Hybrid Attention Densely Connected Ensemble Framework for Lesion Segmentation From Magnetic Resonance Images

BEIBEI HOU[1,3], (Member, IEEE), **XIN XU**[2], **GUIXIA KANG**[1,3],
YUAN TANG[1], (Graduate Student Member, IEEE), AND **CHUAN HU**[1]

[1]Key Laboratory of Universal Wireless Communications, Beijing University of Posts and Telecommunications, Beijing 100876, China
[2]Department of Neurosurgery, Chinese PLA General Hospital, Beijing 100853, China
[3]Wuxi BUPT Sensory Technology and Industry Institute Company Ltd., Wuxi 214001, China

Corresponding author: Guixia Kang (gxkang@bupt.edu.cn)

ABSTRACT White matter hyperintensities (WMHs) are associated with various neurological and aging diseases, and morphological analysis plays a crucial role in the assessment of disease progression. In this article, a novel hybrid attention densely connected ensemble framework is deployed for WMH segmentation from multi-modality magnetic resonance imaging (MRI). On the one hand, hybrid attention densely convolutional network (HA-DCN) is designed with a novel hybrid attention module embedded. The hybrid attention module can further improve the precision of lesion localization by extracting the complementary information of high-level features and low-level features from the spatial domain and channel domain. On the other hand, the focal Tversky loss function and generalized dice loss function derived from dice similarity coefficient are ensembled into the proposed framework, which achieves a trade-off between specificity and sensitivity. As a result, the volume of automation lesion segmentation is more agreeable to the manual lesion segmentation by experts. The proposed framework was evaluated online and offline in the MICCAI 2017 WMH segmentation challenge. A quantitative experiment has further demonstrated the effect of multi-modality and the effectiveness of the proposed hybrid attention densely connected ensemble framework. Furthermore, the challenge dataset consists of three scanners, reflecting the flexibility and robustness of the model. It also exhibits its potential for real-world clinical practice.

INDEX TERMS White matter hyperintensities, segmentation, magnetic resonance imagings, hybrid attention densely convolutional networks, ensembles.

I. INTRODUCTION

White matter hyperintensities (WMHs) are one of the main consequences of small blood vessel disease, which plays a vital role in the assessment of dementia, stroke, and aging [1]. Clinically, conventional magnetic resonance imaging (MRI) scans behave as the most prevalent technique for structural analysis in the brain [2]. It not only provides images with high spatial resolution and high contrast for brain tissue but also exhibits the morphology of lesions well [2]. As shown in Fig.1, WMHs are visible as hyperintense regions on the fluid attenuated inversion recovery (FLAIR) modality and

The associate editor coordinating the review of this manuscript and approving it for publication was Yudong Zhang.

hypointense regions on the T1-weighted (T1) modality within the white matter.

Nowadays, manual annotation of WMH regions remains the gold standard for lesion annotation [3], which is an observer-dependent and time-consuming procedure. Especially, spatial multipleness is a typical property of WMH lesions, which poses a significant challenge to quantize WMH. With the development of three-dimensional imaging technique and the functional and physiological imaging increasing, manual segmentation methods will face enormous challenges.

The development of machine learning technology provides an explorable space for the automatic diagnosis of disease through medical imaging [4]. Especially, the machine
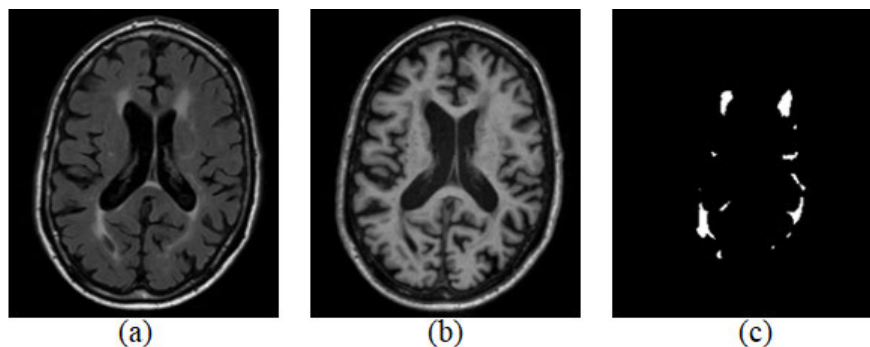
FIGURE 1. A sample of MRI slice from FLAIR modality (a), T1 modality (b) and its corresponding manual annotation of WMH by a neuroradiologist (c).

learning method autonomously mines association rules between images and lesions based on indicated characteristics of imaging, which is a representative algorithmic of data-driven technique. Anbeek *et al.* [5] built a probability feature space for WMH lesion segmentation, extracting characteristics from signal intensities and spatial information based on the k-nearest neighbour (KNN) method. Geremia *et al.* [6], [7] presented a discriminative random decision forest framework for multiple sclerosis lesion segmentation. It is a voxel-level probabilistic classification with context-rich, local spatial and symmetry characteristics. Tomas *et al.* [8] achieved lesion segmentation by estimating the spatially over-all intra-subject intensity distribution and spatial local intensity distribution obtained from a healthy reference population simultaneously. However, the superiority of these methods depends on the extraction of imaging indicated features, resulting in limited versatility.

A. RELATED WORK

Convolutional neural networks (CNNs) express flexibility by autonomously learning the features from the available dataset. In medical image segmentation, CNNs achieve impressive achievements as well as demonstrate the generalizability by training available medical images [12], [13], especially in three-dimensional space. It can be divided into the following three types generally.

Firstly, the voxel-based lesion segmentation framework simplifies the segmentation problem into classification. It predicts whether the central voxel belongs to lesions, and then determines the region of lesions. Valverde *et al.* [14] designed a creative framework for multiple sclerosis lesions segmentation based on a cascade of two 6-layers CNN. Its first network was used to screen possible candidate lesion voxels, behaving sensitively to lesions. The second network was designed to reduce the number of misclassified voxels, decreasing the false positive rate [15]. Secondly, the patch-based lesion segmentation framework restores the input resolution through upsampling operation, thereby reduce data redundancy. Ghafoorian *et al.* [16] integrated a variety of deep neural network architectures that considers

explicit location features or multi-scale patches for WMH lesion segmentation. Kamnitsas *et al.* [17] proposed a dual pathway, incorporating both local and more abundant contextual information and employing three-dimensional conditional random field (CRF) to decrease the false positive rate in the post-processing stage. Thirdly, the volume-based segmentation framework represented by the encoder-decoder network is proposed, which realizes an end-to-end output from images to lesions. Brosch *et al.* [18] conducted a preliminary study of deep 3D convolutional encoder networks based on the pre-training of Boltzmann machine theory. Li *et al.* [19] used deep fully convolutional networks and integrated them to detect WMH using multi-modalities MRIs.

However, above convolutional neural networks are difficult to further optimize owing to its poor interpretability [20]. It is noteworthy that people pay more and more attention to the transparency, accountability, and fairness of deep learning models. Consequently, attention mechanism (AM) has been vigorously promoted in the artificial intelligence community drawing on the selective attention mechanism of human vision, which has been employed in text recognition [21], recommendation system [22], semantic segmentation [23] and so on. AM allows the model to dynamically focus on certain parts of inputs by incorporating the concept of correlation [24], [25]. Generally, AM mainly calculates the association from the spatial and channel domains to focus on characteristics. In the spatial domain, Max Jaderberg *et al.* [26] proposed a spatial transformer network (STN) that can transform the spatial information in the original image into feature maps as well as retain the critical characteristics. In the channel domain, the characteristics of each channel are the components of the image on different convolution kernels. It presents the correlation between channels and the segmentation result by adding weights to the signals on each channel. Squeeze-and-Excitation (SE) Network [27] explicitly model the inter-dependence between feature channels and learn the importance of each feature channel automatically by adopting a "feature re-calibration" strategy. In summary, neural networks have the ability to focus on its input features and focus more on the detailed information of the lesion area when AM is embedded in neural networks.
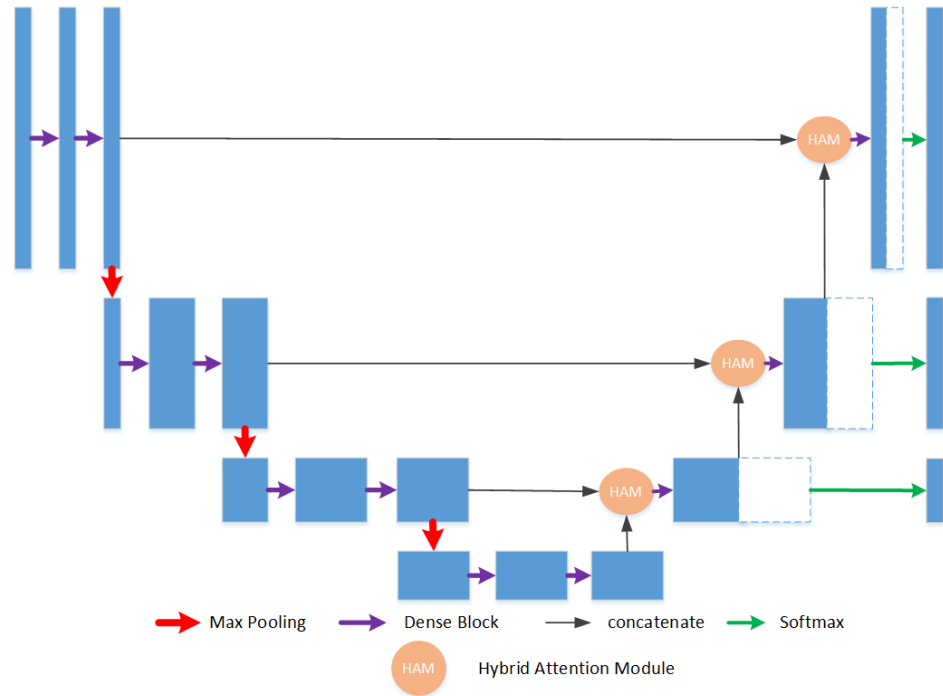
FIGURE 2. An overview of the proposed hybrid attention densely connected neural network.

