only when sufficient amount of training data is available. However, in medical imaging field, the data acquisition is expensive. Besides, there are some other factors that impact the data availability, like the privacy and regulations/ethical concerns, etc. Recently, U-net was proposed to extend the FCN network by increasing and recycling more feature maps [29], so that U-net can perform well with relatively small amounts of training samples. The earliest U-net [29] was designed for labelling cells from light microscopic images and to yield very accurate segmentation results. Such network usually produces high-precision segmentation mainly because its architecture includes more constructions that copy and overlay down-sampling features.

Motivated by the aforementioned research, in this paper, we addressed a novel fusion framework called GMMU, which incorporates the Gaussian-Dirichlet mixture model (GMM) and modified U-net to accurately analyse brain MRIs in terms of tissue types. The proposed algorithm differs from conventional clustering or FMM algorithms with respect to the following considerations. First, the proposed method merges the advantages of deep learning so that our algorithm has ability to use convolutional network to precisely recognize the uncertainty regions. Second, this

paper improves the classical U-net by attaching a padding operation and batch normalization in order to improve the convergence speed. More clearly, CSF makes up a rather low proportion in the brain and its greyscale is very close to the pure black background in most MRIs; the modified U-net should be able to learn and to distinguish the shape of CSF as fast as possible. Third, this paper develops a novel FMM, Gaussian-Dirichlet mixture model that is a modified version of the classical GMM. Comparing with our previous work [30], the proposed GMMU takes the local spatial and intensity information into consideration through Dirichlet distribution so that the performance of the proposed GMMU is insensitive to noise and outliers. Four, in the proposed framework, the majority of pixels belonging to WM and GM can be accurately determined by the proposed GMMU module. The modified U-net is utilized to predict the CSF, as well as the error-prone region done by GMMU or other traditional methods. The experimental results in the Internet Brain Segmentation Repository (IBSR) dataset demonstrate that the proposed framework is superior to several existing unsupervised and supervised models.

The rest of this article is organized as follows. Section 2 addresses the construction of the modified GMM with spatial constraints and details the proposed fusion framework. The experimental results and discussion on the algorithm's performance are given in Section 3. The final concluding remarks are provided in Section 4.

2. Methodology

The proposed fusion architecture consists of two fully convolutional networks: modified U-net and a Gaussian-Dirichlet mixture model with spatial constraints. In this paper, the finite-mixture model technology is adopted in our fusion scheme. It is mainly because it provides a statistical-based approach to model observed data in a probabilistic manner. Another advantage is that it classifies every pixel of an image into certain labels, while U-net network structure could only label object from the background. More precisely, the output of U-net has only two labels, which means WM, GM, and CSF could not be extracted at once. The aim of improving U-net framework in this study is to correct wrong label regions raised from conventional approaches. Therefore, the proposed approach incorporates finite-mixture model scheme with improved U-net network to extract brain tissue more accurately. We start by introducing modified GMM, which follows a Dirichlet distribution for brain MRI segmentation.

2.1. Gaussian-Dirichlet Mixture Model with Spatial Constraints. To consider the local spatial information between neighbouring pixels, this subsection presents a novel Gaussian-Dirichlet mixture model which is a modified version of GMM in terms of Dirichlet distribution. Assume that each pixel x_i , $i = (1, 2, \dots, N)$, is an independent and identically distributed value in a grayscale image. Then, GMM assumes that x_i is independent of the label Ω_j , $j = (1, 2, \dots, K)$. To partition an image consisting of N

pixels into K labels, GMM assumes the density function $f(x_i | \Pi, \Theta)$ at pixel x_i is as follows:

$$f(x_i | \Pi, \Theta) = \sum_{j=1}^K \pi_j G(x_i | \Theta_j) \quad (1)$$

where $\Theta_j = \{\pi_j, \mu_j, \sigma_j^2\}$ denotes the GMM parameters and μ_j and σ_j^2 are the mean and covariance of Gaussian distribution, respectively. $\Pi = \{\pi_j\}$ is the set of prior probabilities of pixel x_i belonging to Ω_j , which satisfies the constraints

$$\begin{aligned} 0 \leq \pi_j \leq 1 \\ \text{and } \sum_{j=1}^K \pi_j = 1. \end{aligned} \quad (2)$$

In (1), $G(x_i | \Theta_j)$ is a component of the Gaussian distribution with the form

$$\begin{aligned} G(x_i | \Theta_j) \\ = \frac{1}{\sigma_j \sqrt{2\pi}} \exp \left\{ -\frac{1}{2\sigma_j^2} (x - \mu_j)^T (x - \mu_j) \right\} \end{aligned} \quad (3)$$

The joint conditional probability density of $X = (x_1, x_2, \dots, x_N)$ is expressed by

$$\begin{aligned} p(X_i | \Pi, \Theta) &= \prod_{i=1}^N f(x_i | \Pi, \Theta) \\ &= \prod_{i=1}^N \sum_{j=1}^K \pi_j G(x_i | \Theta_j) \end{aligned} \quad (4)$$

Taking the natural logarithm of (4), we have the following maximum logarithmic likelihood function:

$$L = \max \sum_{i=1}^N \log \left(\sum_{j=1}^K \pi_j G(x_i | \Theta_j) \right) \quad (5)$$

Next, we introduce the Dirichlet distribution [33] into the classical GMM in order to model the spatial information between neighbouring pixels. First, the probability label is defined as

$$p(z_i | a_i) = \int_0^1 p(z_i | \xi_i) p(\xi_i | a_i) d\xi_i \quad (6)$$

where $z_i = (z_{i1}, z_{i2}, \dots, z_{iK})$ is the i th pixel's discrete label taking the following form:

$$z_{ij} = \begin{cases} 1, & x_i \in \Omega_j \\ 0, & \text{Otherwise} \end{cases} \quad (7)$$

Corresponding to (6), $\xi_i = (\xi_{i1}, \xi_{i2}, \dots, \xi_{iK})$ is the Dirichlet parameters. It also satisfies the constraints

$$\begin{aligned} \xi_{ij} \geq 0 \\ \text{and } \sum_{j=1}^K \xi_{ij} = 1 \end{aligned} \quad (8)$$

Define $a_i = (a_{i1}, a_{i2}, \dots, a_{iK})$ with $a_{ij} \geq 0$ in (6) as the vector format of the Dirichlet parameters. Then, according to method introduced in [31, 34], the probability density function $p(z_i | \xi_i)$ can be written in the polynomial form

$$p(z_i | \xi_i) = \frac{M!}{\prod_{j=1}^K (z_{ij})!} \prod_{j=1}^K (\xi_{ij})^{z_{ij}} \quad (9)$$

Hence, the probability density function $p(\xi_i | a_i)$ takes the form

$$p(\xi_i | a_i) = \frac{\Gamma(\sum_{j=1}^K a_{ij})}{\prod_{j=1}^K \Gamma(a_{ij})} \prod_{j=1}^K \xi_{ij}^{a_{ij}-1} \quad (10)$$

This paper defines the Dirichlet parameters a_{ij} to incorporate the neighbourhood spatial information by

$$a_{ij} = \exp \left(\frac{\beta}{N_i} \sum_{m \in \delta_i} y_{mj}^{(t-1)} \right) \quad (11)$$

where N_i represents the number of neighbour pixels that include x_i and points around x_i in a certain window. $y_{ij}^{(t)}$ represents the posterior probability at iteration step t .

$$y_{ij}^{(t)} = \frac{\pi_{ij} G(x_i | \Theta_j)}{\sum_{k=1}^K \pi_{ik} G(x_i | \Theta_k)} \quad (12)$$

