RFDCR

Chen, Gaoxiang; Li, Qun; Shi, Fuqian; Rekik, Islem; Wang, Li; Pan, Zhifang

Published in:
NeuroImage

DOI:
10.1016/j.neuroimage.2020.116620

Publication date:
2020

Document Version
Publisher's PDF, also known as Version of record

NeuroImage 211 (2020) 116620

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

RFDCR: Automated brain lesion segmentation using cascaded random forests with dense conditional random fields

Gaoxiang Chen [a], Qun Li [a], Fuqian Shi [b], Islem Rekik [c,d], Zhifang Pan [a,e,*]

[a] The First Affiliated Hospital of Wenzhou Medical University, Wenzhou 325000, China
[b] Rutgers Cancer Institute of New Jersey, Rutgers University, NJ 08903, USA
[c] BASIRA Lab, Faculty of Computer and Informatics, Istanbul Technical University, 34469 Istanbul, Turkey
[d] School of Science and Engineering, Computing, University of Dundee, Dundee DD1HN, UK
[e] Information Technology Center, Wenzhou Medical University, Wenzhou 325035, China

ARTICLE INFO

Keywords:
MRI
Lesions segmentation
Brain tumor
Ischemic stroke
Random forests
Conditional random fields

ABSTRACT

Segmentation of brain lesions from magnetic resonance images (MRI) is an important step for disease diagnosis, surgical planning, radiotherapy and chemotherapy. However, due to noise, motion, and partial volume effects, automated segmentation of lesions from MRI is still a challenging task. In this paper, we propose a two-stage supervised learning framework for automatic brain lesion segmentation. Specifically, in the first stage, intensity-based statistical features, template-based asymmetric features, and GMM-based tissue probability maps are used to train the initial random forest classifier. Next, the dense conditional random field optimizes the probability maps from the initial random forest classifier and derives the whole tumor regions referred as the region of interest (ROI). In the second stage, the optimized probability maps are further intergraded with features from the intensity-based statistical features and template-based asymmetric features to train subsequent random forest, focusing on classifying voxels within the ROI. The output probability maps will be also optimized by the dense conditional random fields, and further used to iteratively train a cascade of random forests. Through hierarchical learning of the cascaded random forests and dense conditional random fields, the multimodal local and global appearance information is integrated with the contextual information, and the output probability maps are improved layer by layer to finally obtain optimal segmentation results. We evaluated the proposed method on the publicly available brain tumor datasets BRATS 2015 & BRATS 2018, as well as the ischemic stroke dataset ISLES 2015. The results have shown that our framework achieves competitive performance compared to the state-of-the-art brain lesion segmentation methods. In addition, contralateral difference and skewness were identified as the important features in the brain tumor and ischemic stroke segmentation tasks, which conforms to the knowledge and experience of medical experts, further reflecting the reliability and interpretability of our framework.

1. Introduction

Segmentation of brain lesions from magnetic resonance images (MRI) plays an important role in clinical applications such as disease diagnosis and surgical planning. Manual segmentation of lesions is prone to inter-expert variability and low efficiency. Clinicians can only use rough measurements for quantification of tumor cells (Bauer et al., 2013). Thus, computer-aided automatic and semi-automatic segmentation methods have been widely proposed for the segmentation of different brain lesions. However, due to the unpredictable location and shape of lesions as well as their fuzzy boundaries, effective and robust automatic segmentation of brain lesions is still a challenging problem.

In recent years, many machine learning algorithms have been widely proposed for medical image analysis and interpretation, including generative probability modeling and discriminative modeling methods (Menze et al., 2016). When applying the generative model, it is often necessary to use brain anatomical information to model the lesion region as an abnormal value relative to other healthy tissues or as a mapping relationship with the inferred prior structure of the lesion. On the other hand, discriminative probability models can exactly learn the difference between the feature space of the lesion and other healthy tissues based on the intensities, which can be leveraged for feature extraction and classifier training. For feature design, the statistical analysis method (Xuan and Liao, 2007), which fused the structural information around image

* Corresponding author. The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China.
E-mail address: panzhifang@wmu.edu.cn (Z. Pan).

https://doi.org/10.1016/j.neuroimage.2020.116620
Received 7 November 2019; Received in revised form 11 January 2020; Accepted 6 February 2020
Available online 11 February 2020

voxels and extracted statistical features, was commonly used. Taking brain tumor segmentation as an example, features based on intensity, texture, and symmetrical information can help classify voxels into 'tumor' and other tissue types. However, these simple statistical features cannot fully capture complex tumor shapes. Gaussian mixture model (GMM) is a powerful multivariate modeling tool and regarded as a classic generation model for image segmentation (Permuter et al., 2006). Previous works (Ashburner and Friston, 2005; Greenspan et al., 2006; Balafar, 2014; Zikic et al., 2012; Tustison et al., 2015) have shown that using GMM to approximate the probability distribution of MRI is an established and effective way. The prior probabilities estimated by a trained GMM can also be used as additional features and integrated with the appearance information (Zikic et al., 2012). These prior features could improve the segmentation accuracy of normal tissues, but when estimating voxel classes for tumor regions with complex geometric shapes, the probabilistic results were sensitive to noise, illumination, and other factors. In addition to the features mentioned above, based on multimodal symmetric templates shown in Fig. 1, the asymmetry features and global context-sensitive features were proposed for brain tumor segmentation (Tustison et al., 2015). The disadvantage of this framework is a big computational burden for joint optimization of the probabilities of tumor regions and healthy tissue.

To fully utilize these different features, Random Forest (RF) (Breiman, 2001), which are inherently suited for multi-class problems, is often adopted for brain MRI segmentation (Wang et al., 2015; Geremia et al., 2011; Song et al., 2016; Mitra et al., 2014; Pustina et al., 2016). Particularly, incorporating contextual information largely boosted the performance of RF models. In (Mitra et al., 2014) , the classification performance of RF models using multimodal context-aware features for training is better than those based on single modality features. Their method was robust to the segmentation of ischemic stroke but has not been extended to multi-class lesions segmentation. An RF-based LINDA software package was proposed for stroke lesion segmentation (Pustina et al., 2016), this algorithm focused on the combination of voxel and its surrounding neighborhood information to achieve hierarchical optimization of the lesion prediction process. A system that was composed of a Restricted Boltzmann Machine for unsupervised feature learning, and a RF classifier, was also successfully applied to the segmentation of brain tumor and stroke lesions (Pereira et al., 2018). In addition to the excellent performance of RF, these models have further shown that models containing regional information can significantly improve the accuracy of segmentation. However, conventional RF classification architecture cannot make full use of global contextual information. RF can be used as the kernel of feature learning and integrated local and contextual information in multimodal images (Ma et al., 2018). Similarly, there was a dynamic multiscale tree-based glioma segmentation framework that embeds structured random forest (SRF) and Bayesian Network (BN) classifiers in a binary tree structure (Amiri et al., 2018a). Although compelling, this work was limited by the lack of prior features extraction methods related to lesions and combination of local and global features. Besides, with the exponential growth of the binary tree, the training of both nested SRF and BN classifiers is computationally expensive.

To address these limitations, inspired by the related works, we propose a novel brain lesion segmentation method by leveraging a cooperative optimization of cascaded RF and dense conditional random fields (CRF). Our proposed architecture can effectively integrate local appearance and global contextual information of multimodal MRI and iteratively improve the segmentation results. Specifically, the framework consists of two stages. In the first stage, GMM-based tissue probability maps are used together with appearance features to train an initial RF. Next, the output probability maps predicted by RF are inputted as prior information to a dense CRF which outputs more refined segmentation maps. In the second stage, dense CRF results are utilized as contextual information to hierarchically train a cascade of RFs in combination with the original statistical features and template-based asymmetric features. The main contributions of our method are listed below:

i) In the first stage, the GMM-based segmentation results obtained by unsupervised method provide prior probability features of the lesion regions. Furthermore, in the subsequent stage, these features are replaced by the posterior probability maps optimized by dense CRF.
ii) In addition to the proposed statistical features and probability map features, template-based asymmetric anatomical features are also embedded into our segmentation framework to train robust classifiers.
iii) Compared with the simple use of dense CRF as post-processing of classifiers, to further boost the performance of RF, through hierarchical integration of cascaded RF and dense CRF, the segmentation result is improved layer by layer.

2. Method

In this section, we introduce our framework for brain lesion segmentation in detail. The flowchart of the proposed method is shown in Fig. 2, which includes the training and testing stages. The implementation details of the overall framework will be discussed in the following subsections.

2.1. Preprocessing

First, due to the magnetic field inhomogeneity, the generated original

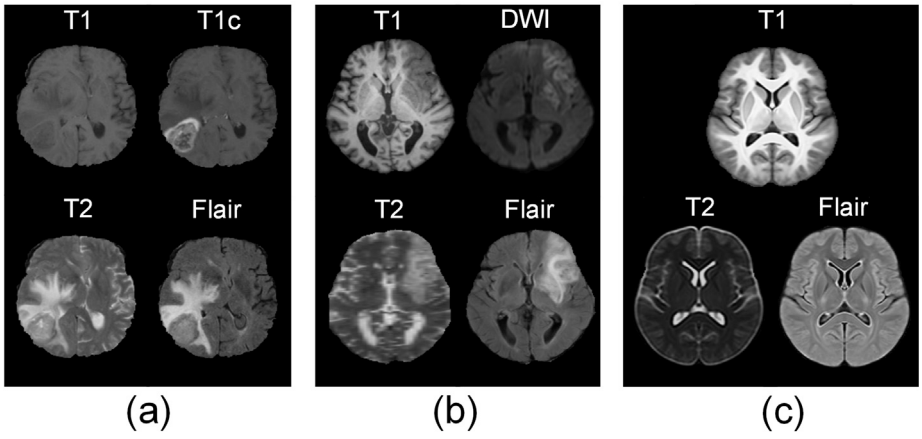

Fig. 1. Examples of multimodal brain lesions and symmetry templates. (a) brain tumor, (b) ischemic stroke, (c) symmetry templates of three modalities, where T1 component of templates is corresponding to both T1 and T1c of lesion images.

MRI suffer from brightness inhomogeneity, also known as bias field. Thus, we applied N4 bias field correction (Tustison et al., 2010) to correct intensity inhomogeneity. In addition, according to (Nyúl et al., 2000), voxel intensities in MRI of different patients are different, thus it is necessary to normalize intensities across all subjects by rescaling to the range [0, 1].

2.2. Feature extraction

Following image preprocessing, for training a robust and effective segmentation model, we first extract diverse image features from the input MRI modalities (T1, T1c, T2, and Flair) including intensity-based statistical features, template-based asymmetry features, and GMM (later CRF) based probability features.