B. CONTRIBUTIONS

In this article, a novel hybrid attention densely convolutional network (HA-DCN) is deployed for WMH lesion segmentation task. In this article, we designed an end-to-end encoder-decoder framework. The framework designs encoder including densely connections [28] and decoder including a hybrid attention module.

In summary, the main contributions of this work are:

(1) A hybrid attention module is designed to exploit the consistency of WMH lesion, combined with the spatial domain attention and channel domain attention to produce complementary weights on features.

(2) A multi-output architecture is constructed to increase the transmission of network features and gradient.

(3) Ensemble models trained with two different loss function and random parameter initializations data are employed, which trade-off in sensitivity and specificity.

II. METHODS

Recently, a large number of studies have proved that U-net [29] is an effective architecture when tackling the segmentation task [30]. In this article, we further deploy an HA-DCN with 3D U-net as the backbone, as shown in Fig.2.

The deployed architecture is a typical encoder-decoder framework. In the stage of encoding, it consists of densely connected block [31] and max pooling operations alternately. In the decoding stage, a hybrid attention module is designed to focus on the sensitive areas in which the lesions are located. Besides, HA-DCN employs the dense block and deconvolutional [32] to realize the restoration of feature maps.

Besides, a multiple output operation is also designed to alleviate the gradient dispersion problem caused by the deepening of the encoding-decoding network. As shown in Fig.2, the probabilistic of various sizes is output via the softmax layer. It can make full use of features as well as transfer features and gradients within the architecture effectively.

A. HYBRID ATTENTION MODULE

In view of the correlation between convolutional feature maps and the segmentation task, AM can effectively capture the distinguishable characteristics of lesions while suppressing the indistinguishable characteristics. It enhances the expression ability of feature maps [34]. Specifically, spatial domain attention module (SAM) encodes global semantic information into local receptive fields, strengthening the representation capability of feature maps and aggregating semantic information of whole brain images [26], [35]. Channel domain attention module (CAM) extracts global context information as a guide to the underlying features by global pooling and filters the detail information of WMH lesions [36], [37]. Based on the above research, this article deploys a hybrid attention module (HAM), whose composition is shown in Fig.3.

As shown in Fig.3 (a), SAM pools the feature map along channels and performs a $1 \times 1 \times 1$ convolutional operation. Then it realizes the nonlinear transformation in the spatial voxel by employing the sigmoid activation function, which computes the correlation W_S reflecting where is the interest on voxel-level. In Fig.3 (b), CAM performs a $1 \times 1 \times 1$ convolutional operation following to global pooling on the
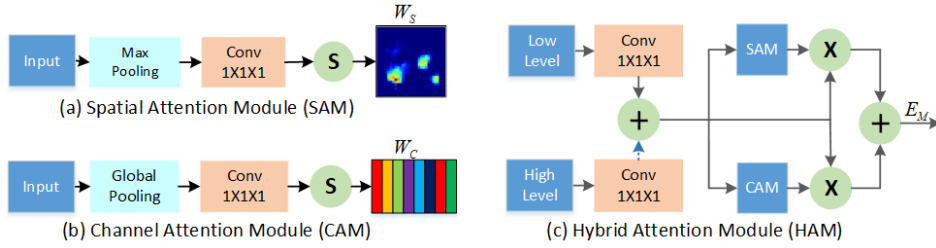
FIGURE 3. Schematic diagram of main components of HA-DCN.

feature map. Then, the sigmoid function is used to calculate the corresponding matrix W_C that measures the impact of each channel on labels.

Furthermore, we design hybrid attention as Fig.3 (c) by combining the SAM and the CAM. Specifically, the number of channels and the spatial dimension of high-level features and low-level features are firstly matched to facilitate feature summation. Secondly, HAM employs SAM and CAM on the summed feature, making the obtained weights and the summed feature multiply element by element. Finally, HAM outputs the final hybrid attention E_M by adding the two attention results:

$$E_M = W_C \odot (F_{HD} + F_{LC}) + W_S \odot (F_{HD} + F_{LC}), \quad (1)$$

where F_{HD} and F_{LC} represent the results of matching high-level features and low-level feature respectively, W_C and W_S denote the matrix calculated by CAM and SAM, respectively, $\odot$ is multiply element by element operation.

Benefiting from the hybrid attention module, HA-DCN can automatically learn the comprehensive attention weight, which captures the relevance between low-level features (in the encoder stage) and high-level features (in the decoder stage). These associations are then used to bulid a spatial matrix E_M that passes through encoding path to the decoding path. Furthermore, high-level features with sufficient category information can guide low-level information to select precise resolution details.

B. LOSS FUNCTION

Dice similarity loss function is the most common loss function that performs comprehensive evaluation for medical image segmentation tasks [38], [39]. It performs fairly well when target voxels behave focus or the distribution of target/background voxels is uniform. However, its performance is not ideal for WMH segmentation task. On the one hand, the number of non-lesion voxels is much higher than that of lesion voxels. On the other hand, the distribution of WMH lesions is relatively scattered. The dice similarity coefficient (DSC) is decreased significantly once the small target has a partial pixel prediction error. Accordingly, this article adopts asymmetric similarity loss functions to alleviate the above problem.

Let R be the manual segmentation label for lesion over $m-th$ image elements r_m and its contrary values $\overline{r_m}$, and P be

the predicted probabilistic map with voxel values p_m and its contrary value $\overline{p_m}$. Then, we define loss functions as follows:

1) FOCAL Tversky LOSS
Focal Tversky loss (FTL) achieves a flexible balance between false positive (FP) and false negative (FN) by introducing the balance factor α manually. Besides, a stable coefficient γ is designed to solve the limitation that the small target is insensitive to the loss function [40]. It takes the form:

$$L_F = (1 - \frac{\sum\limits_{m} r_m p_m + \varepsilon}{\sum\limits_{m} r_m p_m + \alpha \sum\limits_{m} \overline{r_m} p_m + (1-\alpha) \sum\limits_{m} r_m \overline{p_m} + \varepsilon})^{\gamma}, \quad (2)$$

where α is an adjustable hyper-parameter that improves recall when the large class imbalance occurs, ε represents a random small value to prevent the denominator of 0, γ denotes a stable coefficient which varies in the range [0, 1]. In practice, the FTL function pays more attention to less accurate predictions that have been misclassified. The smaller γ, the larger L_F when the voxel is misclassified.

2) GENERALIZED DICE LOSS
Generalized dice loss (GDL) realizes the asymmetry loss function by automatically calculating the balance factor w_l [41]. The correction is achieved by the reciprocal of the category volume, thereby decreasing the correlation between target size and DSC. It is defined as follows:

$$w_l = \frac{1}{(\sum\limits_{m=1}^{M} r_{lm})^2}, \quad (3)$$

$$L_G = 1 - 2\frac{\sum\limits_{l=1}^{2} w_l \sum\limits_{m} r_{lm} p_{lm}}{\sum\limits_{l=1}^{2} w_l \sum\limits_{m} r_{lm} + p_{lm}}, \quad (4)$$

where r_{lm} is the manual annotation of the category l at the $m - th$ voxel, while p_{lm} is the corresponding predicted probability value, w_l is used to provide invariance to different label set properties, M denotes the total number of voxels.

