A 3D CROSS-HEMISPHERE NEIGHBORHOOD DIFFERENCE CONVNET FOR CHRONIC STROKE LESION SEGMENTATION

Yan-Ran Wang, Hengkang Wang[1]
Sophia Chen[1]*, Aggelos K. Katsaggelos*

Northwestern University
Department of EECS

Adam Martersteck[2]*, James Higgins*[2]*,*
Virginia B Hill, Todd B Parrish

Northwestern University
Department of Radiology

ABSTRACT

Chronic stroke lesion segmentation on magnetic resonance imaging scans plays a critical role in helping physicians to determine stroke patient prognosis. We propose a convolutional neural network (CNN) segmentation network - a 3D Cross-hemisphere Neighborhood Difference ConvNet - which utilizes brain symmetry. The main novelty of this network lies on a 3D cross-hemisphere neighborhood difference layer which introduces robustness to position and scale in brain symmetry. Such robustness is important in helping the CNN distinguish between minute hemispheric differences and the asymmetry caused by a lesion. We compared our model with the state-of-the-art method using a chronic stroke lesion segmentation database. Our results demonstrate the effectiveness of the proposed model and the benefit of a CNN that combines the physiologically based information, that is, the brain symmetry property.

Index Terms— stroke lesion segmentation, brain symmetry, convolutional neural networks

1. INTRODUCTION

Stroke is a disease that affects the arteries supplying the brain parenchyma. It is the fifth highest cause of death and a leading cause of disability globally in the last 15 years. The precise delineation of an infarct is critical for understanding the impact of the lesion on behavior and developing biomarkers for stroke patient treatment and rehabilitation. Manual tracing on brain magnetic resonance imaging (MRI) scans, the gold standard for stroke lesion identification, can vary between experts and is a time consuming process. Therefore, there is a real need for an automated segmentation algorithm for stroke lesions. In this paper, we focus on the chronic stroke lesion.

A prominent feature in manual chronic stroke lesion identification is *the symmetry of the brain across the midline axis (left/right) on axial and coronal planes*. For healthy brains, the right and left hemispheres are approximate mirror images of each other, as shown in Figure 1a. Strokes are either ischemic or hemorrhagic, with occlusion of an artery resulting in an infarct or rupture of an arterial aneurysm or arteriovenous malformation resulting in intracranial hemorrhage. As the resulting lesion becomes chronic, the sequela of both ischemic and hemorrhagic strokes is encephalomalacia and gliosis. In either case, the location of the lesion is usually limited to a single hemisphere due to the majority of blood being delivered to the brain via the left or right carotid artery. As a result, chronic infarcts often are asymmetric, as shown in Figure 1b. Large infarcts involving both hemispheres concurrently are less common; therefore, in this paper we only focus on the unilateral cases.

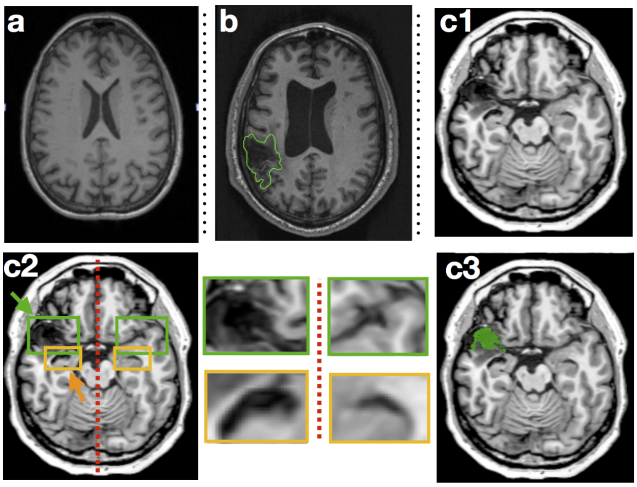

Fig. 1. (a) a MRI scan of a healthy brain. (b) a MRI scan of a chronic stroke lesion patient. The green contour delineates the lesion area. (c) An illustration of the brain symmetric property. c1 shows one slice from a 3D MRI scan of a stroke patient. In c2, the yellow arrows point to healthy areas, and the green arrow points to lesion areas. The four patches in the middle are enlarged crops of c2. c3 shows the groundtruth lesions, masked by color green.

