Biomedical Signal Processing and Control 55 (2020) 101641

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

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

RescueNet: An unpaired GAN for brain tumor segmentation

Shubhangi Nema [a,b], Akshay Dudhane [a], Subrahmanyam Murala [a,*], Srivatsava Naidu [b]

[a] Computer Vision and Pattern Recognition Lab, Indian Institute of Technology Ropar, Rupnagar 140001, India
[b] Center of Biomedical Engineering, Indian Institute of Technology Ropar, Rupnagar 140001, India

ARTICLE INFO

Article history:
Received 21 February 2019
Received in revised form 11 April 2019
Accepted 7 August 2019
Available online 3 September 2019

Keywords:
Computer aided diagnosis
Brain tumor
BraTS
Deep learning

ABSTRACT

Even with proper acquisition of brain tumor images, the accurate and reliable segmentation of tumors in brain is a complicated job. Automatic segmentation become possible with development of deep learning algorithms that brings plethora of solutions in this research prospect. In this paper, we designed a network architecture named as residual cyclic unpaired encoder-decoder network (RescueNet) using residual and mirroring principles. RescueNet uses unpaired adversarial training to segment the whole tumor followed by core and enhance regions in a brain MRI scan. The problem in automatic brain tumor analysis is preparing large scale labeled data for training of deep networks which is a time consuming and tedious task. To eliminate this need of paired data we used unpaired training approach to train the proposed network. Performance evaluation parameters are taken as DICE and Sensitivity measure. The experimental results are tested on BraTS 2015 and BraTS 2017 [1] dataset and the result outperforms the existing methods for brain tumor segmentation. The combination of domain-specific segmentation methods and general-purpose adversarial learning loomed to leverage huge advantages for medical imaging applications and can improve the ability of automated algorithms to assist radiologists.

1. Introduction

Gliomas, which originated from the glial cells and the surrounding infiltrative tissues, are the most malignant brain tumors in adults [2]. They are divided into low grade gliomas (LGG) and high-grade gliomas (HGG), out of which HGG are the most aggressive one. At present, over 130 different types of high grade and low grade brain tumors are known and the average survival is varied between 12 to 15 months. Brain tumor segmentation is a challenging task because of their varied behavior both in terms of structure and function [3]. Also, the tumor intensity of one person differs significantly with other. Magnetic resonance imaging (MRI) is preferred over other imaging modalities for diagnosis and treatment of brain tumor because of its non-invasive property without the exposure to ionizing radiations and superior image contrast in soft tissues. Different types of tissue contrast images have been produced by MRI modalities, which allow extraction of valuable structural information therefore enabling diagnosis and treatment of tumors along with their sub-regions. MRI scans produced on applying different pulse sequences are: (1) T1 weighted scans which distinguish healthy tissues from tumorous one. (2) T2 weighted scans to delineate the edema region which produce bright image region. (3) T1-Gd scans use contrast agent that give bright signal at tumor border due to its accumulation. (4) FLAIR scans uses signal of water molecule suppression which distinguish cerebrospinal fluid (CSF) from edema region. For a radiologist, these scans are useful to annotate different regions of brain tumor. However, slice by slice annotation of brain tumor from MRI scans is a laborious and time-consuming task. This manual burden can be replaced by automatic segmentation with the help of computer vision algorithms [4].

Computer vision techniques and development of computer-aided tools are evolving as the areas of research for automatic segmentation of brain tumors. Some of these techniques showed good results but there is no winning technique as these approaches have often not used practically in hospitals. In recent years, convolution neural networks (CNN) have become the technique of choice for state-of-the-art implementations in various image segmentation tasks [5]. BraTS [1] in conjunction with MICCAI were the first to apply CNN for the brain tumor segmentation. The layers of convolutions take raw image intensities as inputs to calculate an output signal. This makes deep learning not to rely on handcrafted features to segment tumor from brain MRI scan. High level of abstraction in learning complex patterns can be solved using algorithms with non-linearities and many degrees of freedom.

With advancement in machine learning techniques, support vector machine [6] and random forest [7] were used widely for the automated brain tumor segmentation. But these methods require hand-crafted features to be extracted for training the respective

* Corresponding author.
E-mail address: subbumurala@iitrpr.ac.in (S. Murala).

https://doi.org/10.1016/j.bspc.2019.101641
1746-8094/

machine learning algorithm. Inspired from deep learning approach, Ronneberger et al. [8] in 2015 first proposed the U-Net architecture with fully convolutional layers. U-Net is designed with contracting path and a symmetric expanding path with skip connections in between. For the border pixel prediction, mirroring concepts have been applied. Inspired from the U-Net architecture, Dong et al. [9] proposed fully convolution network for brain tumor detection and segmentation. The network was less successful in segmenting enhancing region in LGG cohort. Further, hybrid Pyramid U-Net Model for brain tumor segmentation is proposed by Kong et al. [10]. They extended the U-Net model which explores the global context information combined different region-based context. Alex et al. [11] proposed brain tumor segmentation from multi-modal MR images using fully convolutional neural network (FCNN). They used voxel-based classification with 23 layers and single forward pass followed by false positive reduction using connected component analysis. Havaei et al. [12] achieved better performance for brain tumor segmentation on BraTS dataset. They used two-pathway and cascaded architectures in their network. The two-pathway architecture is used to capture two receptive fields both for local and global features. The label of a pixel is predicted based on two aspects, the visual details around the pixel and where the patch is located in the brain. Wang et al. [13] extended cascaded architecture and used anisotropic CNN for automatic brain tumor segmentation. Further, Hussain et al. [14] proposed patch-based approach and utilized cascaded deep CNN for brain tumor segmentation.