C. AUTO-CONTEXT INFORMATION FUSION

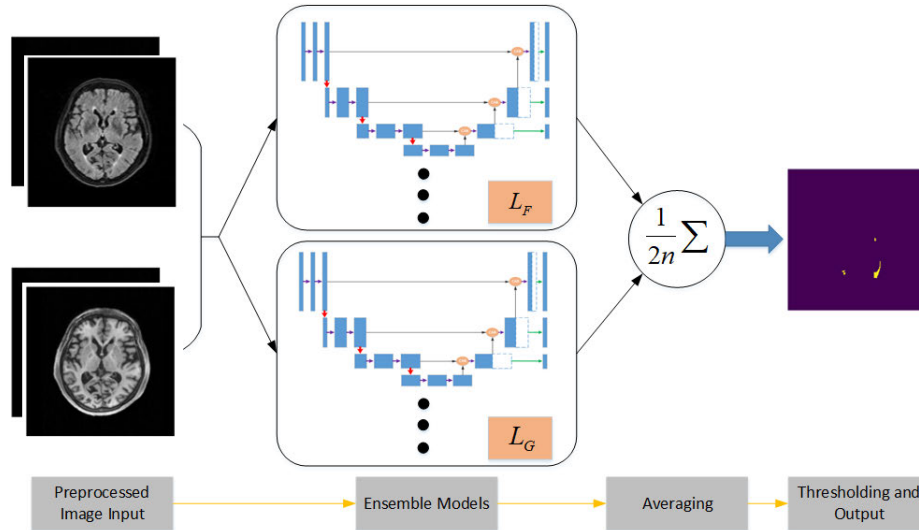Ensemble techniques achieve more accurate and robust results than any single learning algorithm by combing

FIGURE 4. Overall framework for the testing stage.

TABLE 1. Detailed information of WMH segmentation challenge. The training set includes 60 patient from 3 scanners while the test set from 3 scanners.

dataset	Scanner	3D T1		FLAIR		#Training	#Test
		Voxel Size(mm^3)	TR/TE(ms)	Voxel Size(mm^3)	TE/TR/TI(ms)		
Utrecht	3T Philips Achieva	1.00×1.00×3.00	7.9/4.5	0.96×0.95×3.00	11000/125/2800	20	30
Singapore	3T Siemens TrioTim	1.00×1.00×3.00	2300/1.9	1.00×1.00×3.00	9000/82/2500	20	30
GE3T	3T GE Signa HDxt	0.94×0.94×1.00	7.8/3.0	0.98×0.98×1.20	8000/126/2340	20	30

multiple models. In particular, most championship results of various medical image analysis are derived from model integration. Tao Song *et al.* [42] won the ischemic stroke lesion segmentation challenge (ISLES) 2018 by aggerating extractor, generator and segmentor. Isensee *et al.* [43] achieved top performance in automatic cardiac segmentation challenge (ACDC) and multi-atlas labeling beyond the cranial vault challenge (abdomen) by ensembling several U-Nets.

In this article, an integrated approach is used to solve the automated WMH segmentation problem, which combines several models with the same architecture in a carefully designed pipeline. The intention of adopting ensemble models includes two aspects: (i) The attributes of the training data that the model can be learnt vary with different initialization value; (ii) Different loss functions have a different emphasis on segmentation results, achieving a good specificity-sensitivity balance. Specifically, L_F performs excellently in sensitivity while L_G behaves universal that neutralizes the low specificity of L_F, thereby improving the accuracy of model segmentation.

As shown in Fig.4, n HA-DCNs with the same architecture are trained with random parameter initialization concerning loss function L_F and L_G. A probability segmentation result will be output by each of the n HA-DCN models with L_F and L_G. Then, a final map is generated by averaging $2n$ maps. Finally, a threshold is set based on experience to transform the probabilistic output into a binary segmentation result. In this article, the threshold is set to 0.5.

III. EXPERIMENTS

A. DATASET

The MICCAI WMH2017 segmentation challenge [45] captured MR images with various degrees of vascular disease related to ageing. It consists of a training dataset and an unseen test dataset. The organizing committee provided a 3D T1-weighted sequence and a 2D multi-slice FLAIR sequence for each patient. Properties of the detailed information are summarized in Table 1. It is worth mentioning that the images of these patients were acquired from three hospitals. Two experts annotated the pathological reslt concerning the STandards for ReportIng Vascular changes on nEuroimaging (STRIVE) criteria [2].

Furthermore, a pre-processed version was also provided, including bias-corrected using statistical parametric mapping 12 (SPM12) [44], image registration using the Elastix toolbox [46], and face removed manually. The detailed information can be seen in https://wmh.isi.uu.nl/data/.

B. INPUT FEATURE

A further preprocessing based on the initial image processing is of great help for invariance and robustness in the automated lesion segmentation system. The input feature is constructed using the following steps:

(i) BET toolbox [47] is employed to obtain a binary brain mask that is multiplied the preprocessed images for skull stripping.

(ii) Each training image is subtracted firstly by its mean and divided from its variance secondly, aiming to reduce the variations of input data and speed-up training process.

(iii) For each provided sequence, all the axial slices are cropped or padded to a uniform size p automatically. We compute 3D axial volumes by stacking images as $V = [n \times p \times p \times a \times c]$, where n represents the number of volumes, a represents the number of slices in the axial direction and c denotes the number of the available input modality.

In order to overcome the limitations of high variance and weak robustness caused by the limited amount of training data, this article carries out three orientation flips for each training volume, which achieves a data volume ratio of before and after data enhancement as 1:3.

C. EVALUATION METRICS AND RANK METHOD

Challenge organizers introduce five different indicators to compare the pros and cons of the submitted methods by participants. We define the manual annotation as G and a predicted segmentation result as P. Then the five evaluation indicators are calculated on the per-voxel basis:

(i) Dice Similarity Coefficient (DSC) measures the overlap percentage between G and P. It is defined as:

$$DSC = \frac{2 \times |G \cap P|}{|G| + |P|} \qquad (5)$$

(ii) Hausdorff Distance (HD) measures the boundary distance between the ground-truth G and automatic segmentation P, which calculates the maximun degree of mismatch between two sets. On the basis, the $K - th$ ranked distance is used to suppress the outlier, obtaining a robustified version by using the 95th percentile instead of the maximun distance.

$$h_{95}(G, P) = {}^{95}K_{g \in \mathcal{G}}^{th} \min_{p \in \mathcal{P}} ||p - g|| \qquad (6)$$

where ${}^{95}K_{g \in G}^{th}$ is the $K - th$ ranked minimum euclidean distance with $K/N_g = 95\%$, $\mathcal{G}$ and $\mathcal{P}$ represent the boundaries set of manual annotation G and predicted result P, respectively. The lower the HD, the higher the proximity between ground truth and automatic segmentation. The HD is defined as:

$$HD = \max\{h_{95}(G, P), h_{95}(P, G)\} \qquad (7)$$

(iii) Average Volume Difference (AVD):

$$AVD = \frac{||G| - |P||}{|G|} \qquad (8)$$

In view of the spatial multipleness of WMH, individual lesions are defined as 3D connected components within an image. We define the 3D connected components within the manual annotation as G_I and the 3D connected components within the predicted segmentation result as P_I. On this basis, N_G is defined as the number of individual lesions in G_I, N_P and N_F are the number of correctly detected individual lesions and wrongly detected individual lesions in P_I after comparing to G_I.

(iv) Sensitivity for individual lesions (Recall):

$$Recall = \frac{N_P}{N_G} \qquad (9)$$

(v) F1-score for individual lesions (F1):

$$Precision = \frac{N_P}{N_P + N_F} \qquad (10)$$

$$F_1 = 2.0 \times \frac{Precision \times Recall}{Precision + Recall} \qquad (11)$$

D. IMPLEMENTATION DETAILS

To optimize the model parameter, a five-fold cross-validation method is employed to optimize the hyper-parameters and validate the proposed architecture. The stochastic gradient descent (SGD) with a learning rate of 0.01, the decay of 1E-6, momentum of 0.9 in our experiments.

