Brain Atlas Guided Attention U-Net for White Matter Hyperintensity Segmentation

Zicong Zhang, MS¹, Kimerly Powell, PhD²,³, Changchang Yin, MS¹, Shilei Cao, MS¹, Dani Gonzalez, Yousef Hannawi, MD⁶,*, Ping Zhang, PhD¹,²,*

¹Computer Science and Engineering, The Ohio State University, Columbus, Ohio, USA
²Biomedical Informatics, The Ohio State University, Columbus, Ohio, USA
³Department of Radiology, The Ohio State University, Columbus, Ohio, USA
⁴Tencent Jarvis Lab, Tencent, Shenzhen, China
⁵Biomedical Engineering, The Ohio State University, Columbus, Ohio, USA
⁶Department of Neurology, The Ohio State University, Columbus, Ohio, USA
*Corresponding authors: yousef.hannawi@osumc.edu; zhang.10631@osu.edu

Abstract

White Matter Hyperintensities (WMH) are the most common manifestation of cerebral small vessel disease (cSVD) on the brain MRI. Accurate WMH segmentation algorithms are important to determine cSVD burden and its clinical consequences. Most of existing WMH segmentation algorithms require both fluid attenuated inversion recovery (FLAIR) images and T1-weighted images as inputs. However, T1-weighted images are typically not part of standard clinical scans which are acquired for patients with acute stroke. In this paper, we propose a novel brain atlas guided attention U-Net (BAGAU-Net) that leverages only FLAIR images with a spatially-registered white matter (WM) brain atlas to yield competitive WMH segmentation performance. Specifically, we designed a dual-path segmentation model with two novel connecting mechanisms, namely multi-input attention module (MAM) and attention fusion module (AFM) to fuse the information from two paths for accurate results. Experiments on two publicly available datasets show the effectiveness of the proposed BAGAU-Net. With only FLAIR images and WM brain atlas, BAGAU-Net outperforms the state-of-the-art method with T1-weighted images, paving the way for effective development of WMH segmentation. Availability: https://github.com/Ericzhang1/BAGAU-Net

Introduction

Cerebral Small Vessel Disease (cSVD) is a major public health burden leading to vascular cognitive impairment, intracerebral hemorrhage, and acute ischemic stroke. Histologically, the white matter in patients with cSVD exhibits areas of demyelination and pallor that are seen as white matter hyperintensities (WMH) on the brain magnetic resonance imaging (MRI). A higher volume of WMH is thought to represent a higher burden of cSVD. Indeed, WMH volume load has been suggested as a potential biomarker for cSVD burden. Clinical studies have shown that a larger WMH volume is associated with cognitive impairment, worse stroke functional outcome, and worse response to acute stroke therapy. Hence, practical methods for accurate segmentation of WMH to determine its volume are currently needed. Manual tracing WMH in MRI brain images is currently the accepted method for segmenting WMH on brain MRI images. However, it is time-consuming and may be associated inter-observer variability which limits its applicability in daily clinical practice and real-time decision making in patients with acute stroke.

Deep neural networks have shown robust performance in WMH segmentation tasks compared to manual segmentation. Most of these methods make use of T1-weighted and FLAIR imaging sequences acquired for each patient without considering their clinical usage. In general, the T1-weighted image provides the spatial location information of white matter (WM) that helps guide the WMH segmentation process. Additionally, research MRI scans are often acquired using standardized protocols in a homogenous group of patients that limit their applicability to routine clinical MRI scans. However, in the real-world clinical setting, rapid MRI acquisition in the setting of acute ischemic stroke without T1 sequences have become the normal protocol to allow for critical time sensitive clinical decision making in this setting. Hence, there is a critical need to develop a WMH segmentation method based on FLAIR sequences only that can be employed in the treatment of acute stroke patients. Brain atlases have been developed in the past to assist in imaging segmentation by improving preprocessing and image registration and they have been utilized in WMH segmentation as well. However, these approaches heavily rely on the presence of T1 sequences for ade-
Qua registration and segmentation. Recently, atlas-based segmentation was implemented in deep learning, which is either (1) employed in a neural network to learn the correspondences between target image and atlas images or (2) used to provide guidance as prior knowledge during model training. Motivated by the latter case, we explored the use of publicly available WM brain atlas to guide our WMH segmentation in the absence of corresponding T1-imaging sequence.

