



Ischemic Stroke Lesion Segmentation Using Adversarial Learning

Mobarakol Islam^{1,2}, N. Rajiv Vaidyanathan^{2,3}, V. Jeya Maria Jose^{2,4},
and Hongliang Ren²(✉)

¹ NUS Graduate School for Integrative Sciences and Engineering (NGS),
National University of Singapore, Singapore, Singapore
`mobarakol@u.nus.edu`

² Department of Biomedical Engineering, National University of Singapore,
Singapore, Singapore

`rajiv.vaidyanathan4@gmail.com`, `jeyamariajose7@gmail.com`, `ren@nus.edu.sg`

³ Department of Mechanical Engineering, NIT, Tiruchirappalli, India

⁴ Department of Instrumentation and Control Engineering, NIT,
Tiruchirappalli, India

Abstract. Ischemic stroke occurs through a blockage of clogged blood vessels supplying blood to the brain. Segmentation of the stroke lesion is vital to improve diagnosis, outcome assessment and treatment planning. In this work, we propose a segmentation model with adversarial learning for ischemic lesion segmentation. We adopt U-Net with skip connection and dropout as segmentation baseline network and a fully connected network (FCN) as discriminator network. Discriminator network consists of 5 convolution layers followed by leaky-ReLU and an upsampling layer to rescale the output to the size of the input map. Training a segmentation network along with an adversarial network can detect and correct higher order inconsistencies between the segmentation maps produced by ground-truth and the Segmentor. We exploit three modalities (CT, DPWI, CBF) of acute computed tomography (CT) perfusion data provided in ISLES 2018 (Ischemic Stroke Lesion Segmentation) for ischemic lesion segmentation. Our model has achieved dice accuracy of 42.10% with the cross-validation of training and 39% with the testing data.

1 Introduction

A stroke occurs due to an interruption in blood flow to the brain. The most common form of stroke is ischemic stroke [13] which occurs due to a reduction of blood flow, a condition known as ischemia when the brain arteries become narrow or get blocked. It is a medical emergency and undergoes in different disease stages (*acute* : $0 - -24h$, *sub - acute* : $24h - 2w$ and *chronic* : $> 2w$) with time [4]. At present, Ischemic stroke is assessed by manually delineating the lesion from computed tomography (CT) or magnetic resonance imaging (MRI) including other factors like age, blood pressure, speech, and headache. However, manual delineation is a time-consuming and tedious task which prone to human

error and inter-rater variability. Given the severity of the stroke, it is necessary to detect it as quickly as possible. With the recent advancement in the field of deep learning, convolutional neural networks can perform real-time predictions way faster than humans which have improved the process of detection of diseases.

There are few works on developing machine learning and deep learning approaches to detect, localize and segment the stroke lesion. Core and penumbra regions have been characterized using machine learning techniques in sub-acute and acute stage [9]. ISLES 2016 and 2017 [14] presents various deep learning models for stroke lesion segmentation from multi-modal MRI data. A comparison study among different machine learning and convolutional neural network (CNN) has done for stroke segmentation [10]. Islam et al. [7] exploits class-balanced PixelNet [1] to segment multi-modal neurological images (MRI) including ischemic stroke in an efficient way. However, these models have not obtained excellent accuracy by considering the visibility of the ischemic stroke in MRI or CT.

In this paper, we propose an architecture for the ischemic stroke lesion segmentation using adversarial learning. The concept of adversarial learning was first introduced in generative adversarial networks (GANs) [5]. It was initially used widely for image synthesis tasks. From then on, the concept of adversarial networks was exploited for usage in a lot of tasks, especially in segmentation. Luc et al. [8] introduce adversarial learning for the semantic segmentation. Consecutively, semi-supervised adversarial learning has been utilized in segmentation which outperforms previous approach [6]. In this work, we have proposed a variant of the adversarial network for segmentation problems by inspiring [6, 8]. We have also discussed its performance for the segmentation of ischemic stroke lesion.

2 Proposed Model

In this section, we discuss about the in detail architecture of the proposed model. The model consists of two deep networks working simultaneously, a concept adopted from the generative adversarial network [5] and adversarial learning for segmentation [6, 8]. We simultaneously train two models: a generative model G that generates the synthesized model, and a discriminative model D that estimates the probability that a sample came from the original data rather than G . The training procedure for G is to maximize the probability of D making a mistake. In our proposed model, the generator is replaced with a segmentor so that it can take in input CT data and produce segmentation label maps as its output. The discriminator is not changed.