We implement the proposed architecture on Python based on the deep learning library of Keras, utilizing Tensorflow backend and two parallel GPU of NVIDIA GTX2080. It takes around 6.5 hours to train a single model for 150 epochs on a training set containing 400 volumes of size $196 \times 196 \times 48$ each.

IV. EXPERIMENTAL RESULTS

The segmentation performances on the unseen test dataset and public training dataset are demonstrated in this section. Additionally, we also discuss the effectiveness of each module in Section V.

A. ONLINE RESULTS ON UNSEEN TEST DATASET

In the WMH 2017 segmentation challenge, each participant packages the required code as a container through docker and uploads the container to the organizing committee, to compare the effectiveness of the proposed framework fairly. Subsequently, the organizing committee evaluates the unseen test dataset by running the uploaded container. The performance of the proposed method on three test sets is shown in TABLE 2.

TABLE 2. Mean index value when the proposed HA-DCN was tested with the unseen test dataset.

Methods	DSC	HD↓	AVD↓	Recall	F1
proposed	0.80	6.2	17.39	0.69	0.74
pgs [48]	0.82	5.07	18.78	0.82	0.79
sysu [49]	0.81	5.16	25.38	0.87	0.77
coroflo [50]	0.79	5.39	23.52	0.77	0.76
neuro [51]	0.81	4.16	16.90	0.84	0.78
cian [52]	0.79	5.20	20.23	0.83	0.70

Regarding the evaluation of the unseen test dataset, our proposed method is also compared with state-of-the-art methods in the competition. It includes that U-Net with a highlighted foreground [48], ensembled U-Net [49], multi-dimensional convolutional gated recurrent units [50], [52], and an adjusted DeepMedic architecture proposed [51]. These methods are derived from deep learning and achieve
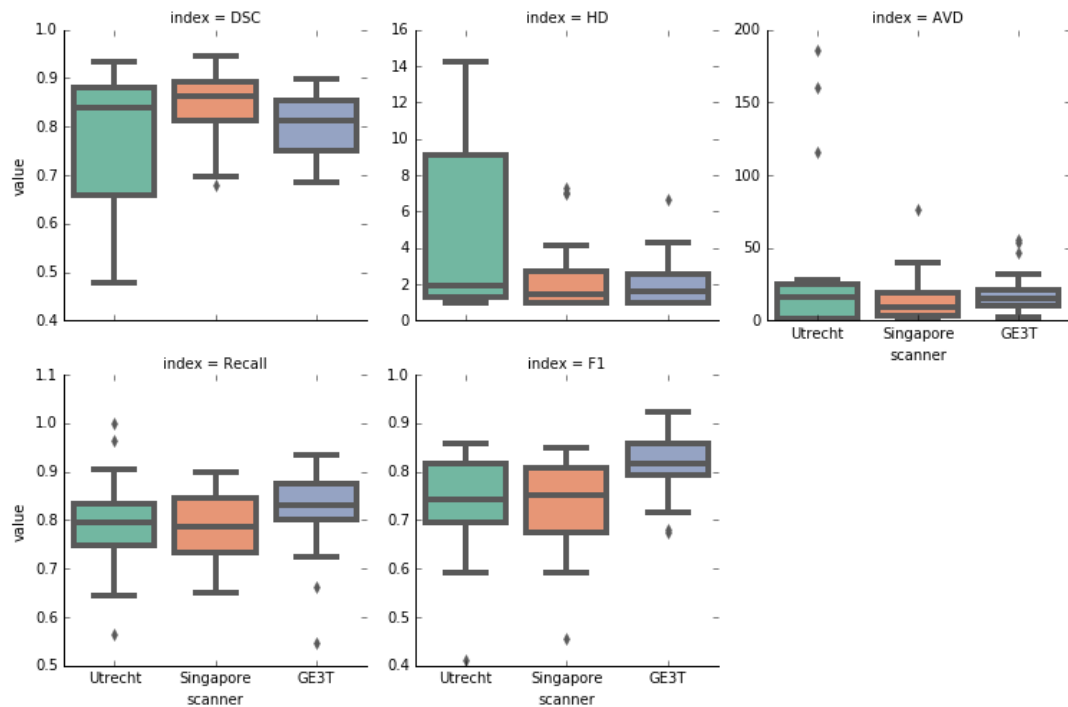
FIGURE 5. The distribution of segmentation performance on public training dataset from three scanners in five-fold cross validation.

the automatic WMH segmentation by incorporating comfortable pre-processing and post-processing. TABLE 2 summarizes and compares the performance of top methods for WMH 2017 segmentation challenge. It can be seen that the proposed method achieves good performance in DSC and AVD, but it has a low rate in terms of recall. Unfortunately, the proposed method lags behind pgs and neuro in most metric values. However, these two methods are based on two-dimensional slice processing, which will be slightly more complicated in pre-processing and post-processing. Overall, the integrated HA-DCN method is competitive and performs well on the test set, especially in AVD, which should benefit from the integration of L_F and L_G.

In addition, The physical meaning of HD is the distance of the edge matching. HD metric values in Table 2 are the highest for the proposed method compared with other state-of-the-art methods. It indicates that the proposed method performs poorly in delineating the boundary of lesions. However, the limitation of the preprocessing method mentioned lies in the boundary. Hence, this phenomenon may be caused by the pre-processing method that cropped volume mentioned in Section III.B.

Nevertheless, the 3D nature of brain makes the analysis of volumes the mainstream trend of current brain image analysis, promoting the development of 3D model. This article provides new ideas for the 3D pre-processing of images of different sizes.

B. OFFLINE RESULTS ON PUBLIC TRAINING DATASET

The public training dataset is also comprehensively evaluated in a random five-fold cross-validation, verfying the generalization performance of our proposed method. Firstly, the 60 training subjects from three scanners are randomly divided into five groups (each group has 4 participants in each scanner). Then, images from one of them are used for validation, and other images from 4 remaining groups are used for training. This procedure is repeated five times until all of the 60 subjects from three scanners are used for validation.

Fig.5 depicts the segmentation performance from three scanners in the five-fold cross-validation. As one can see that their performance varies although the same number of the public training dataset is provided. Specifically, the performance of GE3T scanner is best, followed by Singapore and Utrecht worst. It manifested in two aspects: The distance between the upper and lower edges of boxplot in the Utrecht scanner is long, meaning a large variance; the discrete values in the figure are mostly concentrated in the Utrecht scanner.

Furthermore, false positive volume function (FPVF) indicates over-segmentation rate and false negative volume function (FNVF) that indicate under-segmentation rate is employed to indicate the over- and under-segmentation. Specifically, $FPVF = N_{FP}/N_G$, $FNVF = N_{FN}/N_G$, where N_{FP} denotes the number of false positive voxels, N_G represents the number of manual annotation WMH, N_{FN} is the

number of false negative voxels. Based on this, FPVF and FNVF of 60 samples on public training dataset are plotted in Fig.6. It can be seen that samples are more concentrated below $y = x$, indicating that the proposed method tends to over-segment samples.
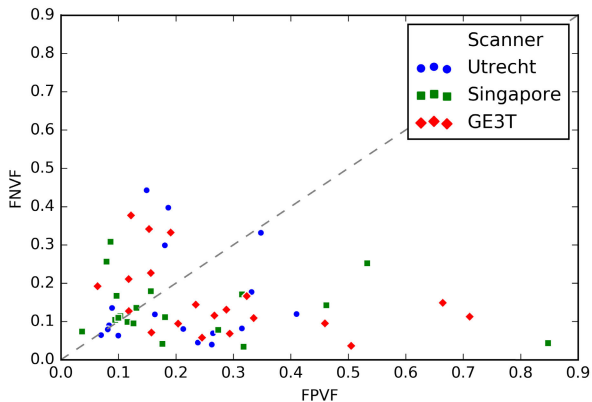

FIGURE 6. Distribution of FPVF and FNVF of 60 samples on public training dataset.

V. DISCUSSION

Quantification of WMH lesion has become an indispensable step for diagnosing and assessing the progression of diseases, whereby measures the morphology and location of lesions. In this article, a novel HA-DCN ensemble framework is deployed for WMH segmentation problem via a combination of the core idea of the densely connected block and attention mechanism. Additionally, a specificity-sensitivity trade-off is achieved by ensembling focal Tversky loss function and generalized dice loss function. As a result, the deployed pipeline is tested in the public challenge, and an intensive analysis is also performed through five-fold cross-validation.