In the field of semantic segmentation of natural scene, Noh et al. [15] proposed encoder–decoder network which is made up of different convolution and pooling layers and gives the representation of high resolution features into low level prominent edges. Badrinarayanan et al. [16] proposed a modified encoder-decoder architecture named as SegNet which overcomes the disadvantages of [15]. The max-pooling indices store all the information efficiently without memorizing feature maps in float precision and the corresponding decoder up samples the input feature map(s) using those indices. Drozdzal et al. [17] extended SegNet for brain tumor segmentation. They utilised advantages of identity mappings (*i.e. short skip connections*) instead of long skip connections as used in U-Net [8]. These identity mappings allow to design a deeper CNN without vanishing gradient and enables a fast network learning with proper recovery of the spatial information lost during down sampling. Researchers [18–20] also used small kernel size to design a deeper network architecture. Pereira et al. [20] applied deep CNN with 3×3 kernels. This model used fewer numbers of weights to obtain feature maps thus reducing over fitting. Different architectures are used for both LGG and HGG. The error caused by the wrong classification of clusters was removed using thresholded volumetric constraint in post processing step. Kamnitsas et al. [21] comes up with 3D CNN with fully connected conditional random field (CRF) for accurate brain lesion segmentation. 3D convolution architecture is 11 layers deep and process adjacent image patches in one pass with proper adaption to inherent class imbalance. For multi-scale processing they used dual pathway architecture. Cui et al. [22] uses tumor localization network (TLN) and an intratumor classification network (ITCN) named as deep cascaded neural network for brain tumor segmentation. Further, Lin et al. [23] uses dense CRF learning along with CNNs for segmentation refinement. For improvement in tumor segmentation task Zhao et al. [24] used integration of FCNN with CRF. The above-mentioned methods were all rely on the patch-based approach in which medical scans were often divided into patches during the training and testing. Thus, there is a need of a model which takes whole images as an input and can be used to solve the class imbalance in data by incorporating both local and global features.

Adversarial approaches are widely used in medical image analysis with their capacity to generate high quality results [25,26]. Xue et al. [26] proposed an end-to-end adversarial network for brain tumor segmentation on MRI scans data. The classic generative adversarial network (GAN) architecture is the basic idea behind this network with the segmentor (*i.e. generator*) as a FCNN to generate segmentation label maps and a critic network (*i.e. discriminator*) with multi-scale L1 loss. Two inputs were fed into the critic network: original images masked by ground truth label maps, and original images masked by predicted label maps from segmentor. The segmentor and critic networks are alternately trained in min-max fashion: segmentor training aims to minimize multi-scale L1 loss, while the training of critic aims to maximize the same loss function. Further, Isola et al. [27] proposed conditional GAN for image-to-image translation network which is modified version of classic GAN. They utilized U-Net based architecture to design the generator network. A convolutional patch GAN classifier is used as the discriminator. Rezaei et al. [28] extended conditional GAN [27] for the brain tumor segmentation application in MRI images. They trained a semantic CNN along with an adversarial network for BraTS 2017 segmentation task.

However, CNN training requires labeled paired data (*i.e. input MRI scan and respective tumor annotation*) with proper delineation which is difficult to achieve on enormously evolved data each day. Thus, there is a need to attain CNN training from scratch even for unpaired data to accurately segment the brain tumor from MRI scan. Also, slice by slice annotation of brain tumor from MRI scan is a time consuming task. For this a network architecture is needed which could be trained on less number of images as well as in an unpaired manner. Zhu et al. [29] comes up with an unpaired training approach for image-to-image style transfer, which was further used for various applications like haze removal [30,31], applying and removing makeup [32], etc.

Inspired from [29–32], we propose the medical image segmentation network namely RescueNet which is built on the principle of identity mapping to overcome the overfitting (caused by less training samples) and mirroring (symmetry) concepts. Moreover, unpaired training approach is utilized to train the proposed RescueNet which limits the need of paired training data. Proposed RescueNet used unpaired training to learn the network parameters and thus overcome the need of paired data.

The main contributions of the proposed work are listed below:

- A novel network architecture RescueNet is proposed for brain tumor segmentation.
- An unpaired GAN based training approach is proposed to train the RescueNet.
- Scale-invariant post-processing algorithm is proposed to enhance the accuracy.
- Performance of the proposed network is tested on BraTS-2015 and BraTS-2017 dataset.

Rest of the paper is organized as follows:

Section 1 illustrate the introduction and literature survey on brain tumor segmentation respectively. Section 2 presents the unpaired training approach. Section 3 depicts the proposed RescueNet. Analysis about the proposed RescueNet, is discussed in Section 4. Further, experimental analysis is discussed in Section 5. Finally, Section 6 concludes the proposed approach for brain tumor segmentation.

2. Proposed unpaired training approach for brain tumor segmentation

Generation of labeled data to train a deep network for brain tumor segmentation in MRI scans is time-consuming and requires dedicated Radiologists. To address this problem, we have proposed

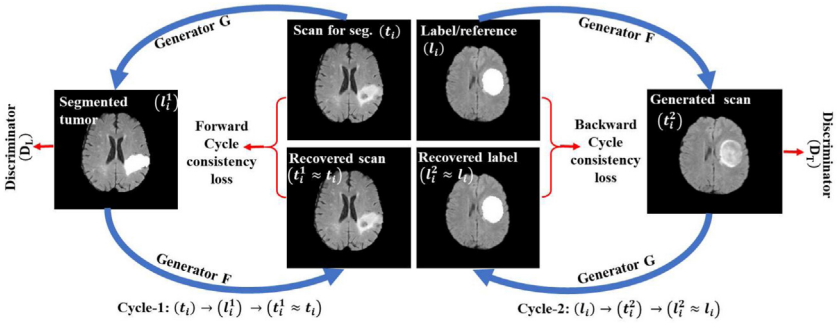

Fig. 1. Unpaired training procedure for RescueNet. *Please magnify the figure for more architecture details.*