2.1 Segmentor

U-Net [3, 12] has been used as the segmentor. The network architecture has been illustrated in Fig. 1. It is of an architecture wherein the convolution layers in the left side contract the input image and the right side expands the feature maps produced by the last convolution layer in the left side. In the left side, the

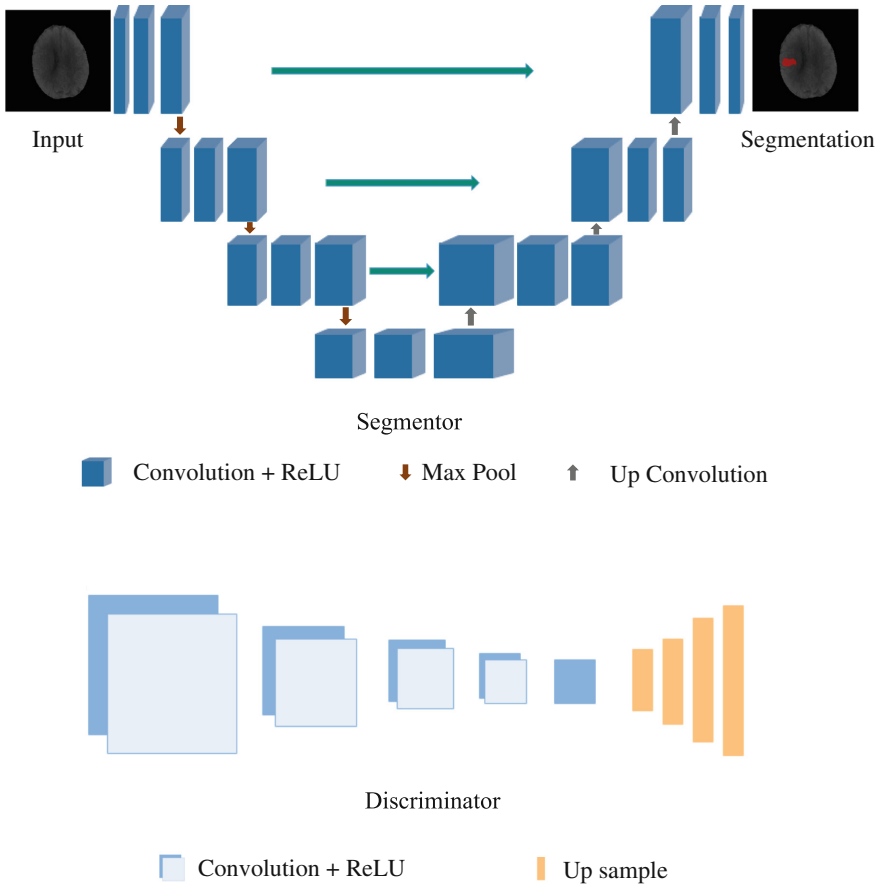


Fig. 1. UNet for adversarial training to segment ischemic stroke from CT images.

convolutional layers are of 3×3 kernel size with stride 1 followed by a rectified linear unit (ReLU) [11]. A 2×2 max pooling layer is used for downsampling with a stride 2. We double the number of feature channels at the same downsampling stage. In the right side of the network, upsampling layers are illustrated. They follow a 2×2 up-convolution which reduces the feature maps by half. A kernel size of 1×1 has been utilized at the final layer to produce feature maps of desire number classes. At the final layer, a 1×1 convolution is used to map each 64 component feature vector to the desired number of classes. In total the network has 23 convolutional layers.

2.2 Discriminator

Fully Convolutional Network [6] is used as the Discriminator. It consists of four layers of 2D Convolution followed by ReLu activation. The number of channels

increases in each convolution. In the final convolution layer, 512 channels are convolved into 2 confidence maps which consist information about the probability of real or fake data. These confidence maps are then up-sampled to the original size where each upsampling layer has a scale factor of 2.