The symmetry of the brain has been mostly ignored in previous studies using convolution neural networks (CNN)

[1,2]These two authors contributed equally to this work

which have revolutionized stroke lesion segmentation performance both in terms of accuracy and computational efficiency [4] [5] [7]. One of the main challenges that limits the use of the brain symmetry property is its complexity: *the brain is not perfectly symmetric at the pixel-level*. Clinicians can easily identify the abnormal area by comparing the hemispheres and outlining the T1 and FLAIR hypointense core infarct, but the symmetry itself cannot simply be outlined. When it comes to the design of the neural network model, some natural questions arise: How to embed the brain symmetry property into the CNN architecture? Specifically, how to teach a CNN to distinguish between minute hemispheric differences and the asymmetry caused by a lesion? Does combining this physiologically based information - brain symmetry property - with a CNN benefit the stroke lesion segmentation?

Before answering these questions, we make an observation of an example case in Figure 1c: the lesion (indicated by the green arrow in 1c2) and the normal crescent-shaped temporal horns of the lateral ventricles (indicated by the yellow arrows in 1c2). The dark chronic infarct with encephalomalacia and gliosis (indicated by the green arrow) breaks brain symmetry and can be clearly identified as a lesion. For the crescent-shaped temporal horns, the situation is more complicated. The left and right areas have a certain degree of similarity, but they do not match perfectly in terms of shape, size, and location because there is ex vacuo dilatation of the temporal horn of the lateral ventricle on the side of the infarct and possibly because of normal asymmetry in the ventricles. Despite this difference, they are labelled as non-lesion regions and represent the lateral ventricular temporal horns. Inspired by this observation, we propose a novel CNN segmentation network - *a 3D Cross-hemisphere Neighborhood Difference ConvNet*. The main novelty of this network lies on a 3D cross-hemisphere neighborhood difference layer which introduces robustness to location, scale and shape in brain symmetry. The effectiveness of this layer is verified in §5. In addition, the middle fusion in our network makes use of the potential semantic information that high-level feature maps present. These novel aspects of our network lead to an important improvement over the previous state of the art method on a chronic stroke lesion database of 70 subjects.

2. RELATED WORK

Recently, some work has made use of CNNs for lesion segmentation. In [5] and [8], the authors proposed a CNN that is trained and tested on 2D MR image slices, but ignored the through-slice linking of lesions. DeepMedic 2017 [7] was one of the first CNN-based methods to make use of the 3D nature of MR images. The algorithm was based on two parallel fully convolutional networks (FCNs) that operate on two different image scales. However, these CNN-based methods ignore a useful feature in stroke lesion segmentation: the symmetry of the brain.

In [10], the brain symmetry property is introduced through early fusion: the bilateral patch descriptors extracted from left and right hemispheres are fed into two input channels and fused right after the first convolution layer. This fusion strategy does not make use of the neighborhood or high-level semantic differences between the bilateral patches. Additional studies that inspired the proposal of our model is [1], where the authors introduce an improved deep learning architecture that adds robustness to positional differences for person re-identification. Our work is different from person re-identification in terms of the data and robustness demand, which require a more complex design in a 3D network model.

3. THE MODEL ARCHITECTURE

We approach the chronic stroke lesion segmentation problem by solving it patch by patch in the 3D space. Therefore, the problem is converted into a patch-wise segmentation problem. The input to the network for one evaluation consists of two patches, a 25^3 original patch and a corresponding symmetric patch extracted from the opposite hemisphere. The output of the network generates independent prediction of the central 9^3 voxels of the original patch. Below we introduce each component of the model in sequential order. The framework is shown in Figure 2.

3.1. Siamese Pathways

A siamese network calculates the similarity between the two images by feeding them into two identical neural network branches, and merging the two branches at the last layer. Inspired by siamese networks, we extract the high-level feature maps of each patch by passing them into two tied convolutional pathways. Each pathway consists of eight convolutional layers with 30 to 50 kernels of size $3 \times 3 \times 3$, which is similar to that described in DeepMedic [7]. At the end of the two feature extraction pathways, 50 feature maps with size 9^3 are generated to represent each input patch.

3.2. 3D Cross-hemisphere Neighborhood Difference Layer (Upscale)

The siamese pathways provide two sets of 50 feature maps for the original patch and the corresponding symmetric patch, respectively. The intuitive next step is to compute the difference (at the brain level) and learn the relationship between the two patches from the left and right hemispheres. Instead of applying a $1 \times 1 \times 1$ convolution (the most commonly used fusing operation) which weighs the whole feature maps (100 feature maps) globally, we propose to add a cross-hemisphere neighborhood difference layer. In this layer we introduce not only neighborhood information, but also scale and location robustness into the calculation of the similarity between the two patches. The operation of this layer is defined as below.

$$\tilde{k}_i = l_i(x, y, z)\mathbb{1}(3, 3, 3) - \mathcal{N}[r_i(x, y, z)] \qquad (1)$$