an unpaired training approach. The proposed RescueNet is the enhanced version of CycleGAN [29] (*which was used for image-to-image style transfer task*). It uses less training data and an unpaired approach to achieve better segmentation results on diverse brain MRI data. For a given set of data having brain tumor MRI scan (T) and labeled image (L) there exists no correlation between them. We have collected $T\left(\left\{t_{i=1,2,\ldots,N}, t_i \in T\right\}\right)$ and $L\left(\left\{l_{i=1,2,\ldots,N}, l_i \in L\right\}\right)$ images by randomly selecting unpaired MRI scans and labeled data from BraTS 2015. Fig. 1 illustrates the flow of the proposed RescueNet in which we train two generator networks (G and F) and two discriminators (D_T and D_L) simultaneously. In cycle-1, the network G is used to generate the tumor segmentation $\left(l_i^1\right)$ from image t_i and the network F synthetically generates an image back $\left(t_i^1\right)$ which is similar to 't_i'. Similarly, cycle-2 uses the network F to generate MRI scan $\left(t_i^2\right)$ from image and G to generate the labeled image $\left(l_i^2\right)$ back which is similar to l_i. The D_T discriminates between t_i and t_i^2 and D_L aims to distinguish l_i and l_i^1. Assume t_i is sampled from T i.i.d. according to some distribution P_T and is sampled from L i.i.d. according to some distribution P_L. The overall procedure involves two different types of losses: *adversarial losses* and *cycle consistency losses*.

2.1. Adversarial losses

2.1.1. Adversarial loss for G:

The adversarial loss is calculated from results of G which look similar to labeled data (L) as given in Eq. (1):

$$L_G(G, D_L) = E_{l_i \sim P_l}[\log D_L(l_i)] + E_{t_i \sim P_T(\hbar)}[\log(1 - D_L(G(t_i)))] \quad (1)$$

2.1.2. Adversarial loss for F

The adversarial loss calculated from results of F which resembles the MRI scan data (T) as given in Eq. (2):

$$L(F, D_T) = E_{t_i \sim P_T}[\log D_T(t_i)] + E_{l_i \sim P_L(\hbar)}[\log(1 - D_T(F(l_i)))] \quad (2)$$

The generator competes with the discriminator and tries to generate the indistinguishable output with the real labeled data. The discriminator tries to approach the accuracy in discriminating the generated data and the real data.

2.2. Cycle consistency losses

The ideal case of training is that the generators G and F should able to transfer the domains $T \rightarrow L$ and $L \rightarrow T$ respectively. Similar to cycleGAN [29], we also use cycle consistency losses as given below for domain transfer which cannot be obtained sufficiently using only adversarial training between real and generated data.

2.2.1. Forward cycle consistency loss (cycle-1):

The L_1 loss is calculated between the t_i and t_i^1 ($t_i \rightarrow G(t_i) \rightarrow F(G(t_i)) \approx t_i$) using Eq. (3):

$$L_{cycle\text{-}1}(G, F) = E_{t_i \sim P_T}\left[\left\|t_i^1 - t_i\right\|_1\right] \quad (3)$$

2.2.2. Backward cycle consistency loss (cycle-2)

The calcuclation of L1 loss is made between the l_i and l_i^2 ($l_i \rightarrow F(l_i) \rightarrow G(F(l_i)) \approx l_i$) using Eq. (4):

$$L_{cycle\text{-}2}(G, F) = E_{l_i \sim P_L}\left[\left\|l_i^2 - l_i\right\|_1\right] \quad (4)$$

The overall loss is given in Eq. (5) and it comprises of all the above four losses in a weighted manner:

$$L_o(G, F, D_T, D_L) = \lambda_1 L_G + \lambda_2 L_F + \lambda_3 L_{cycle\text{-}1} + \lambda_4 L_{cycle\text{-}2} \quad (5)$$

where λ_i; $i = 1, 2, 3, 4$ controls the importance of the losses.

3. The proposed network architecture

Proposed approach for brain tumor segmentation from MRI scan is discussed in this section. The basic network architecture RescueNet followed by its subsequent varients RescueWNet, RescueCNet and RescueENet for whole, core and enhance tumor segmentation respectively are explained here.

The choice of network for segmentation task is a very tedious job. There are some existing networks [15,16] in the literature which are used for segmentation purpose. However, low-level dependencies are failed to render using these networks. U-Net architecture [8] overcomes this drawback by adding skip connections between the down-sampling and up-sampling path. The popularly used term for such kind of arrangement is mirroring. Vanishing gradient is another problem in training of deep networks. Identity mappings proposed in residual architecture [19] disables vanishing gradients problem during back-propagation and helps to learn the network parameter.

Inspired from the advantages of above discussed network architecture, a novel RescueNet is designed for brain tumor segmentation. Proposed RescueNet comprises both mirroring and residual concepts. Fig. 2 gives the pictorial description of the proposed RescueNet. It consists of an encoding path (from 256×256 to 32×32) to retrieve the features of interest and a symmetric decoding path (from 64×64 to 256×256) that enables the prediction of segmented tumor. Four types of basic building blocks are used to construct the proposed network. Each green block is a regular convolutional layer followed by instant normalization layer and ReLU [33]. Each violet block is a residual block that consists of two convolutional layers followed by identity mapping (residual skip connection). Another residual block with first convolution layer of stride 2 is located between levels in the encoding path to perform downsampling for feature compression is represented by blue blocks. These blocks are having identity mappings (skip

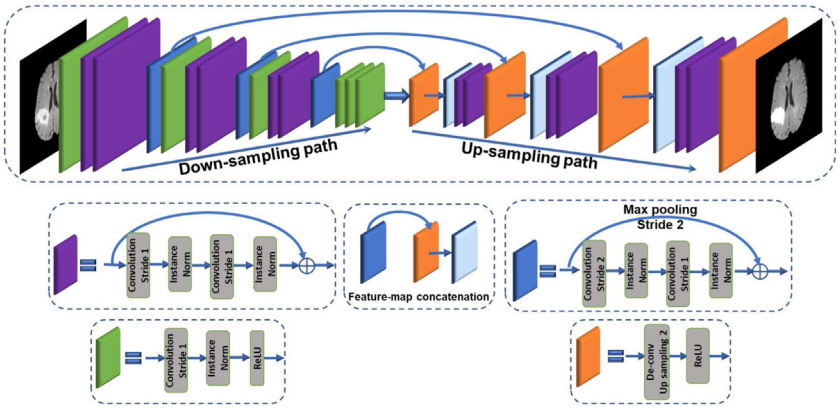

Fig. 2. Generator architecture of proposed RescueNet.