2.2.1. Intensity based statistical features

The statistical features based on normalized intensities are calculated in each region of interest (ROI). Specifically, ROI is referred as the whole brain in stage 1 and lesion regions in stage 2, respectively. These features encode intensity differences between the tumor and normal tissues,

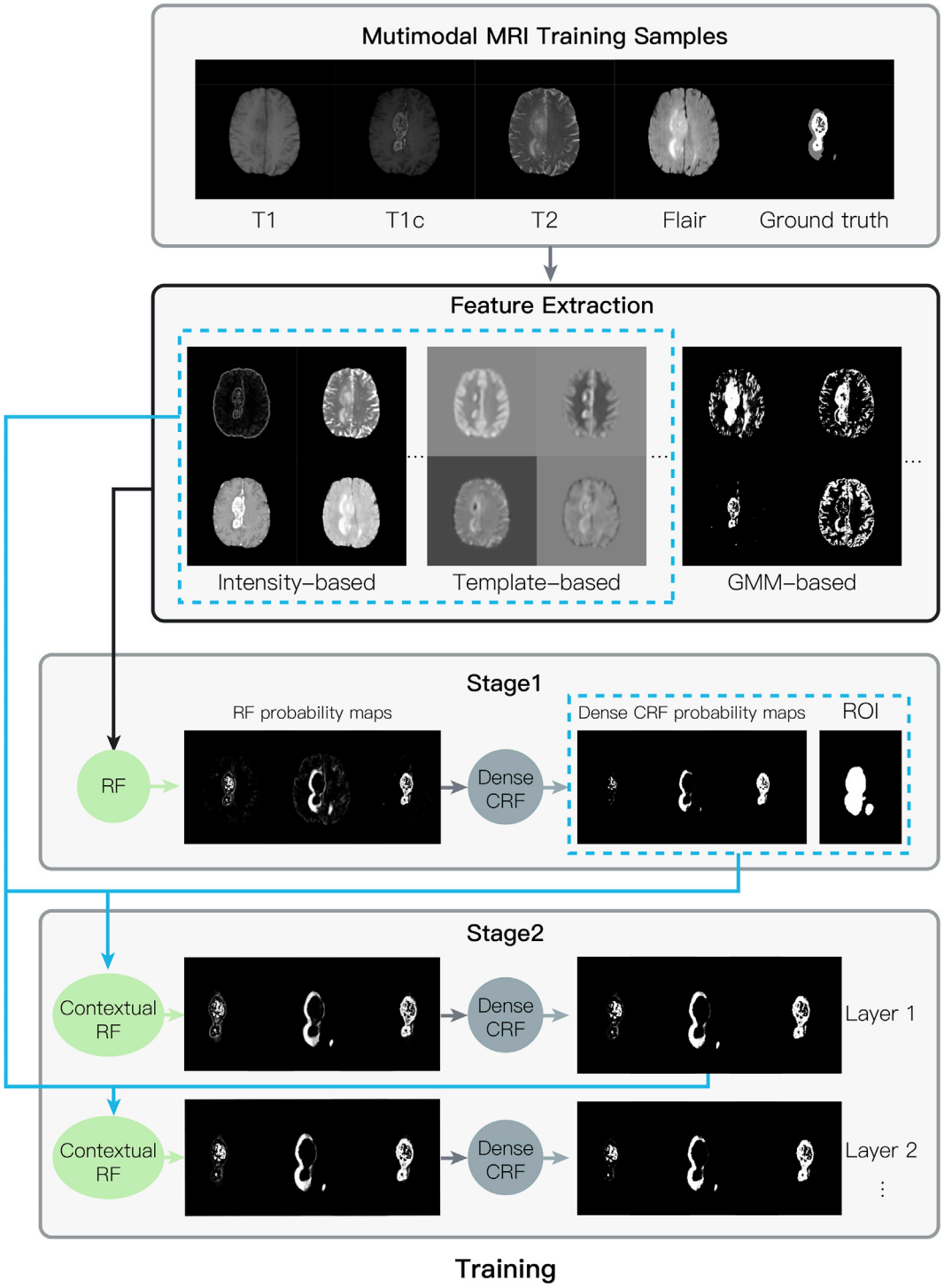

Fig. 2. Overview of the proposed cooperative training and testing framework integrating cascaded random forests (RF) and dense conditional random fields (Dense CRF) for brain tumor segmentation.

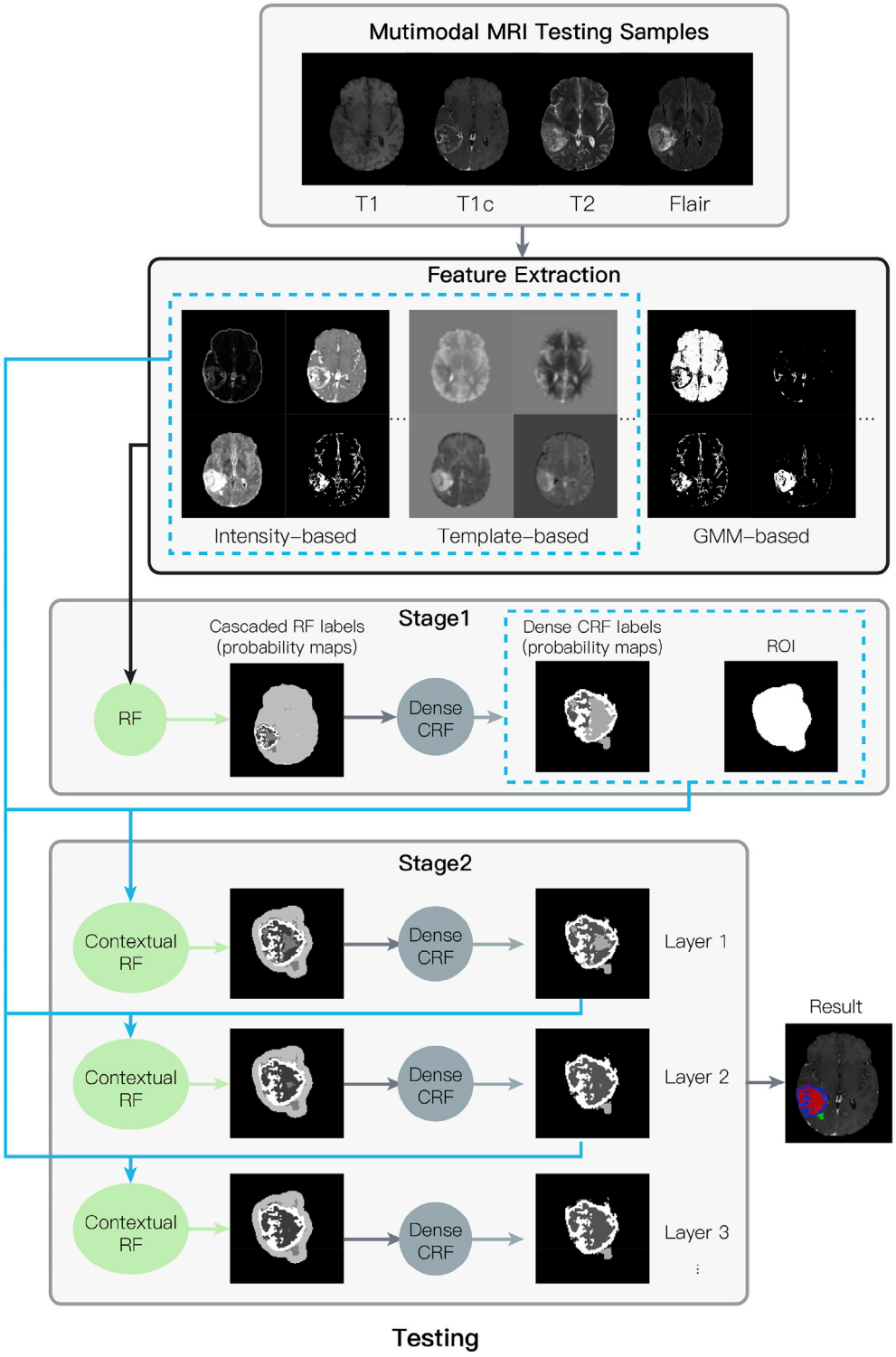

Fig. 2. (*continued*).

including the mean, minimum, maximum, variance, standard deviation, skewness, kurtosis, and entropy. Inspired by (Geremia et al., 2011), based on the multimodal MRI, we can also generate the intensity difference between one modality with a better intensity visualization of lesions and another modality with a better anatomical structure visualization, which can provide complementary features for the classifier learning.

2.2.2. *Template based asymmetry features*

Brain lesions alters not only the local shape of healthy brain tissues but also creates a global hemispheric asymmetry, where the intensities of voxels in the affected hemisphere differ from those in the contralateral hemisphere. Hence, brain symmetric templates (Tustison et al., 2015), can be used to extract additional template-based difference features. Specifically, a symmetric template X_{temp} for a multimodal normal human brain is constructed from the "MMRR" dataset (Landman et al., 2011) as in (Tustison et al., 2015), X_{temp} is registered to the same space as target patient to obtain an asymmetric warped template image X_{asym} by Advanced Normalization Tools (ANTs) (Avants et al., 2014). Next, the

template difference feature f_{temp_diff} is defined as the difference between the original lesion image X_{lesion} and the corresponding asymmetric template image X_{asym} to quantify the brain asymmetry caused by the present lesions in the acquired MR images. It is given as $f_{temp\_diff} = X_{lesion} - X_{asym}$, $X_{asym} = \varphi(X_{temp})$, where φ denotes the mapping from X_{temp} to X_{lesion} space.

As another feature based on asymmetry, the contralateral difference feature f_{contra_diff} is calculated as $f_{contra\_diff} = X_{lesion} - X_{contra}$, where X_{contra} denotes the corresponding contralateral image which can be defined as $X_{contra} = \varphi_b(X_{lesion})$, where φ_b consists of two components that X_{lesion} is first flipped with reflection transformation, then warped it to the space of original lesion image X_{lesion} via the affine transformation. Further details of these asymmetry features we used can be found in (Tustison et al., 2015).

2.2.3. GMM based features

Previous work (Zikic et al., 2012; Tustison et al., 2015) have shown that it is useful to model the appearance information as additional input for training a classifier. For instance, GMM can be used to model the probability of each tissue class. Taking brain tumors as an example, we apply the GMM to MRI of all modalities and design features as in (Tustison et al., 2015). Each lesion image is divided into four classes: edema (ED), enhancing tumor (ET), necrosis (NCR), and background (BG). The intensity distribution of each tissue class is initialized as a Gaussian distribution, and then the expectation–maximization (EM) algorithm (Moon, 1996) is used to estimate all the target distribution's parameters, and the initial probability maps of each tissue are used as a new set of image features to train the classifier.

2.3. Classification

Random Forest (RF), as a classic and effective classifier for brain lesion segmentation, presents a fundamental component of our lesion segmentation framework. In addition, dense conditional random fields (CRF) which are often used as a postprocessing step to achieve more spatially contiguous segmentations, is embedded into our framework as an optimizer.

For RF-based multilayer cascaded RFs, the features used for training are still based on independent voxels and their neighbors, while for the dense CRF, the inference process implicitly assumes the conditional independence between voxels, which may lose correlation constraints between directly adjacent voxels. In order to address the limitations of these two models, our proposed multilayer collaborative learning method can further integrate voxel-level appearance information and high-level context in cascaded RF with global-level constraints in dense CRF.