By (12), one can find that (11) can be regarded as a linear filter, where pixel x_i is replaced by the mean value of its corresponding window. To eliminate the effect of noise, generally, the size of the filter template can be 3×3 or 5×5 . Taking (9) and (10), (6) can be expressed as

$$\begin{aligned} p(z_i | a_i) \\ = \int_0^1 \frac{M!}{\prod_{j=1}^K (z_{ij})!} \prod_{j=1}^K \xi_{ij}^{z_{ij}} \frac{\Gamma(\sum_{j=1}^K a_{ij})}{\prod_{j=1}^K \Gamma(a_{ij})} \\ \times \prod_{j=1}^K (\xi_{ij})^{a_{ij}-1} d\xi_i \\ \times \frac{M!}{\prod_{j=1}^K (z_{ij})} \frac{\Gamma(\sum_{j=1}^K a_{ij})}{\prod_{j=1}^K \Gamma(a_{ij})} \frac{\prod_{j=1}^K \Gamma(a_{ij} + z_{ij})}{\Gamma(\sum_{j=1}^K (a_{ij} + z_{ij}))} \\ \times \int_0^1 \frac{\Gamma(\sum_{j=1}^K (a_{ij} + z_{ij}))}{\prod_{j=1}^K \Gamma(a_{ij} + z_{ij})} \prod_{j=1}^K (\xi_{ij})^{a_{ij}+z_{ij}-1} d\xi_i \end{aligned} \quad (13)$$

According to the property of the probability density function, one has

$$\begin{aligned} \int_0^1 p(\xi_i | a_i) d\xi_i &= \int_0^1 \frac{\Gamma(\sum_{j=1}^K a_{ij})}{\prod_{j=1}^K \Gamma(a_{ij})} \prod_{j=1}^K (\xi_{ij})^{(a_{ij}-1)} d\xi_i \\ &= 1 \end{aligned} \quad (14)$$

Substituting (11) and (14) into (13) leads to

$$\begin{aligned}
p(z_i | a_i) &= \frac{M!}{\prod_{j=1}^K (z_{ij})!} \frac{\Gamma(\sum_{j=1}^K a_{ij})}{\prod_{j=1}^K \Gamma(a_{ij})} \frac{\prod_{j=1}^K \Gamma(a_{ij} + z_{ij})}{\Gamma(\sum_{j=1}^K (a_{ij} + z_{ij}))} \\
&= \frac{M!}{\prod_{j=1}^K (z_{ij})!} \frac{\Gamma(\sum_{j=1}^K \exp((\beta/N_i) \sum_{m \in \partial_i} y_{mj}))}{\prod_{j=1}^K \Gamma(\exp((\beta/N_i) \sum_{m \in \partial_i} y_{mj}))} \\
&\quad \times \frac{\prod_{j=1}^K \Gamma(\exp((\beta/N_i) \sum_{m \in \partial_i} y_{mj}) + z_{ij})}{\Gamma(\sum_{j=1}^K (\exp((\beta/N_i) \sum_{m \in \partial_i} y_{mj}) + z_{ij}))}
\end{aligned} \quad (15)$$

The Dirichlet-based constraints consider the spatial information of neighbouring pixels in the form of linear filtering. More specifically, if considering discrete label z_{ij} in (7) and the Gamma function $\Gamma(t+1) = t\Gamma(t)$, the prior probability of (15) with $M = 1$, $z_{ij} = 1$ for each pixel x_i can be written by

$$\begin{aligned}
\pi_{ij} = p(z_{ij} = 1 | a_i) &= \frac{\exp((\beta/N_i) \sum_{m \in \partial_i} y_{mj}^{(t-1)})}{\sum_{j=1}^K \exp((\beta/N_i) \sum_{m \in \partial_i} y_{mj}^{(t-1)})}
\end{aligned} \quad (16)$$

In this case, we have the following new log-likelihood function:

$$L(\Xi) = \sum_{i=1}^N \log \left(\sum_{j=1}^K \pi_{ij} G(x_i | \Theta_j) \right) \quad (17)$$

where the parameter set Ξ is defined as $\Xi = \{\mu_j, \sigma_j^2, \beta_j\}$. It is reasonable to optimize the negative likelihood function because the logarithm is monotonically increasing. Thus, the new loss function is expressed by

$$J(\Xi) = -L(\Xi) = -\sum_{i=1}^N \log \left(\sum_{j=1}^K \pi_{ij} G(x_i | \Theta_j) \right) \quad (18)$$

$J(\Xi)$ can also be supposed as the difference expressions

$$\begin{aligned}
&J(\Xi^{(t+1)}) - J(\Xi^{(t)}) \\
&= -\sum_{i=1}^N \log \left\{ \frac{\sum_{j=1}^K \pi_{ij}^{(t+1)} G^{(t+1)}(x_i | \Theta_j)}{\sum_{j=1}^K \pi_{ij}^{(t)} G^{(t)}(x_i | \Theta_j)} \times \frac{y_{ij}^{(t+1)}}{y_{ij}^{(t)}} \right\}
\end{aligned} \quad (19)$$

Considering $y_{ij}^{(t)} \geq 0$, $\sum_{j=1}^K y_{ij}^{(t)} = 1$, and the Jensen inequality [35], expression (19) can be rewritten as

$$\begin{aligned}
&J(\Xi^{(t+1)}) - J(\Xi^{(t)}) \\
&\leq -\sum_{i=1}^N \sum_{j=1}^K y_{ij}^{(t)} \log \left\{ \frac{\pi_{ij}^{(t+1)} G^{(t+1)}(x_i | \Theta_j)}{y_{ij}^{(t)} \sum_{j=1}^K \pi_{ij}^{(t)} G^{(t)}(x_i | \Theta_j)} \right\}
\end{aligned} \quad (20)$$

Thus, minimizing the negative log-likelihood function in (18) is equivalent to minimizing the following error function:

$$\begin{aligned}
E(\Xi) &= -\sum_{i=1}^N \sum_{j=1}^K y_{ij}^{(t)} \left\{ \log \pi_{ij}^{(t+1)} + \log G^{(t+1)}(x_i | \Theta_j) \right\} \\
&= -\sum_{i=1}^N \sum_{j=1}^K y_{ij}^{(t)} \left\{ \log G^{(t+1)}(x_i | \Theta_j) \right. \\
&\quad \left. + \frac{\beta^{(t+1)}}{N_i} \sum_{m \in \partial_i} y_{mk}^{(t)} \right. \\
&\quad \left. - \log \left(\sum_{j=1}^K \exp \left(\frac{\beta^{(t+1)}}{N_i} \sum_{m \in \partial_i} y_{mk}^{(t)} \right) \right) \right\}
\end{aligned} \quad (21)$$

Next, we apply the gradient descent method [36] for parameter learning.

$$\Xi^{(t+1)} = \Xi^{(t)} - \eta \nabla E(\Xi^{(t)}) \quad (22)$$

where η is the learning rate and $\nabla E(\Xi) = (\partial E / \partial \mu_j, \partial E / \partial \sigma_j^2, \partial E / \partial \beta)$. The partial derivatives of the parameter $E(\Xi)$ with respect to μ are calculated by

$$\frac{\partial E(\Xi)}{\partial \mu} = -\sum_{i=1}^N y_{ij}^{(t)} \left(\frac{x_i - \mu_j}{\sigma_j^2} \right) \quad (23)$$

Similarly, we take the partial derivative of $E(\Xi)$ with respect to σ^2 as

$$\frac{\partial E(\Xi)}{\partial (\sigma^2)} = -\frac{1}{2} \sum_{i=1}^N y_{ij}^{(t)} \left\{ \frac{(x_i - \mu_j)^2}{(\sigma_j^2)^2} - \frac{1}{\sigma_j^2} \right\} \quad (24)$$

Next, considering the partial derivative of $E(\Xi)$ with respect to β yields

$$\begin{aligned}
\frac{\partial E(\Xi)}{\partial \beta} &= -\sum_{i=1}^N \sum_{j=1}^K y_{ij}^{(t)} \times \left\{ \frac{1}{N_i} \sum_{m \in \partial_i} y_{mj}^{(t)} \right. \\
&\quad \left. - \frac{\sum_{j=1}^K \left[\left((1/N_i) \sum_{m \in \partial_i} y_{mk}^{(t-1)} \right) \exp \left((\beta/N_i) \sum_{m \in \partial_i} y_{mj}^{(t-1)} \right) \right]}{\sum_{j=1}^K \left[\left((\beta/N_i) \sum_{m \in \partial_i} y_{mj}^{(t-1)} \right) \right]} \right\}
\end{aligned} \quad (25)$$

In detail, the computation process of parameters is summarized as follows.