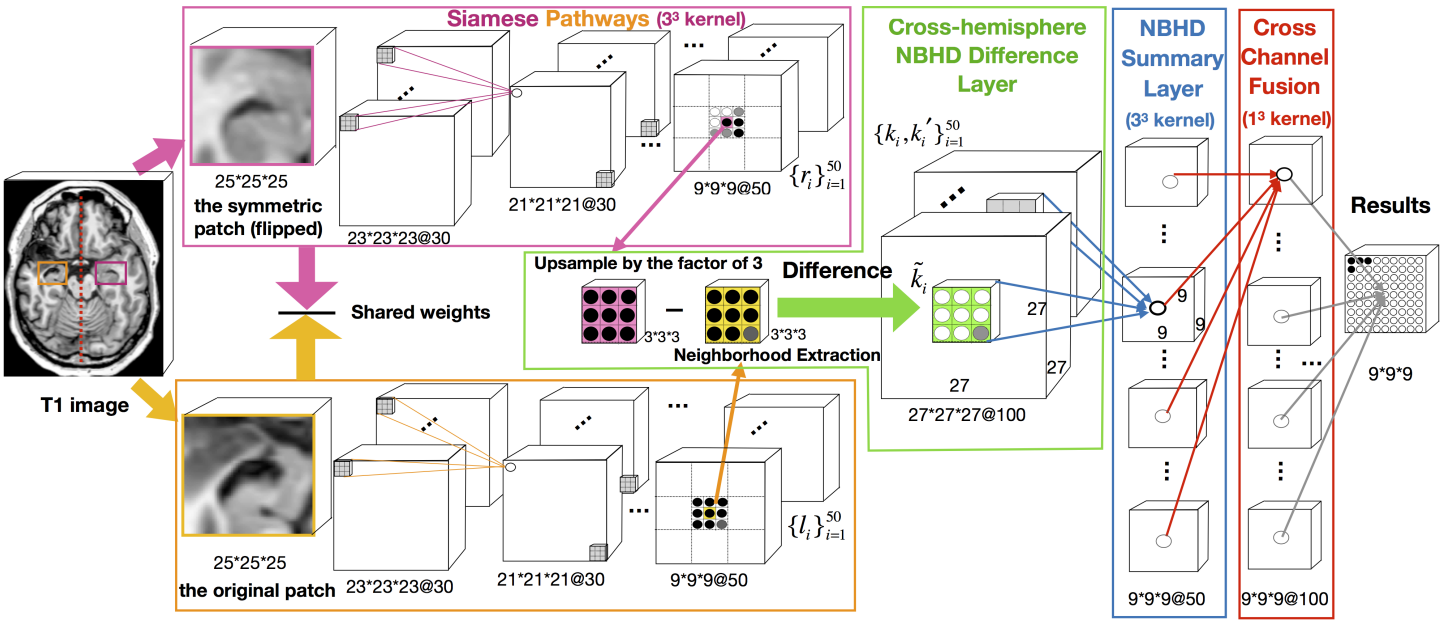

Fig. 2. An overview of the 3D Cross-hemisphere Neighborhood Difference ConvNet. NBHD is short for neighborhood. The size of feature maps and the number of channels in each layer are depicted in the format (featuremap@channel).

Where

- $l_i \in \mathbb{R}^{9\times9\times9}$ represents the ith $(1 \leq i \leq 50)$ feature map of the original patch
- $r_i \in \mathbb{R}^{9\times9\times9}$ represents the ith $(1 \leq i \leq 50)$ feature map of the symmetric patch
- (x, y, z) is a coordinate $((1 \leq x \leq 9), (1 \leq y \leq 9), (1 \leq z \leq 9))$
- $l_i(x, y, z)$ and $r_i(x, y, z)$ represent pixels in l_i and r_i feature at location (x, y, z)
- $\mathbb{1}(3, 3, 3) \in \mathbb{R}^{3\times3\times3}$ is a $3 \times 3 \times 3$ matrix filled with ones
- $\mathcal{N}[r_i(x, y, z)] \in \mathbb{R}^{3\times3\times3}$ is the $3 \times 3 \times 3$ neighborhood centered at $r_i(x, y, z)$
- $\tilde{k}_i \in \mathbb{R}^{3\times3\times3}$ is a neighborhood difference map

In other words, the 3^3 matrix $\tilde{k}_i$ is the difference of two 3^3 matrices, where the first matrix is filled with the same scalar $l_i(x, y, z)$ (i.e., upsample $l_i(x, y, z)$ by a factor of 3) and the second matrix is the 3^3 neighborhood in r_i centered at (x, y, z). The motivation of operation 1 is to incorporate scale and location robustness into brain symmetry. By concatenating $\tilde{k}_i \in \mathbb{R}^{3\times3\times3}$ produced at each location of the ith feature map, we get one output feature map $k_i \in \mathbb{R}^{27\times27\times27}$. Since there are 50 feature maps (l_i) extracted from the original patch, we get 50 output feature maps, denoted as $\{k_i\}_{i=1}^{50}$. By exchanging the order of l_i and r_i, another set of the output feature maps is produced, denoted as $\{k_i{}'\}_{i=1}^{50}$. The 3D cross-hemisphere neighborhood difference layer upscales the feature maps.

