IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received September 29, 2019, accepted October 15, 2019, date of publication October 21, 2019, date of current version November 5, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2948476

Skip Connection U-Net for White Matter Hyperintensities Segmentation From MRI

JIONG WU[1], YUE ZHANG[2,3], KAI WANG[1], AND XIAOYING TANG[2]

[1]School of Electronics and Information Technology, Sun Yat-sen University, Guangzhou 510006, China
[2]Department of Electrical and Electronic Engineering, Southern University of Science and Technology, Shenzhen 518055, China
[3]Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong

Corresponding authors: Kai Wang (wangkai23@mail.sysu.edu.cn) and Xiaoying Tang (tangxy@sustech.edu.cn)

This work was supported in part by the National Key Research and Development Program of China under Grant 2017YFC0112404, and in part by the National Natural Science Foundation of China under Grant NSFC 81501546.

ABSTRACT White matter hyperintensity (WMH) is associated with various aging and neurodegenerative diseases. In this paper, we proposed and validated a fully automatic system which integrates classical image processing and deep neural network for segmenting WHM from fluid attenuation inversion recovery (FLAIR) and T1 magnetic resonance (MR) images. In this system, a novel skip connection U-net (SC U-net) was proposed. In addition, an atlas-based method was introduced in the preprocessing stage to remove non-brain tissues (namely skull-stripping) and thus to improve the segmentation accuracy. Effectiveness of the proposed system was validated on a dataset of 60 paired images based on cross-scanner validation. Our experimental results revealed the effectiveness of the skull-stripping strategy. More importantly, compared to two existing state-of-the-art methods for segmenting WHM, including a U-net-like method and another deep learning method, the proposed SC U-net had a faster convergence, a lower loss and a higher segmentation accuracy. Both quantitative and qualitative analyses (via visual examinations) revealed the superior performance of our proposed SC U-net. The mean dice score of the proposed SC U-net was 78.36% which was much higher than those of a U-net-like method (74.99%) and an alternative deep learning method (74.80%). The software environment and model of the proposed system were made publicly accessible at Dockerhub.

INDEX TERMS Deep learning, white matter hyperintensity, skip connection, U-net.

I. INTRODUCTION

White matter hyperintensities (WMH), also known as leukoaraiosis, are brain areas of increased signal intensities, indicating macroscopic changes of brain tissues induced by white matter damages [1].

A. CLINICAL MOTIVATION

WMH are characteristics of aging and neurodegenerative diseases targeting the white matter, including stroke, dementia and multiple sclerosis (MS) [2]–[5]. These diseases may cause irreversible damages to the human brain [6], [7]. Although etiologies of these diseases are not yet fully understood, there is considerable evidence suggesting that they are related to WMH [3], [8]. The white matter is primarily composed of myelinated axons. Trapp and colleagues concluded that MS is an immune-mediated demyelinating disease [9]. Also a systematic review and meta-analysis have been conducted to demonstrate that WMH may be an important biomarker for predicting the risk of stroke, dementia and mortality [3]. In addition, since WMH often occur in the preclinical stage of dementia, the presence of WMH may increase the likelihood of developing from mild cognitive impairment (MCI) to dementia [10]. Most of the risk factors associated with WMH have been shown to be similar to those associated with cerebrovascular and cardiovascular diseases, including elevated blood pressure, diabetic atherosclerosis and homocysteine levels [11]. WMH can be observed on magnetic resonance images (MRIs), especially on images scanned by the fluid attenuated inversion recovery (FLAIR) sequence [6]. The volume of WMH has been demonstrated to correlate with symptoms' severity, disability progression and clinical outcomes [12]–[14]. The WMH volume can be measured non-invasively and quantitatively to aid the design of treating plans.

The associate editor coordinating the review of this manuscript and approving it for publication was Muhammad E. H. Chowdhury.

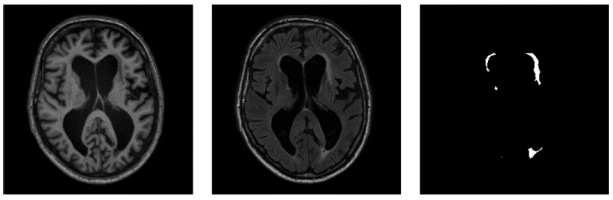

FIGURE 1. A sample of MR slice from T1 (left) and FLAIR modality (middle), and its corresponding WMH mask (right).

Accordingly, quantifying WMH volume has become a focus in clinical research, of which a prerequisite is to segment out the WMH (see Fig. 1). Manually delineating WMH is reliable and accurate, but it is labor intensive, time consuming and subjective [15]. Therefore, it is urgent to develop an automatic segmentation method for WMH.

B. RELATED WORK

Over the last few years, various automatic WMH segmentation methods have been proposed. These methods can be broadly classified into unsupervised methods and supervised methods, depending on whether expertly annotated data is available or not.

Unsupervised methods do not require annotated WMH masks in the entire segmentation process. Since WMH are brighter than healthy brain tissues when revealed on FLAIR images, many unsupervised methods used thresholding to segment WMH. For instance, Jack *et al.* [16] used an optimal FLAIR intensity threshold to segment WMH based on image histogram. Gibson *et al.* [17] proposed a method using a conservative FLAIR threshold followed by two-class fuzzy C-means clustering (FCM). Yoo *et al.* [18] introduced an optimal threshold intensity which varies with WMHs volume for segmenting WMH.

Most other unsupervised methods are probabilistic models and are usually designed as hybrid models [19]. For example, Van Leemput *et al.* [20] proposed a weighted Expectation-Maximization (EM) model to detect MS lesions from a large dataset of T1-, T2- and proton density (PD) weighted scans. This was an early attempt to develop an atlas-based technique and the total lesion load correlation to the automated segmentation results is higher than that to the manual delineations. Schmidt *et al.* [21] proposed a lesion growth algorithm (LGA) for WMH segmentation using a parametric mixture model in which a Markov random field (MRF) was embedded. A Bayesian MRF relief was employed by Schwarz *et al.* [22] using lognormal distributions for detection white matter and WMH.Wu *et al.* [23] proposed a multi-atlas based method for simultaneously detecting and localizing WMH on FLAIR data. Although these unsupervised methods obtained some promising results, some parameters of these unsupervised models may be inadvertently over-tuned across different datasets.

