

# Deep Learning Based Detection and Localization of Cerebral Aneurysms in Computed Tomography Angiography

Ziheng Duan\*

duanziheng@zju.edu.cn

College of Control Science and  
Engineering  
Zhejiang University

Daniel Montes\*

Javier M. Romero

Ramon Gilberto Gonzalez  
Department of Neuroradiology

Yangsibo Huang, Dufan Wu

Quanzheng Li†

Center for Advanced Medical  
Computing and Analysis  
Gordon Center for Medical Imaging

## ABSTRACT

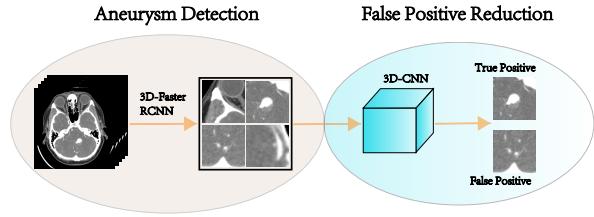
Detecting cerebral aneurysms is an important clinical task of brain computed tomography angiography (CTA). However, human interpretation could be time consuming due to the small size of some aneurysms. In this work, we proposed DeepBrain, a deep learning based cerebral aneurysm detection and localization algorithm. The algorithm consisted of a 3D faster region-proposal convolution neural network for aneurysm detection and localization, and a 3D multi-scale fully convolutional neural network for false positive reduction. Furthermore, a novel hierarchical non-maximum suppression algorithm was proposed to process the detection results in 3D, which greatly reduced the time complexity by eliminating unnecessary comparisons. DeepBrain was trained and tested on 550 brain CTA scans and achieved sensitivity of 93.3% with 0.3 false positives per patient on average.

## KEYWORDS

cerebral aneurysm, deep learning, object detection, computer-aided diagnosis

## 1 INTRODUCTION

A cerebral aneurysm is a swelling in the cerebral blood vessels, which may leak or rupture and cause subarachnoid hemorrhage (SAH). Aneurysm accounts for 85% of all SAHs, which has an average mortality rate of 51%, and one third of the survivors have long-term disabilities [8]. Early detection of cerebral aneurysms is essential for the treatment of cerebral aneurysms. Brain computed tomography angiography (CTA) is an effective approach for cerebral aneurysm detection by providing contrast-enhanced imaging of brain vascular, especially in the emergency room. Cerebral aneurysms vary in sizes from a few millimeters to several centimeters, and they can appear in various locations insides the brain. Hence, finding aneurysms in brain CTA is a non-trivial and time-consuming task for human readers. Unfortunately, there is no existing automatic tool for aneurysm detection in CTA yet. Recently, there are some works that incorporate deep learning with aneurysm detection. Park et al. [6] trained a 3D UNet to segment aneurysms and demonstrated augmented performance of physicians with the aid of the deep learning model. However, UNet needs segmentation during training which require huge amount of human labor for annotation. Ueda et al. [9] used the Inceptionv3 to further reduce the false positive (FP) rate of a morphology-based aneurysm



**Figure 1: The framework of DeepBrain. It used 3D Faster R-CNN to generate candidate aneurysms, followed by a 3D multi-scale CNN for FP reduction.**

detection software for time-of-flight magnetic resonance (MR). Unfortunately, such software does not exist for CTA. Region proposal network, such as faster region-proposal convolutional network (Faster R-CNN) [7], has been proved to be effective for object detection in natural images. Instead of semantic segmentation, it gives the central locations and bounding boxes of objects, which require far less efforts for annotation. Faster R-CNN has been extended to 3D and successfully applied in lung nodule detection. In this paper, we proposed DeepBrain, a deep learningbased aneurysm detection and localization algorithm for CTA. We used a two-step approach, where 3D Faster R-CNN was used for aneurysm detection followed by a multi-scale CNN for FP reduction. A novel 3D non-maximum suppression (NMS) algorithm was also proposed to postprocess the detection results, which greatly accelerated the 3D NMS step. To the best of our knowledge, this is the first work for cerebral aneurysm detection and localization in brain CTA without the need for the labor-intensive segmentation annotation. The method was trained and tested on 550 CTA scans and demonstrated comparable performance to reported values of existing studies on both MR and CT.