A detailed explanation of how to design the proposed ensemble classifier framework of cascaded RF and dense CRF for brain lesion segmentation will be given in the following sections.

2.3.1. Cascaded random forests

For each given image $I = \{I_1, I_2, \ldots, I_n\}$ $y = \{y_1, y_2, \ldots, y_n\}$, n represents the number of voxels and I_i represents the intensity of voxel i. y_i is the label corresponding to voxel i, and is defined in the label set $L = \{l_1, l_2, \ldots, l_m\}$, where m represents the number of class labels. The high-dimensional feature set corresponding to training images is denoted as $f(I, A)$, where A represents the image set from different modalities.

In the construction process of RF, the samples used for training each tree are obtained by the bootstrap method from the original samples, not all the training samples. The random selection of features is a hallmark of RF. When splitting the nodes of each decision tree, the feature subspace is selected instead of the best attributes in all feature sets to split the input data into subsequent nodes. The randomness introduced by the bootstrap method and random feature selection alleviate the overfitting issue of RF model. When applied to test data, the feature vector f of voxel i is extracted in the same way and input to the trained model. As each decision tree in RF is independent of each other, the final prediction probability will generally be calculated by simply averaging the respective classification probabilities of all trees T defined as:

$$P(y_i|I_i) = \frac{1}{T}\sum_{t=1}^{T} P_t(y_i|f(I_i, A)), \quad t \in [1, T] \tag{1}$$

To improve classification performance, we integrate the idea of auto-context (Tu and Bai, 2010) in traditional RF. Auto-context iteratively learns the low-level image appearance information and rich context information by constructing a plurality of discriminative classifiers. First, an initial classifier is trained to classify voxels of each training image, generating a batch of probability maps of training set. In addition to the features directly calculated from the images, these probability maps containing context information are also used to train new classifiers. In each learning process of context information, the classifier can select a different candidate context location in the probability maps, recalculate the posterior marginal distribution, and increase or reduce the classification probability of the current voxel to make it converge to the ground truth. The image appearance information and high-level context are implicitly integrated, the balance between them is automatically learned during the training process.

By introducing this iterative learning into RF, a series of cascaded RFs can be learned. As defined before, we can train a traditional RF classifier by the training set:

$$K_0 = \left\{ \left(y_{h,i},\ f_h(I_i) \right), h = 1, 2, \ldots, H; i = 1, 2, \ldots, n \right\} \tag{2}$$

where H is the number of all images in the training set. To further simplify the representation, we assume that there is only one training image in the training set:

$$K_0 = \{ (y_i,\ f(I_i)), i = 1, 2, \ldots, n \} \tag{3}$$

By adding auto-context algorithm to RF, we apply the trained initial RF to the training set K_0 to generate the probability maps of lesion tissue including the three tumor tissue types: $P^0 = \{P^0_{ED}, P^0_{ET}, P^0_{NCR}\}$. With these maps, a new training set $K_1 = \{(y_i, (f(I_i), P^0(i))), i = 1, 2, \ldots, n\}$ can be constructed, where $P^0(i)$ represents the posterior probability corresponding to voxel i. Next, we can train a new RF with the training set K_1. During the iterative training of each layer, RF learns and integrates more spatial contextual information from voxels' neighborhoods. Once the probability maps are obtained, the same process is repeated as before to train the new classifier until convergence. Note that the only difference between later iterations and the first iteration is that we start the tissue probability maps from a rough estimate of GMM as mentioned in 2.2.3. Thus, in our method, the initial training set can be defined as:

$$K_0 = \left\{ \left(y_i, \left(f(I_i), P^{GMM}_{ED}(i), P^{GMM}_{ET}(i), P^{GMM}_{NCR}(i) \right) \right), i = 1, 2, \ldots, n \right\} \tag{4}$$

2.3.2. Dense conditional random fields

In the classification stage, RF independently predicts the corresponding label of each input voxel. However, such predictions at the voxel-scale can be very noisy since the splitting of each individual tree in RF is inherently noisy. This can result in incorrect labels and holes. Hence, a post-processing step of the output label probability maps is highly desired in RF-based segmentation frameworks.

CRF, as a classic probabilistic graph discrimination model, can remove noise in segmentation results and provide a more contiguous image segmentation with a smoother boundary (Zhao et al., 2016). It is well known that adjacent CRF structures generated by basic CRF model only consider the correlation between neighborhood voxels, which usually excessively smoothes the segmentation maps. As a variant of CRF, dense CRF considers each voxel as a node, and each node has a connection relationship with all other nodes. By converting binary constraints in dense CRF into a fast filtering of CRF, the energy functions containing all voxel labels and location information can be learned and inferred with $O(N)$ complexity (Krähenbühl and Koltun, 2012).

In our research, the extended 3D dense CRF model (Kamnitsas et al., 2017) is adopted, which has higher computational efficiency and is more suitable for 3D multimodal medical images than simple CRF. It was mostly used in post-processing for deep learning frameworks to eliminate small isolated regions (Kamnitsas et al., 2017; Liu et al., 2018; Zhao et al., 2018).

We represent the conditional random field as $\{I, y\}$, which can be modeled as a Gibbs distribution:

$$P(y|I) = \frac{1}{Z(I)}\exp\left(-\sum_{c\in C_G}\psi_c(y_c|I)\right) \tag{5}$$

where $Z(I)$ is a normalization term, G represents an undirected map of the random field, C_G means the cliques that represent all subsets of nodes with edges connected between any two nodes in G, and ψ_c is the potential function of clique C.

For label combination $y \in L$ for an input image I, the Gibbs energy function in the dense CRF model is defined as:

$$E(y|I) = \sum_{c\in C_G}\psi_c(y_c|I) \tag{6}$$

By finding the optimal label combination y^* to satisfy $y^* = argmin_{y\in L} E(y|I)$, the segmentation problem is transformed into an energy minimization problem.

Assuming that the voxel intensity distribution of the input image is a fixed condition, the representation of the condition I can be omitted for convenience of presentation, thus total energy function $E(y)$ is given as:

$$E(y) = \sum_i \psi_u(y_i) + \sum_{ij}\psi_p(y_i, y_j) \tag{7}$$

where $\psi_u(y_i)$ is the unary energy potential function, that represents the energy calculated independently for each voxel i by the result of the classifier when the relationship between voxels is not considered, showing the difference between voxel and the assigned label class. Thus, the unary potential is defined as $\psi_u(y_i) = -\log P(y_i)$. Note that $P(y_i)$ is computed by the coarse segmentation probability maps of RF in our framework.

And $\psi_p(y_i, y_j)$ is the pairwise energy potential function, which represents the energy of the voxels i, j simultaneously assigned to the labels y_i, y_j. Each voxel class is related to all other voxel classes and feature vectors describing the difference in voxel intensity and spatial distance. This potential function imposes mutual constraints on the classification of any voxel i, j and the constraints are also generally seen as a penalty when two voxels are assigned different labels. Especially when the two voxels are more similar but the corresponding labels are different, the closer their distance is, the larger the penalty. Therefore, the pairwise potential often has the following form:

$$\psi_p(y_i, y_j) = \mu(y_i, y_j)k(f_i, f_j) \tag{8}$$

where f_i and f_j are the feature vector of the voxel i and j in any feature space, and a label compatibility function $\mu(y_i, y_j)$ is given as:

$$\mu(y_i, y_j) = \begin{cases} 1, & y_i = y_j \\ 0, & y_i \neq y_j \end{cases} \tag{9}$$

To apply the model to multimodal scans of medical images, two Gaussian filter kernels, the smoothness kernel $k^{(1)}(f_i, f_j)$ and the appearance kernel $k^{(2)}(f_i, f_j)$, were proposed, both depending on the intensities of different modal images for voxel i (denoted as $I_{i,a}$, a represents the a-th modality) and the voxel coordinates (denoted as $p_{i,r}$):

$$k(f_i, f_j) = \omega_1 k^{(1)}(f_i, f_j) + \omega_2 k^{(2)}(f_i, f_j) \tag{10}$$

$$k^{(1)}(f_i, f_j) = \exp\left(-\sum_{r=\{x,y,z\}}\frac{|p_{i,r} - p_{j,r}|^2}{2\theta_{\alpha,r}^2}\right) \tag{11}$$

$$k^{(2)}(f_i, f_j) = \exp\left(-\sum_{r=\{x,y,z\}}\frac{|p_{i,r} - p_{j,r}|^2}{2\theta_{\beta,r}^2} - \sum_{a=1}^{A}\frac{|I_{i,a} - I_{j,a}|^2}{2\theta_{\gamma,r}^2}\right) \tag{12}$$

The hyperparameters $\theta_{\alpha,r}$ and $\theta_{\beta,r}$ control the similarity of neighborhoods that can have homogeneous labels. In the channels A of the input modality, $\theta_{\gamma,r}$ quantifies the scale of uniform appearance. ω_1 and ω_2 are combined weight of the two kernels.

The smoothness kernel $k^{(1)}(f_i, f_j)$ is only related to the locational voxel relationship, focusing on spatial proximity during smoothing, and is used to remove small isolated regions. The appearance kernel $k^{(2)}(f_i, f_j)$ depends on both location and intensity relationship between voxels by measuring the similarity between voxels, and forcing nodes of similar strength and location to have a consistent classification.

In the case of full connectivity, the number of pairwise energy terms is excessively large, which makes it difficult to infer such random fields using traditional algorithms until an effective inference algorithm was proposed (Krähenbühl and Koltun, 2012). First, a simple distribution $R(y)$, where y labels are independent of each other, is used to replace the original CRF voxel distribution $P(y)$:

$$R(y) = \prod_i R_i(y_i) \tag{13}$$

Then, the Kullback–Leibler divergence of R and P is minimized:

$$KL(R||P) = \sum_y R(y)ln\left(\frac{R(y)}{P(y)}\right) \tag{14}$$

This leads to the following iterative update equation:

$$R_i(y_i = l) = \frac{1}{Z_i}exp\left\{-\psi_u(y_i) - \sum_{l'}\sum_{i\neq j}\psi_p(y_i, y_j)R_j(l')\right\} \tag{15}$$

The inference and learning of the dense CRF model mainly take as inputs both multimodal MRI and predicted probability maps of RF to finally output the target label y.

2.3.3. Combination of cascaded RF and dense CRF

In the following, we introduce our framework with cascaded RF and dense CRF for brain lesion segmentation, as illustrated in Fig. 2. After preprocessing the original multimodal images as described in Section 2.1, the proposed segmentation process consists of two stages.

In the first stage, appearance features (statistical intensity information of the whole brain region and template-based features) and contextual features (GMM-based lesion tissue probability maps) are used for training the first initial RF classifier. The prediction probability map obtained by the classifier is input as a prior into the dense CRF model for optimization, the small irrelevant and isolated regions around the lesion region are removed. Therefore, the approximate location of the whole lesion can be detected in the first stage, and the initial segmentation result is output.