Different from the aforementioned unsupervised methods, many supervised methods utilizing various classifiers based on manually delineated data to segmenting WMH were proposed. In the early stage, researchers usually used K-Nearest Neighbours (K-NN) as the classifiers for voxel-wise WMH classification. After that, many researchers used some other models to segment WMH. For instance, Dadar *et al.* [24] proposed a linear model trained by using spatial and intensity features from multiple MRI contrasts and manually labeled data to predict the lesion.

Recently, with the development of deep learning methods especially the Convolutional Neural Networks (CNNs) in computer vision, CNNs have been introduced into biomedical image processing [25]–[27]. In terms of WMH segmentation, many CNN-based methods have been proposed and have yielded promising results [28]–[35]. To be specific, Xu *et al.* [31] applied transfer learning by using pre-trained Visual Geometry Group (VGG) model on Image-Net for natural image classification and fine-tuned model on WMH data. Guerrero *et al.* [34] used a CNN trained with large high resolution image patches and differentiated between WMH-related pathology and stroke. Among these CNN methods, one of the best methods for WMH is a U-net-like method proposed elsewhere [30]. U-net has been widely used in segmenting biomedical images because it can work efficiently even with limited training samples [36]. In the work of [30] three identical U-net-like models with different randomly-initialized weights were adopted and ensembled and ranked the highest in a WMH segmentation challenge [37]. Although this U-net-like model in [30] has good performance, the connections were only designed between some layers of the down-convolutional part and those of the up-convolutional part resulting in a loss of segmentation accuracy and more time for convergence. To alleviate these two issues, based on the U-net-like model proposed in [30], we propose a new U-net-like architecture, namely skip connection U-net model (SC U-net), by utilizing a novel skip connection strategy.

C. CONTRIBUTION

In this paper, we propose a novel fully automatic system for WHM segmentation. The main contribution of this work is that we propose a new skip connection U-net model, SC U-net, to classify and localize WMH. This architecture consists of a shrinking part that aims to capture context, a symmetric expansive part that gradually combines features to enable a precise localization, and a skip connection part that alleviates the vanishing gradient problem and hence improve the optimization convergence speed. Another contribution is the introduction of a preprocessing step, skull-stripping, to improve WMH segmentation. In addition, we encapsulate our method and all related dependencies into a docker image which is made publicly available at DockerHub [1]. To make our neural network stable, we apply the MRSA initialization [38] in the training stage.

[1] https://hub.docker.com/r/skipconnection/wmhchallenge_scunet/

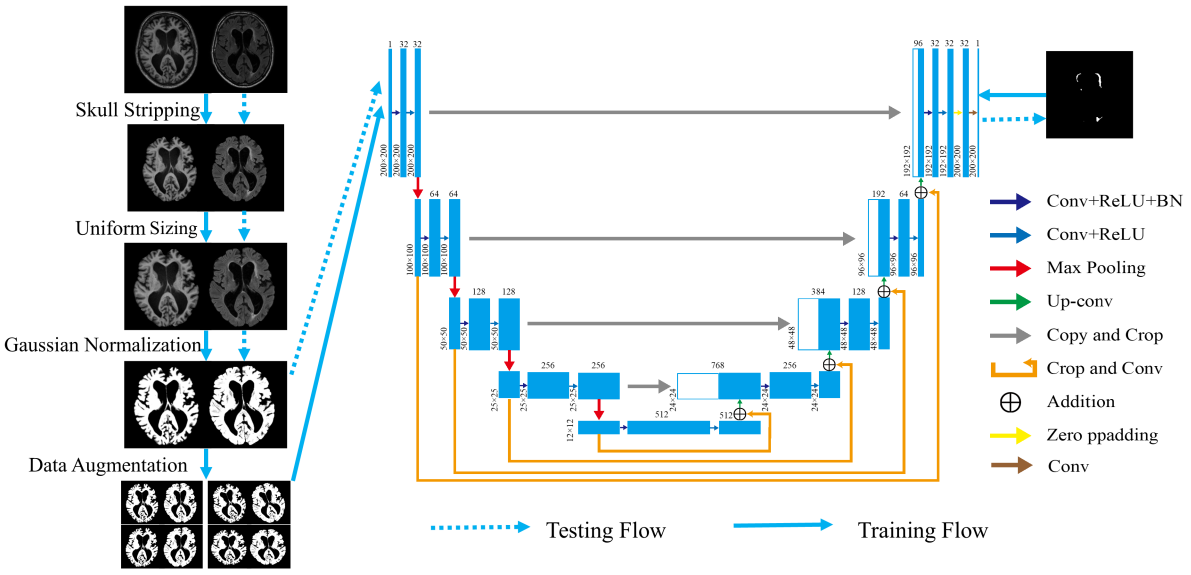

FIGURE 2. The flowchart of the proposed automatic WMH segmentation algorithm, including preprocessing, training, and testing.

This paper is organized as follows. In Section II, the flowchart of our method is depicted and the data processing procedures, as well as SC U-net, are detailed. The datasets, evaluation criteria and five metrics on assessing segmentation performance are presented in Section III. Evaluations of the proposed method on a public training dataset employing leave-one-out testing and cross scanner testing are presented in Section IV. Finally, Section V discusses the advantages of the proposed method and how the key parameters are optimized.

II. METHOD

As shown in Fig. 2, our pipeline for WMH segmentation consists of a data preprocessing procedure, a SC U-net training procedure and a testing procedure. In the data preprocessing, four steps including skull stripping, uniform sizing, gaussian normalization and data augmentation are conducted step by step. After that, the SC U-net is trained by the preprocessed training data. In the testing procedure, after the same preprocessing steps except data augmentation, the testing data are fed into the trained SC U-net to get the final segmentation results.

A. DATA PREPROCESSING

To reduce the false positive caused by scalp fat as it has high intensity in both T1 and FLAIR MRI, to normalize the image size and voxel intensity, and to equip the model with desired invariance and robustness, we conduct the following data preprocessing step by step.

1) SKULL STRIPPING

To separate brain from non-brain tissues, such as skull and eyes, we use the Brain Extraction Tool (BET) in FMRIB Software Library (FSL) [2] to generate brain masks from T1 images and the masks are used to extract the brain [39].