In this paper, we propose a novel brain atlas guided attention U-Net (BAGAU-Net) for WMH segmentation to address the primary limitations mentioned above. BAGAU-Net consists of two segmentation paths that take the FLAIR image and the spatially registered WM brain atlas separately to provide accurate segmentation results. The two segmentation paths are combined using two novel attention-based connecting mechanisms. Our contribution in this work is four fold: (1) We propose to use only FLAIR and publicly available WM brain atlas to achieve robust performance on WMH segmentation compared to using T1-weighted image; (2) we propose an end-to-end dual path model called brain atlas guided attention U-Net (BAGAU-Net) that leverage an additional path, namely atlas encoding path, to effectively capture the prior knowledge from WM brain atlas to help improve segmentation performance; (3) we introduce multi-input attention module (MAM) and attention fusion module (AFM) to combine the information from two paths to further improve the segmentation performance; (4) we evaluate our model on two publicly available datasets, the 2017 MICCAI WHM segmentation challenge and the Aging Brain: Vasculature, Ischemia, and Behavior Study (ABVIB) dataset. The results show that BAGAU-Net has out-performed previously proposed state-out-the-art method on both datasets.

**Method**

As aforementioned, T1-weighed images are not acquired frequently for WMH segmentation due to time constraints in most the clinical settings. In contrast, FLAIR images are expected in the treatment of acute stroke patients, but does not possess detailed information as T1-weighted image. To relieve such a dilemma, we propose to explore extra prior knowledge to improve the performance with only FLAIR images to match the performance with T1-weighted images. Inspired by the classical concepts of atlas-based segmentation, we propose to exploit prior knowledge hidden in a WM brain to guide the segmentation of the FLAIR images. To this end, we introduce the BAGAU-Net architecture for WMH segmentation. We first describe the generation process of standard WM brain atlas. Then we go into details about the dual-path architecture, including (1) the segmentation path, (2) the atlas encoding path, (3) the multi-input attention module (MAM), and (4) the attention fusion module (AFM).

![Figure 1: Examples of MR images, from left to the right are: (a) T1-weighted atlas image from ICBM152, (b) corresponding WM brain atlas, (c) FLAIR image, (d) manual WMH segmentation](image)

**Atlas Generation**

Brain atlases are often used to identify neuroanatomical structures of the brain. We used the ICBM152 (2009c Non-linear Symmetric, 1 × 1 × 1 in mm) and the corresponding WM brain atlas for our model. Elastix was used to spatially register the T1-weighted ICBM152 atlas to each of our target images with the following parameters: A multi-resolution pyramid with three levels using advanced mattes mutual information, standard gradient descent optimizer. 

---

https://wmh.isi.uu.nl
https://ida.loni.usc.edu/
maximum number of iterations 2,000, and a third order B-spline interpolator. The corresponding ICBM152 WM brain atlas was then transformed to match the target image. An example of spatially registered ICBM152 T1-weighted and WM brain atlas, target FLAIR and manually segmented WMH are shown in Fig. 1.

Model Architecture

Convolution neural networks have shown robust performance in many segmentation tasks. As inspired by Li et al., we proposed brain atlas guided attention U-Net (BAGAU-Net) that consists of two separate encoding-decoding paths. As shown in Fig 2, the upper path is a U-Net like architecture designed to extract semantic information from the image itself. The lower path is the atlas encoding path where the spatially registered atlas image is input to help guide the decoding process in the segmentation path. Moreover, we designed a multi-input attention module (MAM) and attention fusion module (AFM) to effectively combine the information between the two paths during the decoding process of segmentation path based on the attention gate (AG) as introduced by Oktay et al. The overview of the proposed architecture can be viewed in Fig. 2.

Figure 2: Overview of the BAGAU-Net for WMH segmentation. The model consists of 3x3, 5x5 encoding and decoding block, multi-input attention module (MAM), attention fusion module (AFM), upsampling, and pooling operation.