A. INFLUENCE OF IMAGING MULTI-MODALITIES

MRI generates multi-modality images reflecting different information of the brain by adjusting imaging parameters. T1 modality reflects brain anatomy by providing high contrast between tissues while FLAIR sequence behaves highly sensitive to pathological changes, especially for white matter lesion. Clinically, a combination of T1 and FLAIR is usually used to extract their complementary information for lesion annotation. Accordingly, this article adopts the combination of T1 and FLAIR modalities to improve the automatic segmentation accuracy of WMH lesions. To further examine the influence of imaging multi-modalities, HA-DCN is trained using FLAIR and FLAIR+T1 as input, respectively. TABLE 3 demonstrates the mean index value when the proposed HA-DCN is tested with different input modality on the public training dataset. It can be seen that fusing T1 and FLAIR modalities can improve the index value of the model performance to a certain extent. In addition, Holm's correction is used to calculate the statistical difference between them.

TABLE 3. Mean index value when the proposed HA-DCN is tested with different input modality with five-fold cross-validation evaluation. P-value denotes the statistical difference using Holm's correction.

Input modality	DSC	HD↓	AVD↓	Recall	F1
FLAIR	0.79	4.47	20.28	0.71	0.70
FLAIR+T1	0.80	3.23	23.46	0.80	0.75
P-value	0.51	0.079	0.68	6e-5	0.0023

The test results do not significantly improve the performance of the method on DSC, HD, and AVD while displaying a significant difference in Recall and F1.

In Fig.7, the segmentation results of random samples from three scanners are visualized. It can be observed that the false negative rate decrease when T1 and FLAIR modality are stacked as the input of the proposed HA-DCN. Indeed, the blue voxels in the green dashed box will be reduced significantly. This phenomenon suggests that T1 modality provides complementary information of FLAIR modality on focusing WMH lesions as well as shows high tissue contrast.

B. INFLUENCE OF L_F AND L_G ON VOLUME CORRELATION

Bland-Altman diagram can assess the consistency of two technologies, which is a commonly used statistical analysis method in clinical practice. Fig.8 plots the volume agreement between the manual annotation and automatic lesion segmentation for L_F, L_G and multi-target architecture. It can be perceived that L_F tends to overestimate lesion volume because of the mean solid line below 0, that is, lesions are usually estimated larger than they are. On the contrary, L_G predicts underestimated lesions. After the integration of L_F and L_G, the multi-target architecture balances the estimated volume and shows a smaller variance. In terms of volumetric correlation, L_F demonstrates the smallest mean value, while the ensemble model shows a relatively small variance.

Fig.9 depicts fitted linear regression models between the volumes of all ground truth lesions and corresponding predicted WMH lesions for each loss function. For each subgraph, 95% confidence interval areas also are provided for comparison. With a quantitative evaluation, L_F overestimates lesion volume because of the slope smaller than 1 while L_G behaves the opposite. The multi-target architecture (red circle) predicts the slope that is closest to $y = x$. Consequently, multi-target architecture provides a stronger global agreement between manual labeling and predicted results, that is, the volume is the most agreeable to the manual segmentation by experts.

C. EFFECTIVENESS OF THE HYBRID ATTENTION BLOCK

The composition of the attention mechanism module is crucial for the coarse screening of lesions [56]. Recently, a rich body of studies focused on designing models with attention mechanisms to improve the performance of the segmentation task [53], [54]. The main innovation research can be divided into the channel domain and the spatial domain. In quest of the influence of different attention modules on
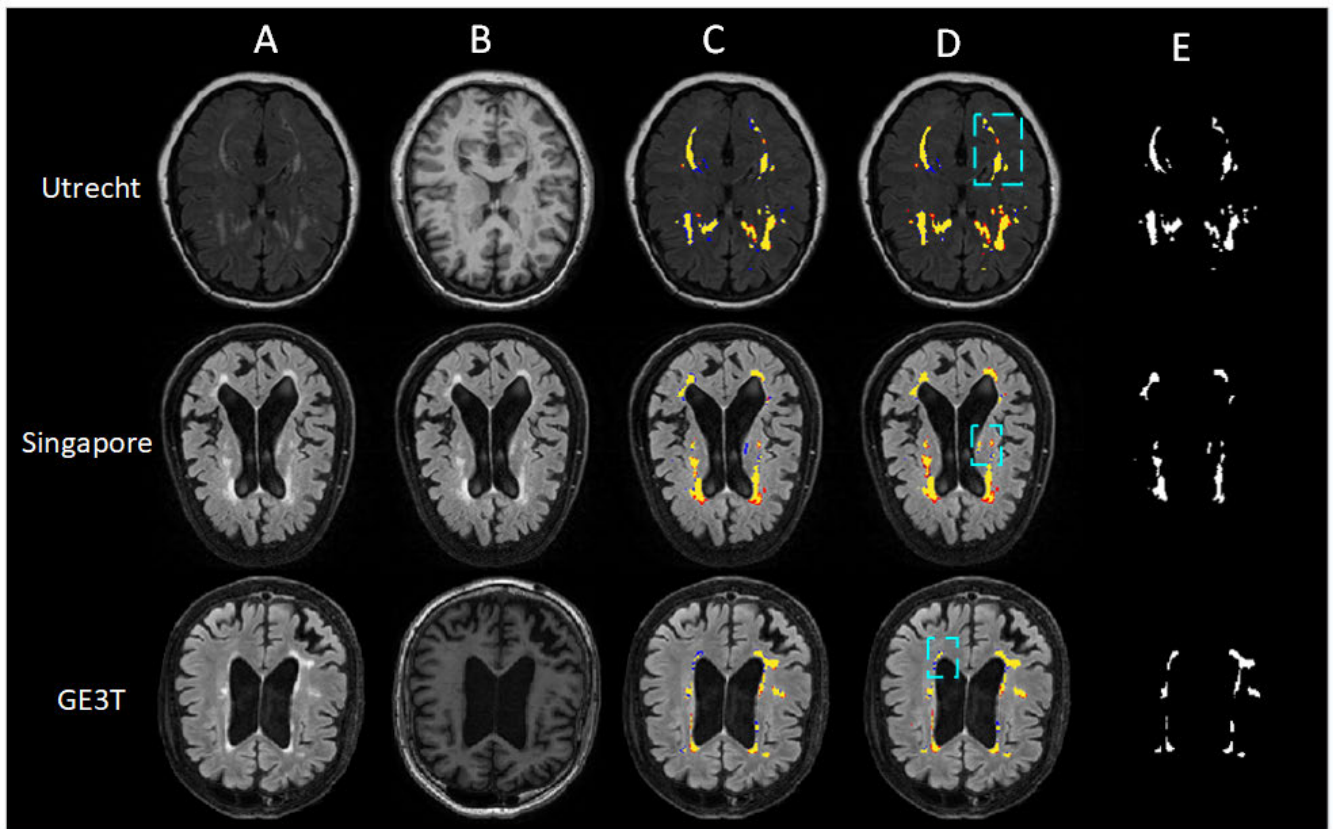
FIGURE 7. Segmentation results from three scanners. A: FLAIR MR images, B: T1 MR images, C: segmentation results only using FLAIR modality, D: segmentation results using FLAIR and T1 modalities, E: the associated ground truth. In the predicted result, the yellow area denotes true positives, the red voxels are the false positives, and the blue voxels are the false negatives.
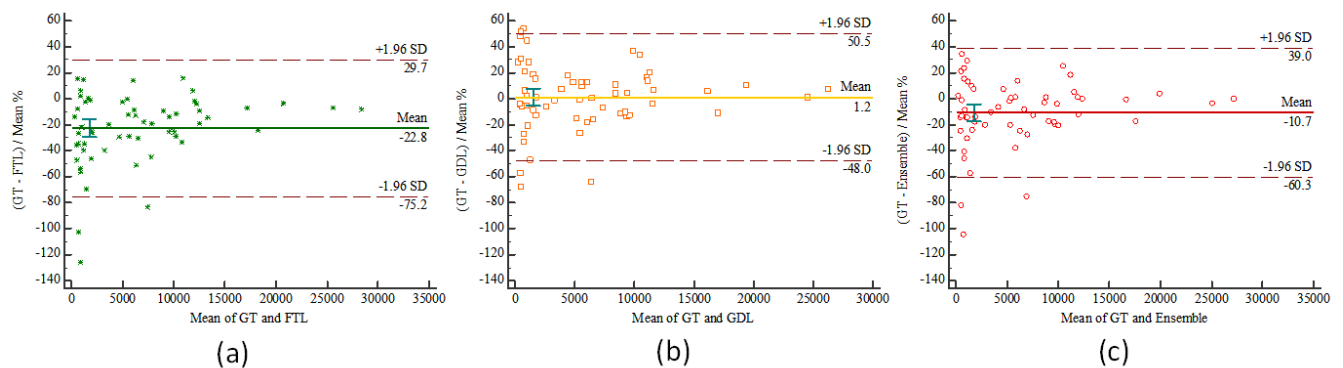

FIGURE 8. Bland-Altman plots for volumetric comparisons between lesion resulting from manual labeling versus the results of each objective function.