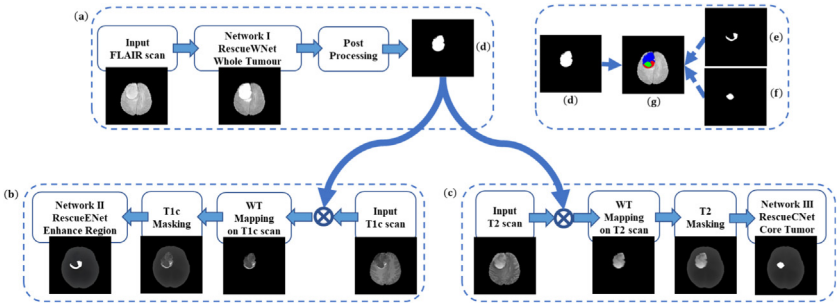

Fig. 3. Overall pipeline of the proposed brain tumor segmentation approach. (a), (b) and (c) gives the procedure for whole, enhance and core tumor region segmentation respectively. (d), (e) and (f) segmented whole, enhance and core tumor region using proposed RescueWNet, RescueENet and RescueCNet respectively. (g) The mapping of segmented region on FLAIR MRI scan. Note: Blue, Red and Green color shows the whole, enhance and core tumor region respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

connections) with maxpooling layers. In the decoding path, the deconvolutional layer (orange blocks) is located between levels to up-sample the input data. The residual blocks solve the problem of vanishing gradient during backpropagation. There is one convolutional layer before and after a set of residual blocks. This convolutional layer serves as a connector to bridge the input feature maps and the residual block because the number of feature maps from the previous layer may differ from that of the residual block. Also, this arrangement gives symmetricity to the proposed network. A concatenation between symmetric feature maps from the encoding path and the output of each deconvolutional layer from the decoding path is provided for eliminating the issue of loss in resolution. The number of feature maps and their sizes which provide the more specified details of proposed network is provided in Table 1. For the discriminator network, we used Patch-GAN [27,29] which can distinguish whether the input coming to it is real or fake. In next subsection, variants of proposed network RescueWNet, RescueCNet and RescueENet for whole, core and enhance tumor segmentation respectively are discussed.

3.1. RescueWNet for whole tumor segmentation

As discussed in Section 1, FLAIR scans are used to distinguish CSF from edema region. Thus, in this study, FLAIR MRI scans are used to segment whole tumor. Fig. 3(a) shows the procedure for whole tumor segmentation. Initially, FLAIR MRI scan is fed into the RescueWNet to generate segmentation results for whole tumor region followed by a novel scale-invariant post processing algorithm. Fig. 3(d) shows the segmented whole tumor. Need of the post processing is explained as follows.

Post-processing: It is observed that the initial and final scans of a FLAIR MRI volume does not contain any tumor still these slices have brighter shades for non-tumorous regions. This could be the major difficulty for any segmentation algorithm (even for CNN-based approaches) to perform segmentation on these regions. Because, every texture-based segmentation algorithm works on the separation between gray shades within different regions of an image which may segment this bright region as a brain tumor. Thus, a

Table 1
Network configuration and parameter details of proposed RescueNet. Note: Re1 to Re9 are the residual blocks on encoder side, Rd10 to Rd15 are residual blocks on decoder side, C1 to C9 are the convolutional layers and D4 to D0 are de-convolutional layers.

Generator building block	Residual block	Layers (stride)	Filter size
Encoder	–	C1 (1)	32
	Re1, Re2	C1-1 (1), C1-2 (1)	32
	Re3	C1-1 (2)	32
		C1-2 (1)	32
	–	C2 (1)	64
	Re4, Re5	C2-1 (1), C2-2 (1)	64
	Re6	C2-1 (2)	64
		C2-2 (1)	64
	–	C3 (1)	128
	Re7, Re8	C3-1 (1), C3-2 (1)	128
	Re9	C3-1 (2)	128
		C3-2 (1)	128
Transformation	–	C4 (1), C5 (1)	128
	–	C6 (1)	256
Decoder	–	D4 (1)	128
	Rd10, Rd11	C7-1 (1), C7-2 (1)	256
	–	D3 (2)	64
	Rd12, Rd13	C8-1 (1), C8-2 (1)	128
	–	D2 (2)	32
	Rd14, Rd15	C9-1 (1), C9-2 (1)	64
	–	D1 (2)	32
	–	D0 (1)	3

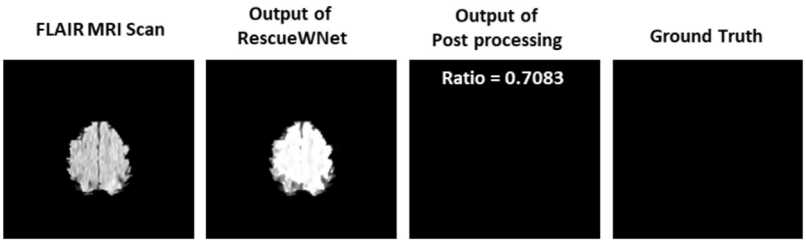

Fig. 4. Result of scale invariant approach (post processing) used in whole tumor segmentation.