In the second stage, a multilayer cooperative optimization architecture of cascaded RF and dense CRF is proposed, which is a refined segmentation process within the region obtained from the results of the first stage. The lesion mask obtained in the first stage is dilated as a new region of interest (ROI). The regional appearance features are extracted from the set of features of the first stage masked by ROI, and the multiclass regional RF is trained using the output segmentation probability map of previous dense CRF instead of GMM-based initial tissue probability map as the new context information. Then, the dense CRF model is used to optimize the results of RF in the same layer, with focus on encouraging voxels with similar locations and gray scales to correspond to consistent labels, and output probability map for improving the RF in

the next layer. The next iteration process is similar to that of the first layer where cascaded RFs are trained by features based on intensity and template along with the probability maps from the output of the previous layer. Algorithm 1 details the training process with cascaded RFs and dense CRF.

Algorithm 1. Training process of cascaded Random Forests integrated with Dense Conditional Random fields (RFDCR).

INPUT: Multimodal MRI lesion images and corresponding ground truth $\{I, y\}$

1. Extract image features and make a training set with GMM-based prior $K_0 = \{(y_i, (f(I_i), P^{GMM}(i)), i = 1, 2, \ldots, n\}$.
2. Train a RF classifier with both appearance features and contextual features from $f(I_i)$ and $P^{GMM}(i)$.
3. Apply the previous RF to the training set to calculate probability images $P_0^{RF}(i)$ and as a prior of dense CRF to output new probability maps $P_0^{dCRF}(i)$, which can be processed as ROI for the next stage.
4. **For** Layer q = 1 …, Q:
   - Make a new training set $K_q = \{(y_i, (f(I_i), P_{q-1}^{dCRF}(i))), i = 1, 2, \ldots, n\}$.
   - Train a new RF with K_q.
   - Apply the trained RF to the training set to calculate probability images $P_q^{RF}(i)$ and as a prior of dense CRF to output new probability maps $P_q^{dCRF}(i)$.
5. A series of RF and dense CRF are trained to predict the segmentation probability maps $P^{RF}(i)$, $P^{dCRF}(i)$ and final segmentation labels in the testing process.

3. Experiments and results

In this section, we first describe the public datasets used to evaluate our framework, then select metrics to evaluate the performance of our proposed method and benchmark against other state-of-the-art methods.

3.1. Datasets and metrics

The proposed framework was evaluated on two brain lesion segmentation tasks using two public multimodal MRI datasets: (1) the brain tumor segmentation dataset (BRATS 2015, 2018) (Menze et al., 2015), and (2) the ischemic stroke segmentation dataset (ISLES 2015) (Maier et al., 2017).

First, for brain tumor segmentation, the experiments were performed on the dataset provided by the Brain Tumor Segmentation Challenge (BRATS) from MICCAI. This series of datasets consist of four MRI modalities acquired from each patient, including T1, T1c, T2, and Flair scans, all sequences have been skull-stripped and rigidly co-registered to T1c scans, and resampled to 1*1*1 mm^3 isotropic resolution. To verify the applicability of our method on multiple datasets, we chose the challenge datasets of 2015 and 2018, which contain MRI scans of 220 and 210 patients respectively with high-grade gliomas, to train and test our framework.

Next, the sub-acute ischemic stroke lesion segmentation (SISS) dataset of the 2015 Ischemic Stroke Lesion Segmentation (ISLES) challenge at MICCAI was used to test ischemic stroke segmentation. In the ISLES 2015 training process, 28 cases of multimodal MRI sequences including T1, DWI, T2 and Flair were available, and the test dataset consisted of 36 subjects. They were acquired from two MRI centers with 3T Phillips systems and have been preprocessed by researchers as follows: all the data were skull-stripped, rigidly co-registered to their Flair scans, and b-spline-resampled to a resolution of 1*1*1 mm^3 isotropic resolution.

The brain tumor and ischemic stroke segmentation performance is evaluated using the segmentation evaluation metrics of BRATS and ISLES, respectively. Both include Dice, precision, and sensitivity (Menze et al., 2015), which can be defined by segmented lesion region S and the manually labeled region M as follows:

$$Dice(S, M) = \frac{2|S \cap M|}{|S| + |M|} \tag{16}$$

Table 1
Number of features extracted for our two-stage tumor segmentation framework.

Features Name	First Stage	Second Stage
Statistical features	80	80
Asymmetry features	11	11
Probability features (GMM/CRF)	40	4 (tumor class + background)
Total	131	95

$$Precision(S, M) = \frac{|S \cap M|}{|S|} \tag{17}$$

$$Sensitivity(S, M) = \frac{|S \cap M|}{|M|} \tag{18}$$

and ISLES has the additional evaluation of Hausdorff distance and average symmetric surface distance (ASSD), which are calculated online on the challenge's website[1] and a detailed definition can be found in (Maier et al., 2017).

3.2. Implementation and parameters

The implementation of the RF was performed using the R package provided by CRAN.[2] Some parameters of the RF classifier used in our experiments were derived from (Lefkovits et al., 2016). The main parameters are set as follows: the number of trees is 200, the number of random sampling regions for each class is 2000. For multi-class brain tumor, the randomly selected features from the input feature set are of 131 and 95 at each stage, respectively, as presented in Table 1. The parameters of the dense CRF were optimized by random searching.

All implementations were carried out on our system (Intel(R) Xeon(R) CPU E5-2676 v3 @ 2.40 GHz). The feature extraction required 7–8 min on average for each patient, the RF training on a single-core CPU (non-parallel implementation) required an average of 8 h in each layer of our framework, whereas the testing of each patient (including the CRF process) required 3–4 min.

3.3. Evaluation on BRATS datasets

In our implementation, we first tested the effectiveness of our framework on BRATS2015 and BRATS 2018. Throughout the evaluation process, we applied 5-fold cross-validation and evaluated the segmentation accuracy in terms of three tumor regions, i.e., the whole tumor (edema + necrosis + enhancing tumor, WT), the tumor core (necrosis + enhancing tumor, TC), and the enhancing tumor (ET).

A series of experiments were carried out on the BRATS datasets. First, we tested the performance of cascaded RF and the optimization effect of the dense CRF on the framework and discussed the influence of the region of interest (ROI) on the segmentation results, followed by the comparison with other top methods.

3.3.1. Impact of cascaded random forests

Table 2 lists the results of the evaluation about the general cascaded RF model with GMM-based tissue features on BRATS 2018. In order to compare with the proposed segmentation framework integrated with dense CRF, the baseline cascaded RF is also trained in two stages. The feature images are extracted in the first stage and the output probability maps of RF are used as the context information to train the first layer model of the second stage. Next, the multilayer classifier is continuously trained until the segmentation results converge. As presented in Table 2 and Fig. 3, with the increase in the cascaded RF levels, the performance of the classifier improves and converges at the third layer.

[1] http://www.isles-challenge.org/.
[2] https://www.r-project.org/.

Table 2
Evaluation of hierarchical segmentation performance of the cascaded RF and dense CRF model trained in three different conditions on BRATS 2018.

Model	Stage		Dice			Precision			Sensitivity		
			WT	TC	ET	WT	TC	ET	WT	TC	ET
cascaded RF	Stage1	RF	0.67	0.65	0.62	0.59	0.58	0.55	0.91	**0.85**	0.85
	Layer1	RF	0.69	0.67	0.64	0.60	0.61	0.57	**0.92**	0.84	**0.86**
	Layer2	RF	0.70	0.69	0.66	0.61	0.64	0.59	**0.92**	0.83	0.85
	Layer3	RF	**0.72**	**0.70**	**0.68**	**0.63**	**0.66**	**0.61**	0.91	0.81	0.85
cascaded RF	Stage1	RF	0.67	0.65	0.62	0.59	0.58	0.55	0.91	0.85	**0.85**
+		dCRF	**0.82**	0.79	**0.75**	**0.83**	0.82	**0.74**	0.84	0.81	0.81
dense CRF	Layer1	RF	0.75	0.78	0.72	0.67	0.74	0.68	0.94	**0.87**	0.84
		dCRF	**0.82**	**0.81**	0.72	0.78	**0.83**	**0.74**	0.90	0.83	0.81
	Layer2	RF	0.77	0.80	0.74	0.68	0.80	0.72	**0.95**	0.86	0.81
		dCRF	0.81	0.80	**0.75**	0.76	**0.83**	**0.74**	0.92	0.83	0.77
	Layer3	RF	0.77	0.80	0.74	0.68	0.80	0.72	**0.95**	0.86	0.81
		dCRF	0.81	0.80	0.72	0.76	**0.83**	**0.74**	0.92	0.83	0.77
cascaded RF	Stage1	RF	0.67	0.65	0.62	0.59	0.58	0.55	**0.91**	0.85	**0.85**
+		dCRF	0.82	**0.79**	**0.75**	0.83	**0.82**	**0.74**	0.84	0.81	0.81
dense CRF	Layer1	RF	0.82	0.73	0.71	0.77	0.68	0.68	**0.91**	**0.86**	0.81
+		dCRF	0.85	0.77	0.73	0.83	0.75	0.71	0.90	**0.86**	0.82
ROI	Layer2	RF	0.84	0.77	0.74	0.81	0.74	0.72	**0.91**	**0.86**	0.82
		dCRF	0.85	0.78	**0.75**	**0.84**	0.77	**0.74**	0.90	**0.86**	0.83
	Layer3	RF	0.85	0.78	**0.75**	0.81	0.76	0.73	**0.91**	**0.86**	0.82
		dCRF	**0.86**	**0.79**	**0.75**	**0.84**	0.78	**0.74**	**0.91**	**0.86**	0.83