## 2 METHODOLOGY

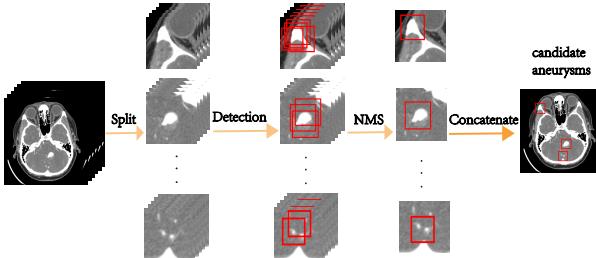
As demonstrated in Figure 1, the proposed DeepBrain consists of two parts: aneurysm detection and false positive (FP) reduction. 3D Faster R-CNN was used for the detection, followed by a novel 3D non-maximum suppression (NMS). A multi-scale 3D CNN [2] was used for the FP reduction.

### 2.1 Aneurysm Detection and Localization

Our detection and localization network followed [11] where a 3D Faster R-CNN was used to predict the spatial location of a cerebral aneurysm. The Faster R-CNN took a  $96 \times 96 \times 96$  patch as input

\*Equal contribution.

†Correspondence to: Quanzheng Li



**Figure 2: The process of 3D NMS.** Firstly, the original input 3D image was split into non-overlapping sub-volumes. Secondly, the sub-volumes were fed into the detection network which generated bounding boxes. Thirdly, NMS was performed for each volume. Finally, the bound boxes from individual volumes were aggregated and went through another NMS for the final localization results.

and output a  $24 \times 24 \times 24$  patch where each voxel had a vectorized value as  $\mathbf{v} = (p, \mathbf{k})$ , where  $p$  is the probability of the presence of an aneurysm and  $\mathbf{k}$  is the normalized relative coordinates of the aneurysm to the voxel:

$$\mathbf{k} = \left( \frac{x - x_a}{d_a}, \frac{y - y_a}{d_a}, \frac{z - z_a}{d_a}, \log\left(\frac{d}{d_a}\right) \right), \quad (1)$$

where  $x, y, z$  is the predicted aneurysm coordinates and  $d$  is its diameter.  $x_a, y_a, z_a$  is the coordinate of the voxel and  $d_a$  is the anchor size defined according to the aneurysm's size for normalization.

Various  $d_a$  are used to account for different sizes of aneurysms, where an independent  $\mathbf{v}$  is given under each  $d_a$ . For each voxel under each anchor, the training loss is:

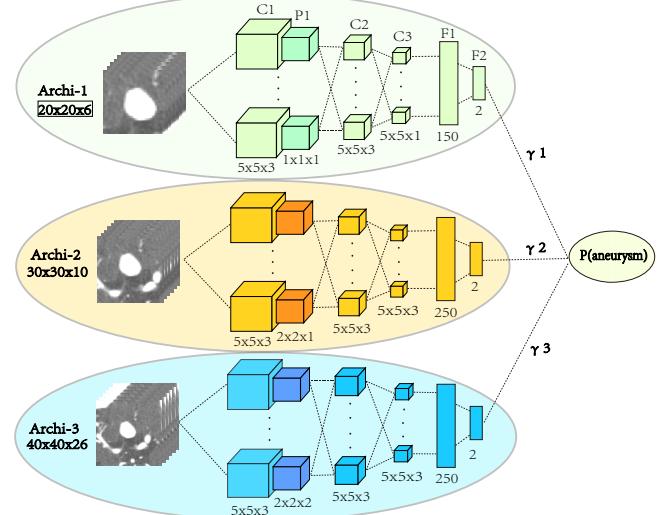
$$L(\mathbf{v}, \mathbf{v}^*) = \lambda L_{cls}(p, p^*) + p^* L_{reg}(\mathbf{k}, \mathbf{k}^*), \quad (2)$$