3.3. Neighborhood Summary Layer (Downscale)

After the cross-hemisphere neighborhood difference layer, the network produces $\{k_i\}_{i=1}^{50}$ and $\{k_i{}'\}_{i=1}^{50}$, which means each pixel in the output 9^3 patch is now represented by one hundred (50+50) neighborhood difference matrices (maps) of size 3^3. We summarize these 100 matrices by producing a holistic representation. This can be easily achieved through a $3 \times 3 \times 3$ kernel convolution (actually the real kernel size is $3 \times 3 \times 3 \times 100$) with a stride of 3 in each of the three directions. By matching the width of the neighborhood difference cubic matrix to the stride, we make sure that each holistic representation is computed locally within the neighborhood difference map (size 3^3). We apply 50 kernels in this layer, resulting in 50 feature maps of size 9^3 .

3.4. Cross-channel Fusion

Two convolutional layers with 100 kernels of size $1 \times 1 \times 1$ are adopted to fuse the 50 channels, generated from the neighborhood summary layer. We then apply a convolutional layer with one $1 \times 1 \times 1$ kernel to generate the final prediction of the central 9^3 patch.

3.5. The Comparison Models

Patch-wise Difference Model (PDM): we replace the *3D Cross-hemisphere Neighborhood Difference Layer* and *Neighborhood Summary Layer* with a $1 \times 1 \times 1$ convolutional layer which weighs the importance of the whole feature maps (from both hemispheres) directly. $1\times1\times1$ convolutional layer is commonly used to fuse multi-channel information [3].

Single Patch Model (SPM): only the original patch is

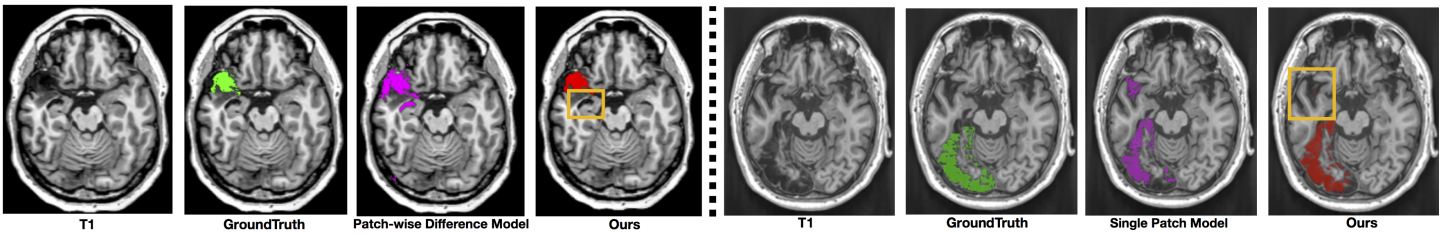

Fig. 3. The left four images show qualitative comparison on subject one with dice 98.3% achieved by our method (ours). The right four images show comparison on subject two with dice 72.3% achieved by our method.

fed into the network. No symmetry information is used. The architectures follows one of the siamese pathways in § 3.1.

4. IMAGE PREPROCESSING AND DATASET

Briefly, the 3D brain data were standardized into the right-anterior-superior (RAS) orientation. Next the subjects brain was linearly warped to MNI152 template space using 3dAllinate in AFNI [2] in order to generate the axis of symmetry. The brain was then skull stripped to extract the signal from the brain only. We used the ROBEX [6] method. Finally, the bias field created by magnetic field inhomogeneities, coil loading, RF transmit errors, and head coil receive characteristics was removed. The N4 Bias Field Correction [9] from the Insight Segmentation and Registration Toolkit (ITK) was used.

Seventy participants (age: 58.4 ± 11.9 yrs, 24 female), with aphasia resulting from a single left-hemisphere stroke, were recruited from three research laboratories. All participants underwent identical imaging that was harmonized across sites, which included a standard T1-weighted 3D MPRAGE scan acquired in the sagittal plane with isotropic resolution of $1mm^3$. Ground truth lesions were generated by experts using manual tracing, with MRIcro on the 3D T1 volume in native space.