2) UNIFORM SIZING

The sizes of the axial slices from different datasets vary. We set the aimed image size to be 200 × 200, and padding (or cropping) is automatically operated if the axial slice size is smaller (or larger) than the aimed image size.

3) GAUSSIAN NORMALIZATION

MRI scans are often collected under different acquisition parameters or protocols, which may result in a large variation in the intensity ranges. Gaussian normalization is applied to rescale the voxel intensities to guarantee each 3D image has a similar intensity range and overall brightness.

$$I(x, y, z) = \frac{I(x, y, z) - \mu}{\sigma}, \tag{1}$$

where $I(x, y, z)$ is the intensity value at location (x, y, z), and μ and σ are mean and standard deviation computed across all voxels of the image of interest.

4) DATA AUGMENTATION

Data augmentation is an effective way to equip a deep network with desired invariance and robustness properties when the training data are limited. For each axial slice, rotation, shearing and scaling are conducted. Finally, the size of our dataset is expanded by four times.

B. SKIP CONNECTION U-NET

1) ARCHITECTURE

We build a SC U-net based on the work of Li *et al.* [30] and Ronneberger *et al.* [36]. As shown in Fig. 2, our SC

[2] https://fsl.fmrib.ox.ac.uk/fsl/

U-net consists of a U-net part and a skip connection part. Similar to the U-net model proposed in [36], it consists of a down-convolution part which aims at extracting features for classifying each voxel into WMH and non-WMH and an up-convolution part which aims at locating WMH more precisely. The down-convolution part consists of two 3×3 convolution layers, each followed by a rectified linear unit (ReLU) and a 2×2 max pooling layer for down-sampling. In each step of the up-convolution part, after an up-sampling of the feature map, there are two 3×3 convolution layer and each followed by a ReLU. Concatenations are performed between the down-convolution part and the up-convolution part as shown in Fig. 2 using the gray lines. Random initialization is used to initialize the model weights.

Inspired by the work of [40], we propose a novel skip connection between the up-convolution part and the down-convolution part to increase convergence speed and improve segmentation performance. It consists of a cropping operation and a convolution operation (orange arrow with an addition sign shown in Fig 2). To be specific, the outputs of the 4th, 7th, 10th and 13th layers in the down-convolution part are respectively connected to the outputs of the 15th, 18th, 21th and 24th layers in the up-convolution part, and considered as the inputs of the 16th, 19th, 22th and 25th layers.

2) THE INFLUENCE OF SKIP CONNECTION

Skip connections are extra connections between nodes in different layers of a neural network [41], [42]. To get a better analytic understanding of skip connections, we consider linear network firstly. In a $L-$layer linear plain network, the input-output mapping is given by

$$\mathbf{x}_L = \mathbf{W}_{L-1}\mathbf{W}_{L-2}\cdots\mathbf{W}_1\mathbf{x}_1, \tag{2}$$

where $\mathbf{x}_1$ denotes the input, $\mathbf{x}_L$ denotes the output and $\mathbf{W}_i$ $(i = 1, \ldots, L-1)$ denotes the weights which are used to perform a linear projection that matches the number of feature maps at layer i to that at layer $i+1$. In a linear network with skip connection between adjacent layers, the input-output mapping can be computed as

$$\mathbf{x}_L^{SC} = (\mathbf{W}_{L-1}+\mathbf{I})(\mathbf{W}_{L-2}+\mathbf{I})\cdots(\mathbf{W}_1+\mathbf{I})\mathbf{x}_1, \tag{3}$$

where $\mathbf{x}_L^{SC}$ denotes the output of the L-th layer under the action of skip connection.

In a nonlinear case, the input-output mapping has the following form

$$\mathbf{x}_L = f\left(\mathbf{W}_{L-1}f\left(\cdots f\left(\mathbf{W}_2 f\left(\mathbf{W}_1\mathbf{x}_1\right)\right)\right)\right), \tag{4}$$

$$\mathbf{x}_L^{SC} = f((\mathbf{W}_{L-1}+\mathbf{I})f(\cdots f((\mathbf{W}_2+\mathbf{I}) f((\mathbf{W}_1+\mathbf{I})\mathbf{x}_1)))), \tag{5}$$

where f denotes the ReLU nonlinearity and $\mathbf{I}$ denotes the identity matrix having the same size as the input. To learn more features across different layers, we introduce skip connections in some of the symmetric layers rather than adjacent layers. The skip connection operation can be computed as

$$\mathbf{x}_l^{SC} = f(\mathbf{W}_{l-1}\mathbf{x}_{l-1}) + \mathbf{x}_{l-n_l}. \tag{6}$$

FIGURE 3. A simple example of a building block in the proposed framework which ignores the max pooling and up convolution operations. The blue rectangle denotes feature maps. The blue arrows and black arrows respectively denote convolution and skip connection.

Skip connections are also quite effective at dealing with the problem of gradient vanishing by creating short paths from the top layers to bottom layers [44]. The neural network is easy to train because skip connections improve the flow of information and gradient. As an example shown in Fig. 3, X denotes the input of the first layer, and after two convolution layers c_1 and c_2, the output is X_1. In the gradient back-propagation process, a layer receives gradients from the layers which it is connected to. To update the parameters represented by θ_2 of c_2, the derivate of loss function J with respect to θ_2 is calculated as

$$\nabla_{\theta_2}J(\theta_2) = \frac{\partial J}{\partial X_1}\frac{\partial X_1}{\partial \theta_2} + \frac{\partial J}{\partial X_2}\frac{\partial X_2}{\partial \theta_2}, \tag{7}$$

where X_1 and X_2 are the same. We can further fomulate (7) as

$$\nabla_{\theta_2}J(\theta_2) = \frac{\partial J}{\partial X_4}\frac{\partial X_4}{\partial X_3}\frac{\partial X_3}{\partial X_1}\frac{\partial X_1}{\partial \theta_2} + \frac{\partial J}{\partial X_4}\frac{\partial X_4}{\partial X_2}\frac{\partial X_2}{\partial \theta_2}. \tag{8}$$

Without skip connection, only $\frac{\partial J}{\partial X_4}\frac{\partial X_4}{\partial X_3}\frac{\partial X_3}{\partial X_1}\frac{\partial X_1}{\partial \theta_2}$ is computed and its magnitude may become very small after back-propagating through many layers. Thus the gradient is more likely to decrease to zero with the first term only. The second term $\frac{\partial J}{\partial X_4}\frac{\partial X_4}{\partial X_2}\frac{\partial X_2}{\partial \theta_2}$ carries a larger component since it goes through less layers. According to these analyses, our proposed skip connection is beneficial for updating the weights of bottoms layers and alleviating the gradient vanishing, furthermore, accelerating the convergence of the neural network.