where  $\mathbf{v}$  is the predicted value whereas  $\mathbf{v}^*$  is the ground truth value.  $p^* = 1$  if the center of the aneurysm falls within the anchor and  $p^* = 0$  otherwise.  $L_{cls}$  is the classification loss where binary cross entropy was used, and  $L_{reg}$  is the regression loss where smooth  $l_1$  loss was used.  $\lambda$  was set to 0.5 in the study.

Dual-path network [11] was selected as the structure for the Faster R-CNN, whose structure is similar to that of UNet.

## 2.2 3D Non-Maximum Suppression

The detection network output 3 bounding boxes for each voxel in the prediction, and many of these bounding boxes may overlap. The goal of non-maximum suppression (NMS) algorithm is eliminating overlapping bounding boxes by leaving the ones with locally largest probability only. Conventional NMS algorithm traverses through all possible pairs of bounding boxes. For each pair, it removes the one with lower probability if the intersection over union (IoU) of them is above certain threshold. It has time complexity of  $O(n^2)$  where  $n$  is the total number of candidates, which leads to very long computational time in 3D due to the large number of candidates. To improve computational efficiency, as shown in Figure 2, a novel hierarchical 3D NMS algorithm was proposed in our study. To reduce computational time, given the fact that most overlapping happened locally, we did NMS in each  $96 \times 96 \times 96$  patch, followed by a



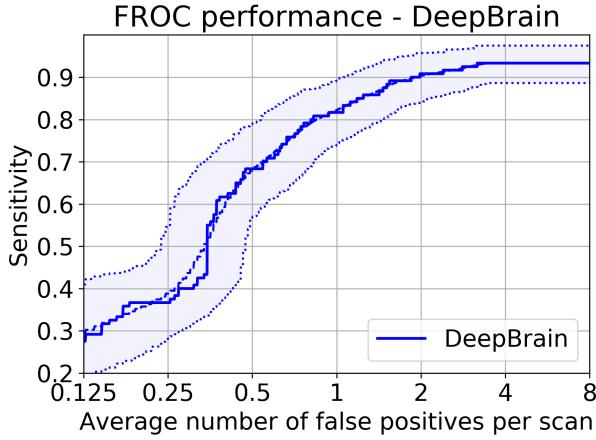
**Figure 3: The 3D multi-scale CNN for FP reduction.** Three CNNs were used to incorporate different levels of contextual information. C1, C2, and C3 are convolutional layers; P1 is a max pooling layer; F1 and F2 are fully connected layers. All the convolutional layers have 64 feature maps and the kernel sizes are given below the boxes. F1 has length of 150, 250, and 250 for the three scales respectively. F2 has length of 2. The predictions of these networks are fused to produce the final classification result.

whole image-level NMS. The first step accelerated the algorithm by a factor of approximately 75, which is the number of patches, but effectively reduced the number of bounding boxes to 1%. The second step was calculated only on the 1% bounding boxes and the acceleration factor could be 10,000.