5. EXPERIMENT

5.1. Classifier and Evaluation

Of the 70 patients with chronic stroke lesions, we randomly selected 50 brains for training and 20 brains for testing. Generally, the lesion area is much smaller than the area of healthy voxels. As a result, there is a class imbalance between the number of lesion and healthy voxels. To alleviate this imbalance, for each training case we extracted an equal number of lesion and healthy patches. For training our model, this resulted in 4000 patches in total for each brain, with 2000 lesion samples and 2000 healthy samples. The Dice similarity index was used to determine the accuracy of the segmentation network output.

5.2. Results and Discussion

We trained four models: our proposed method, PDM, and SPM, as well as the state of the art method DeepMedic [7].

Method	Dice(%)	Variance
Ours	**80.0**	1.1
PDM	77.4	1.5
SPM	75.6	2.3
DeepMedic	76.9	2.1

Table 1. Performance on 20 testing subjects. PDM stands for the patch-wise difference model. SPM is the single patch model. DeepMedic is the state-of-the-art method

The test results are shown in Table 1. We can see that our proposed method outperforms PDM by 2.6 points, SPM by 4.4 points, and DeepMedic (No symmetric information is used in this work) by 3.1 points. While these improvements may seem modest, at these levels of Dice (0.80) it is often hard to make large improvements and these small changes have a large impact on the quality of the lesion segmentation.

Figure 3 depicts two example subjects. For subject one (the left four images), the 3D cross-hemisphere neighborhood difference layer adds location and scale robustness in measuring the brain symmetry, thus preventing a false positive (the crescent-shape area - lateral ventricular temporal horns - in the yellow rectangle) compared to PDM. The benefit of adding brain symmetry is clearly shown in subject two (the right four images), where the symmetric hypo-intensity sylvian fissure (the dark region within the yellow rectangle) is misclassified as a lesion by SPM due to its limited local view.

6. CONCLUSION

In this paper, we verify that brain symmetry is beneficial for chronic stroke lesion segmentation. Most importantly, we propose a novel CNN segmentation network - *a 3D Cross-hemisphere Neighborhood Difference ConvNet*, where the novel 3D cross-hemisphere neighborhood difference layer introduces the position and scale robustness to brain symmetry. The proposed idea is not confined to chronic stroke lesion segmentation, but to any brain MRI segmentation tasks influenced by the brain symmetric property.

7. ACKNOWLEDGE

This study was supported by NIDCD P50 DC012283 and the Center for Translational Imaging, Northwestern University.

8. REFERENCES

[1] E. Ahmed, M. Jones, and T. K. Marks. An improved deep learning architecture for person re-identification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3908–3916, 2015.

[2] R. W. Cox. Afni: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical research*, 29(3):162–173, 1996.

[3] C. Feichtenhofer, A. Pinz, and A. Zisserman. Convolutional two-stream network fusion for video action recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1933–1941, 2016.

[4] R. Guerrero, C. Qin, O. Oktay, C. Bowles, L. Chen, R. Joules, R. Wolz, M. Valdés-Hernández, D. Dickie, J. Wardlaw, et al. White matter hyperintensity and stroke lesion segmentation and differentiation using convolutional neural networks. *NeuroImage: Clinical*, 17:918–934, 2018.

[5] M. Havaei, F. Dutil, C. Pal, H. Larochelle, and P.-M. Jodoin. A convolutional neural network approach to brain tumor segmentation. In *International Workshop on Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, pages 195–208. Springer, 2015.

[6] J. E. Iglesias, C.-Y. Liu, P. M. Thompson, and Z. Tu. Robust brain extraction across datasets and comparison with publicly available methods. *IEEE transactions on medical imaging*, 30(9):1617–1634, 2011.

[7] K. Kamnitsas, C. Ledig, V. F. 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. *Medical image analysis*, 36:61–78, 2017.

[8] O. Maier, C. Schröder, N. D. Forkert, T. Martinetz, and H. Handels. Classifiers for ischemic stroke lesion segmentation: a comparison study. *PloS one*, 10(12):e0145118, 2015.

[9] N. J. Tustison, B. B. Avants, P. A. Cook, Y. Zheng, A. Egan, P. A. Yushkevich, and J. C. Gee. N4itk: improved n3 bias correction. *IEEE transactions on medical imaging*, 29(6):1310–1320, 2010.

[10] Y. Wang, A. K. Katsaggelos, X. Wang, and T. B. Parrish. A deep symmetry convnet for stroke lesion segmentation. In *Image Processing (ICIP), 2016 IEEE International Conference on*, pages 111–115. IEEE, 2016.