3 Experiments

3.1 Dataset

We use ISLES 2018 (Ischemic Stroke Lesion Segmentation) CT dataset to perform all the experiments. ISLES 2018 consists 6 modalities of CT perfusion images. Such as CT, DPWI, CBF, CBV, MTT, and Tmax. Each modality has 3D CT scan of 256×256 axial dimension and a variable slice of 2 to 18 (approximately). There are 94 and 62 cases for training and testing set respectively. The ground-truth for 94 training cases is given in the training phase. The annotation contains 2 classes where 1 and 0 denote stroke lesion and the healthy tissue respectively.

3.2 Training

We exploit only 3 modalities namely CT, DPWI and CBF. We split the training data into train and valid data in an 80/20 ratio to find the best epoch for segmentation. For testing, another 62 cases are used. The 3D CT image is first converted into stacks of 2D CT images. These images are given as input to the segmentor which gives probability maps of segmented image as its output. The synthesized segmented probability maps from generator is fed to the discriminator as one of the inputs along with the ground truth. The discriminator is simultaneously trained on the ground truth which is labelled as 1 corresponding to real. 0 corresponds to fake data which is the initial data produced by the segmentor. So, when the probability maps from generator are fed to the discriminator, it gives confidence maps from which the cross-entropy loss is found with respect to two classes. With this adversarial loss along with the segmentor's loss, we seek to train the segmentation network to fool the discriminator by maximizing the probability of the segmentation prediction [6]. The total loss function is denoted as

$$\chi = \chi_{seg} + \lambda_{adv}\chi_{adv}$$

where χ is the total loss that is used to backpropagate through both the segmentor and the discriminator. χ_{seg} is the loss that is got in the segmentor by comparing the output probability map directly with the ground truth. χ_{adv} is the loss that is got in the discriminator which is actually the cross-entropy loss calculated between the two classes. λ_{adv} is the factor introduced to increase or decrease the amount by which adversarial loss will affect the total loss. λ_{adv} for ischemic stroke lesion data set is taken as 0.1. The total loss is backpropagated in both the segmentor and discriminator.

3.3 Testing

While testing, the discriminator network is removed and the segmentation labels are got by inputting the test image to the segmentor network. 2D slices have been fed to the model to predict lesion. Finally, all the predicted slices have converted to 3D CT to measure the performances.

4 Results

The cross-validation dice accuracy of different segmentation architectures has been compared with our proposed model in Table 1. Table 2 gives details about the performance of our model on the test data set. Figures 2 and 3 show the predictions of the proposed model for the validation and test data set respectively.

Table 1. Cross validation performance of different models with training dataset

Models	Ours	PixelNet [7]	U-Net [12]	Deeplab v2 [2]	ICNet [15]	PSPNet [16]
Dice	0.421	0.409	0.419	0.373	0.387	0.319

Table 2. Performance of our model with testing dataset

Dice	Hausdorff	Avg distance	Precision	Recall	AVD
0.39	17741954.64	17741938.19	0.55	0.36	10.90

5 Discussion

From the above results, it is evident that even though the network gives a good segmentation prediction, there is still room for improvement. With the same architecture, the network would perform better with a more consistent data set. This data set is inconsistent and has certain defects as some data had just two 2D slices and some images were of poor contrast or shaken. The data set is still less in number as it is difficult to find the best-converged network with such fewer cross-validation data. Even the performance gap in different folds of the cross-validation is very high.

From Table 1, it can be noted that U-Net, when trained in an adversarial way, gave a better dice accuracy when compared to normal U-Net. Training a segmentation network along with an adversarial network can detect and correct higher order inconsistencies between the segmentation maps produced by ground-truth and the Segmentor. This is due to the extra addition to the loss