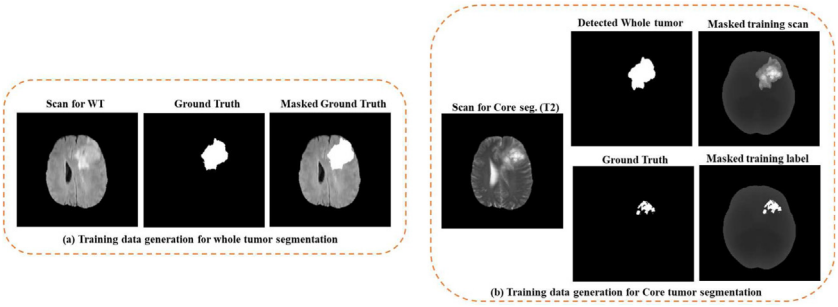

Fig. 5. Proposed training data preparation. (a) For whole tumor segmentation and (b) for core tumor segmentation. Note: we have used similar procedure as (b) for training data preparation of the enhance region.

post-processing algorithm is needed to make sure the accurate segmentation in initial and final slices.

In this work, we propose a novel scale-invariant post processing algorithm to filter-out the initial and end MRI scans and thus improves the reliability of the proposed approach. Proposed post processing algorithm is based on the ratio of the segmented tumor region from MRI scan (*i.e. number of pixels corresponds to the segmented brain tumor*) to actual brain region from MRI scan (*i.e. number of pixels belongs to entire brain*). Thus, a sample segmented scan for which this ratio exceeds the predefined threshold is discarded and an empty mask is considered as a segmentation result. Proposed post processing algorithm is scale-invariant (*as it is based upon the ratio*) and valid when area of the entire brain in MRI scan is nearly equals to the area of segmented brain tumor. Ideal value for the ratio is 1 however, after extensive analysis, we have decided ratio = 0.4 which is used as a threshold in our experiments for whole tumor segmentation. The effect of proposed post-processing algorithm is shown in Fig. 4.

3.2. RescueCNet for core tumor segmentation

For the core tumor segmentation we used T2 MRI scans. The whole procedure of training is shown in Fig. 3(c). A cascaded approach is being used in which whole tumor region segmented using RescueWNet is mapped onto the T2 scan and region is extracted from T2 scan. Further, to avoid the false detection we mapped extracted region on a dummy mask[1] (see Fig. 3(c)). The mapping procedure is important because T2 scans give bright signals for cerebrospinal fluid (CSF) region also. Finally, this T2 masked image is provided as the input to the RescueCNet to generate segmented core tumor region. Fig. 3(f) depicts the segmented core tumor.

3.3. RescueENet for enhance tumor segmentation

The T1c scans that shows completely brighter enhanced tumor region due to the contrast enhancing agent are used for segmenting enhance tumor region. Using the similar procedure of core tumor segmentation, we generate masked T1c scans for RescueENet to segment enhance tumor region. Fig. 3(e) shows the segmented enhance tumor region.

4. Experiments and discussion

4.1. Training pipeline

Image-to-image style transfer is the conversion of the data from one domain (style) to other domain (style). It can be understood by correlating it to transferring of one type of texture pattern (horse) to other texture patterns (zebra) using semantic texture analysis. In this paper, the tumor segmentation from brain MRI data is considered as the style transfer problem. The proposed RescueNet transfers the original brain MRI data style to labeled (segmented tumor) style.

4.1.1. Whole tumor segmentation

The labeled data is prepared using masking operation as illustrated in Fig. 5(a), in order to perform the whole tumor segmentation on brain MRI which is similar to style transfer. For a given FLAIR MRI scan, the labeled data (ground truth) is collected and masked with the MRI scan in such a way that tumor region is represented by labeled image content and rest of the region from the MRI scan itself. The resultant image is considered as the labeled image for the given MRI scan. We have randomly selected 10% slices from HGG and 10% from LGG patients from BraTS 2015 dataset for training operation and ground truths are generated by following the procedure discussed above. Further, these scans along with the masked ground truths are shuffled to achieve the unpaired training using proposed RescueWNet for whole tumor segmentation.

4.1.2. Enhance and core tumor segmentation

For making our training process more accurate towards core tumor segmentation and its convergence in the correct direction, we generated the masked input T2-weighted MRI scan and its corresponding masked ground truth as shown in Fig. 5(b). We extracted a region from T2-weighted MRI scan using the segmented whole tumor segmentation result and mapped it to a synthetically

[1] Dummy mask is generated by normalized summation of all the FLAIR MRI scans from BraTS 2015 database.

Table 2
Performance comparison between the proposed method and existing state-of-the-art methods for the whole, core and enhance tumor segmentation on BraTS 2015 database.

Method (testing data %)	Dice			Sensitivity		
	Whole	Core	Enhance	Whole	Core	Enhance
Pereira et al. [20] (15%)	0.78	0.65	0.75	0.87	0.75	0.75
Yi et al. [34] (–)	0.89	0.76	0.8	–	–	–
Dong et al. [9] (–)	0.86	0.86	0.65	–	–	–
Zhao et al. [24] (18%)	0.8	0.68	0.61	0.81	0.71	0.68
Kamnitsas et al.[21] (40%)	0.85	0.67	0.63	0.88	0.6	0.67
Cui et al. [22] (10%)	0.89	0.77	0.8	0.87	0.84	0.76
Xue et al. [26] (10%)	0.85	0.7	0.66	0.8	0.65	0.62
Kong et al. [10] (30%)	0.8993	0.6839	0.771	0.9581	0.7769	0.6067
Proposed method (90%)	0.9401	0.9429	0.8732	0.88	0.9489	0.9124

Table 3
Performance comparison between the proposed method and existing state-of-the-art methods for the whole, core and enhance tumor segmentation on BraTS 2017 database.

Method (testing data %)	Dice			Sensitivity		
	Whole	Core	Enhance	Whole	Core	Enhance
Alex et al. [11] (–)	0.83	0.69	0.69	–	–	–
Beers et al. [35] (–)	0.88	0.73	0.73	–	–	–
Wang et al. [13] (50%)	0.9	0.83	0.78	–	–	–
Kong et al. [10] (40%)	0.908	0.7353	0.6867	0.9624	0.9368	0.6413
Proposed method (90%)	0.9463	0.856	0.9354	0.9127	0.86	0.9587

generated scan. Also we mapped ground truth of core tumor region on the similar synthetically generated scan. Similar to whole tumor segmentation, we consider 10% masked T2-weighted MRI scans and random 10% masked ground truths for the training of RescueCNet. Here input MRI scans and ground truths are shuffled to achieve the unpaired training. The same procedure is applied for enhancing region segmentation using RescueENet (T1c-weighted MRI scans are used for enhancing region segmentation). Weight parameters of networks are updated in 200 epochs on NVIDIA DGX station with processor 2.2 GHz, Intel Xeon E5-2698, NVIDIA Tesla V100 4 × 16 GB GPU.