Segmentation Path

The segmentation path is designed to extract semantic features from the MRI scans. We adopt a U-Net like architecture that consists of four consecutive encoding blocks with feature channel equaling to 64, 96, 128, 256, 512 followed by four decoding blocks connected by skip connections. Each encoding block consists of two consecutive 3x3 convolutions followed by batch normalization and a rectified linear unit (ReLU). The max-pooling operation is applied for the down-sampling process to extract high-level features. Each decoding block are constructed in a similar fashion as the encoding architecture, except that an up-sampling operation is applied at the end of the block.
The structure of Multi-input Attention Module (MAM)

![Diagram of MAM](image)

The structure of attention fusion module (AFM)

![Diagram of AFM](image)

Figure 3: The structure of components in BAGAU-Net

**Atlas Encoding Path**

The atlas encoding path is used to encode spatial information from atlas images and serves as supplementary features during the decoding process of the segmentation path. We have adopted a similar architecture as our segmentation path with two modifications: (1) each 3×3 convolution in substitute with 5×5 convolution as inspired by Peng et al. to capture multi-scale features. (2) we use addition instead of concatenation as inspired by Long et al. to combine lower and higher level features.

**Multi-input Attention Module**

We implemented an attention mechanism that takes multi-input to compute target-aware features. Attention Gate, as proposed by Oktay et al., is an efficient way of extracting salient features and contextual information, which allows the model to learn to focus on a subset of target features. As shown in Fig. 3 (a), MAM takes input from both segmentation path and atlas encoding path at each level and computes the combined feature based on the re-weighted summation of results computed by the attention gate.

\[
q_{att} = W_{x}^{T} x_{i} + W_{g}^{T} g_{i} + b_{g},
\]

\[
\alpha_{i} = \sigma(q_{att}(x_{i}, g_{i}, \theta_{att})).
\]

As shown in Fig. 3 (a), MAM takes input from both segmentation path and atlas encoding path at each level and computes the combined feature based on the re-weighted summation of results computed by the attention gate.

**Attention Fusion Module**

Given the output from the last layer of both the segmentation path and atlas encoding path, we implemented a channel-attention mechanism to refine the feature of the last layer. The output from the atlas encoding path is used to compute the attention mask, where, as shown in Fig. 3 (b), the result is then fused with the output from the segmentation path using element-wise multiplication. A final convolution layer is applied to produce the segmentation result based the concatenation of the fused features.

**Experiment**

**Datasets**

The 2017 MICCAI WMH segmentation challenge dataset is publicly available. The dataset consists of 60 training subjects collection from 3 different types of scanners. A more detailed description of the challenge dataset can be found in Table. 1. For each subject, a FLAIR, T1, and the ground truth (manual segmentation of WMH) are provided. All images were bias-corrected using SPM12. We performed image registration through Elastix. The ABVIB dataset is a publicly available dataset originally established to study aging brains and cognitive consequences secondary to cardiovascular risk factors. 30 FLAIR subject images were selected and bias-corrected without associated
Table 1: Overview of the 2017 MICCAI WMH segmentation challenge dataset

<table>
<thead>
<tr>
<th>Institute</th>
<th>Scanner</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>UMC Utrecht</td>
<td>3 T Philips Achieva</td>
<td>20</td>
</tr>
<tr>
<td>NUHS Singapore</td>
<td>3 T Siemens TrioTim</td>
<td>20</td>
</tr>
<tr>
<td>VU Amsterdam</td>
<td>3 T GE Signa HDxt</td>
<td>20</td>
</tr>
</tbody>
</table>

T1 sequences. Manual segmentation of WMH for ground truth was performed by a trained student and independently verified by two investigators including a board certified neurointensivist and neurologist.

Model Training and Implementation Details

We split both WMH challenge and ABVIB dataset into training set, validation set and testing set with a ratio of 8:1:1. Instead of using DSC loss for segmentation, we adopt the Tversky loss metric\(^\text{15}\) for all model training to provide more balanced weighting between false positives (FPs) and false negatives (FNs). Let P and G be the predict and true label correspondingly. The Tversky loss can be expressed as:

\[
T(P, G; \alpha) = \frac{|PG|}{|PG| + \alpha|P \setminus G| + (1 - \alpha)|G \setminus P|},
\]

where \(\alpha\) in the above equation is a hyper-parameter that controls the trade off between FP and FN. We set the value of \(\alpha\) to 0.7 based on the results of extensive search by Salehi et al.\(^\text{15}\). We use Adam optimizer with learning rate of 0.0002 and batch size of 32 for all models; the number of epochs is set to 200. To demonstrate the effectiveness of the atlas encoding path, we compared our proposed model to the winning team’s solution\(^\text{9}\) as well as the same architecture but simply adding the atlas knowledge as a third channel input. We implemented our model using Pytorch platform\(^\text{7}\) and trained the model on single Nvidia Volta V100 GPU with 16GB memory. We adopt gradient accumulation when dealing with out of memory problems. We employed data augmentation to both datasets, including mirroring, rotation, shearing and scaling. Gaussian normalization is further applied in addition to bias-correction to reduce variation across images. A threshold of 0.5 was applied to transform the output of the model into a binary segmentation map.