3) DICE LOSS

In general, a loss function maps the values of one variable or multiple variables into a real number and calculates a penalty for every incorrect prediction. In the task of WMH segmentation, the numbers of positives and negatives are highly unbalanced. In the work of [45], they proposed a novel loss function based on the dice coefficient and obtained good performance on a highly unbalanced dataset. As such, we use it as our loss function to train the proposed SC U-net. Let $G = \{g_1, \cdots, g_N\}$ be the ground truth masks over N slices, and $P = \{p_1, \cdots, p_N\}$ be the corresponding predicted maps

TABLE 1. Characteristics of the *MICCAI WMH challenge* dataset used in this study.

Dataset	Scanner	Voxel Size mm^3	Size of FLAIR Scans	Number
Utrecht	3T Philips Achieva	$0.96 \times 0.95 \times 3.00$	$240 \times 240 \times 48$	20
Singapore	3T Siemens TrioTim	$1.00 \times 1.00 \times 3.00$	$252 \times 232 \times 48$	20
GE3T	3T GE Signa HDxt	$0.98 \times 0.98 \times 1.20$	$132 \times 256 \times 83$	20

over those N slices. The dice loss function can be expressed as

$$J = -\frac{\sum_{n=1}^{N} |p_n \circ g_n| + s}{\sum_{n=1}^{N} |p_n| + |g_n| + s}, \tag{9}$$

where $\circ$ denotes entrywise product of two matrices and $|\cdot|$ denotes the sum of matrix entries. The s term denotes a smoothing factor used to avoid a division by 0 when G and P are all zeros.

C. IMPLEMENTATION

The proposed method is implemented in Python language, using Keras with Tensorflow backend. All experiments are conducted on a Linux machine running Ubuntu 16.04 with 32 GB RAM. The model is trained using two GTX 1080 Ti GPUs. Adam optimizer is employed with an initial learning rate of 0.001 and the batch size is set to be 30×30. MSRA initialization proposed in [38] is used to initialize the model weights.

III. MATERIAL AND EVALUATION CRITERIA

A. MICCAI WMH SEGMENTATION CHALLENGE DATASET

In this work, all dataset came from a MICCAI WMH challenge [3]. Sixty cases from three centers were released as a publicly-available benchmarking dataset for researchers to use. Characteristics of this data are summarized in Table 1.

B. EVALUATION

1) DICE SIMILARITY COEFFICIENT (DSC)

DSC is a statistic used for comparing the similarity of two sets, which is defined as

$$DSC(G, S) = \frac{2\,|G \cap S|}{|G| + |S|}, \tag{10}$$

where G denotes the gold standard segmentation of WMH, S denotes the corresponding automatic segmentation and $|G \cap S|$ denotes the overlap of G and S.

2) HAUSDORFF DISTANCE USING 95^{TH} PERCENTILE (H95)

Hausdorff distance measures how far two subsets in a metric space are from each other. It is defined as

$$H(G, S) = max\left\{\sup_{x \in G} \inf_{y \in S} d\,(x, y), \sup_{y \in S} \inf_{x \in G} d\,(x, y)\right\}, \tag{11}$$

where $d\,(x, y)$, sup and inf respectively denote the distance between x and y, supremum and infimum. To avoid potential issues induced by noisy segmentations, it is modified to be a robust version by using 95^{th} percentile, namely H95, instead of the maximum distance.

[3] https://wmh.isi.uu.nl/

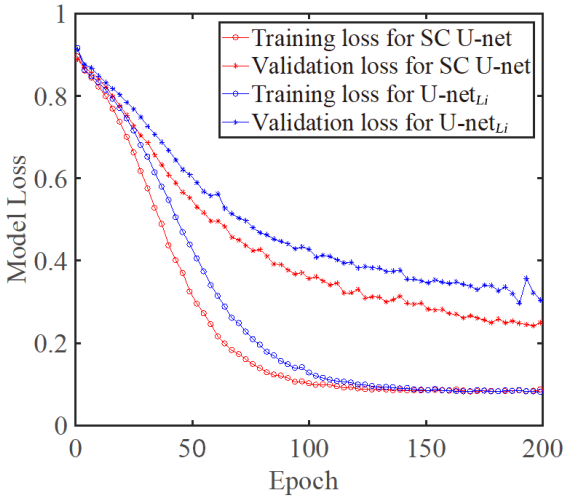

FIGURE 4. Training and validation loss curves for U-net$_{Li}$ and SC U-net.

3) AVERAGE VOLUME DIFFERENCE (AVD)

AVD quantifies the difference of volumes between two subsets. It is defined as

$$AVD\,(G, S) = \frac{|V_G - V_S|}{V_G}, \tag{12}$$

where V_G and V_S respectively denote the volumes of WMH in G and S.