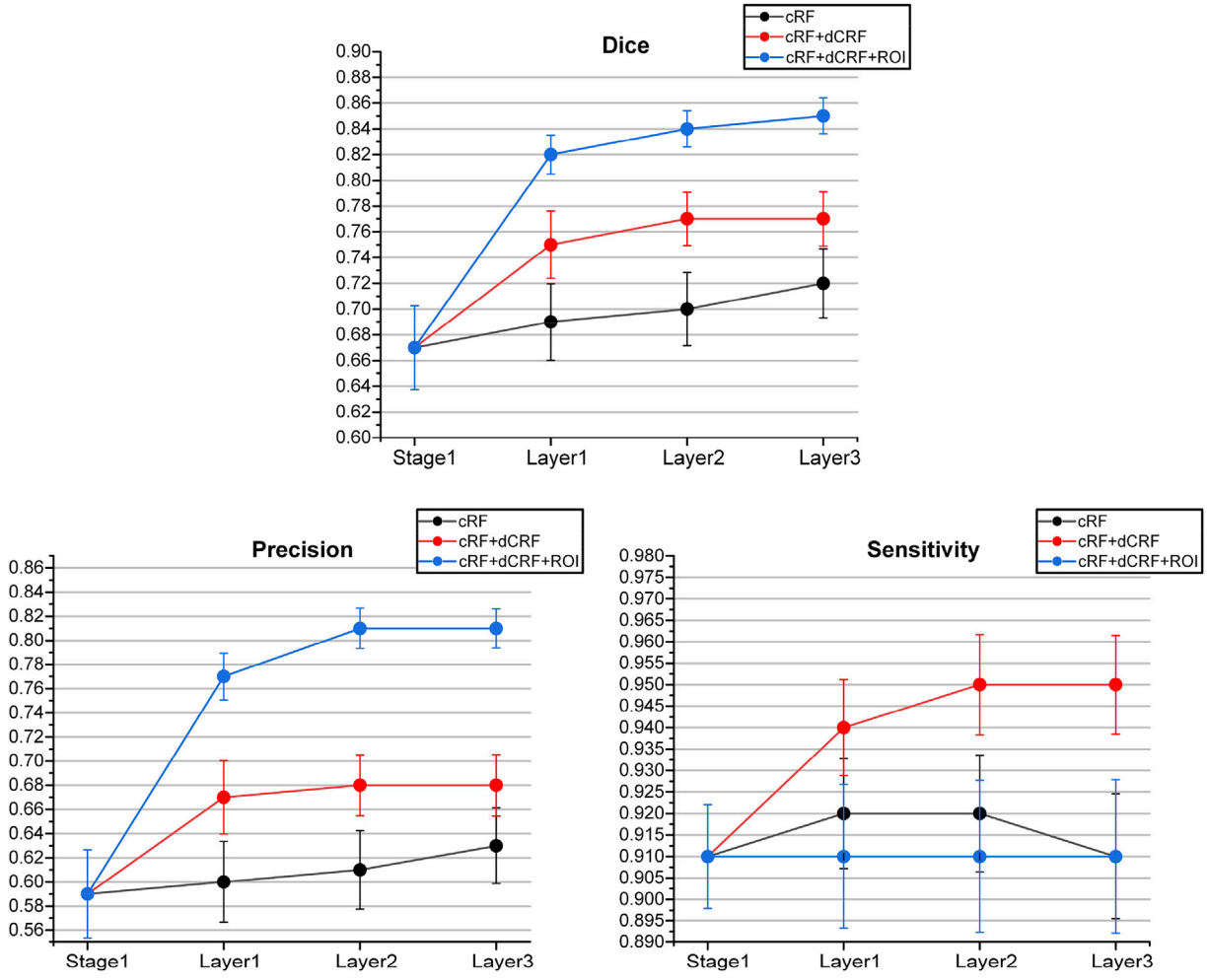

Fig. 3. Hierarchical RF segmentation of the whole tumor regions trained in three different conditions (cRF: cascaded RF, dCRF: dense CRF).

3.3.2. *Impact of the dense CRF*

To investigate the optimization effect of the dense CRF on the segmentation results, the dense CRF model of Section 2.3.2 is applied to enhance the output of each RF, and the results are used as better contextual features to train the next RF model. In the first stage, dense CRF model focus on removing small regions to locate the whole lesion area, and to refine the local segmentation in the second stage. As shown in Fig. 3, the classification performance of cascaded RF with the dense

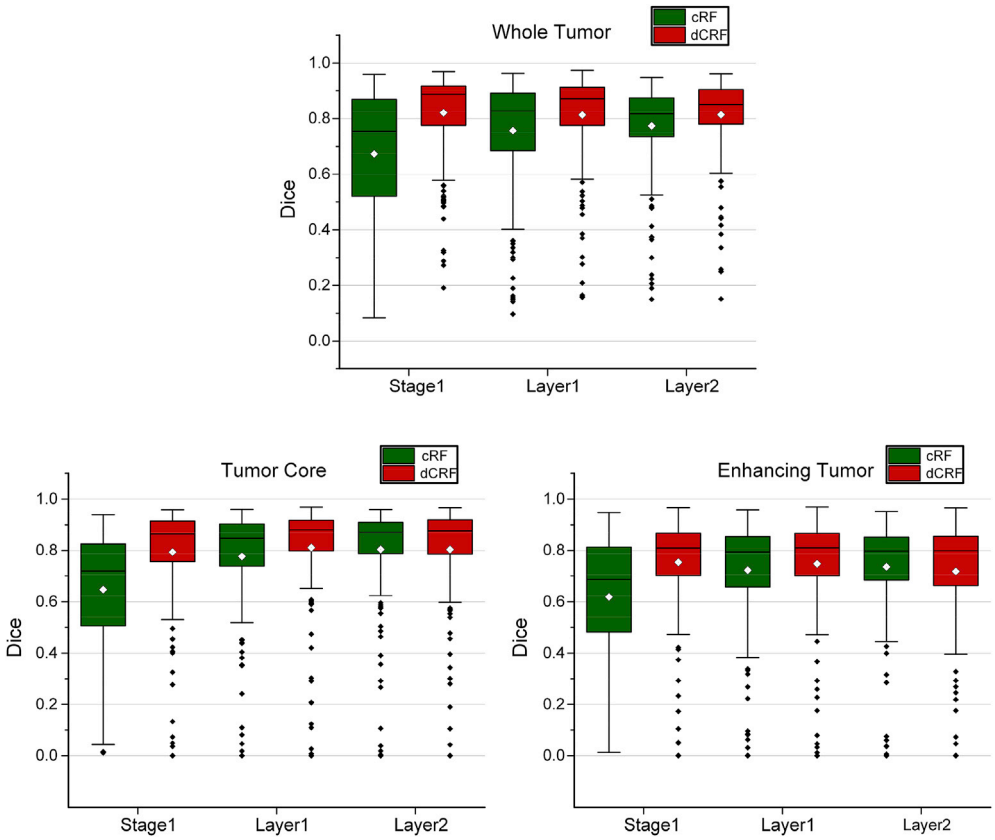

Fig. 4. Dice ratio of baseline model integrating the dense CRF with cascaded RF (cRF) evaluated on three tumor regions.

CRF is better than that of the conventional RF without CRF. The classification effect of the RF has been improved hierarchically. In addition, the detailed results of the baseline cascade RF with dense CRF are presented in Table 2 and Fig. 4, In the first stage, after the dense CRF optimized the probability image, the three tumor regions have been significantly improved with respect to both Dice and precision. In the second stage, for this type of basic cascade RF model trained and tested in the whole brain region, the dense CRF improvement becomes negligible.

3.3.3. *Effectiveness of the ROI*

In order to further improve the segmentation accuracy of the results in the first stage, we define a specific ROI for subsequent training and testing instead of the whole brain region. As shown in Fig. 2, the segmentation result of the first stage was dilated by 5 voxels and considered as a 'new ROI'. We multiply the statistical feature maps and asymmetry feature maps by the ROI at the voxel level, for training or testing of classifiers in the second stage. The results in Table 2 and Fig. 3 show that with features extracted from ROI, the performances of RF classifier in the second stage is obviously improved, and its segmentation performance in the first layer of the whole tumor is better than that of the third layer of RF trained and tested on the whole brain region.

Fig. 5 displays the segmentation results following CRF optimization and iteration over three layers. For the whole tumor and enhancing tumor, the final segmentation performance based on the dilated ROI improved the labeling accuracy as to when using the whole brain as a region of interest.

The hierarchical evaluation results of our framework are shown in Fig. 6, as the number of layers increases, performances in terms of tumor core and enhancing tumor is gradually optimized. Due to the localization effect of ROI, the segmentation performance on the whole tumor in the second stage is improved. The qualitative comparison between our segmentation results and the ground truth (GT) is shown in Fig. 7. We also measured the importance of the features used in the second stage based

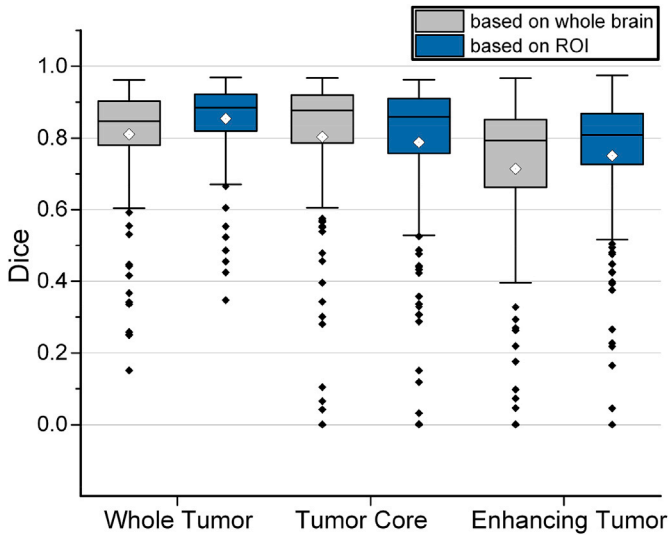

Fig. 5. Comparison of our final results based on the whole brain and a dilated regions of interest (ROI) respectively. The ROI-based method differs from the whole brain method in the sense that it uses the segmentation results of the first stage as a 'new ROI' with a slight morphological dilation for training and testing in the second stage.

mean decrease accuracy, which quantitatively described the effect of each feature on the accuracy of multilayer model. Part of top ranked features are shown in Fig. 8.

3.3.4. *Comparison with other methods on BRATS2015 and BRATS2018*

Our model produced accurate and stable segmentation results in

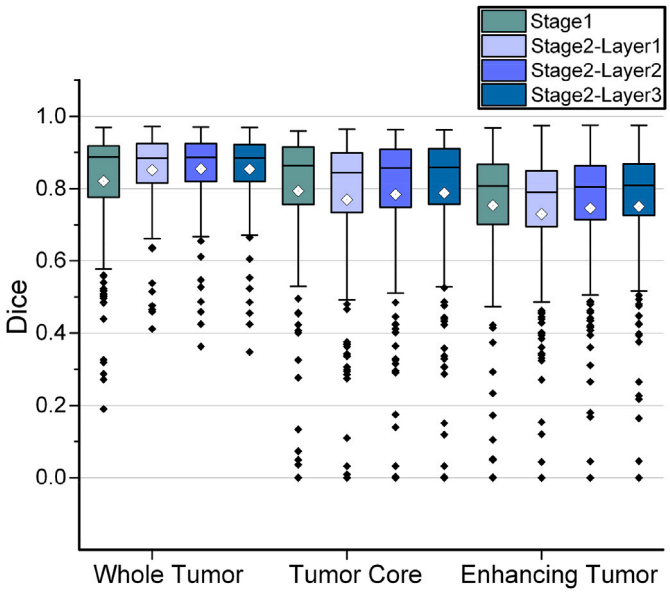

Fig. 6. Hierarchical evaluation results of our framework for cooperative optimization of cascaded RF and dense CRF.

different regions of two datasets and compared with several benchmark methods based on RF, CRF, and deep learning. Specifically, Zhao et al. (2016) applied the CRF as a post-processing step of the fully convolutional neural network to refine the segmentation results, and Kamnitsas et al. (2017) recently implemented a 3D CNN architecture for brain lesion segmentation named DeepMedic, and the dense CRF was employed as a post-processing step. Both RF and Markov random field based framework (Tustison et al., 2015) and cooperative learning of SRF and Bayesian Networks (BN) integrating contextual information with super-pixel features (Amiri et al., 2018b) are similar to our framework with a good classification performance on the BRATS dataset.