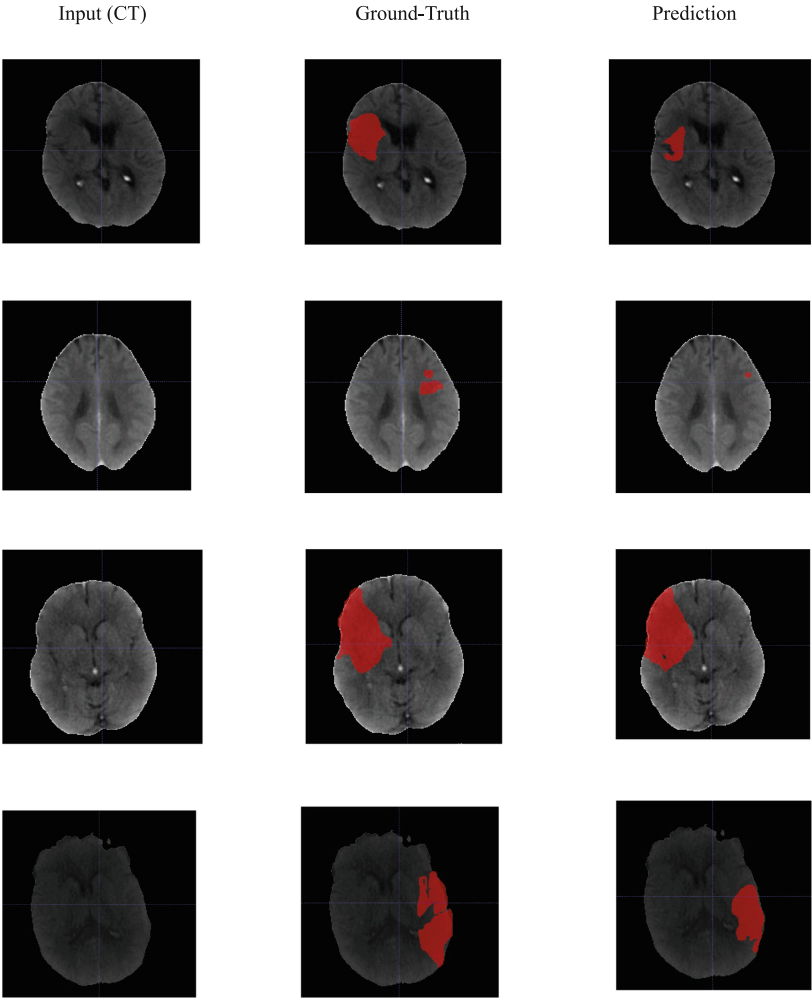


Fig. 2. Qualitative results of cross validation on training data

function that is given due to adversarial learning. Adversarial loss improves the performance of both the segmentor and discriminator. As the training is similar to a min-max game and as the segmentor does not only have its loss but also the extra loss from the discriminator, adversarial training of any segmentation architecture gives a better result when compared to its normal training.

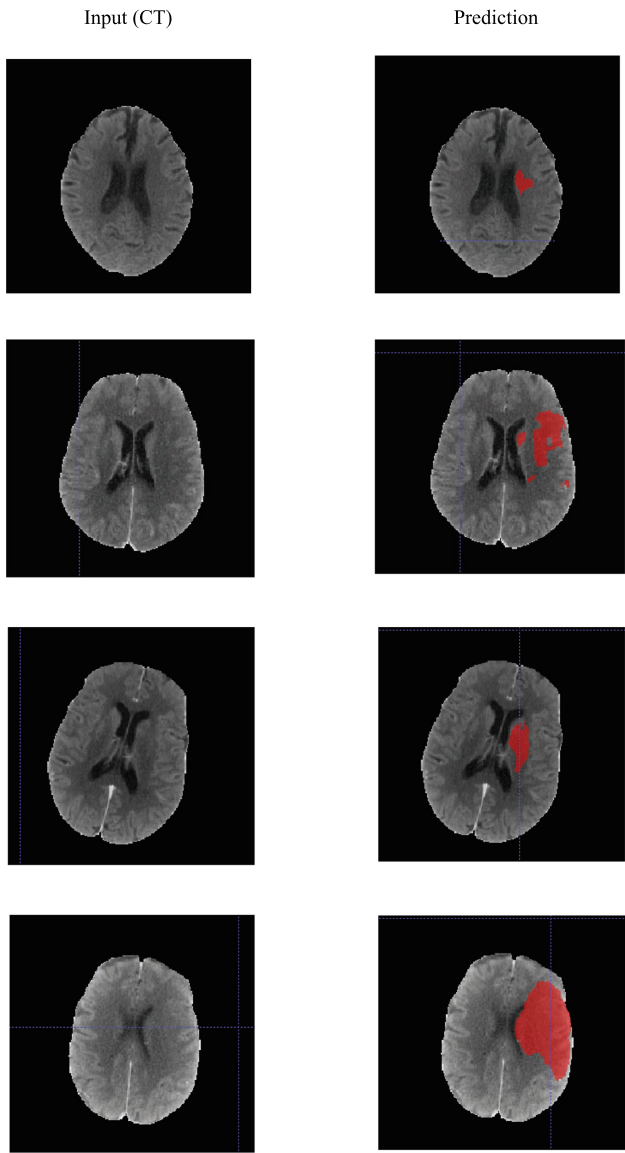


Fig. 3. Qualitative results of on testing data

6 Conclusion

In this paper, we have proposed an automatic ischemic stroke lesion segmentation using adversarial learning scheme. We have demonstrated than adversarial learning can improve the performance of the segmentation architecture like U-Net [12]. Our model has achieved better performance comparing to other state