4) SENSITIVITY AND F1-SCORE FOR INDIVIDUAL LESIONS

Let N_G be the number of individual lesions in G, N_P and N_F be the number of correctly detected lesions and the number of wrongly detected lesions. Each individual lesion is defines as a 3D connected component. Then the recall is defined as

$$Recall\,(G, S) = \frac{N_P}{N_G}, \tag{13}$$

and the F1-score is defined as

$$F1\,(G, S) = \frac{N_P}{N_P + N_F}. \tag{14}$$

IV. EXPERIMENTAL RESULTS

For the validation experiment, we split all datasets into a training set and a validation set by randomly picking 80% and 20% cases. For cross-scanner experiments, we used 40 subjects

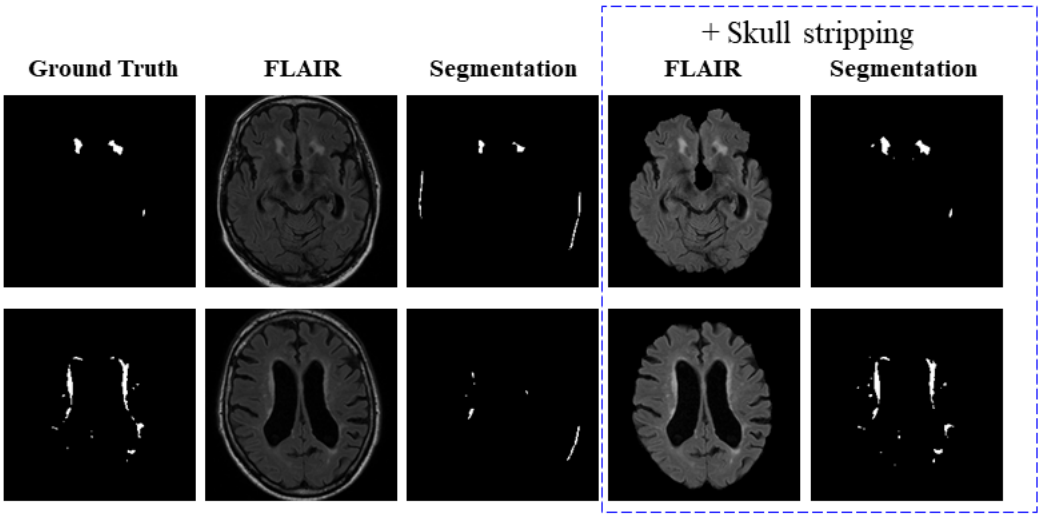

FIGURE 5. Comparisons between segmentation results with and without skull stripping, revealed on two cases (top and bottom).

from two scanners as the training data and 20 subjects from the other scanner as the testing data. Please note, we used 2D images at the axial view as our samples rather than the entire 3D images. U-net is the most widely used convolutional network in biomedical image segmentation. Li et al. used a U-net-like model in the same WMH segmentation challenge and obtained superior performance over other deep learning based methods [30]. Thus we compared our proposed method with the U-net-like model proposed in [30] and for convenience, we named it as U-net$_{Li}$. To conduct a fair comparison, the ensemble method adopted in [30] was not considered. In addition, we also compared our proposed method with another deep learning based segmentation approach proposed in [35].

A. VALIDATION EXPERIMENTAL RESULTS

The validation experiment was used to determine the hyper-parameters in a network and compare the convergence rates of U-net$_{Li}$ and SC U-net. The corresponding loss curves over 200 epochs in the validation experiments are demonstrated in Fig. 4.

The red and blue lines respectively denote the results of SC U-net and U-net$_{Li}$. The circle and star markers respectively denote the results on the training data and the validation data. We notice that the training loss does not significantly decrease any more after about 100 epochs. As such, we use 100 epochs in our subsequent cross-scanner experiment. It can be observed that the loss of SC U-net decreases faster than that of U-net$_{Li}$. And the validation loss of SC U-net is lower than that of U-net$_{Li}$. A lower loss means a lower difference between the outputs of the network and the ground truth.

Fig. 5 depicts representative segmentation results with or without skull-stripping in the data preprocessing. It is clearly shown that the segmentation results are largely affected by the skull fat. The reason is that the skull fat and WMH have similar intensity profiles. Without removing the skull in the FLAIR image, there will be some false positive in the segmentation results. And there will also be some false negative because the neural network will focus more on the skull fat. After applying skull-stripping in the data preprocessing, the segmentation performance is greatly improved.

B. CROSS-SCANNER EXPERIMENTAL RESULTS

The cross-scanner experiments are used to qualitatively and quantitatively analyze the segmentation performance of U-net$_{Li}$ and SC U-net. The first two rows of Fig. 6 show representative segmentation results on *Utrecht* when training the model using data from the other two datasets *Singapore* and *GE3T*.

In the segmentation results of Fig. 6, the green, red and blue areas respectively represent true positive, false negative and false positive. Evidently, the results of SC U-net suffer less false positive (blue area) and false negative (red area) than those of U-net$_{Li}$.

C. QUANTITATIVE RESULTS

Table 2 shows the segmentation performance of U-net$_{Li}$ [30] and SC U-net in terms of five evaluation criteria described in section III-B. DSC is the main evaluation criterion to quantitatively compare the segmentation performance of different methods. Clearly, the proposed SC U-net has a higher mean and a lower standard deviation of DSC than U-net$_{Li}$, suggesting that SC U-net is more accurate and more stable than U-net$_{Li}$. The average DSC, Recall and F1 achieved by the proposed SC U-net are respectively 78.36%, 81.49% and 70.86%, which are higher than those of U-net$_{Li}$ (74.99%, 76.06% and 69.25%). The average H95 and AVD achieved by SC U-net are respectively 7.36 mm and 28.23%, which are lower than U-net$_{Li}$ (10.51 mm and 36.18 %).

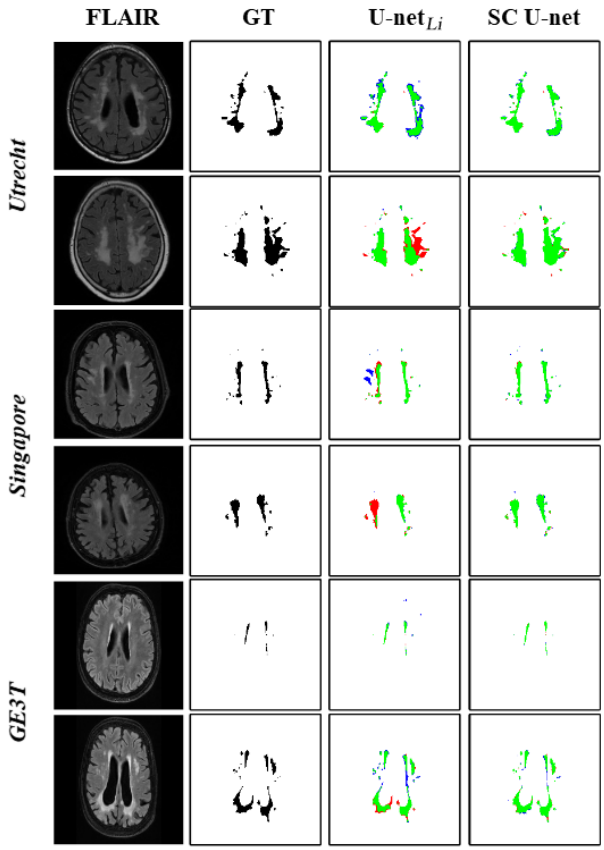

FIGURE 6. Representative segmentation results obtained from the cross-scanner experiment. From left to right, FLAIR MR images, the associated ground truth, segmentation results using U-net$_{Li}$ and SC U-net.