Particularly, the method proposed by Pereira et al. (2016a) ranked the second on the BRATS2013 Challenge dataset and the third on the BRATS2013 Leaderboard dataset, and it has also obtained the second position on the BRATS 2015 (Pereira et al., 2016b). Isensee et al. (2018) is the state of the art on the BRATS 2015 dataset and one of the leading methods on the BRATS2017 dataset. The comparison results with these methods (Tustison et al., 2015; Zhao et al., 2016; Kamnitsas et al., 2017; Amiri et al., 2018b; Pereira et al., 2016a; Isensee et al., 2018; Le Folgoc et al., 2016; Chen et al., 2018) are summarized in Table 3 and 4.

3.4. Evaluation on ISLES2015

We also investigated the segmentation performance of the framework for ischemic stroke lesions. First, our method was evaluated on ISLES2015 training data with 5-fold cross validation. An example of segmentation results is shown in Fig. 9. The top 10 important features in the proposed model are shown in Fig. 10. In addition, testing datasets were also used to evaluate the effectiveness of our framework. Table 5 lists the performance of the ISLES2015 test data and compares the results with other top six ranking teams participating in the MICCAI 2015 ISLES workshop (Kamnitsas et al., 2017; Feng et al., 2016; Halme et al., 2016; Reza et al., 2015; Robben et al., 2016; Maier et al., 2015). Particularly, Kamnitsas et al. (2017) ranked third on the ISLES2015 Challenge dataset.

As the ground truth of the testing cases is not available for participants, all segmentation results were uploaded and the average and standard deviation of metrics were calculated by the online evaluation system. As shown in Table 5, the performance of our framework is very competitive on all evaluation aspects. It is worth noting that our proposed method performed better in sensitivity than all current methods, ranking first in the online evaluation system of ISLES2015 SISS testing leaderboard. The evaluation results by other metrics were better than those of approaches (Halme et al., 2016; Reza et al., 2015; Robben et al., 2016; Maier et al., 2015), which are based on the classical RF or cascaded extremely randomized forest classifiers.

4. Discussion

We have proposed a novel framework for automatic brain lesion

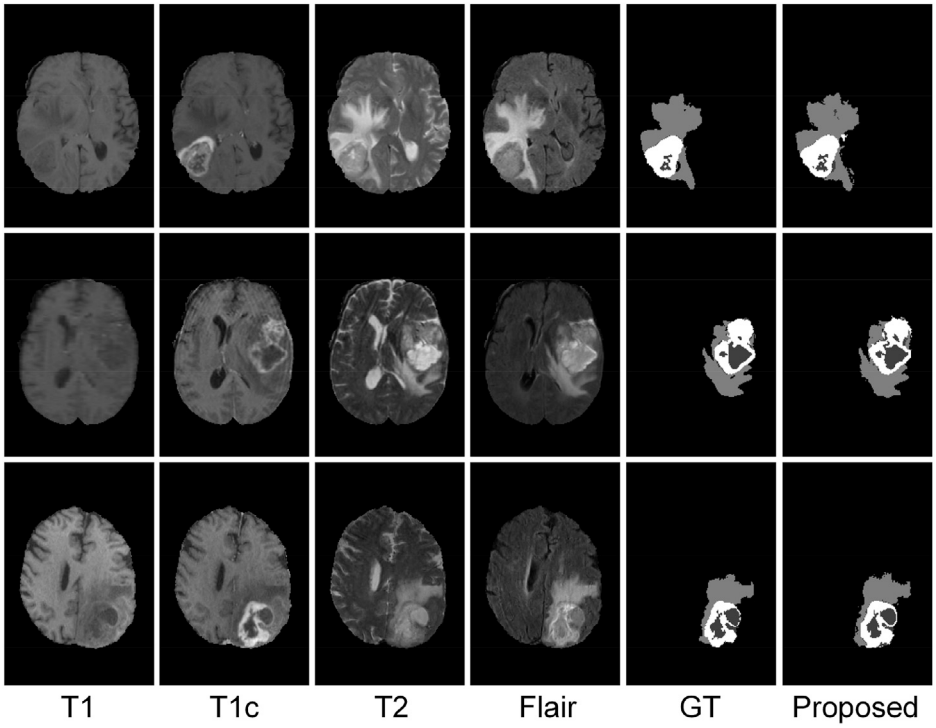

Fig. 7. Three cases of qualitative segmentation results on BRATS dataset using the proposed method.

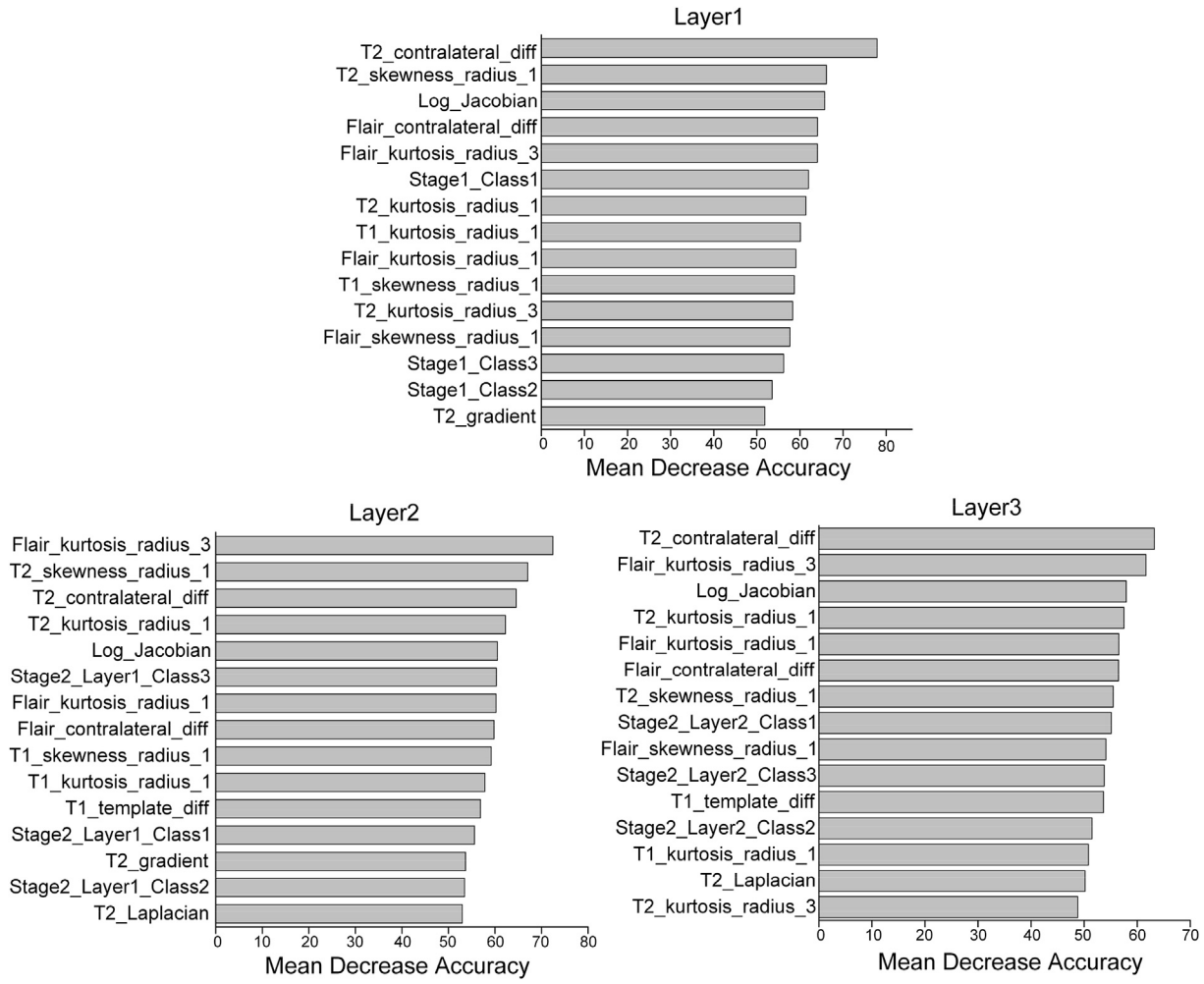

Fig. 8. Feature importance ranking measured by mean decrease accuracy of models for BRATS 2018.

Table 3
Average Dice scores of our proposed framework on BRATS2015 and BRATS2018 training datasets and comparisons with other methods.

	WT	TC	ET	Dataset
Pereira et al., (2016a)	0.87	0.75	**0.75**	BRATS2015
FCNN + dCRF (Zhao et al., 2016)	0.80	0.68	0.65	
Le Folgoc et al. (2016)	0.82	0.73	**0.75**	
SRF + BN (Amiri et al., 2018b)	**0.88**	**0.78**	0.68	
Ours	0.84	0.75	0.72	
Pereira et al., (2016a)	0.84	0.72	0.62	BRATS2018
Deepmedic + dCRF (Kamnitsas et al., 2017)	**0.90**	0.75	0.73	
Chen et al. (2018)	0.72	**0.83**	**0.81**	
ANTsR + dCRF (Tustison et al., 2015)	0.82	0.78	0.75	
Ours	0.86	0.79	0.75	

segmentation from multimodal 3D MRI sequences. The proposed method integrates cascaded RF and dense CRF within on a multilayer cooperative learning architecture.

Although the cascaded RF based on automatic context can make full use of the rich spatial information and low-level appearance information output by the previous classifier to construct a new classifier, its drawback is the independently estimation of voxel probability that is blind to the local neighborhood. Moreover, since the RF model distributes features independently across trees, simple RF architecture cannot well capture the global information and might not be able to handle outliers and noisy observations. We can also confirm this from Table 2. In the first stage, the sensitivity of the RF segmentation result is very high, but the Dice is very low. Therefore, it can be inferred that the segmentation result contains many 3D small regions with noise.

As an extension of the CRF, the dense CRF is designed to model the

Table 4
Comparison results with one of state-of-the-art methods on the BRATS2015 and BRATS2018 dataset.