4.2. Patient-wise data acquisition

The proposed method is trained on BraTS 2015 dataset. It consists of 220 HGG and 54 LGG patient scans. HGG patients include edema, enhancing, non-enhancing and necrotic part while in LGG enhancing part may or may not be there. The multimodal MRI scans are taken with different pulse sequences for each patient using T1-weighted (T1), T1-weighted imaging with gadolinium-enhancing contrast (T1c) shows the visibility of active tumor region containing only the enhancing core structures, T2-weighted (T2) shows the visibility of tumor core which includes all tumor structures excepts edema and FLAIR shows whole tumor region including all tumor structures, edema, non-enhancing (solid) core, enhancing core and necrotic or fluid filled core. For each patient, the T1, T2 and FLAIR images were co-registered into the T1c data, which had the finest spatial resolution, and then resample and interpolated into $1 \times 1 \times 1\ mm^3$ with an image size of $240 \times 240 \times 155$. Each scan comprises of 155 brain slices. The HGG has a short life expectancy and has limited treatment options. The aggressive nature of this illness necessitates efficient diagnosis and treatment planning to improve the quality of and extend patient life.

We extracted 2D slices for each image. As most of the slices do not contain any data, we choose a minimum set threshold for whole tumor region, core tumor region and enhancing tumor region. We extract separate binary segmentation mask for each region and use them against different MRI sequences.

For the testing, BraTS 2017 database is also considered along with BraTS 2015 database. The BraTS 2017 database contains images and annotations of BraTS 2012 and BraTS 2013. Due to the fact that BraTS 2014 and BraTS 2016 datasets contains a mixture of pre- and post-operative scans, also their ground truth labels have been annotated by the fusion of segmentation results from algorithms, they have been discarded in this issue. This dataset contains 210 HGG scans and 75 LGG scans with the same pulse sequences as used for BraTS 2015.

5. Evaluation

The experiment is performed on BraTS 2015 and BraTS 2017 dataset consists of HGG and LGG divided into four different classes. The dataset contains some images for which ground truth is not available. Therefore, we consider random data to train our model. The BraTS 2015 dataset is divided randomly into training and testing sets in 10:90 ratio. BraTS 2017 dataset is tested on the same network which is trained on BraTS 2015 dataset.

The evaluation has been done on HGG and LGG data for segmentation of three tumor regions namely: (1) whole tumor comprises of all regions edema, advancing and non-advancing tumor (label 1, 2 and 4); (2) core tumor region comprises of necrotic core and non-advancing region (label 1); (3) enhancing region (label 4).

The method opted for evaluating these segmented regions against the ground truth is calculation of Dice similarity coefficient (DSC), which gives the ratio between the intersection of actual label with the predicted label to the sum of actual and predicted label. Sensitivity measure is used to detect number of true positive and false negative.

5.1. Results analysis

In this paper, we designed a very elegant symmetric neural network named RescueNet for brain tumor segmentation. Superior accuracy has been achieved by the proposed method for segmentation using unpaired medical data. Proposed network is feasible to train on very less number of data from BraTS 2015 database and tested for nearly 90% scans from both BraTS 2015 and BraTS 2017 database. It is an end-to-end network with concatenations in between DownBlock and UpBlock. We specifically designed the DownBlock structure and the UpBlock structure to adopt these connections. Also the use of residual blocks in both decoder and encoder side to reduce the problem of vanishing gradient during