TABLE 2. Quantitative comparisons between U-net$_{Li}$ and SC U-net, shown as mean ± standard deviation. ↑ means a higher value is better. ↓ means a lower value is better.

Method	U-net$_{Li}$	SC U-net
DSC[%] ↑	74.99 ± 12.40	**78.36 ± 10.30**
H95[mm] ↓	10.51 ± 9.91	**7.36 ± 7.86**
AVD [%] ↓	36.18 ± 47.87	**28.23 ± 42.44**
Recall [%] ↑	76.06 ± 9.71	**81.49 ± 9.31**
F1[%] ↑	69.25 ± **11.16**	**70.86** ± 11.21

These quantitative results reveal the effectiveness of SC U-net in respect of localization accuracy and overall lesion detection.

Compared to the method proposed in [35], the mean DSC of our proposed method is higher than that in [35] (78.36% versus 74.80%), which also demonstrates the superior performance of our proposed segmentation method.

V. CONCLUSION

In this work, we proposed and validated a novel WMH segmentation method by combining image preprocessing and deep learning techniques. Skip connection was used to make U-net capture more features and converge to a better optimum. Both qualitative and quantitative analyses revealed the superior performance of the proposed pipeline compared to U-net$_{Li}$. The proposed method is fully automatic and it has a great potential to become an end-to-end Computer Aided Diagnosis application. In the future, we will explore more efficient and more accurate skull-stripping methods given that FSL based skull-stripping is not very accurate nor computationally efficient. A plausible direction is to develop CNN based methodology to remove non-brain tissues in the preprocessing stage and then use the proposed SC U-net to automatically segment WMH.

ACKNOWLEDGMENT

This study was supported by the National Key R&D Program of China (2017YFC0112404) and the National Natural Science Foundation of China (NSFC 81501546). We thank the MICCAI-2017 WMH Challenge organizer Dr. Hugo J. Kuijf for making the dataset for this research available.

REFERENCES

[1] J. M. Wardlaw, M. C. V. Hernández, and S. Muñoz-Maniega, "What are white matter hyperintensities made of? Relevance to vascular cognitive impairment," *J. Amer. Heart Assoc.*, vol. 4, no. 6, 2015, Art. no. e001140.
[2] J. C. de Groot, F.-E. de Leeuw, M. Oudkerk, A. Hofman, J. Jolles, and M. M. Breteler, "Cerebral white matter lesions and depressive symptoms in elderly adults," *Arch. Gen. Psychiatry*, vol. 57, no. 11, pp. 1071–1076, 2000.
[3] S. Debette and H. Markus, "The clinical importance of white matter hyperintensities on brain magnetic resonance imaging: Systematic review and meta-analysis," *BMJ*, vol. 341, Jul. 2010, Art. no. c3666.
[4] R. Schmidt, P. Scheltens, T. Erkinjuntti, L. Pantoni, H. S. Markus, A. Wallin, F. Barkhof, and F. Fazekas, "White matter lesion progression: A surrogate endpoint for trials in cerebral small-vessel disease," *Neurology*, vol. 63, no. 1, pp. 139–144, 2004.
[5] P. Schmidt, V. Pongratz, P. Küster, D. Meier, J. Wuerfel, C. Lukas, B. Bellenberg, F. Zipp, S. Groppa, P. G. Sämann, F. Weber, C. Gaser, T. Franke, M. Bussas, J. Kirschke, C. Zimmer, B. Hemmer, and M. Mühlau, "Automated segmentation of changes in FLAIR-hyperintense white matter lesions in multiple sclerosis on serial magnetic resonance imaging," *NeuroImage, Clin.*, vol. 23, May 2019, Art. no. 101849.
[6] J. M. Wardlaw *et al.*, "Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration," *Lancet Neurol.*, vol. 12, no. 8, pp. 822–838, 2013.
[7] A. A. Gouw, A. Seewann, W. M. Van Der Flier, F. Barkhof, A. M. Rozemuller, P. Scheltens, and J. J. Geurts, "Heterogeneity of small vessel disease: A systematic review of MRI and histopathology correlations," *J. Neurol., Neurosurgery Psychiatry*, vol. 82, no. 2, pp. 126–135, 2010.
[8] J. Conklin, F. L. Silver, D. J. Mikulis, and D. M. Mandell, "Are acute infarcts the cause of leukoaraiosis? Brain mapping for 16 consecutive weeks," *Ann. Neurol.*, vol. 76, no. 6, pp. 899–904, Dec. 2014.
[9] B. D. Trapp and K.-A. Nave, "Multiple sclerosis: An immune or neurodegenerative disorder?" *Annu. Rev. Neurosci.*, vol. 31, pp. 247–269, Jul. 2008.
[10] M. Wu, C. Rosano, M. Butters, E. Whyte, M. Nable, R. Crooks, C. C. Meltzer, C. F. Reynolds, III., and H. J. Aizenstein, "A fully automated method for quantifying and localizing white matter hyperintensities on MR images," *Psychiatry Res., Neuroimag.*, vol. 148, nos. 2–3, pp. 133–142, 2006.
[11] L. J. Launer, "Epidemiology of white matter lesions," *Topics Magn. Reson. Imag.*, vol. 15, no. 6, pp. 365–367, 2004.
[12] K. Bendfeldt, J. O. Blumhagen, H. Egger, P. Loetscher, N. Denier, P. Kuster, S. Traud, N. Mueller-Lenke, Y. Naegelin, A. Gass, J. Hirsch, L. Kappos, T. E. Nichols, E.-W. Radue, and S. J. Borgwardt, "Spatiotemporal distribution pattern of white matter lesion volumes and their association with regional grey matter volume reductions in relapsing-remitting multiple sclerosis," *Hum. Brain Mapping*, vol. 31, no. 10, pp. 1542–1555, 2010.