WMH segmentation, three additional neural networks are designed with spatial attention (shown in Fig. 3a), with channel attention (shown in Fig.3b) and without attention embedded. Accordingly, SAM represents the model of employing spatial attention module instead of HAM in HA-DCN, CAM represents the model of using channel attention module instead of HAM in HA-DCN, and WAM represents the model without attention module. In this article, the above networks are trained three times and averaged their network output to eliminate the effects of random initialization of parameters.

Fig.10 reveals the performance on the DSC, AVD, HD, Recall, and F1 evaluation metrics obtained from the above-mentioned different models on the public training dataset. Among all the models, the performance of the neural network will be slightly improved when the attention mechanism is embedded. When HAM is designed in a neural network, its performance is superior in DSC, HD, AVD, F1 compared with other models. Clinically, multiple comparison correction with the conservation Holm's method is commonly used for easily discovering statistical differences between groups. Fig.10 depicts the statistically significant
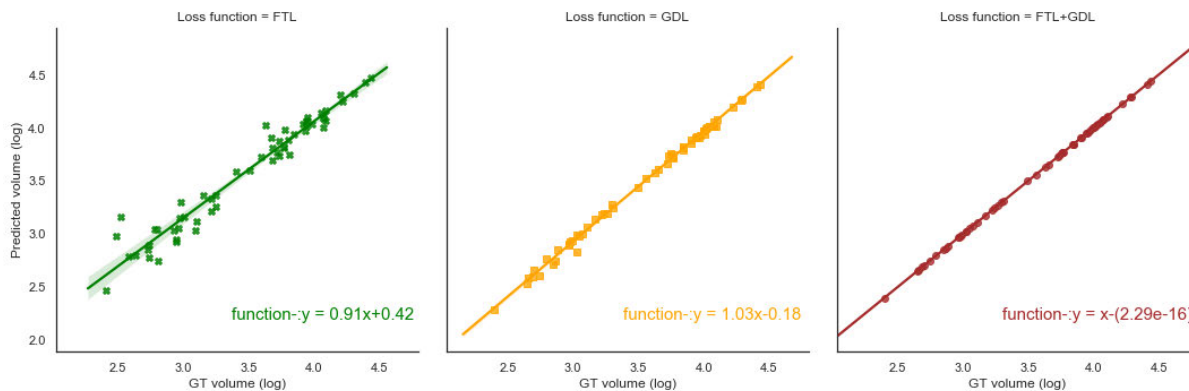
FIGURE 9. Correlation between manual lesion annotations volume and predicted WMH lesion volume for each objective function. The solid lines represents the linear regression function along with confidence intervals at 95%.
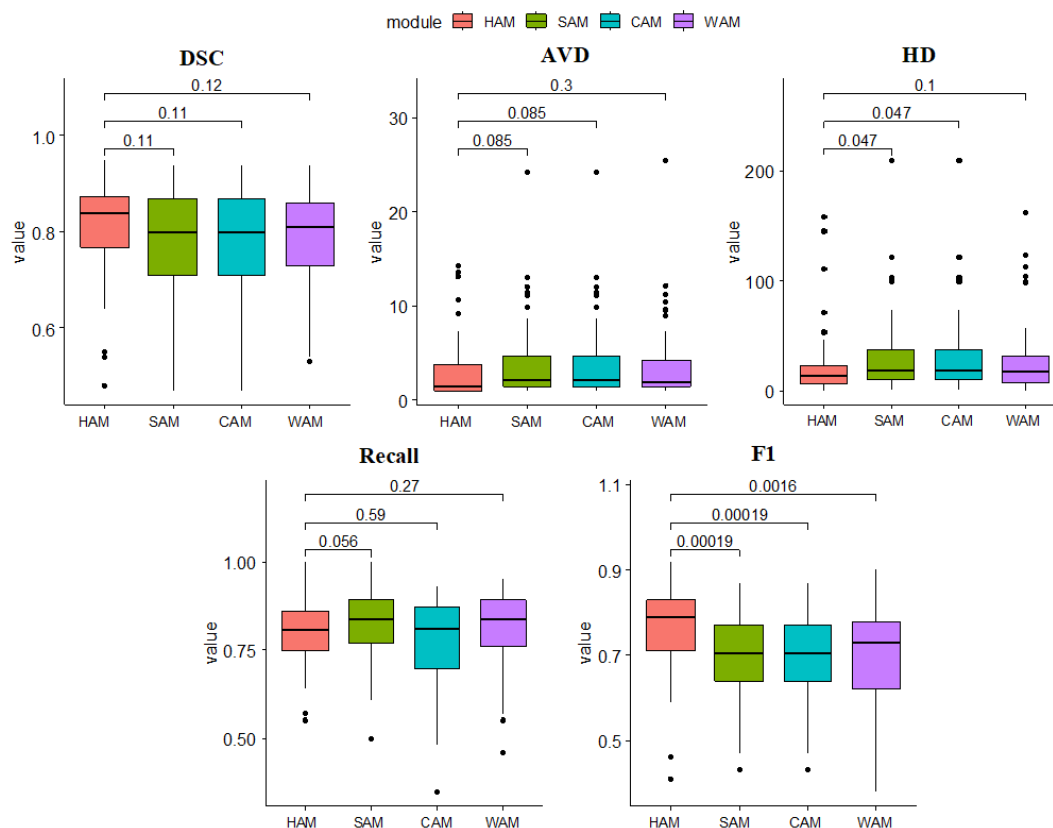

FIGURE 10. Boxplots showing the performance of different model embeded various attention mechanism on the public training data. Statistically significant differences *p* between HAM and other attention mechanism are assessed using resampling statistics with multiple comparison correction.

differences between model performances for most metrics when compared to HAM.

D. EFFECT ON THE SIZE OF ENSEMBLES

Ensemble learning boosts segmentation performance and enhances the stability of the framework by aggregating different models. The characteristics of random initialization of parameters and random batch gradient descent in model training make each trained model unique. However, it may

make the segmentation performance closer to saturation as the property of the training dataset is fully explored gradually. Therefore, how many models in the ensemble learning are needed is still an open issue. To this end, we further analyze the effect of the size of the integrated model on the WMH segmentation task.

Fig.11 illustrates the curves of segmentation performance on five metrics with respect to different ensemble size. Representing the mean and standard deviation of the data in the same graph can compare the differences in data
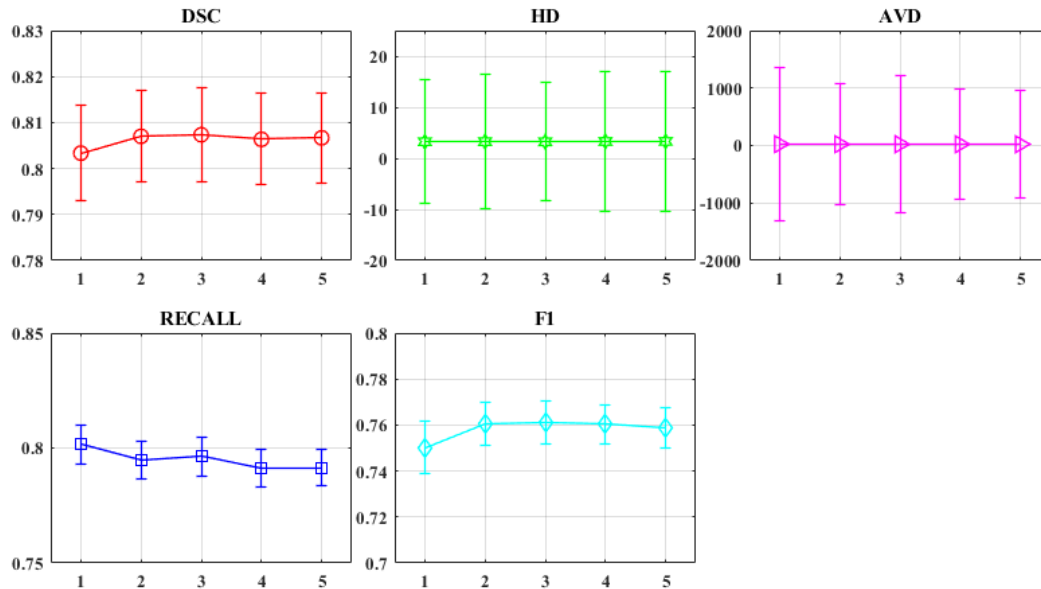
FIGURE 11. The errorbar graph as a function of the size of the ensemble.

distribution well. It could be seen that: (1) WMH segmentation performance that the ensemble size n is 2 outperforms the ensemble of only one model, especially in DSC and F1. At this time, the improvement of the integrated model is most obvious; (2) With the continual increase of the ensemble size n, the segmentation performance tends to saturate and feeble promotion in some of the indicators boost at the cost of small decreased in others. Overall, when the size of the ensemble is 3, the performance of the ensemble model tends to be stable and saturated. Therefore, we choose the ensemble size of $n = 3$ to obtain better segmentation performance.