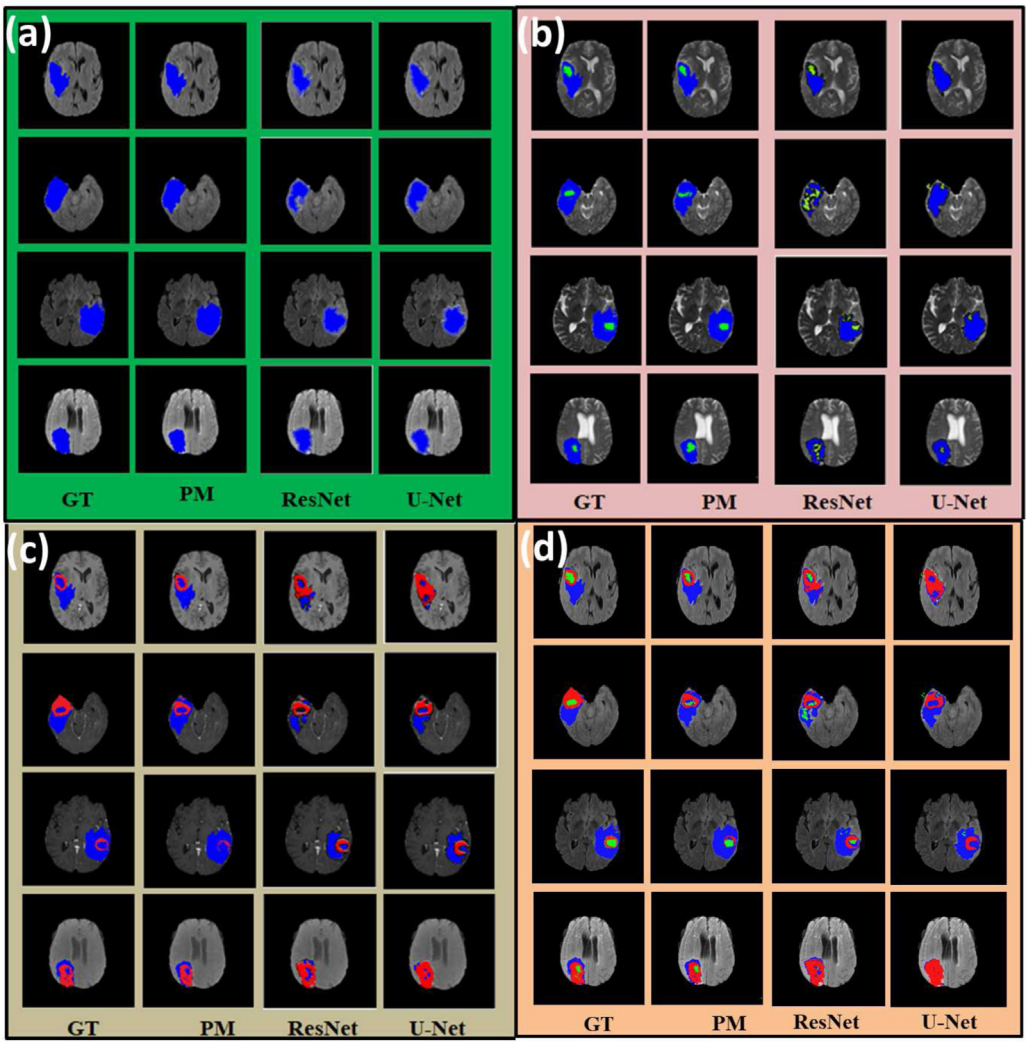

Fig. 6. Segmentation results for MRI scans from BraTS 2015 and 2017 database for whole, core and enhance tumor region. Sample MRI scans GT (ground truth) from BraTS database and their respective results using the PM (proposed approach), ResNet as generator and U-Net as generator are given row-wise. (a) Result of whole tumor segmentation on FLAIR MRI scan. (b) Result of whole and core tumor segmentation on T2 MRI scan. (c) Result of whole and enhance tumor segmentation on T1c MRI scan. (d) Combined result on FLAIR MRI scan. Note: Blue, Red and Green color shows the whole, enhance and core tumor region respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

back propagation. We used DICE and sensitivity as two similarity measurement metrics for evaluation purpose. The illustration in Tables 2 and 3 demonstrate the comparison of proposed method with other state-of-art methods for brain tumor segmentation on BraTS 2015 and BraTS 2017 database. The false positives are removed using scale invariant ratio approach based on threshold. The performance of the proposed RescueNet is tested on 90% of total MRI scans from both BraTS 2015 and 2017 dataset. Whereas, other existing methods considered less percentage of scans for testing (highest is 50%). The DICE value reached to 0.9401% on BraTS 2015 dataset and 0.9463% on BraTS 2017 dataset for whole tumor segmentation which is more than any other existing method. *Results have shown that the testing DICE value for BraTS 2017 dataset is high even with network training on BraTS 2015.* This gives our network a global approach for different kind of datasets. The post processing approach and the use of masking increases the DICE value of segmented enhance tumor region and core tumor region to 0.8732% and 0.9429% for BraTS 2015 datset and 0.9354% and 0.8560% for BraTS 2017 dataset. The sensitivity measure obtained using proposed approach for BraTS 2015 database are 0.8800%, 0.9489% and 0.9124% for whole, core and enhance tumor region respectively and 0.9127%, 0.8600% and 0.9587% for BraTS 2017 dataset. Fig. 6 shows the visual results of brain tumor segmentation on sample MRI scans. Also, we have shown results obtained for whole, core and enhance tumor region using U-Net [8] and ResNet [19] as generator network with unpaired training. We observed that the visual results obtained using the proposed network are better than other two approaches. From Tables 2, 3 and Fig. 6 we can observe that the proposed approach is superior both quantitatively and qualitatively.

6. Conclusion

The end-to-end network architecture is proposed in this paper for brain tumor segmentation namely RescueNet. Need of large-scale annotated brain tumor MRI data for training of deep networks is overcome by proposing an unpaired training approach. Unpaired training can be used to train MRI images for which labeled annotations are not available by utilizing labels from some known dataset. Thus, can be used to produce better results on any dataset. Separate training for whole, core and enhance tumor is performed using three different networks. Proposed RescueNet designed using residual and mirroring principles. It requires very less amount of data for training and is used to produce better segmentation

results on large amount of testing data. Also, the introduction of scale-invariant approach for post processing improves the network performance for precise whole tumor segmentation. BraTS 2015 and BraTS 2017 brain tumor database are used to evaluate the performance of proposed network. Sensitivity and DICE coefficient measure is adopted for the performance measure, which shows better performance when compared to other methods. In future, we intend to combine the multi-task learning technique to proposed architecture to further enhance accuracy. The network can be used for inter-conversion of modalities in future to combine more features to improve allocation of tumor more accurately. Also, we are planning to use the similar architecture for other imaging modalities and general purpose segmentation tasks.

Declaration of Competing Interest

None declared.

References

[1] B.H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, Y. Burren, N. Porz, J. Slotboom, R. Wiest, et al., The multimodal brain tumor image segmentation benchmark (brats), IEEE Trans. Med. Imaging 34 (2015) 1993.

[2] N.R. Smoll, K. Schaller, O.P. Gautschi, Long-term survival of patients with glioblastoma multiforme (GBM), J. Clin. Neurosci. 20 (2013) 670–675.

[3] Z. Li, Y. Wang, J. Yu, Y. Guo, W. Cao, Deep learning based radiomics (DLR) and its usage in noninvasive IDH1 prediction for low grade glioma, Sci. Rep. 7 (2017) 5467.

[4] Z. Wu, K.D. Paulsen, J.M. Sullivan, Adaptive model initialization and deformation for automatic segmentation of T1-weighted brain MRI data, IEEE Trans. Biomed. Eng. 52 (2005) 1128–1131.