Step 1. Initialize the prior probability π_{ij} , mean μ , and variance σ^2 of each label by using K-means. $\beta = 20$; $\eta = 10^{-6}$.

Step 2 (E-step). Calculate the posterior probability y_{ij} using (12).

Step 3 (M-step). Evaluate the mean μ , variance σ^2 , and β in terms of (22), (23), (24), and (25). Update the prior probability π_{ij} using (16).

Step 4. Check the convergence of log-likelihood in (17). If the convergence criterion is not satisfied, increase the iteration $t = t + 1$, and repeat Steps 2–4.

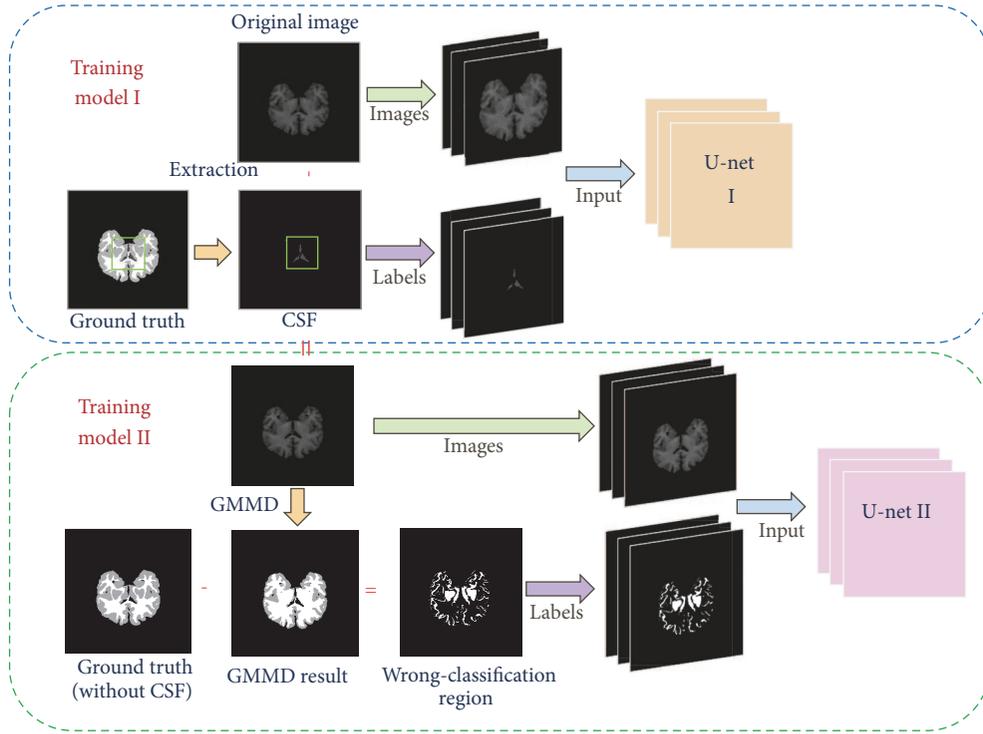


FIGURE 1: The architecture of the proposed algorithm in the training phase.

2.2. Training Procedure. In this paper, all experiments are performed on the IBSR 18 dataset [37], which provides manually guided expert segmentation results along with magnetic resonance brain image data. This dataset consists of MRIs and 3D ground truth volume of 18 brains of size $256 \times 256 \times 128$ with 1.5 mm slice thickness. These volumes are provided after skull-stripping, normalization and bias field correction. The ground truth is provided with manual segmentation by experts with tissue labels as 0, 1, 2, 3 for background, CSF, GM, and WM, respectively. Each MRI volume is read, via 256 number of axial brain slices of size 256×128 each, in the proposed model. The proposed network structure of the training procedure can be divided into two modified U-net subnetworks and GMMD, as shown in Figure 1. To acquire the training set, we define the ground truth of CSF as the binary mask and then obtain the brain MRI without CSF by the following relation:

$$O_{noCSF} = O \times (1 - \text{mask}_{CSF}), \quad (26)$$

where O is the original brain MRI and mask_{CSF} is the ground truth of CSF. All original brain MRIs and mask_{CSF} images will be fed into the first modified U-net network as the training dataset. This procedure is summarized in Training Model I, shown in Figure 1. To clarify this process, Figure 2 presents several examples to show the visually extracted CSF and O_{noCSF} . After that, the brain MRIs (without CSF) are fed into the GMMD as original input images. After obtaining the segmentation results of GM and WM using the proposed GMMD, as shown in the schematic representation of Training Model II, the training set of wrong-classification

regions is then acquired from the difference between the regions labelled by GMMD and the ground truth without CSF. For the purpose of clarifying this process, we present four examples for extracting wrong-classification areas in Figure 3. The term ‘wrong-classification area’ indicates the pixels that are not usually sufficiently distinguished by the classical segmentation methods. In the training phase, as can be seen from Figure 4, we found that the wrong-classification regions are almost always GM. More specifically, it can be found that almost 8 – 9% of GM pixels are mistakenly classified as WM (see Figure 5). In most cases, sample images that are segmented by GMMD have few wrong-classification pixels, except for GM pixels. Due to the small proportion of WM area in the residual region, in this paper, the prediction of the second modified U-net module is regarded as the GM detector to correct the wrong-classification regions caused by GMMD or other classical segmentation algorithms.

2.3. Modified U-Net Framework. U-net [29] is a fully convolutional network that performs excellent in image segmentation. Classical U-net network does not require a huge amount of training sets. What is more, the training time of U-net is relatively short, having a simple structure, and demanding less parameters compared with other network. In this paper, we improve the classical U-net by attaching a padding operation and batch normalization in order to improve the convergence speed. This is important because batch normalization can allow one to utilize higher learning rates and require less intensive initialization. The whole architecture of this modified U-net network is depicted in

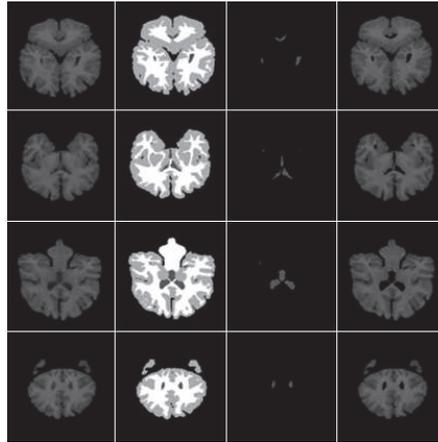


FIGURE 2: CSF extraction. From left to right: the original images, ground truth, extracted CSF, and brain images (without CSF).