VI. CONCLUSION

In this article, a novel automated lesion segmentation method with application to WMH patients is deployed. Specifically, a HAM that combines the advantages of the spatial domain and the channel domain is designed to improve the overall performance of segmentation task, increasing the interpretability of deep learning and facilitating further optimization. Comprehensive experimental analysis on the WMH segmentation challenge demonstrates the effectiveness of HAM and the robustness of HA-DCN. The obtained results are encouraging, yielding our CNN architecture closer to human expert inter-rater variability. However, we still make an effort to analysis diseases with no obvious structural changes on MRI in the clinical. In the future, we will continue to explore the application of HA-DCN to other lesion segmentation tasks, such as focal cortical dysplasia.

REFERENCES

[1] N. D. Prins and P. Scheltens, "White matter hyperintensities, cognitive impairment and dementia: An update," *Nature Rev. Neurol.*, vol. 11, no. 3, p. 157, 2015.

[2] J. M. Wardlaw, E. E. Smith, G. J. Biessels, C. Cordonnier, F. Fazekas, R. Frayne, R. I. Lindley, J. T. O'Brien, F. Barkhof, O. R. Benavente, and S. E. Black, "Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration," *Lancet Neurol.*, vol. 12, no. 8, pp. 822–838, Aug. 2013.

[3] J. Grimaud, M. Lai, J. Thorpe, P. Adeleine, L. Wang, G. J. Barker, D. L. Plummer, P. S. Tofts, W. I. McDonald, and D. H. Miller, "Quantification of MRI lesion load in multiple sclerosis: A comparison of three computer-assisted techniques," *Magn. Reson. Imag.*, vol. 14, no. 5, pp. 495–505, Jan. 1996.

[4] A. Danelakis, T. Theoharis, and D. A. Verganelakis, "Survey of automated multiple sclerosis lesion segmentation techniques on magnetic resonance imaging," *Computerized Med. Imag. Graph.*, vol. 70, pp. 83–100, Dec. 2018.

[5] P. Anbeek, K. L. Vincken, M. J. P. van Osch, R. H. C. Bisschops, and J. van der Grond, "Probabilistic segmentation of white matter lesions in MR imaging," *NeuroImage*, vol. 21, no. 3, pp. 1037–1044, Mar. 2004.

[6] E. Geremia, B. H. Menze, O. Clatz, E. Konukoglu, A. Criminisi, and N. Ayache, "Spatial decision forests for MS lesion segmentation in multichannel MR images," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Berlin, Germany, 2010, pp. 111–118.

[7] E. Geremia, O. Clatz, B. H. Menze, E. Konukoglu, A. Criminisi, and N. Ayache, "Spatial decision forests for MS lesion segmentation in multi-channel magnetic resonance images," *NeuroImage*, vol. 57, no. 2, pp. 378–390, Jul. 2011.

[8] X. Tomas-Fernandez and S. K. Warfield, "A model of population and subject (MOPS) intensities with application to multiple sclerosis lesion segmentation," *IEEE Trans. Med. Imag.*, vol. 34, no. 6, pp. 1349–1361, Jun. 2015.

[9] H. Li, "Deep learning for natural language processing: Advantages and challenges," *Nat. Sci. Rev.*, vol. 5, no. 1, pp. 24–26, Jan. 2018.

[10] C. Hazirbas, L. Ma, C. Domokos, and D. Cremers, "Fusenet: Incorporating depth into semantic segmentation via fusion-based cnn architecture," in *Proc. Asian Conf. Comput. Vis.*, Taipei, Taiwan, 2016, pp. 213–228.

[11] C. Vicas, I. Rusu, N. A. Hajjar, and M. Lupsor-Platon, "Deep convolutional neural nets for objective steatosis detection from liver samples," in *Proc. 13th IEEE Int. Conf. Intell. Comput. Commun. Process. (ICCP)*, Stanford, CA, USA, Sep. 2017, pp. 385–390.

[12] P. Moeskops, J. de Bresser, H. J. Kuijf, A. M. Mendrik, G. J. Biessels, J. P. W. Pluim, and I. Išgum, "Evaluation of a deep learning approach for the segmentation of brain tissues and white matter hyperintensities of presumed vascular origin in MRI," *NeuroImage, Clin.*, vol. 17, pp. 251–262, Jan. 2018.

[13] Z. Akkus, A. Galimzianova, A. Hoogi, D. L. Rubin, and B. J. Erickson, "Deep learning for brain MRI segmentation: State of the art and future directions," *J. Digit. Imag.*, vol. 30, no. 4, pp. 449–459, Aug. 2017.

[14] 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.

[15] R. S. Gill, S.-J. Hong, F. Fadaie, B. Caldairou, B. C. Bernhardt, C. Barba, A. Brandt, V. C. Coelho, L. d'Incerti, M. Lenge, and M. Semmelroch, "Deep convolutional networks for automated detection of epileptogenic brain malformations," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, Granada, Spain, 2018, pp. 490–497.

[16] 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, Dec. 2017.

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

[18] T. Brosch, L. Y. W. Tang, Y. Yoo, D. K. B. Li, A. Traboulsee, and R. Tam, "Deep 3D convolutional encoder networks with shortcuts for multiscale feature integration applied to multiple sclerosis lesion segmentation," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1229–1239, May 2016.

[19] 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.

[20] S. Chaudhari, G. Polatkan, R. Ramanath, and V. Mithal, "An attentive survey of attention models," 2019, *arXiv:1904.02874*. [Online]. Available: http://arxiv.org/abs/1904.02874

[21] Z. Wojna, A. N. Gorban, D.-S. Lee, K. Murphy, Q. Yu, Y. Li, and J. Ibarz, "Attention-based extraction of structured information from street view imagery," 2017, *arXiv:1704.03549*. [Online]. Available: http://arxiv.org/abs/1704.03549

[22] J. Xiao, H. Ye, X. He, H. Zhang, F. Wu, and T.-S. Chua, "Attentional factorization machines: Learning the weight of feature interactions via attention networks," 2017, *arXiv:1708.04617*. [Online]. Available: http://arxiv.org/abs/1708.04617

[23] H. Li, P. Xiong, J. An, and L. Wang, "Pyramid attention network for semantic segmentation," 2018, *arXiv:1805.10180*. [Online]. Available: http://arxiv.org/abs/1805.10180

[24] F. Wang, M. Jiang, C. Qian, S. Yang, C. Li, H. Zhang, X. Wang, and X. Tang, "Residual attention network for image classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 3156–3164.

[25] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," 2014, *arXiv:1409.0473*. [Online]. Available: http://arxiv.org/abs/1409.0473

[26] M. Jaderberg, K. Simonyan, and A. Zisserman, "Spatial transformer networks," in *Proc. Adv. Neural Inf. Process. Syst.*, Montreal, QC, Canada, 2015, pp. 2017–2025.

[27] J. Hu, L. Shen, and G. Sun, "Squeeze-and-Excitation networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Salt Lake City, UT, USA, Jun. 2018, pp. 7123–7141.

[28] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *Proc. Eur. Conf. Comput. Vis.*, Amsterdam, The Netherlands, 2016, pp. 630–645.

[29] 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.* Munich, Germany, 2015, pp. 234–241.

[30] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. Van Der Laak, B. Van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017.

[31] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 4700–4708.