	Dice			Precision			Sensitivity			Dataset
	WT	TC	ET	WT	TC	ET	WT	TC	ET	
Isensee et al. (2018)	**0.88**	**0.78**	**0.77**	0.89	**0.81**	**0.81**	0.88	0.77	**0.78**	BRATS
cRF + dCRF	0.80	0.72	0.62	0.74	0.78	0.74	**0.92**	0.75	0.64	2015
Ours	0.84	0.75	0.72	**0.90**	0.76	0.75	0.82	**0.80**	**0.78**	
Isensee et al. (2018)	**0.87**	**0.85**	**0.78**	**0.88**	**0.88**	**0.78**	0.86	0.82	0.80	BRATS
cRF + dCRF	0.81	0.80	0.72	0.76	0.83	0.74	**0.92**	0.83	0.77	2018
Ours	0.86	0.79	0.75	0.84	0.78	0.74	0.91	**0.86**	**0.83**	

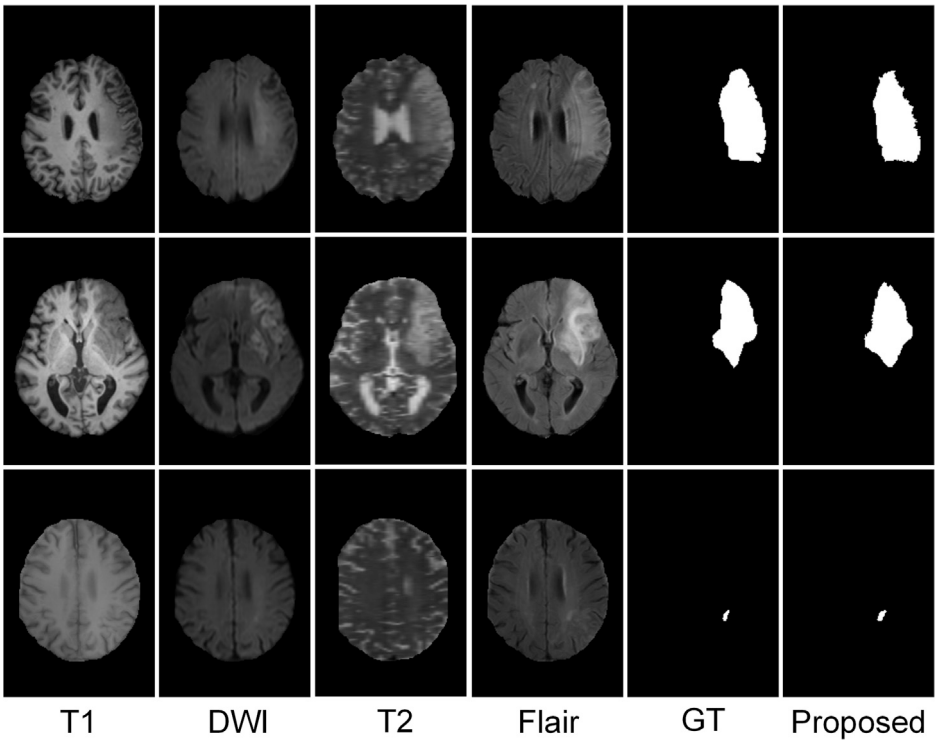

Fig. 9. Examples of segmentation results on ISLES(SISS) 2015 dataset using the proposed framework.

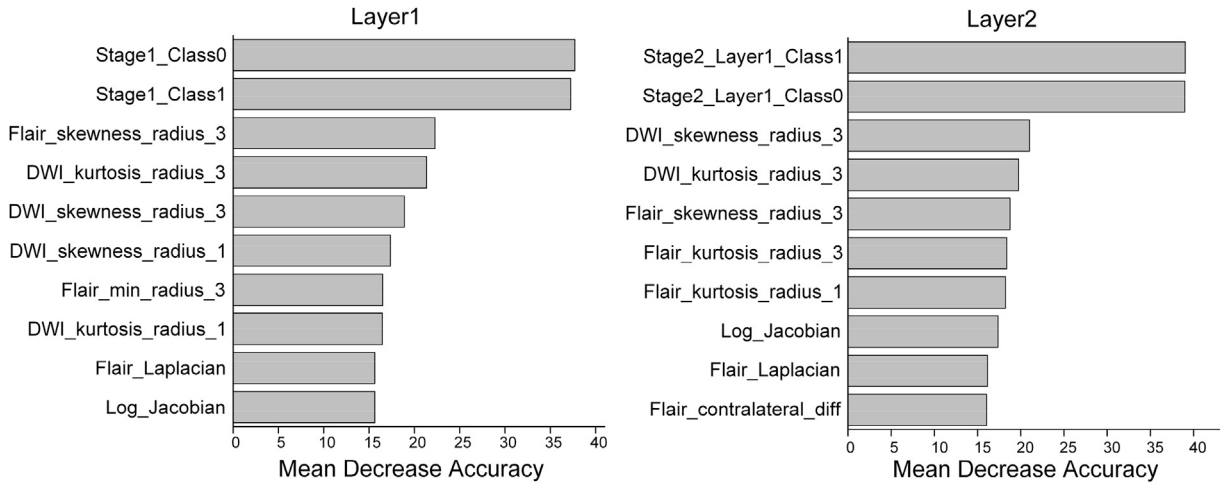

Fig. 10. Feature importance ranking measured by mean decrease accuracy of models for ISLES 2015.

Table 5
Our mean segmentation results and standard deviation on ISLES2015 testing dataset, and comparisons with the performance of other six top ranking participants in the testing stage of the ISLES-SISS 2015 challenge.

	Dice	Precision	Sensitivity	ASSD (mm)	Hausdorff (mm)	Case
Kamnitsas et al. (2017)	0.59±0.31	0.68±0.33	0.60±0.27	7.87±12.63	39.61±30.68	34/36
Feng et al. (2016)	0.55±0.30	0.64±0.31	0.57±0.33	8.13±15.15	25.02±22.02	32/36
Halme et al. (2016)	0.47±0.32	0.47±0.34	0.56±0.33	14.61±20.17	46.26±34.81	31/36
Reza et al. (2015)	0.43±0.28	0.37±0.29	0.72±0.30	11.90±20.50	46.38±29.32	33/36
Robben et al. (2016)	0.43±0.30	0.41±0.34	0.67±0.31	14.22±14.41	62.58±30.61	33/36
Maier et al. (2015)	0.42±0.33	0.45±0.38	0.55±0.34	17.59±21.06	56.39±30.65	31/36
Ours (Layer2)	0.51±0.29	0.49±0.32	0.73±0.30	11.23±16.02	48.24±29.99	33/36
Ours (Layer3)	0.48±0.29	0.43±0.29	**0.75±0.32**	11.66±15.98	45.89±29.84	33/36

relationship between all voxels in the image and predict labels according to the intensity and position information of each voxel. It is commonly used for processing 3D multimodal medical images. For computational efficiency, the inference process of this modeling implicitly assumes the conditional independence between voxels. This strong hypothesis has no effect on long-range voxels, but may lose the correlation constraint between direct or short-range neighbors.

Our algorithm overcomes the shortcomings of both models and is evaluated on the publicly available BRATS2015 and BRATS2018 datasets and the ISLES2015 dataset to verify the segmentation performance for brain tumor and ischemic stroke lesions, respectively. The results of brain tumor segmentation in Table 3 and 4 have shown that our framework achieves competitive results on both BRATS datasets compared with top performing and similar methods. In particular, the segmentation performance on BRATS2018 was competitive with most of the compared deep learning methods and the sensitivity was also higher than one of state-of-the-art methods (Isensee et al., 2018).

The results on the sub-acute ischemic stroke lesion dataset in Table 5 indicate that the proposed method has the best performance in sensitivity and may be also effective for diseases other than brain tumors. Besides, since the feature extraction steps are simple and convenient, the optimization of the dense CRF accelerates the convergence speed of cascaded RF, reducing the number of iteration layers. The above-mentioned advantages make our framework easy to implement.

In general, our goal was to further improve the RF and our results showed the effectiveness of our method. Although our proposed method cannot outperform the latest deep learning methods, compared to deep neural networks that require extensive and ad-hoc hyper-parameter tuning, our RF-based framework is easier to train. It can achieve performances that are better than the existing RF frameworks and are close to those of deep learning methods. Moreover, our method is explainable and interpretable whereas the deep learning methods are not. Specifically, Figs. 8 and 10 show weighted ranking based on the importance of each feature in the target brain tumor and ischemic stroke segmentation. We found that features with greater influence in the model learning vary across layers. However, it is worth noting that our proposed features based on the probability of target classes were important in each layer, especially for the ischemic stroke segmentation. As one of the used asymmetric features, contralateral difference feature quantifying the disparity between lesion image and corresponding flipped image, stood out in the target brain tumor segmentation task. Similarly, in addition to the proposed probability features, the skewness was more important than other features in each layer of the ischemic stroke segmentation models. Detecting and segmenting brain lesions based on image asymmetry features conforms to the knowledge and experience of medical experts, further reflecting the reliability and interpretability of our framework.

The limitation of our framework lies in the choice of relatively simple features and the need for labeled data for training. Given the difficulty of manual labeling MRI clinical data, in our future work, we will incorporate semi-supervised learning into our framework while leveraging a small-sized training set, extracting higher-order feature representations, and evaluating our extended method across a wide range of segmentation tasks such as traumatic brain injury and multiple sclerosis. In our brain tumor segmentation framework, we found that the features from T1c modality are less important than those from other modalities, which will be further investigated in our future studies.

5. Conclusion

In this paper, we presented a two-stage automated brain lesion segmentation framework integrating cascaded RF and dense CRF (RFDCR). We evaluated our model on two segmentation tasks: brain tumor and ischemic stroke segmentations. Experimental results have shown that our proposed method achieves competitive performance in two different brain tumor segmentation datasets of BRATS, especially in BRATS 2018, the segmentation performances of the whole tumor regions are comparable to state-of-the-art deep learning methods. In addition, our segmentation technique using the ischemic stroke dataset ISLES2015 outperformed comparison methods based on RF and our segmentation sensitivity was better than existing methods.

Ethics statement

The data used in this work is downloaded from publicly available BRATS and ISLES websites.

CRediT authorship contribution statement

Gaoxiang Chen: Methodology, Software, Writing - original draft. **Qun Li:** Visualization, Investigation. **Fuqian Shi:** Software, Validation. **Islem Rekik:** Writing - review & editing, Validation. **Zhifang Pan:** Software, Validation, Supervision, Project administration, Funding acquisition.

Acknowledgements

The author would like to thank Dr. Li Wang for help with editing. This research was supported by Zhejiang Provincial Natural Science Foundation of China under Grant No. LY16F030010 and Wenzhou Science & Technology Bureau under Grant No. 2018ZG016.

References

Amiri, S., Ali Mahjoub, M., Rekik, I., 2018. Tree-based ensemble classifier learning for automatic brain glioma segmentation. Neurocomputing 313, 135–142. https://doi.org/10.1016/j.neucom.2018.05.112. HYPERLINK.

Amiri, S., Mahjoub, M.A., Rekik, I., 2018. Bayesian network and structured random forest cooperative deep learning for automatic multi-label brain tumor segmentation. ICAART 2018 183–190. https://doi.org/10.5220/0006629901830190.

Ashburner, J., Friston, K.J., 2005. Unified segmentation. Neuroimage 26 (3), 839–851.

Avants, B.B., Tustison, N.J., Stauffer, M., Song, G., Wu, B., Gee, J.C., 2014. The Insight ToolKit image registration framework. Front. Neuroinf. 8, 44. https://doi.org/10.3389/fninf.2014.00044. HYPERLINK.

Balafar, M., 2014. Gaussian mixture model based segmentation methods for brain MRI images. Artif. Intell. Rev. 41 (3), 429–439. https://doi.org/10.1007/s10462-012-9317-3.