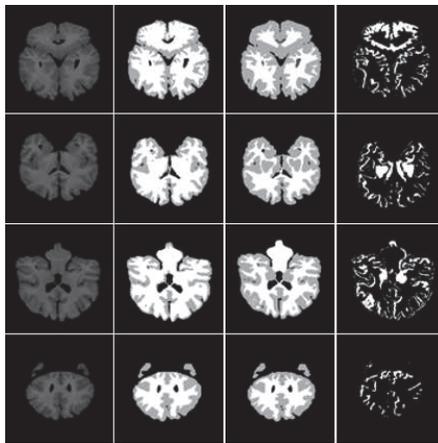


FIGURE 3: Examples for extracting wrong-classification areas. From left to right: the brain images (without CSF), segmentation results for WM and GM using GMM, ground truth without CSF, and wrong-classification regions.

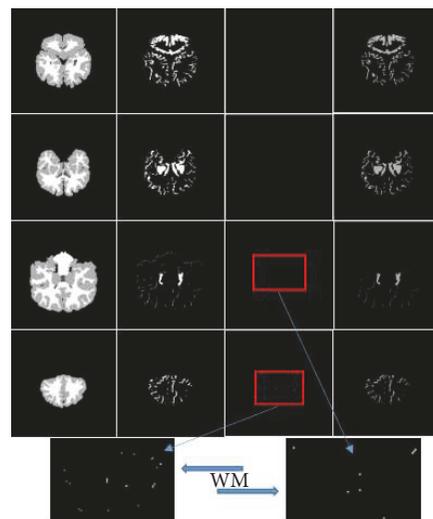


FIGURE 4: The results of WM and GM from wrong-classification areas. From left to right: the ground truth without CSF, wrong-classification areas, WM part of wrong-classification areas, and GM part of wrong-classification areas.

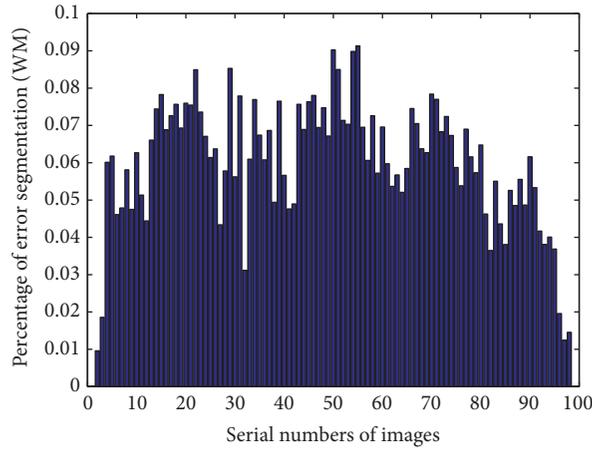


FIGURE 5: Percentage of error segmentation (WM).

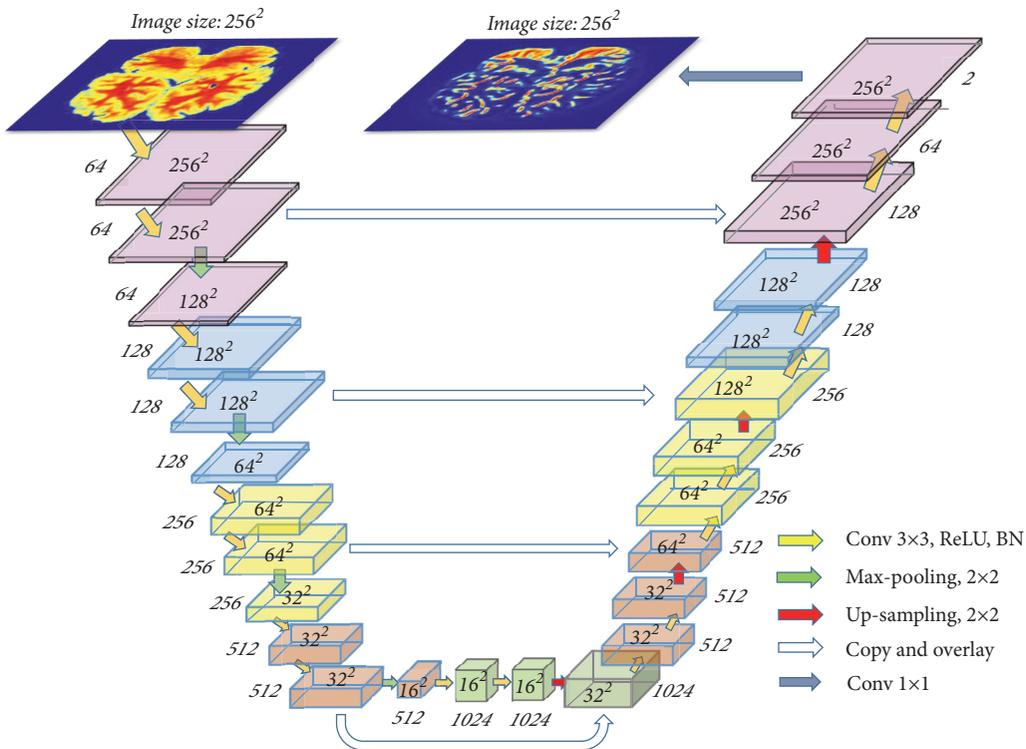


FIGURE 6: The architecture of the improved U-net. The number within the block demonstrates the image size and the outside number stands for the quantity of feature maps.

Figure 6. This network combines a feature-extracting path for collecting global features and an expanding method to locate pixels belonging to the features. More specifically, the network structure in Figure 6 contains batch normalization next to each 3×3 convolution through zero-padding. The ReLU layer, which is the nonlinear activation function, is used instead of the traditional Sigmoid function. The ReLU function only needs a single threshold to activate itself, and moreover can eliminate the complexity of calculation [26]. After two continuous convolution operations, max-pooling

with stride 2 is attached in the upper structure. Correspondingly, upsampling with stride 2 is operated in the lower structure. Before each max-pooling, the feature maps will be copied and then be transferred to the same location as the upsampling part. At the output layer, a 1×1 convolution with the generalized linear soft-max function is applied to calculate the probability of each pixel in the classification. Cross entropy is adopted as the loss function in this network, which describes the distance between the prediction and real values of the proposed model [38].

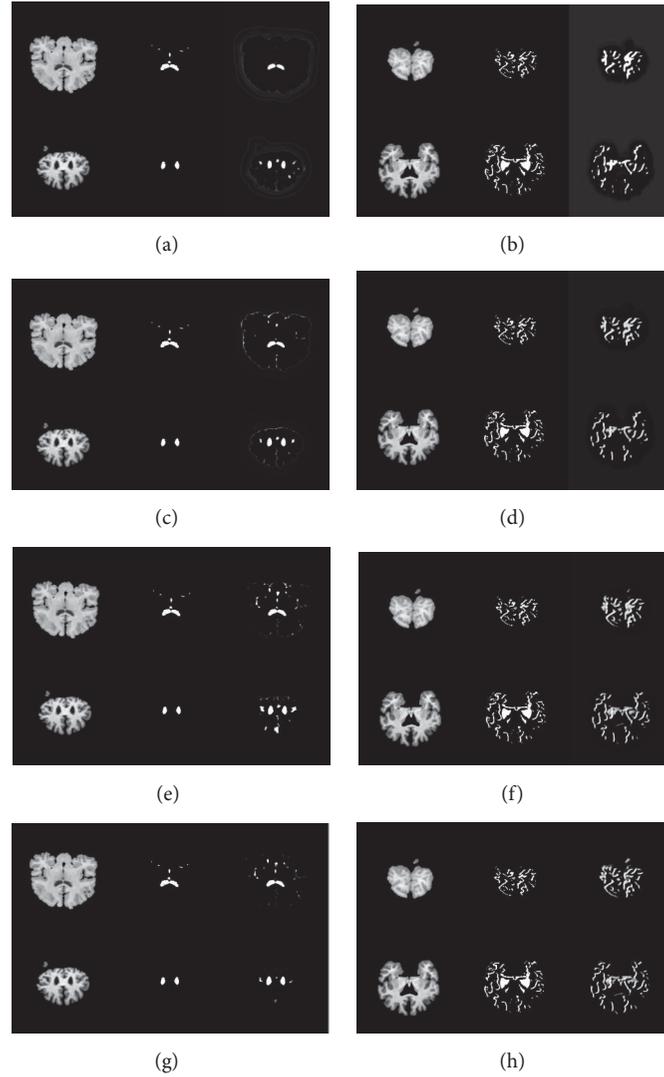


FIGURE 7: Network training output. The first column is CSF training epoch (Training Model I), and the second column is GM region training epoch (Training Model II). For each picture, from left to right are the original images, CSF/wrong-classification region, and prediction. (a) and (b) Epoch 10, (c) and (d) epoch 40, (e) and (f) epoch 70, and (g) and (h) epoch 90.