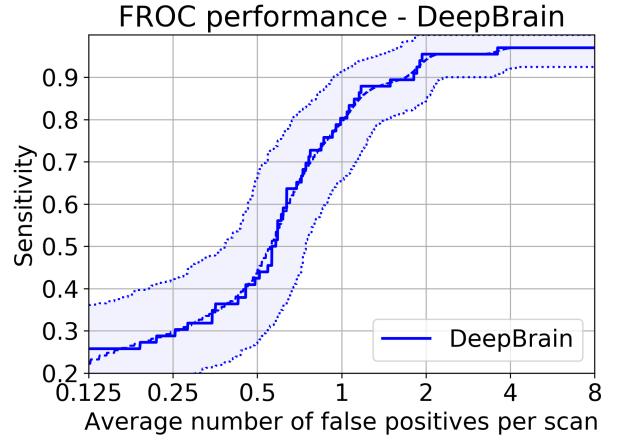
## 2.3 False Positive Reduction

Because aneurysms are very small compared to the whole images, positive samples only composed a very small portion of the training set for the detection and localization task. As the consequence, despite of the low training loss, number of FPs can be as high as several hundred per patient, making the prediction meaningless. FP reduction network further classifies the predicted bounding boxes into true positives and false positives. One of the major challenges for FP reduction is the wide range of aneurysms' size. Too small reception field of the network leads to incomplete inclusion of large aneurysms, whereas too large reception field leads to noisy background for small aneurysms. A 3D multi-scale CNN as demonstrated in Figure 3 was proposed to address this challenge. As shown in Figure 3, the network consisted of 3 subnetworks, A1, A2, and A3 with different receptive fields covering 86%, 95% and 98% of all the aneurysms in the training dataset, respectively. The networks shared similar structures of 3 convolutional layers and 2 fully connected layers. Weighted binary cross entropy was used as the training loss:

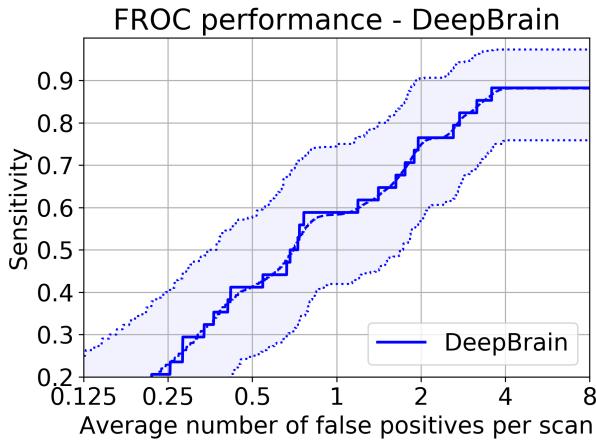
$$L = \gamma_1 L_{cls1} + \gamma_2 L_{cls2} + \gamma_3 L_{cls3}, \quad (3)$$



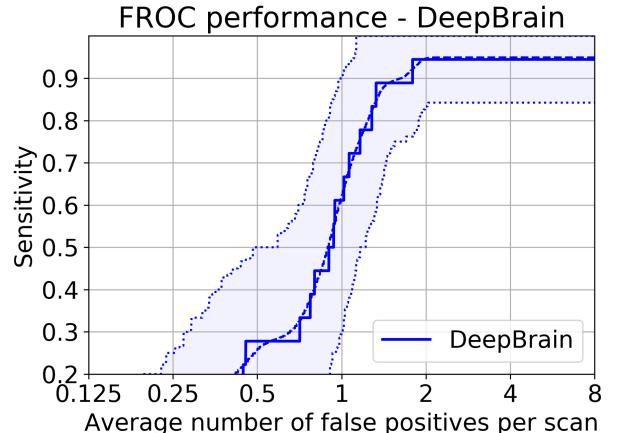
**Figure 4: FROC on the test set.**  
Only the detection network is used.



**Figure 6: FROC on the test set.**  
Only the detection network is used and only aneurysms with a radius of 5-10 mm are given.



**Figure 5: FROC on the test set.**  
Only the detection network is used and only aneurysms with a radius of 3-5 mm are given.



**Figure 7: FROC on the test set.**  
Only the detection network is used and only aneurysms with a radius more than 10 mm are given.

where  $L_{cls_i}$  is the binary cross entropy loss for sub network , and the weighting factors were selected as  $\gamma_1 = 0.3$ ,  $\gamma_2 = 0.4$ , and  $\gamma_3 = 0.3$ .