Bauer, S., Wiest, R., Nolte, L.-P., Reyes, M., 2013. A survey of MRI-based medical image analysis for brain tumor studies. Phys. Med. Biol. 58 (13), 97–129. https://doi.org/10.1088/0031-9155/58/13/R97.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324. HYPERLINK.

Chen, L., Wu, Y., DSouza, A.M., Abidin, A.Z., Wismuller, A., Xu, C., 2018. MRI tumor segmentation with densely connected 3D CNN. In: Medical Imaging 2018: Image Processing, vol. 10574. International Society for Optics and Photonics, pp. 32–39. https://doi.org/10.1117/12.2293394.

Feng, C., Zhao, D., Huang, M., 2016. Segmentation of ischemic stroke lesions in multi-spectral MR images using weighting suppressed FCM and three phase level set. In: Lecture Notes in Computer Science, vols. 233–245. https://doi.org/10.1007/978-3-319-30858-6_20.

Geremia, E., Clatz, O., Menze, B.H., Konukoglu, E., Criminisi, A., Ayache, N., 2011. Spatial decision forests for MS lesion segmentation in multi-channel magnetic resonance images. Neuroimage 57 (2), 378–390. https://doi.org/10.1016/j.neuroimage.2011.03.080.

Greenspan, H., Ruf, A., Goldberger, J., 2006. Constrained Gaussian mixture model framework for automatic segmentation of mr brain images. IEEE Trans. Med. Imag. 25 (9), 1233–1245. https://doi.org/10.1109/tmi.2006.880668.

Halme, H.L., Korvenoja, A., Salli, E., 2016. ISLES (SISS) challenge 2015: segmentation of stroke lesions using spatial normalization, random forest classification and contextual clustering. In: Lecture Notes in Computer Science, vols. 211–221. https://doi.org/10.1007/978-3-319-30858-6_18.

Isensee, F., Kickingereder, P., Wick, W., Bendszus, M., Maier-Hein, K.H., 2018. Brain tumor segmentation and radiomics survival prediction: contribution to the BRATS 2017 challenge. Lect. Notes Comput. Sci. 287–297. https://doi.org/10.1007/978-3-319-75238-9_25.

Kamnitsas, K., Ledig, C., Newcombe, V.F.J., Simpson, J.P., Kane, A.D., Menon, D.K., Rueckert, D., Glocker, B., 2017. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. Med. Image Anal. 36, 61–78. https://doi.org/10.1016/j.media.2016.10.004.

Krähenbühl, P., Koltun, V., 2012. Efficient inference in fully connected CRFs with Gaussian edge potentials. Adv. Neural Inf. Process. Syst. 24. Proceedings NIPS).

Landman, B.A., Huang, A.J., Gifford, A., Vikram, D.S., Lim, I.A.L., Farrell, J.A.D., Bogovic, J.A., Hua, J., Chen, M., Jarso, S., Smith, S.A., Joel, S., Mori, S., Pekar, J.J., Barker, P.B., Prince, J.L., van Zijl, P.C.M., 2011. Multi-parametric neuroimaging

reproducibility: a 3-T resource study. Neuroimage 54 (4), 2854–2866. https://doi.org/10.1016/j.neuroimage.2010.11.047.

Le Folgoc, L., Nori, A.V., Ancha, S., Criminisi, A., 2016. Lifted auto-context forests for brain tumour segmentation. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). https://doi.org/10.1007/978-3-319-55524-9_17.

Lefkovits, L., Lefkovits, S., Szilágyi, L., 2016. Brain tumor segmentation with optimized random forest. Lect. Notes Comput. Sci. 88–99. https://doi.org/10.1007/978-3-319-55524-9_9.

Liu, Y., Zhang, Pengcheng, Song, Q., Li, A., Zhang, Peng, Gui, Z., 2018. Automatic segmentation of cervical nuclei based on deep learning and a conditional random field. IEEE Access 6, 53709–53721. https://doi.org/10.1109/ACCESS.2018.2871153.

Ma, C., Luo, G., Wang, K., 2018. Concatenated and connected random forests with multiscale patch driven active contour model for automated brain tumor segmentation of MR images. IEEE Trans. Med. Imag. 37, 1943–1954. https://doi.org/10.1109/TMI.2018.2805821.

Maier, O., Wilms, M., Handels, H., 2015. Random forests with selected features for stroke lesion segmentation. In: Proceedings of ISLES 2015 Challenge, pp. 17–22.

Maier, O., Menze, B.H., von der Gablentz, J., Häni, L., Heinrich, M.P., Liebrand, M., Winzeck, S., Basit, A., Bentley, P., Chen, L., et al., 2017. ISLES 2015 - a public evaluation benchmark for ischemic stroke lesion segmentation from multispectral MRI. Med. Image Anal. 35, 250–269. https://doi.org/10.1016/j.media.2016.07.009. HYPERLINK.

Menze, B.H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., Burren, Y., Porz, N., Slotboom, J., Wiest, R., et al., 2015. The multimodal brain tumor image segmentation benchmark (BRATS). IEEE Trans. Med. Imag. 34, 1993–2024. https://doi.org/10.1109/TMI.2014.2377694. HYPERLINK.

Menze, B.H., Van Leemput, K., Lashkari, D., Riklin-Raviv, T., Geremia, E., Alberts, E., Gruber, P., Wegener, S., Weber, M.A., Székely, G., Ayache, N., Golland, P., 2016. A generative probabilistic model and discriminative extensions for brain lesion segmentation - with application to tumor and stroke. IEEE Trans. Med. Imag. 35, 933–946. https://doi.org/10.1109/TMI.2015.2502596.

Mitra, J., Bourgeat, P., Fripp, J., Ghose, S., Rose, S., Salvado, O., et al., 2014. Lesion segmentation from multimodal MRI using random forest following ischemic stroke. Neuroimage 98, 324–335. https://doi.org/10.1016/j.neuroimage.2014.04.056.

Moon, T.K., 1996. The expectation-maximization algorithm. IEEE Signal Process. Mag. 13, 47–60. https://doi.org/10.1109/79.543975.

Nyúl, L.G., Udupa, J.K., Zhang, X., 2000. New variants of a method of MRI scale standardization. IEEE Trans. Med. Imag. 19, 143–150. https://doi.org/10.1109/42.836373.

Pereira, S., Pinto, A., Alves, V., Silva, C.A., 2016. Brain tumor segmentation using convolutional neural networks in MRI images. IEEE Trans. Med. Imag. 35, 1240–1251. https://doi.org/10.1109/TMI.2016.2538465.

Pereira, S., Pinto, A., Alves, V., Silva, C.A., 2016. Deep convolutional neural networks for the segmentation of gliomas in multi-sequence MRI. Lect. Notes Comput. Sci. 131–143. https://doi.org/10.1007/978-3-319-30858-6_12.

Pereira, S., Meier, R., McKinley, R., Wiest, R., Alves, V., Silva, C.A., Reyes, M., 2018. Enhancing interpretability of automatically extracted machine learning features: application to a RBM-Random Forest system on brain lesion segmentation. Med. Image Anal. 44, 228–244. https://doi.org/10.1016/j.media.2017.12.009.

Permuter, H., Francos, J., Jermyn, I., 2006. A study of Gaussian mixture models of color and texture features for image classification and segmentation. Pattern Recogn. 39 (4), 695–706. https://doi.org/10.1016/j.patcog.2005.10.028. HYPERLINK.

Pustina, D., Coslett, H.B., Turkeltaub, P.E., Tustison, N., Schwartz, M.F., Avants, B., 2016. Automated segmentation of chronic stroke lesions using LINDA: lesion identification with neighborhood data analysis. Hum. Brain Mapp. 37, 1405–1421. https://doi.org/10.1002/hbm.23110.

Reza, S.M., Pei, L., Iftekharuddin, K.M., 2015. Ischemic stroke lesion segmentation using local gradient and texture features. Proc. ISLES 2015 Chall. 23–26.

Robben, D., Christiaens, D., Rangarajan, J.R., Gelderblom, J., Joris, P., Maes, F., Suetens, P., 2016. A voxel-wise, cascaded classification approach to ischemic stroke lesion segmentation. In: Lecture Notes in Computer Science, pp. 254–265. https://doi.org/10.1007/978-3-319-30858-6_22.

Song, B., Chou, C.-R., Chen, X., Huang, A., Liu, M.-C., 2016. Anatomy guided brain tumor segmentation and classification. In: Proceedings of MICCAI-BRATS 2016 Multimodal Brain Tumor Image Segmentation Benchmark: "Change Detection", pp. 61–64. https://doi.org/10.1007/978-3-319-55524-9_16.

Tu, Z., Bai, X., 2010. Auto-context and its application to high-level vision tasks and 3D brain image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 32, 1744–1757. https://doi.org/10.1109/TPAMI.2009.186. HYPERLINK.

Tustison, N.J., Avants, B.B., Cook, P.A., Zheng, Y., Egan, A., Yushkevich, P.A., Gee, J.C., 2010. N4ITK: improved N3 bias correction. IEEE Trans. Med. Imag. 29, 1310–1320. https://doi.org/10.1109/TMI.2010.2046908. HYPERLINK.

Tustison, N.J., Shrinidhi, K.L., Wintermark, M., Durst, C.R., Kandel, B.M., Gee, J.C., Grossman, M.C., Avants, B.B., 2015. Optimal symmetric multimodal templates and concatenated random forests for supervised brain tumor segmentation (simplified) with ANTsR. Neuroinformatics 13, 209–225. https://doi.org/10.1007/s12021-014-9245-2.

Wang, L., Gao, Y., Shi, F., Li, G., Gilmore, J.H., Lin, W., Shen, D., 2015. LINKS: learning-based multi-source IntegratioN frameworK for Segmentation of infant brain images. Neuroimage 108, 160–172. https://doi.org/10.1016/j.neuroimage.2014.12.042.

Xuan, X., Liao, Q., 2007. Statistical structure analysis in MRI brain tumor segmentation. In: Proceedings of the 4th International Conference on Image and Graphics, ICIG 2007, pp. 421–426. https://doi.org/10.1109/ICIG.2007.181. HYPERLINK.

Zhao, X., Wu, Y., Song, G., Li, Z., Fan, Y., Zhang, Y., 2016. Brain tumor segmentation using a fully convolutional neural network with conditional random fields. Lect. Notes Comput. Sci. 75–87. https://doi.org/10.1007/978-3-319-55524-9_8.

Zhao, X., Wu, Y., Song, G., Li, Z., Zhang, Y., Fan, Y., 2018. 3D brain tumor segmentation through integrating multiple 2D FCNNs. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 10670 LNCS 191–203. https://doi.org/10.1007/978-3-319-75238-9_17.

Zikic, D., Glocker, B., Konukoglu, E., Criminisi, A., Demiralp, C., Shotton, J., Thomas, O.M., Das, T., Jena, R., Price, S.J., 2012. Decision Forests for Tissue-specific Segmentation of High-Grade Gliomas in Multi-Channel MR, pp. 369–376. https://doi.org/10.1007/978-3-642-33454-2_46.