[32] M. D. Zeiler, D. Krishnan, G. W. Taylor, and R. Fergus, "Deconvolutional networks," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, San Francisco, CA, USA, Jun. 2010, pp. 2528–2535.

[33] H. Chen, Q. Dou, L. Yu, J. Qin, and P.-A. Heng, "VoxResNet: Deep voxelwise residual networks for brain segmentation from 3D MR images," *NeuroImage*, vol. 170, pp. 446–455, Apr. 2018.

[34] K. Xu, J. Ba, R. Kiros, K. Cho, A. Courville, R. Salakhudinov, R. Zemel, and Y. Bengio, "Show, attend and tell: Neural image caption generation with visual attention," in *Proc. Int. Conf. Mach. Learn.*, Lille, France, 2015, pp. 2048–2057.

[35] O. Oktay, J. Schlemper, L. Le Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, B. Kainz, and B. Glocker, "Attention U-net: Learning where to look for the pancreas," 2018, *arXiv:1804.03999*. [Online]. Available: http://arxiv.org/abs/1804.03999

[36] Z. Huang, X. Wang, L. Huang, C. Huang, Y. Wei, and W. Liu, "CCNet: Criss-cross attention for semantic segmentation," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Seoul, South Korea, Oct. 2019, pp. 603–612.

[37] H. Tang, D. Xu, N. Sebe, Y. Wang, J. J. Corso, and Y. Yan, "Multi-channel attention selection GAN with cascaded semantic guidance for cross-view image translation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Long Beach, CA, USA, Jun. 2019, pp. 2417–2426.

[38] 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)*, Oct. 2016, pp. 565–571.

[39] B. Hou, G. Kang, N. Zhang, and K. Liu, "Multi-target interactive neural network for automated segmentation of the hippocampus in magnetic resonance imaging," *Cognit. Comput.*, vol. 11, no. 5, pp. 630–643, Oct. 2019.

[40] N. Abraham and N. M. Khan, "A novel focal tversky loss function with improved attention U-Net for lesion segmentation," in *Proc. IEEE 16th Int. Symp. Biomed. Imag. (ISBI)*, Venice, Italy, Apr. 2019, pp. 683–687.

[41] C. H. Sudre, W. Li, T. Vercauteren, S. Ourselin, and M. J. Cardoso, "Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations," in *Proc. Deep Learn. Med. Image Anal. Multimodal Learn. Clin. Decision Support*, Quebec City, QC, Canada, Sep. 2017, pp. 240–248.

[42] T. Song and N. Huang, "Integrated extractor, generator and segmentor for ischemic stroke lesion segmentation," in *Proc. Int. MICCAI Brainlesion Workshop*, Granada, Spain, 2018, pp. 310–318.

[43] F. Isensee, J. Petersen, S. A. A. Kohl, P. F. Jäger, and K. H. Maier-Hein, "nnU-net: Breaking the spell on successful medical image segmentation," 2019, *arXiv:1904.08128*. [Online]. Available: http://arxiv.org/abs/1904.08128

[44] J. Ashburner and K. J. Friston, "Unified segmentation," *NeuroImage*, vol. 26, no. 3, pp. 839–851, Jul. 2005.

[45] H. J. Kuijf, J. M. Biesbroek, J. De Bresser, R. Heinen, S. Andermatt, M. Bento, M. Berseth, M. Belyaev, M. J. Cardoso, A. Casamitjana, and D. L. Collins, "Standardized assessment of automatic segmentation of white matter hyperintensities and results of the WMH segmentation challenge," *IEEE Trans. Med. Imag.*, vol. 38, no. 11, pp. 2556–2568, Nov. 2019.

[46] S. Klein, M. Staring, K. Murphy, M. A. Viergever, and J. Pluim, "Elastix: A toolbox for intensity-based medical image registration," *IEEE Trans. Med. Imag.*, vol. 29, no. 1, pp. 196–205, Jan. 2010.

[47] S. M. Smith, "Fast robust automated brain extraction," *Hum. Brain Mapping*, vol. 17, no. 3, pp. 143–155, Nov. 2002.

[48] G. Park, J. Hong, and J. M. Lee, "White matter hyperintensities segmentation using UNet with highlighted foreground," Tech. Rep., 2020. [Online]. Available: https://wmh.isi.uu.nl/wp-content/uploads/2019/08/pgs.pdf

[49] 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.

[50] R. Camarasa, C. Doué, M. de Bruijne, and F. Dubost, "Segmentation of white matter hyperintensities with an ensemble of multi-dimensional convolutional gated recurrent units," Tech. Rep., 2018. [Online]. Available: https://wmh.isi.uu.nl/wp-content/uploads/2018/08/coroflo.pdf

[51] S. Boris and M. Belyaev, "WMH segmentation using an adjusted DeepMedic architecture and an improved learning approach," Tech. Rep., 2018. [Online]. Available: https://wmh.isi.uu.nl/wp-content/uploads/2019/02/neuroml_2.pdf

[52] S. Andermatt, S. Pezold, and P. Cattin, "Multi-dimensional gated recurrent units for the segmentation of white matter hyperintensites," Tech. Rep., 2018. [Online]. Available: https://wmh.isi.uu.nl/wp-content/uploads/2017/09/cian.pdf

[53] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "Cbam: Convolutional block attention module," in *Proc. Eur. Conf. Comput. Vis. (ECCV),* Munich, Germany, 2018, pp. 3–19.

[54] J. Fu, J. Liu, H. Tian, Y. Li, Y. Bao, Z. Fang, and H. Lu, "Dual attention network for scene segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 3146–3154.

[55] M. Liu and H. Yin, "Feature pyramid encoding network for real-time semantic segmentation," 2019, *arXiv:1909.08599*. [Online]. Available: http://arxiv.org/abs/1909.08599

[56] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," 2017, *arXiv:1706.03762*. [Online]. Available: http://arxiv.org/abs/1706.03762

GUIXIA KANG received the M.S. degree from Tianjin University, Tianjin, China, and the Ph.D. degree in electrical engineering from the Beijing University of Posts and Telecommunications (BUPT), Beijing.

She is currently a Professor of BUPT and the Director of Beijing International S and T Cooperation Base of Smart Medicine. She has expertise in the physical layer of 5G wireless systems and in the wireless e-Health systems. From 2002 to 2004, she was a Research Scientist with the Future Radio Concept Department of Siemens, Munich, Germany. She is/was the Project Manager of several national projects such as Important National Science and Technology Specific Project, National 863 project, National Natural Science Foundation of China, and several international cooperation projects. She has authored one English book (Shaker Verlag, Germany), three Chinese books, and authored or coauthored over 100 journal articles and conference papers.

Dr. Kang was a recipient of the New Century Talent of Ministry of Educations and the Beijing New Star of Science and Technology. She was also a recipient of the First Prize of Science and Technology Award of China Communications Association and the Second Prize of Beijing Science and Technology Progress Award.

BEIBEI HOU (Member, IEEE) was born in Henan, China, in 1992. She received the B.S. degree from Henan Normal University, Henan, in 2015. She is currently pursuing the M.S. and Ph.D. degrees with the Information and Communication Engineering College, Beijing University of Posts and Telecommunication (BUPT). Her research interests include medical image analysis and pattern recognition in brain science.

YUAN TANG (Graduate Student Member, IEEE) was born in Henan, China, in 1991. She received the B.S. degree from Xiangnan University, Hunan, in 2014, and the M.S. degree from Zhengzhou University, Henan, in 2017. She is currently pursuing the Ph.D. degree with the Information and Communication Engineering College, Beijing University of Posts and Telecommunications. Her research interests include medical image analysis and pattern recognition in brain science.

XIN XU received the M.D. degree from the Hubei University of Medicine, Hubei, China, and the M.S. degree from the 6th University of Paris, France.

She is currently an Associate Professor of medicine with the Chinese PLA General Hospital, Beijing, China. She has expertise in neuromodulation, magnetic resonance imaging interpretation and electrophysiological analysis in the field of functional neurosurgery. From 2004 to 2010, she was a Physician with the Beijing Sanbo Brain Hospital, Beijing. Since 2012, she has been the Associate Professor of medicine with the Neurosurgery Department, Chinese PLA General Hospital.

CHUAN HU was born in Hunan, China. He received the B.Eng. degree from the College of Information Engineering, Xiangtan University, Xiangtan, China, in 2018. He is currently pursuing the M.S. degree with the Beijing University of Posts and Telecommunications. His research interests include non-linear circuit, biomedical signal processing, image analysis, and deep learning.

• • •