SUPPLEMENTARY MATERIAL FOR "FAN-NET: FOURIER-BASED ADAPTIVE NORMALIZATION FOR CROSS-DOMAIN STROKE LESION SEGMENTATION"

Weiyi Yu[1] *Yiming Lei*[2,†] *Hongming Shan*[1]

[1] The Institute of Science and Technology for Brain-inspired Intelligence
[2] Shanghai Key Lab of Intelligent Information Processing, School of Computer Science
Fudan University, Shanghai 200433, China

1. MORE DETAILS ABOUT DATASET AND IMPLEMENTATION

All the experiments are performed on the benchmark stroke lesion dataset ATLAS. As Table 1 shows, this dataset consists of 229 patients' T1-weighted MR images, involving three countries and eight cities. What's more, there is no test-retest scan inter-site or intra-site, except for site 8 including 9 test-retest, which is no impact on the "leave-one-site-out" validation.

In our experiment, the backbone of the model is U-Net, and one convolutional block consists of a 3×3 convolution, batch normalization, and ReLU activation. The structural details of it are represented in Table 2.

Table 1. The information of the 9 sites from ATLAS dataset in our experiment.

	Site	Location	# Patients
1	University of Southern California	Los Angeles, USA	55
2	University of California	Irvine, USA	34
3	University of Tübingen	Tübingen, Germany	27
4	Sunnaas Rehabilitation Hospital	Nesodden, Norway	12
5	Oslo University Hospital	Oslo, Norway	27
6	University of Oslo	Oslo, Norway	14
7	Nathan S. Kline Institute for Psychiatric Research	Orangeburg, USA	11
8	University of Texas Medical Branch	Galveston, USA	35
9	University of Michigan	Ann Arbor, USA	14

†: Corresponding Author.

(a) (b) (c)

Fig. 1. The comparison results of one MR image through FAN various α. (**a**) Origin MR image; (**b**) α = 0.1; (**c**) α = 0.2.

2. MORE ABLATION STUDY AND TRAINING CURVE

We randomly selected one slice through FAN with various α, and the comparison results are shown in Fig. 1. Compared with Fig. 1(b), some texture features related to the high-frequency amplitude component are missed in Fig. 1(c). Consequently, a suitable value for α is essential.

As for the weight of Domain loss, the ablation study is shown in Table 3. Though Domain loss can definitely improve to segmentation performance on all the metrics, different weights correspond to different degrees of improvement. Thus, λ was set to 1 in the comparison experiments.

Taking site 5 as an example, we repeat the training process five times and display the averaged training curves with a standard deviation as error bound in Fig. 2, which shows that the model converges robustly.

In addition, we also tried different values of λ for the weight of Domain loss, and set λ to 1 finally. The experimental results are listed in the supplementary materials.

Table 2. Details on the structure of U-Net.

	Feature size	Parameters
Input	$1 \times 224 \times 96$	
Conv 1	$64 \times 224 \times 96$	[3×3, 64 conv] $\times$ 2[a]
Pooling	$64 \times 112 \times 48$	[2×2, max pooling][b]
Conv 2	$128 \times 112 \times 48$	[3×3, 128 conv] $\times$ 2
Pooling	$128 \times 56 \times 24$	[2×2, max pooling]
Conv 3	$256 \times 56 \times 24$	[3×3, 256 conv] $\times$ 2
Pooling	$256 \times 28 \times 12$	[2×2, max pooling]
Conv 4	$512 \times 28 \times 12$	[3×3, 512 conv] $\times$ 2
Pooling	$512 \times 14 \times 6$	[2×2, max pooling]
Conv 5	$1024 \times 14 \times 6$	[3×3, 1024 conv] $\times$ 2
Upsampling	$1024 \times 28 \times 12$	[2×2, upsampling]-[Conv 4][c]
Conv 6	$512 \times 28 \times 12$	[3×3, 512 conv] $\times$ 2
Upsampling	$512 \times 56 \times 24$	[2×2, upsampling]-[Conv 3]
Conv 7	$256 \times 56 \times 24$	[3×3, 256 conv] $\times$ 2
Upsampling	$256 \times 112 \times 48$	[2×2, upsampling]-[Conv 2]
Conv 8	$128 \times 112 \times 48$	[3×3, 128 conv] $\times$ 2
Upsampling	$128 \times 224 \times 96$	[2×2, upsampling]-[Conv 1]
Conv 9	$64 \times 224 \times 96$	[3×3, 64 conv] $\times$ 2
Output	$1 \times 224 \times 96$	[1×1, 1 conv]+Sigmoid

[a] [3×3, 64 conv] corresponds to a convolution with a kernel size of 3×3 and 64 feature maps.
[b] [2×2, max pooling] denotes max pooling with a kernel size of 2×2.
[c] [2×2, upsampling] indicates that the height and width of the feature map are twice as large as the original via upsampling, following a convolution with a kernel size of 1×1 and a half as many feature maps; [ ]-[ ] denotes the concatenation of two feature maps.

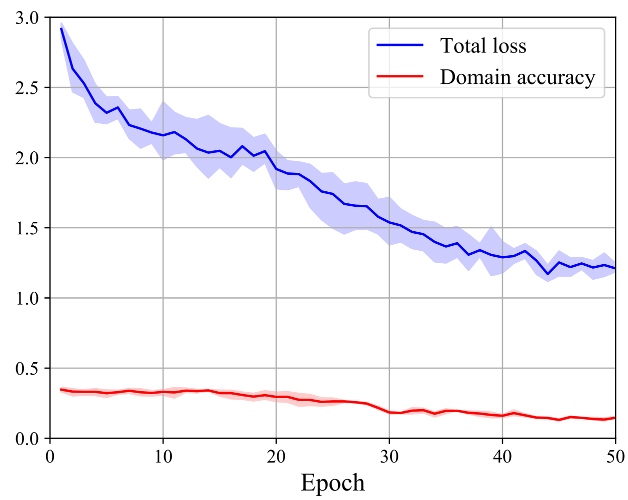

Fig. 2. Training curves of the total loss for the proposed FAN-Net and the domain accuracy for the domain classifier during training.

Table 3. An ablation study for the adjusted range of the amplitude component. As λ varies, the evaluation metrics are presented. We try $\lambda = 0.2, 0.5, 1, 1.5$.

λ	**Dice**	**Recall**	**F1-score**
0.2	0.4454	0.4529	0.4851
0.5	0.4823	0.4353	0.5070
1	**0.5098**	**0.5117**	**0.5484**
1.5	0.4597	0.4439	0.5003