Model Evaluation

The results of the BAGAU-Net segmentation were evaluated both by visual comparison of the automated segmentation and the manually segmentations and by using quantitative metrics. Five quantitative metrics were used for the quantitative assessment: Dice coefficient (DSC), average volume difference (AVD), Recall, and F1 score were calculated for each segmentation. Dice coefficient (DSC) is a measure of the amount of overlap between two segmented structures. The average volume difference measures the total volume difference between ground truth and the predicted segmentations. The average volume difference measures the total volume difference between ground truth and the predicted segmentations. Recall and F1 were calculated and based on individual number of lesions.

Results and Discussion

Here we show our evaluation of the proposed BAGAU-Net on the 2017 MICCAI WMH segmentation challenge and the ABVIB datasets. A quantitative evaluation of the results is summarized in Table 2 and Table 3 respectively. Results showed that incorporate the knowledge of atlas can improve segmentation performance, and our proposed BAGAU-Net yields competitive results by achieving the best performance across all metric comparisons. More importantly, our proposed method has achieved better DSC and AVD scores compared to segmentation performance using standard FLAIR and T1-weighted images, which are important metrics for evaluating WMH segmentation in clinical images. The recall and F1 statistics were lower in the ABVIB data set than in the WMH Challenge data set. This was primarily due to the presence of a number of small WMH lesions in the ABVIB dataset that were not as prevalent in the WMH Challenge dataset. However the DSC and AVD metrics were similar for both data sets, suggesting that the same WMH lesion load for the whole brain was segmented correctly using the BAGAU-Net model. This was observed to be true in the visual review.
Table 2: Results of BAGAU-Net and other methods on WMH challenge datasets. ↓ indicates that the smaller the better (0=best, and 100=worst). ↑ indicates that the greater the better (0=worst, and 100=best).

<table>
<thead>
<tr>
<th>Model</th>
<th>DSC (%) ↑</th>
<th>AVD (%) ↓</th>
<th>Recall (%) ↑</th>
<th>F1 (%) ↑</th>
</tr>
</thead>
<tbody>
<tr>
<td>U-Net³ (FLAIR + T1)</td>
<td>81.39</td>
<td>16.63</td>
<td>81.41</td>
<td>78.12</td>
</tr>
<tr>
<td>U-Net³ (FLAIR)</td>
<td>80.58</td>
<td>17.13</td>
<td>81.72</td>
<td>76.75</td>
</tr>
<tr>
<td>U-Net³ (FLAIR + Atlas)</td>
<td>80.61</td>
<td>17.99</td>
<td>80.27</td>
<td>76.07</td>
</tr>
<tr>
<td>BAGAU-Net</td>
<td>82.02</td>
<td>14.19</td>
<td>82.58</td>
<td>78.42</td>
</tr>
</tbody>
</table>

Table 3: Results of BAGAU-Net and other methods on ABVIB datasets.

<table>
<thead>
<tr>
<th>Model</th>
<th>DSC (%) ↑</th>
<th>AVD (%) ↓</th>
<th>Recall (%) ↑</th>
<th>F1 (%) ↑</th>
</tr>
</thead>
<tbody>
<tr>
<td>U-Net³ (FLAIR)</td>
<td>77.48</td>
<td>21.18</td>
<td>51.50</td>
<td>56.13</td>
</tr>
<tr>
<td>U-Net³ (FLAIR + Atlas)</td>
<td>76.56</td>
<td>23.87</td>
<td>52.12</td>
<td>56.14</td>
</tr>
<tr>
<td>BAGAU-Net</td>
<td>80.27</td>
<td>16.17</td>
<td>58.51</td>
<td>60.48</td>
</tr>
</tbody>
</table>