$$P_{softmax} = \frac{e^{l_n}}{\sum_n e^{l_n}} \quad (27)$$

where l_n is the logit value from the image-prediction matrix and n is the size of the matrix. Then, substituting $P_{softmax}$ into the equation of cross entropy yields

$$\begin{aligned} & \text{cross entropy} \\ &= \frac{1}{n} \sum_x [-y \times \ln(\alpha) - (1 - y) \times \ln(1 - \alpha)] \end{aligned} \quad (28)$$

where a represents $P_{softmax}$ and y is the ground-truth value.

As illustrated in Figure 1, the proposed model utilizes two modified U-net modules for training the CSF and wrong-classification areas based on the Tensorflow framework [39]. The first training model is for CSF detection and the second model is attributed to wrong-classification region prediction.

Here, the wrong-classification region is some pixels that cannot be precisely classified using GMMD. Figure 7 shows some training output epochs of Training Models I and II. As indicated by the figure, our proposed training model achieves higher accuracy with increasing epoch. The curves of loss function versus the number of iterations corresponding to each training procedure are depicted in Figure 8. It can be clearly observed that the loss function presents a steady decrease with increasing iterations, which confirms the effectiveness of the proposed training models.

3. Experimental Results

Figure 9 displays the flowchart of the test procedure using the proposed GMMD-U framework, which contains three parts corresponding to CSF detection (Part I), wrong-classification prediction (Part II), and remaining region clustering (Part

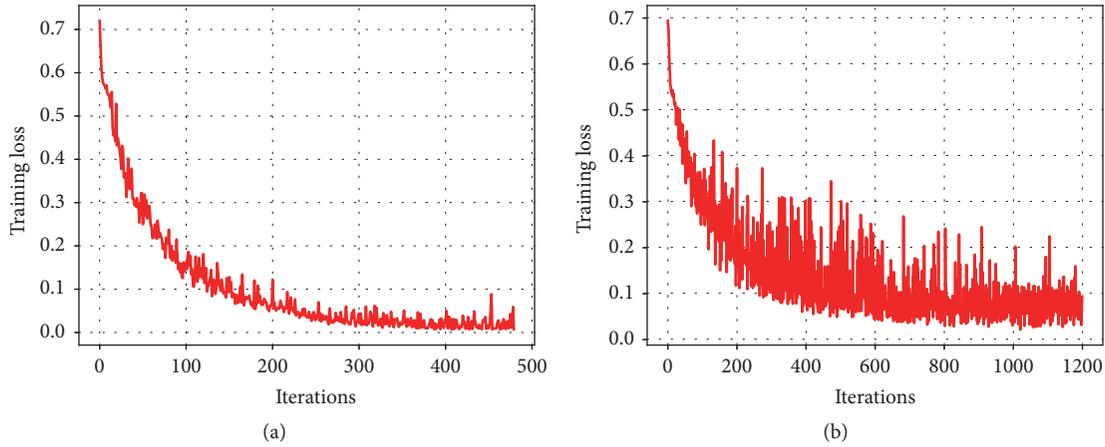


FIGURE 8: The training loss of two network modules. The horizontal axis represents the iterations and the vertical axis indicates the cross entropy. (a) The loss of Training Model I. (b) The loss of Training Model II.

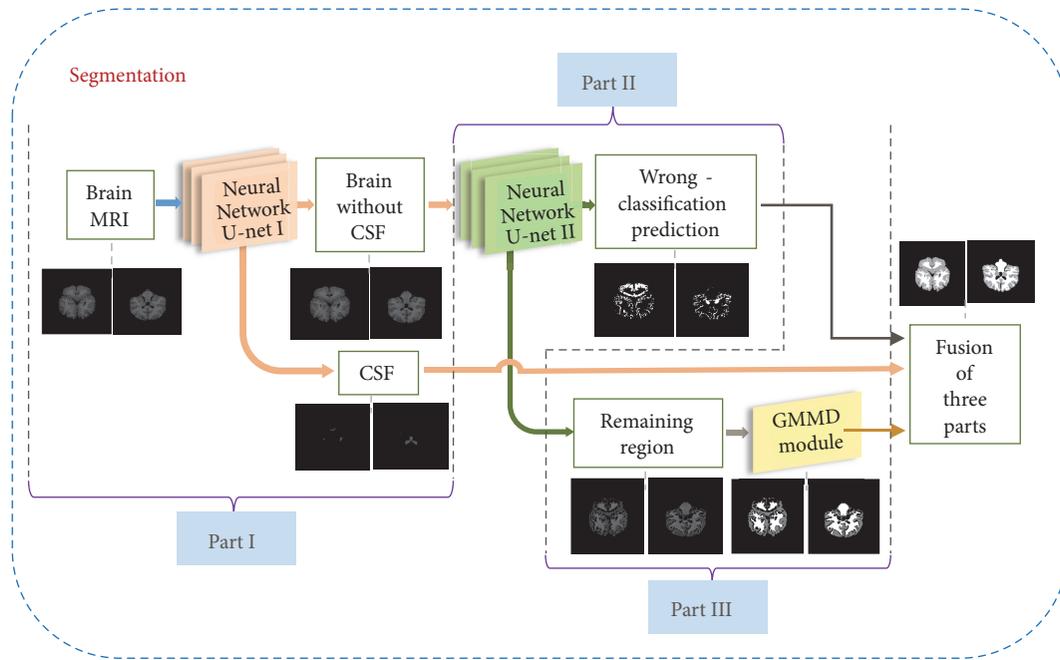


FIGURE 9: Detailed framework of segmentation using GMMD in combination with fully convolutional networks.

III). We run two trained U-net models on the testing set and produce heat maps to illustrate the probabilities of prediction. In this paper, all testing samples are randomly selected from the IBSR 18 dataset, and the test results are presented in Figure 10. Obviously, the results of prediction are visually similar to the ground truth. This means some regions that GMMD never labels precisely can be predicted by the proposed model. Furthermore, comparing the ground truth (labels) and predicted areas reported in Figure 10(b), one can observe that almost all of the predicted pixels belong to GM because the prediction areas are basically the same as the ground truth of GM. GMMD has an ability to model the shape of the pixel regions better. This is significant because these regions have fewer blur boundaries and patches. Finally,

for each MRI, the final segmentation results come from three parts: (1) CSF detected by modified neural network U-net I, (2) GM and WM labelled using the posterior probability of the i th pixel of the MRI to class j , and (3) GM and WM obtained from uncertain pixel areas via modified neural network U-net II and the wrong-classification prediction module.

We have evaluated the proposed fusion algorithm on the public brain datasets from IBSR 18. Specifically, we focus on the segmentation of clinical MRI and aim to separate the three parts of the real brain image data that contain CSF, GM, and WM. For experimentation purposes, all MR images from the IBSR 18 dataset are preprocessed by removing blank images and extracting the main brain parts from the skull.