[13] D. T. Chard, C. M. Griffin, G. J. M. Parker, R. Kapoor, A. J. Thompson, and D. H. Miller, "Brain atrophy in clinically early relapsing–remitting multiple sclerosis," *Brain*, vol. 125, no. 2, pp. 327–337, 2002.
[14] K.-O. Löuvbld, A. E. Baird, G. Schlaug, A. Benfield, B. Siewert, B. Voetsch, A. Connor, C. Burzynski, R. R. Edelman, and S. Warach, "Ischemic lesion volumes in acute stroke by diffusion-weighted magnetic resonance imaging correlate with clinical outcome," *Ann. Neurol., Off. J. Amer. Neurological Assoc. Child Neurol. Soc.*, vol. 42, no. 2, pp. 164–170, 1997.
[15] M. E. Caligiuri, P. Perrotta, A. Augimeri, F. Rocca, A. Quattrone, and A. Cherubini, "Automatic detection of white matter hyperintensities in healthy aging and pathology using magnetic resonance imaging: A review," *Neuroinformatics*, vol. 13, no. 3, pp. 261–276, Feb. 2015.
[16] C. R. Jack, Jr., P. C. O'Brien, D. W. Rettman, M. M. Shiung, Y. Xu, R. Muthupillai, A. Manduca, R. Avula, and B. J. Erickson, "FLAIR histogram segmentation for measurement of leukoaraiosis volume," *J. Magn. Reson. Imag., Off. J. Int. Soc. Magn. Reson. Med.*, vol. 14, no. 6, pp. 668–676, 2001.
[17] E. Gibson, F. Gao, S. E. Black, and N. J. Lobaugh, "Automatic segmentation of white matter hyperintensities in the elderly using FLAIR images at 3T," *J. Magn. Reson. Imag.*, vol. 31, no. 6, pp. 1311–1322, 2010.
[18] B. I. Yoo, J. J. Lee, J. W. Han, S. Y. W. Oh, E. Y. Lee, J. R. MacFall, M. E. Payne, T. H. Kim, J. H. Kim, and K. W. Kim, "Application of variable threshold intensity to segmentation for white matter hyperintensities in fluid attenuated inversion recovery magnetic resonance images," *Neuroradiology*, vol. 56, no. 4, pp. 265–281, Apr. 2014.
[19] A. S. Keçeli, A. B. Can, and A. Kaya, "A GPU-based approach for automatic segmentation of white matter lesions," *IETE J. Res.*, vol. 63, no. 4, pp. 461–472, 2017.
[20] K. V. Leemput, F. Maes, D. Vandermeulen, A. Colchester, and P. Suetens, "Automated segmentation of multiple sclerosis lesions by model outlier detection," *IEEE Trans. Med. Imag.*, vol. 20, no. 8, pp. 677–688, Aug. 2001.
[21] P. Schmidt, C. Gaser, M. Arsic, D. Buck, and A. Förschler, A. Berthele, M. Hoshi, R. Ilg, V. J. Schmid, C. Zimmer, B. Hemmer, and M. Mühlau, "An automated tool for detection of FLAIR-hyperintense white-matter lesions in multiple sclerosis," *NeuroImage*, vol. 59, no. 4, pp. 3774–3783, 2012.
[22] C. Schwarz, E. Fletcher, C. DeCarli, and O. Carmichael, "Fully-automated white matter hyperintensity detection with anatomical prior knowledge and without FLAIR," in *Proc. Int. Conf. Inf. Process. Med. Imag.* Cham, Switzerland: Springer, 2009, pp. 239–251.
[23] D. Wu, M. Albert, A. Soldan, C. Pettigrew, K. Oishi, Y. Tomogane, C. Ye, T. Ma, M. I. Miller, and S. Mori, "Multi-atlas based detection and localization (MADL) for location-dependent quantification of white matter hyperintensities," *NeuroImage, Clin.*, vol. 22, Mar. 2019, Art. no. 101772.
[24] M. Dadar, T. A. Pascoal, S. Manitsirikul, K. Misquitta, V. S. Fonov, M. C. Tartaglia, J. Breitner, P. Rosa-Neto, O. T. Carmichael, C. Decarli, and D. L. Collins, "Validation of a regression technique for segmentation of white matter hyperintensities in alzheimer's disease," *IEEE Trans. Med. Imag.*, vol. 36, no. 8, pp. 1758–1768, Aug. 2017.
[25] J. Wu, Y. Zhang, and X. Tang, "A multi-atlas guided 3D fully convolutional network for MRI-based subcortical segmentation," in *Proc. IEEE 16th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2019, pp. 705–708.
[26] Y. Zhang, J. Wu, W. Chen, Y. Chen, and X. Tang, "Prostate segmentation using Z-Net," in *Proc. IEEE 16th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2019, pp. 11–14.
[27] S. Andermatt, S. Pezold, and P. Cattin, "Multi-dimensional gated recurrent units for the segmentation of biomedical 3D-data," in *Deep Learning and Data Labeling for Medical Applications*. Cham, Switzerland: Springer, 2016, pp. 142–151.
[28] M. Ghafoorian, N. Karssemeijer, T. Heskes, I. W. M. van Uden, C. I. Sanchez, G. Litjens, F.-E. de Leeuw, B. van Ginneken, E. Marchiori, and B. Platel, "Location sensitive deep convolutional neural networks for segmentation of white matter hyperintensities," *Sci. Rep.*, vol. 7, no. 1, p. 5110, 2017.
[29] S. Valverde, M. Cabezas, E. Roura, S. González-Villà, D. Pareto, J. C. Vilanova, L. Ramió-Torrentà, À. Rovira, A. Oliver, and X. Lladó, "Improving automated multiple sclerosis lesion segmentation with a cascaded 3D convolutional neural network approach," *NeuroImage*, vol. 155, pp. 159–168, Jul. 2017.
[30] H. Li, G. Jiang, J. Zhang, R. Wang, Z. Wang, W.-S. Zheng, and B. Menze, "Fully convolutional network ensembles for white matter hyperintensities segmentation in MR images," *NeuroImage*, vol. 183, pp. 650–665, Dec. 2018.
[31] Y. Xu, T. Géraud, É. Puybareau, I. Bloch, and J. Chazalon, "White matter hyperintensities segmentation in a few seconds using fully convolutional network and transfer learning," in *Proc. Int. MICCAI Brainlesion Workshop*. Cham, Switzerland: Springer, 2017, pp. 501–514.
[32] J. Hong, B.-Y. Park, M. J. Lee, C.-S. Chung, J. Cha, and H. Park, "Two-step deep neural network for segmentation of deep white matter hyperintensities in migraineurs," *Comput. Methods Programs Biomed.*, vol. 183, Jan. 2020, Art. no. 105065.
[33] M. F. Rachmadi, M. del C. Valdés-Hernández, M. L. F. Agan, C. Di Perri, T. Komura, and Alzheimer's Disease Neuroimaging Initiative, "Segmentation of white matter hyperintensities using convolutional neural networks with global spatial information in routine clinical brain MRI with none or mild vascular pathology," *Computerized Med. Imag. Graph.*, vol. 66, pp. 28–43, Jun. 2018.
[34] R. Guerrero, C. Qin, O. Oktay, C. Bowles, L. Chen, R. Joules, R. Wolz, M. Valdés-Hernández, D. Dickie, J. Wardlaw, and D. Rueckert, "White matter hyperintensity and stroke lesion segmentation and differentiation using convolutional neural networks," *NeuroImage, Clin.*, vol. 17, pp. 918–934, Dec. 2018.
[35] H. Kervadec, J. Bouchtiba, C. Desrosiers, É. Granger, J. Dolz, and I. B. Ayed, "Boundary loss for highly unbalanced segmentation," in *Proc. Int. Conf. Med. Imag. Deep Learn.*, 2019, pp. 285–296.
[36] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Springer, 2015, pp. 234–241.
[37] H. J. Kuijf *et al.*, "Standardized assessment of automatic segmentation of white matter hyperintensities; Results of the WMH segmentation challenge," *IEEE Trans. Med. Imag.*, to be published.
[38] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2015, pp. 1026–1034.
[39] S. M. Smith, "Fast robust automated brain extraction," *Hum. Brain Mapping*, vol. 17, no. 3, pp. 143–155, 2002.
[40] M. Drozdzal, E. Vorontsov, G. Chartrand, S. Kadoury, and C. Pal, "The importance of skip connections in biomedical image segmentation," in *Deep Learning and Data Labeling for Medical Applications*. Cham, Switzerland: Springer, 2016, pp. 179–187.
[41] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 770–778.
[42] A. E. Orhan and X. Pitkow, "Skip connections eliminate singularities," Jan. 2017, *arXiv:1701.09175*. [Online]. Available: https://arxiv.org/abs/1701.09175
[43] X. Mao, C. Shen, and Y.-B. Yang, "Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections," in *Proc. Adv. Neural Inf. Process. Syst.*, 2016, pp. 2802–2810.
[44] T. Tong, G. Li, X. Liu, and Q. Gao, "Image super-resolution using dense skip connections," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 4809–4817.
[45] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-Net: Fully convolutional neural networks for volumetric medical image segmentation," in *Proc. 4th Int. Conf. 3D Vis. (3DV)*, 2016, pp. 565–571.