Figure 4: Test images from the 2017 WMH challenge dataset (1st column), ground truth (2nd column), the segmentation results obtained by U-Net with FLAIR and T1 (3rd column), FLAIR only (4th column), our proposed BAGAU-Net (5th column) with FLAIR and WM atlas.
Table 4: Results of ablation study on MAM and AFM on WMH challenge dataset

<table>
<thead>
<tr>
<th>Model</th>
<th>DSC (%)</th>
<th>AVD (%)</th>
<th>Recall (%)</th>
<th>F1 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BAGAU-Net (without MAM and AFM)</td>
<td>80.52</td>
<td>17.33</td>
<td>81.64</td>
<td>74.19</td>
</tr>
<tr>
<td>BAGAU-Net (without MAM)</td>
<td>81.29</td>
<td>15.72</td>
<td>82.13</td>
<td>77.31</td>
</tr>
<tr>
<td>BAGAU-Net (without AFM)</td>
<td>80.93</td>
<td>16.19</td>
<td>81.87</td>
<td>76.63</td>
</tr>
<tr>
<td>BAGAU-Net</td>
<td>82.02</td>
<td>14.19</td>
<td>82.58</td>
<td>78.42</td>
</tr>
</tbody>
</table>

Segmentation results are shown in Fig. 4. It is clear that the absence of T1-weighted images can result in less accurate segmentation results as shown in column 4 (FLAIR only) versus columns 3 (FLAIR + T1). It can be observed that BAGAU-Net, using only FLAIR and WM brain atlas, yield similar segmentation results compared to segmentation results using FLAIR and T1-weighted image as shown in Fig. 4 column 5. Both experimental results and visual comparison showed that BAGAU-Net can effectively capture useful information from WM brain atlas and use it as prior guidance to yield more accurate segmentation results. Compared to the state-of-the-art method developed for WMH segmentation that requires the presence of T1-weighted image, our proposed method can provide even better results with only FLAIR and publicly available WM brain atlas. Furthermore, our model performs better compared to a recent review of the automated methods for WMH segmentation in which the calculated average DSC across studies was 0.73\(^{20}\).

To illustrate the effectiveness of different components in BAGAU-Net, we perform ablation studies on the WMH challenge datasets as shown in Table 4. For our baseline method, we took out both MAM and AFM from the model and replace it with simple concatenation. Compared to our baseline, we observed that adding either attention modules improves segmentation performance across all metrics. Moreover, including both MAM and AFM further improves the results.

Model Limitation

The main limitation of our model is its application in a relatively small sample size of patients with cSVD (total sample size is 90 subjects). However, more importantly, to address this limitation and increase the external validity of our model, our study subjects were derived from two different datasets (MICCAI and ABVIB). Furthermore, we have only utilized FLAIR sequences in this study to simulate real life scans in this study. The second limitation of our model is that we did not segment separately other lesion types of cSVD such as lacunes since lacunes may have different clinical and prognostic significance than WMH. Our aim in this study, however, was to first develop a successful paradigm for WMH since it is the most prevalent type of cSVD. Current paradigms are under development to subsequently segment lacunes within WMH regions. Finally, our current model was developed in cohorts of patients without other acute or chronic brain etiologies such as acute stroke or brain tumors which can introduce further heterogeneities in brain MRI scans in different clinical settings. The aim of this study, however, was to focus on the method development for combining brain atlas knowledge with deep learning for WMH alone. We will subsequently validate this approach in patients with other brain pathologies.

Conclusion and Future Work

In this paper, we proposed brain atlas guided attention U-Net to improve performance on WMH segmentation. Our model combined domain knowledge of WM brain atlas with deep convolution neural networks by introducing the segmentation path and the atlas encoding path. The proposed MAM and AFM mechanisms were applied at each decoding steps and final prediction step accordingly to help to capture more comprehensive features as well as increase model interpretability by incorporating prior knowledge. We showed that WMH segmentation with T1-weighted image could be replaced and even improved by using publicly available WM brain atlas with our proposed BAGAU-Net. Results showed that our proposed model has out-performed state-out-of-the-art for WMH segmentation on both datasets. Datasets used in this study were collected from elderly group (>65), but the brain atlas we used were generated from a younger group (44 ± 7) (currently, there is no public available brain atlas that focus on aging groups). Ventricles and sulci spaces in aging brains are generally larger than that of young brains\(^{16}\). Hence, our model can be further improved by the implementation of an aging white matter brain atlas.
References