[5] J. Liu, F. Chen, C. Pan, M. Zhu, X. Zhang, L. Zhang, H. Liao, A cascaded deep convolutional neural network for joint segmentation and genotype prediction of brainstem gliomas, IEEE Trans. Biomed. Eng. (2018).

[6] R. Ayachi, N.B. Amor, Brain tumor segmentation using support vector machines, in: European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty, Springer, 2009, pp. 736–747.

[7] A. Liaw, M. Wiener, et al., Classification and regression by randomforest, R News 2 (2002) 18–22.

[8] O. Ronneberger, P. Fischer, T. Brox, U-net: convolutional networks for biomedical image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2015, pp. 234–241.

[9] H. Dong, G. Yang, F. Liu, Y. Mo, Y. Guo, Automatic brain tumor detection and segmentation using U-Net based fully convolutional networks, in: Annual Conference on Medical Image Understanding and Analysis, Springer, 2017, pp. 506–517.

[10] X. Kong, G. Sun, Q. Wu, J. Liu, F. Lin, Hybrid pyramid U-Net model for brain tumor segmentation, in: International Conference on Intelligent Information Processing, Springer, 2018, pp. 346–355.

[11] V. Alex, M. Safwan, G. Krishnamurthi, Brain tumor segmentation from multi modal MR images using fully convolutional neural network, Medical Image Computing and Computer Assisted Intervention – MICCAI (2017) 1–8.

[12] M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P.-M. Jodoin, H. Larochelle, Brain tumor segmentation with deep neural networks, Med. Image Anal. 35 (2017) 18–31.

[13] G. Wang, W. Li, S. Ourselin, T. Vercauteren, Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks, in: International MICCAI Brainlesion Workshop, Springer, 2017, pp. 178–190.

[14] S. Hussain, S.M. Anwar, M. Majid, Brain tumor segmentation using cascaded deep convolutional neural network, in: 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2017, pp. 1998–2001.

[15] H. Noh, S. Hong, B. Han, Learning deconvolution network for semantic segmentation, Proceedings of the IEEE International Conference on Computer Vision (2015) 1520–1528.

[16] V. Badrinarayanan, A. Kendall, R. Cipolla, Segnet: a deep convolutional encoder–decoder architecture for image segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 39 (2017) 2481–2495.

[17] M. Drozdzal, E. Vorontsov, G. Chartrand, S. Kadoury, C. Pal, The importance of skip connections in biomedical image segmentation, in: Deep Learning and Data Labeling for Medical Applications, Springer, 2016, pp. 179–187.

[18] K. Simonyan, A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition, 2014 arXiv:1409.1556.

[19] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2016) 770–778.

[20] S. Pereira, A. Pinto, V. Alves, C.A. Silva, Brain tumor segmentation using convolutional neural networks in MRI images, IEEE Trans. Med. Imaging 35 (2016) 1240–1251.

[21] K. Kamnitsas, C. Ledig, V.F. Newcombe, J.P. Simpson, A.D. Kane, D.K. Menon, D. Rueckert, B. Glocker, Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation, Med. Image Anal. 36 (2017) 61–78.

[22] S. Cui, L. Mao, J. Jiang, C. Liu, S. Xiong, Automatic semantic segmentation of brain gliomas from MRI images using a deep cascaded neural network, J. Healthcare Eng. 2018 (2018).

[23] G.-C. Lin, W.-J. Wang, C.-M. Wang, S.-Y. Sun, Automated classification of multi-spectral MR images using linear discriminant analysis, Comput. Med. Imaging Graph. 34 (2010) 251–268.

[24] X. Zhao, Y. Wu, G. Song, Z. Li, Y. Zhang, Y. Fan, A deep learning model integrating FCNNS and CRFS for brain tumor segmentation, Med. Image Anal. 43 (2018) 98–111.

[25] D. Nie, R. Trullo, J. Lian, L. Wang, C. Petitjean, S. Ruan, Q. Wang, D. Shen, Medical image synthesis with deep convolutional adversarial networks, IEEE Trans. Biomed. Eng. (2018).

[26] Y. Xue, T. Xu, H. Zhang, L.R. Long, X. Huang, Segan: adversarial network with multi-scale l 1 loss for medical image segmentation, Neuroinformatics (2018) 1–10.

[27] P. Isola, J.-Y. Zhu, T. Zhou, A.A. Efros, Image-to-image translation with conditional adversarial networks, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2017, pp. 5967–5976.

[28] M. Rezaei, K. Harmuth, W. Gierke, T. Kellermeier, M. Fischer, H. Yang, C. Meinel, A conditional adversarial network for semantic segmentation of brain tumor, in: International MICCAI Brainlesion Workshop, Springer, 2017, pp. 241–252.

[29] J.-Y. Zhu, T. Park, P. Isola, A.A. Efros, Unpaired image-to-image translation using cycle-consistent adversarial networks, IEEE International Conference on Computer Vision (2017).

[30] X. Yang, Z. Xu, J. Luo, Towards perceptual image dehazing by physics-based disentanglement and adversarial training, The Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18) (2018).

[31] H. Chang, J. Lu, F. Yu, A. Finkelstein, Pairedcyclegan: asymmetric style transfer for applying and removing makeup, 2018 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2018).

[32] D. Engin, A. Genc, H. Kemal Ekenel, Cycle-dehaze: Enhanced cyclegan for single image dehazing, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (2018) 825–833.

[33] D. Ulyanov, A. Vedaldi, V. Lempitsky, Instance Normalization: The Missing Ingredient for Fast Stylization, 2016 arXiv:1607.08022.

[34] D. Yi, M. Zhou, Z. Chen, O. Gevaert, 3-D Convolutional Neural Networks for Glioblastoma Segmentation, 2016 arXiv:1611.04534.

[35] A. Beers, K. Chang, J. Brown, E. Sartor, C. Mammen, E. Gerstner, B. Rosen, J. Kalpathy-Cramer, Sequential 3D U-Nets for Biologically-Informed Brain Tumor Segmentation, 2017 arXiv:1709.02967.