JIONG WU received the B.S. degree in information and computing science and the M.S. degree in computer technology from Lanzhou University, Lanzhou, China, in 2013 and 2016. He is currently pursuing the Ph.D. degree with the School of Electronics and Information Technology, Sun Yat-sen University, Guangzhou, China.

His research interests include medical image segmentation and registration, MRI analysis, machine learning, and deep learning.

YUE ZHANG was born in Shandong, China, in 1994. He received the B.S. degree in communication engineering from the Southern University of Science and Technology (SUSTech), in 2016. He is currently pursuing the Ph.D. degree in electrical and electronic engineering with The University of Hong Kong, Hong Kong, and the Ph.D. degree in electrical and electronic engineering with the SUSTech. He has published nine conference articles, including MICCAI, ISBI, and EMBC in his research related areas. His research interests include medical image segmentation and reconstruction.

KAI WANG received the B.S. degree in electrical engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2000, the M.S. degree in microelectronics and solid state electronics from the Dalian University of Technology, Dalian, China, in 2003, and the Ph.D. degree in electrical and computer engineering from the University of Waterloo, Waterloo, Canada, in 2008.

He was thereafter appointed as a NSERC Postdoctoral Fellow of the Thunder Bay Health Science Centre, Thunder Bay, Canada, where he was conducting research in biomedical X-ray imaging until he joined Apple Inc., Cupertino, CA, in August 2011. At Apple Inc., he was involved in touch sensor development for various Apple products such as iPad and Apple Watch. He is currently a Full Professor with the School of Electronics and Information Technology, Sun Yat-Sen University, in China, and also an Adjunct Professor with the Department of Electrical and Computer Engineering, Carnegie Mellon University, USA. His current research interests include emerging applications of thin-film transistors including flat-panel X-ray imaging, fingerprint biometrics, tactile sensors, energy harvesting, and sensor interfaces.

XIAOYING TANG received the B.S. degree in control system engineering and foreign language education from the Huazhong University of Science and Technology, Wuhan, China, in 2009, the M.S. degrees in electrical and computer engineering and in applied mathematics and statistics in 2011 and 2014, respectively, and the Ph.D. degree in electrical and computer engineering in 2014, from Johns Hopkins University, USA.

From 2014 to 2015, she was a Visiting Professor in electrical and computer engineering with Carnegie Mellon University. From 2015 to 2018, she was a Tenure-Track Assistant Professor with the SYSU-CMU Joint Institute of Engineering, Sun Yat-sen University. She is currently an Assistant Professor and an Associate Researcher with the Department of Electrical and Electronic Engineering, Southern University of Science and Technology. Her research interests include medical image segmentation and registration, diffusion tensor image analysis, statistical shape analysis, manifold learning and clustering, spatiotemporal analysis, multi-modality MRI analysis, pattern recognition, machine learning, and big data in medicine.