of the art models. However, the overall performance of the model is not excellent because of poor visibility of the lesion area in the CT imaging. Moreover, 3D deep learning models cannot be suitable for this dataset where dept slices are inconsistency and very small. As a future work, we can reslice or resample the CT into a higher dimension and try to exploit 3D models directly on 3D CT images.

Acknowledgement. This work is supported by the Singapore Academic Research Fund under Grant R-397-000-227-112, NUSRI China Jiangsu Provincial Grant BK20150386 and BE2016077 and NMRC Bedside & Bench under grant R-397-000-245-511 awarded to Dr. Hongliang Ren.

References

1. Bansal, A., Chen, X., Russell, B., Gupta, A., Ramanan, D.: PixelNet: representation of the pixels, by the pixels, and for the pixels. arXiv preprint [arXiv:1702.06506](https://arxiv.org/abs/1702.06506) (2017)
2. Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L.: DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Trans. Pattern Anal. Mach. Intell.* **40**(4), 834–848 (2018)
3. Drozdal, M., Vorontsov, E., Chartrand, G., Kadoury, S., Pal, C.: The importance of skip connections in biomedical image segmentation. In: Carneiro, G., et al. (eds.) *LABELS/DLMIA -2016*. LNCS, vol. 10008, pp. 179–187. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46976-8_19
4. González, R.G., Hirsch, J.A., Koroshetz, W.J., Lev, M.H., Schaefer, P.W.: *Acute Ischemic Stroke*. Springer, Heidelberg (2011)
5. Goodfellow, I., et al.: Generative adversarial nets. In: *Advances in Neural Information Processing Systems*, pp. 2672–2680 (2014)
6. Hung, W.-C., Tsai, Y.-H., Liou, Y.-T., Lin, Y.-Y., Yang, M.-H.: Adversarial learning for semi-supervised semantic segmentation. arXiv preprint [arXiv:1802.07934](https://arxiv.org/abs/1802.07934) (2018)
7. Islam, M., Ren, H.: Class balanced pixelnet for neurological image segmentation. In: *Proceedings of the 2018 6th International Conference on Bioinformatics and Computational Biology*, pp. 83–87. ACM (2018)
8. Luc, P., Couprie, C., Chintala, S., Verbeek, J.: Semantic segmentation using adversarial networks. arXiv preprint [arXiv:1611.08408](https://arxiv.org/abs/1611.08408) (2016)
9. Maier, O., et al.: ISLES 2015-a public evaluation benchmark for ischemic stroke lesion segmentation from multispectral MRI. *Med. Image Anal.* **35**, 250–269 (2017)
10. Maier, O., Schröder, C., Forkert, N.D., Martinetz, T., Handels, H.: Classifiers for ischemic stroke lesion segmentation: a comparison study. *PloS One* **10**(12), e0145118 (2015)
11. Nair, V., Hinton, G.E.: Rectified linear units improve restricted boltzmann machines. In: *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, pp. 807–814 (2010)
12. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) *MICCAI 2015*. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28

13. World Health Organization (WHO), et al.: Cause-specific mortality estimates 2000–2012 (2014)
14. Winzeck, S., et al.: ISLES 2016 and 2017-benchmarking ischemic stroke lesion outcome prediction based on multispectral MRI. *Front. Neurol.* **9** (2018)
15. Zhao, H., Qi, X., Shen, X., Shi, J., Jia, J.: ICNet for real-time semantic segmentation on high-resolution images. arXiv preprint [arXiv:1704.08545](https://arxiv.org/abs/1704.08545) (2017)
16. Zhao, H., Shi, J., Qi, X., Wang, X., Jia, J.: Pyramid scene parsing network. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2881–2890 (2017)