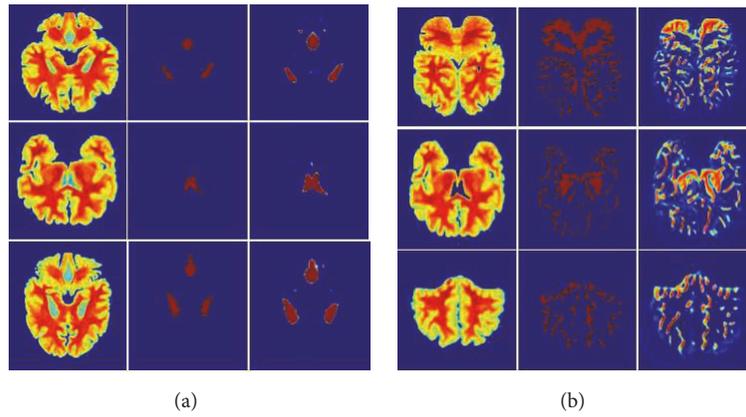


FIGURE 10: Testing procedure; from left to right are the brain MRI, labels from ground truth, and prediction. (a) CSF prediction result using U-net I; (b) wrong-classification area prediction result using U-net II.

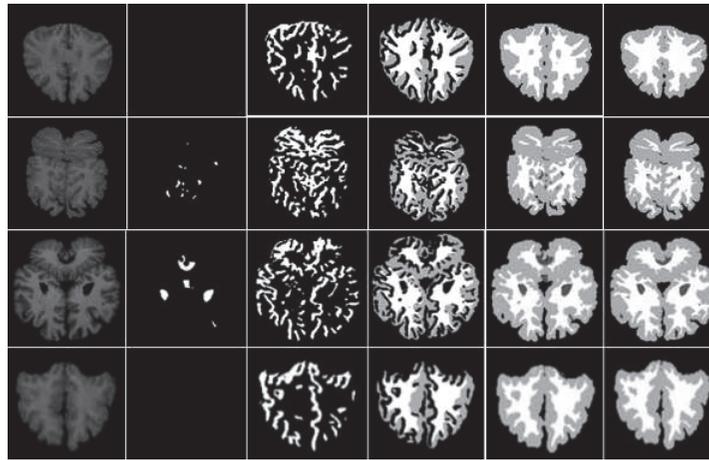


FIGURE 11: The results of each part of the whole algorithm architecture. From left to right are the original images, the segmentation results of CSF, prediction of the wrong-classification region, segmentation results using GMMD, proposed method, and ground truth.

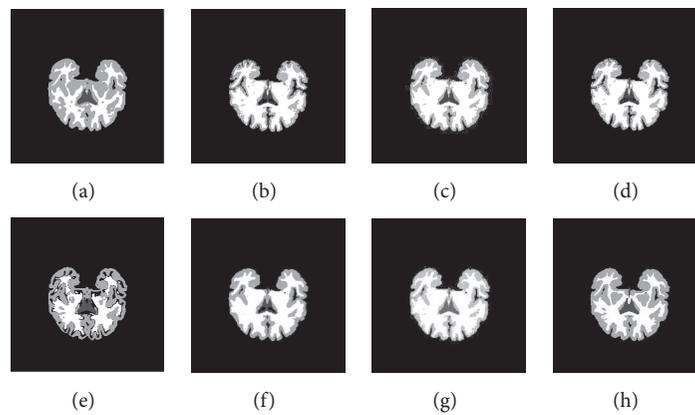


FIGURE 12: Comparison of different methods for brain MRI (slice 70 in ISBR 18). (a) Proposed method, (b) FCM, (c) GMM, (d) K-means, (e) classical U-net, (f) SMM-SC (see [31]), (g) CBCLO (see [32]), and (h) ground truth.

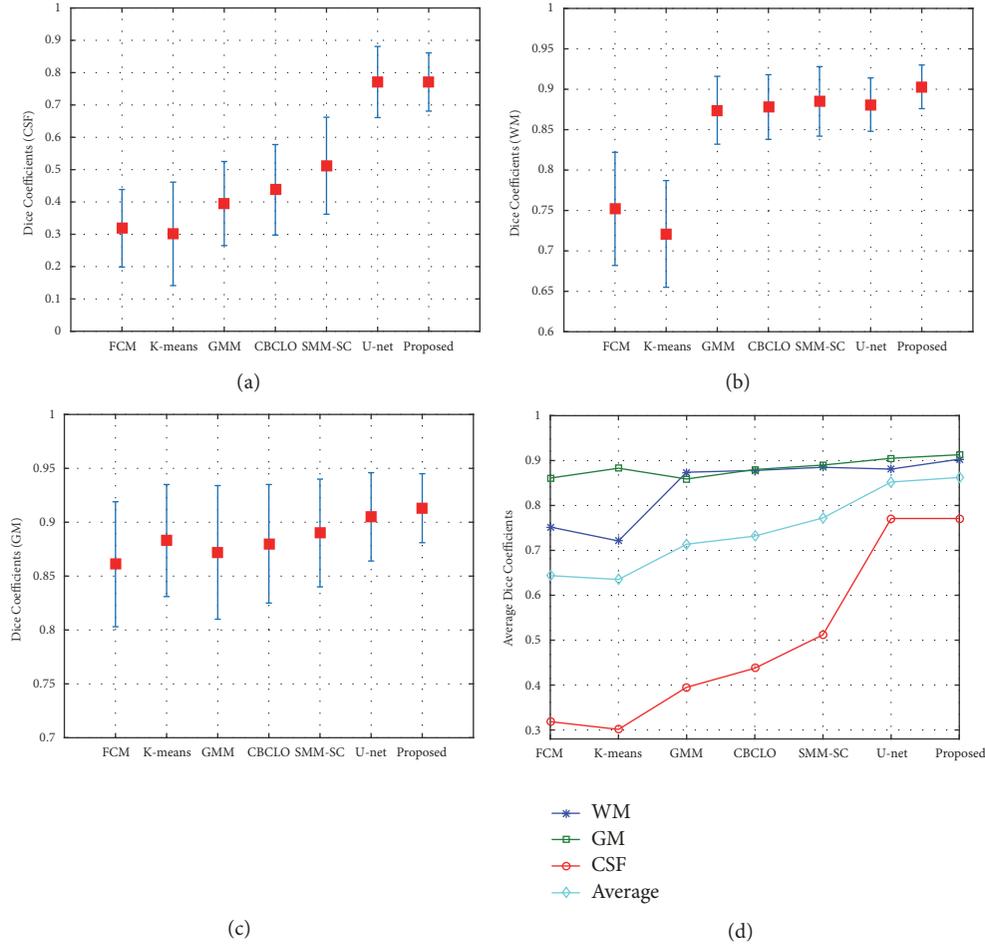


FIGURE 13: Comparison of segmentation performance on the ISBR 18 dataset (one patient). (a) CSF, (b) WM, (c) GM, and (d) average Dice coefficients.

Thus, there are 1000 images of size 255×255 that are selected from ten random subjects for training, and 300 images are adopted from other three subjects for testing. The ground-truth images have been divided into four parts: background, CSF, GM, and WM. Our experiments have been developed under the environment of MATLAB2016a and the Tensorflow library.

The experiment first runs the proposed framework on several brain MRIs (slice 92, slice 23, slice 40, and slice 96 in ISBR 18). The corresponding results of the four different patients are presented in Figure 11. These results indicate that our supervised network performs well in detecting the CSF and error-classification pixels in a visual medium. In addition, we observe, regardless of which slices are used, that the proposed model can obtain rich details. In the next experiment, the performance of GMMD-U is compared to that of FCM, GMM, K-means, SMM-SC (see [31]), CBCLO (see [32]), and classical U-net by using the brain MRI (slice 70 of one patient) in ISBR 18, and the results are provided in Figure 12. The classical U-net algorithm used here is divided into three parts for dealing with the segmentation task, which contains CSF, GM, and WM. Based on visual comparison with the ground truth, GMMD-U performs better than

several other comparing methods. SMM-SC method achieves better segmentation results compared to CBCLO.

In order to make a quantitative comparison of the different algorithms, this paper considers the Dice similarity coefficient [40] to evaluate the performance of segmentation models. The Dice coefficient is extensively accepted to evaluate the performance of segmentation algorithms in the image-processing field. The Dice coefficient of each part's prediction and ground truth can be calculated by

$$Dice(P_{gt}, P_{seg}) = \frac{2|P_{gt} \cap P_{seg}|}{|P_{gt}| + |P_{seg}|} \quad (29)$$

where P_{gt} means all pixel labels that belong to the ground truth of the brain image and P_{seg} represents the result of the predicted labels of each tissue. The Dice coefficient lies in the range $[0, 1]$, and a high value indicates high performance of the segmentation approach. The experiment determines the performance of five algorithms on the testing set. Figure 13 shows the Dice coefficient of the segmentation performance of FCM, K-means, GMM, SMM-SC (see [31]), CBCLO (see [32]), classical U-net, and GMMD-U on the brain MRI of one patient (ISBR 18). It is clear from this figure

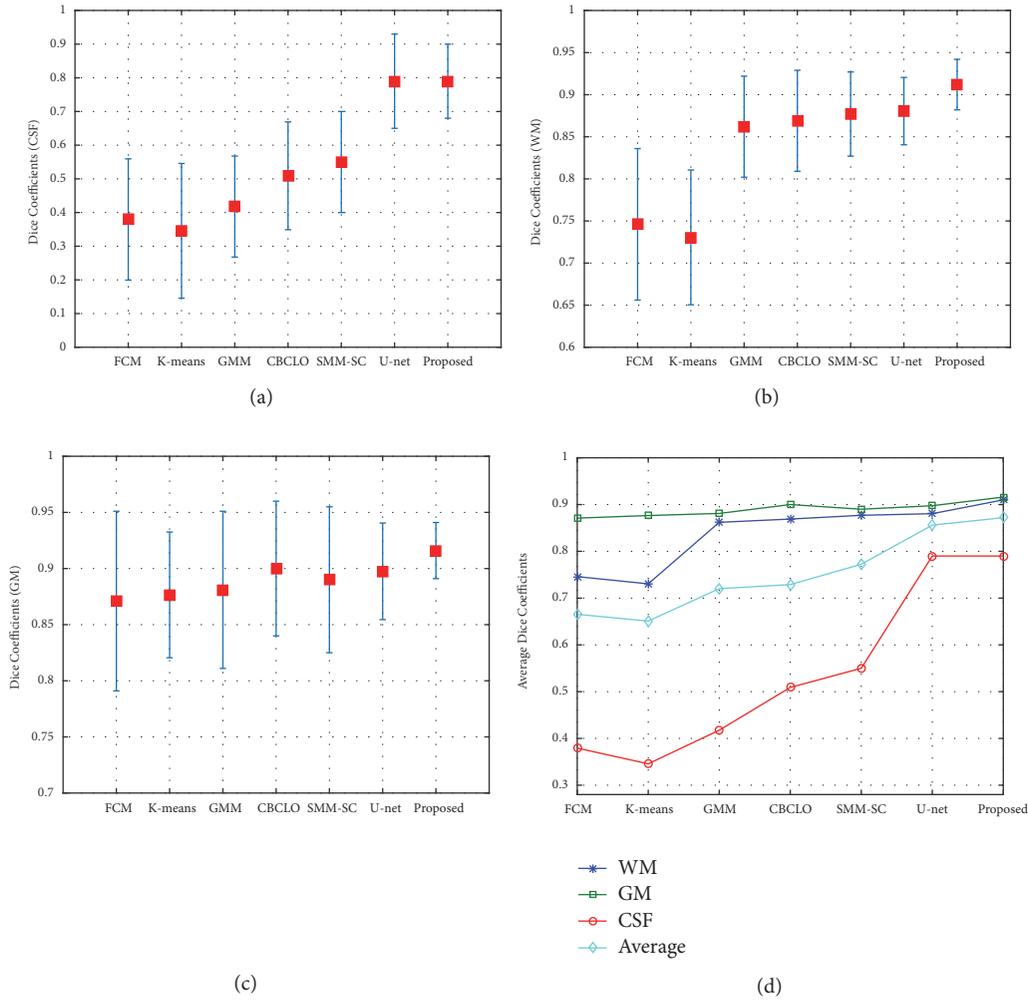


FIGURE 14: Comparison of segmentation performance on ISBR 18 (whole dataset). (a) CSF, (b) WM, (c) GM, and (d) average Dice coefficients.

that GMMD-U shows the highest Dice coefficient values among all algorithms and the smallest fluctuation range in each part of tissues. These results indicate that GMMD-U is more accurate than several related methods, such as classical U-net, SMM-SC (see [31]), CBCLO (see [32]), K-means, GMM, and FCM. Figure 14 shows the Dice coefficient obtained for the whole testing image dataset (ISBR 18) after applying the five methods. From this figure, consistently high values of Dice coefficients can be easily observed for the GMMD-U algorithm, which indicates that brain tissues are better recognized. The classical U-net model has higher Dice coefficients and achieves better accuracy than FCM, K-means, GMM, SMM-SC, and CBCLO, respectively.

4. Conclusions

In this paper, we proposed a new fusion framework for brain MRI segmentation based on two fully convolutional networks, modified U-net, and the Gaussian-Dirichlet mixture model with spatial constraints. Compared with other finite-mixture models, our proposed method can accurately

acquire different brain tissue, such as WM, GM, and CSF. The advantage of the proposed framework is that it incorporates the local spatial constraints through Dirichlet distribution. Furthermore, the modified U-net could accurately distinguish different tissue, which could be categorized with some fault decisions. The experimental results of the brain MRI from the ISBR 18 dataset illustrated the effectiveness of GMMD-U on segmentation tasks.

Data Availability

The data used to support the findings of this study have been deposited in the Internet Brain Segmentation Repository, <https://www.nitrc.org/projects/ibsr>. Please find them in [36] of this paper.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the National Nature Science Foundation of China (Grant no. 61872143).

References

- [1] S. Bonte, I. Goethals, and R. Van Holen, "Machine learning based brain tumour segmentation on limited data using local texture and abnormality," *Computers in Biology and Medicine*, vol. 98, pp. 39–47, 2018.
- [2] M. Soltaninejad, G. Yang, T. Lambrou et al., "Supervised learning based multimodal MRI brain tumour segmentation using texture features from supervoxels," *Computer Methods and Programs in Biomedicine*, vol. 157, pp. 69–84, 2018.
- [3] 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," *Journal of Neuroscience Methods*, vol. 270, pp. 111–123, 2016.
- [4] B. Zhao, Z. Cao, and S. Wang, "Lung vessel segmentation based on random forests," *IEEE Electronics Letters*, vol. 53, no. 4, pp. 220–222, 2017.
- [5] K. Kim, C. Oh, and K. Sohn, "Non-parametric human segmentation using support vector machine," *IEEE Transactions on Consumer Electronics*, vol. 62, no. 2, pp. 150–158, 2016.
- [6] R. S. Kamathe and K. R. Joshi, "A novel method based on independent component analysis for brain MR image tissue classification into CSF, WM and GM for atrophy detection in Alzheimer's disease," *Biomedical Signal Processing and Control*, vol. 40, pp. 41–48, 2018.
- [7] I. Cabria and I. Gondra, "MRI segmentation fusion for brain tumor detection," *Information Fusion*, vol. 36, pp. 1–9, 2017.
- [8] N. Mahata, S. Kahali, S. K. Adhikari, and J. K. Sing, "Local contextual information and Gaussian function induced fuzzy clustering algorithm for brain MR image segmentation and intensity inhomogeneity estimation," *Applied Soft Computing*, vol. 68, pp. 586–596, 2018.
- [9] J. Aparajeeta, S. Mahakud, P. K. Nanda, and N. Das, "Variable Variance Adaptive Mean-Shift and possibilistic fuzzy C-means based recursive framework for brain MR image segmentation," *Expert Systems with Applications*, vol. 92, pp. 317–333, 2018.
- [10] M. Bruneau, T. Mottet, S. Moulin et al., "A clustering package for nucleotide sequences using laplacian eigenmaps and gaussian mixture model," *Computers in Biology and Medicine*, vol. 93, pp. 66–74, 2018.
- [11] L. Kong, H. Zhang, Y. Zheng, Y. Chen, J. Zhu, and Q. M. Wu, "Image segmentation using a hierarchical student's-t mixture model," *IET Image Processing*, vol. 11, no. 11, pp. 1094–1102, 2017.
- [12] X. Li and K. N. Plataniotis, "Circular mixture modeling of color distribution for blind stain separation in pathology images," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 1, pp. 150–161, 2017.
- [13] H. Zhu, X. Pan, and Q. Xie, "Merging Student's-t and Rayleigh distributions regression mixture model for clustering time-series," *Neurocomputing*, vol. 266, pp. 247–262, 2017.
- [14] T. M. Nguyen and Q. M. Jonathan Wu, "Bounded asymmetrical Students-t mixture model," *IEEE Transactions on Cybernetics*, vol. 44, no. 6, pp. 857–869, 2014.
- [15] S. Shoham, "Robust clustering by deterministic agglomeration EM of mixtures of multivariate t-distributions," *Pattern Recognition*, vol. 35, no. 5, pp. 1127–1142, 2002.
- [16] M. Kwon, H.-H. Kwon, and D. Han, "A spatial downscaling of soil moisture from rainfall, temperature, and AMSR2 using a Gaussian-mixture nonstationary hidden Markov model," *Journal of Hydrology*, vol. 564, pp. 1194–1207, 2018.
- [17] J. Liu, X. Zhuang, L. Wu et al., "Myocardium segmentation from de mri using multicomponent gaussian mixture model and coupled level set," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 11, pp. 2650–2661, 2017.
- [18] S. S. Mohseni Salehi, D. Erdogmus, and A. Gholipour, "Auto-context convolutional neural network (auto-net) for brain extraction in magnetic resonance imaging," *IEEE Transactions on Medical Imaging*, vol. 36, no. 11, pp. 2319–2330, 2017.
- [19] P. Liu, J.-M. Guo, C.-Y. Wu, and D. Cai, "Fusion of deep learning and compressed domain features for content-based image retrieval," *IEEE Transactions on Image Processing*, vol. 26, no. 12, pp. 5706–5717, 2017.
- [20] F. Qin, N. Gao, Y. Peng, Z. Wu, S. Shen, and A. Grudtsin, "Fine-grained leukocyte classification with deep residual learning for microscopic images," *Computer Methods and Programs in Biomedicine*, vol. 162, pp. 243–252, 2018.
- [21] J. Han, D. Zhang, G. Cheng, N. Liu, and D. Xu, "Advanced deep-learning techniques for salient and category-specific object detection: a survey," *IEEE Signal Processing Magazine*, vol. 35, no. 1, pp. 84–100, 2018.
- [22] P. Moeskops, M. A. Viergever, A. M. Mendrik, L. S. De Vries, M. J. N. L. Benders, and I. Isgum, "Automatic segmentation of MR brain images with a convolutional neural network," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1252–1261, 2016.
- [23] S. Zahia, D. Sierra-Sosa, B. Garcia-Zapirain, and A. Elmaghraby, "Tissue classification and segmentation of pressure injuries using convolutional neural networks," *Computer Methods and Programs in Biomedicine*, vol. 159, pp. 51–58, 2018.
- [24] D. C. Ciresan, A. Giusti, L. M. Gambardella, and J. Schmidhuber, "Deep neural networks segment neuronal membranes in electron microscopy images," *Advances in Neural Information Processing Systems*, vol. 25, pp. 2843–2851, 2012.
- [25] Q. N. Vo, S. H. Kim, H. J. Yang, and G. S. Lee, "Text line segmentation using a fully convolutional network in handwritten document images," *IET Image Processing*, vol. 12, no. 3, pp. 438–446, 2018.
- [26] H. R. Roth, H. Oda, X. Zhou et al., "An application of cascaded 3D fully convolutional networks for medical image segmentation," *Computerized Medical Imaging and Graphics*, vol. 66, pp. 90–99, 2018.
- [27] E. Shelhamer, J. Long, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 4, pp. 640–651, 2017.
- [28] L. Bi, J. Kim, E. Ahn, A. Kumar, M. Fulham, and D. Feng, "Dermoscopic Image Segmentation via Multistage Fully Convolutional Networks," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 9, pp. 2065–2074, 2017.
- [29] O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention and Medical Image Computing and Computer-Assisted Intervention (MICCAI '15)*, vol. 9351 of *Lecture Notes in Computer Science*, pp. 234–241, November 2015.
- [30] J. Lai and H. Zhu, "A fusion algorithm: fully convolutional networks and Student's-T mixture model for brain magnetic

- resonance imaging segmentation,” in *Proceedings of the 25th IEEE International Conference on Image Processing (ICIP '18)*, pp. 1598–1602, October 2018.
- [31] T. M. Nguyen and Q. M. J. Wu, “Robust student’s-t mixture model with spatial constraints and its application in medical image segmentation,” *IEEE Transactions on Medical Imaging*, vol. 31, no. 1, pp. 103–116, 2012.
- [32] Y. Zhou and H. Zhu, “Image segmentation using a trimmed likelihood estimator in the asymmetric mixture model based on generalized gamma and gaussian distributions,” *Mathematical Problems in Engineering*, vol. 2018, Article ID 3468967, 17 pages, 2018.
- [33] N. Bouguila, D. Ziou, and J. Vaillancourt, “Unsupervised learning of a finite mixture model based on the Dirichlet distribution and its application,” *IEEE Transactions on Image Processing*, vol. 13, no. 11, pp. 1533–1543, 2004.
- [34] C. Nikou, A. C. Likas, and N. P. Galatsanos, “A Bayesian framework for image segmentation with spatially varying mixtures,” *IEEE Transactions on Image Processing*, vol. 19, no. 9, pp. 2278–2289, 2010.
- [35] W. Rudin, *Real and Complex Analysis*, McGraw-Hill Book, New York, NY, USA, 3rd edition, 1987.
- [36] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, New York, NY, USA, 2006.
- [37] The Internet Brain Segmentation Repository, <https://www.nitrc.org/projects/ibsr>.
- [38] P. de Boer, D. P. Kroese, and S. Mannor, “A tutorial on the cross-entropy method,” *Annals of Operations Research*, vol. 134, pp. 19–67, 2005.
- [39] Tensorflow library <https://github.com/tensorflow/tensorflow>.
- [40] L. R. Dice, “Measures of the amount of ecologic association between species,” *Ecology*, vol. 26, pp. 297–302, 1945.



Hindawi

Submit your manuscripts at
www.hindawi.com