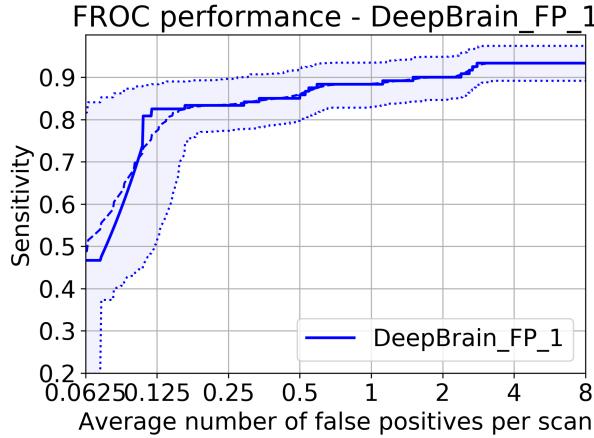
### 3 EXPERIMENTS

DeepBrain was retrospectively trained and validated on 550 emergency room CTA scans acquired at Massachusetts General Hospital under patients' consent. The study was reviewed and approved by our Institutional Review Board. Each patient had 1 to 3 aneurysms whose diameters are mostly between 3mm and 30mm. The location and diameter of each aneurysm were annotated by a neuroradiological resident onsite. Among all the patients, 165 had SAH. All the images had axial resolution of  $0.43 \times 0.43 \text{ mm}^2$  and slice thickness between 0.625 mm and 2.5 mm. We randomly split the data and used

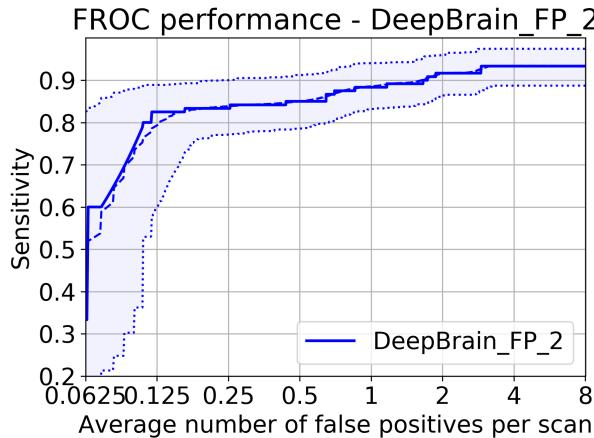
440 for training and 110 for testing. Each CTA image was scaled to Hounsfield Units (HU) / 1000 before input to the network. The trained detection network was applied on the training dataset, and the bounding boxes after NMS were extracted as the training set for FP reduction. Both networks were trained with Adam optimizer [3] with learning rate of 0.0001. The detection network is trained for 1000 epochs and all three FP reduction networks are trained for 80 epochs. Each patch was augmented with random flipping, rotation and zooming during the training.

### 4 RESULTS

We changed the decision threshold of the false positive reduction module and obtained the free-response receiver operating curve

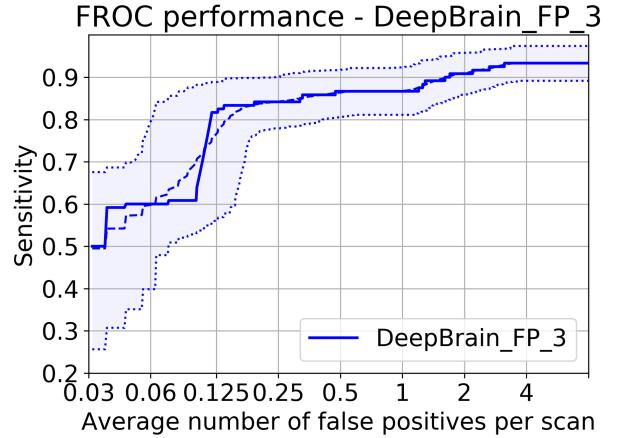
**Figure 8: FROC on the test set.**

The detection network and FP reduction network Archi-1 are used.

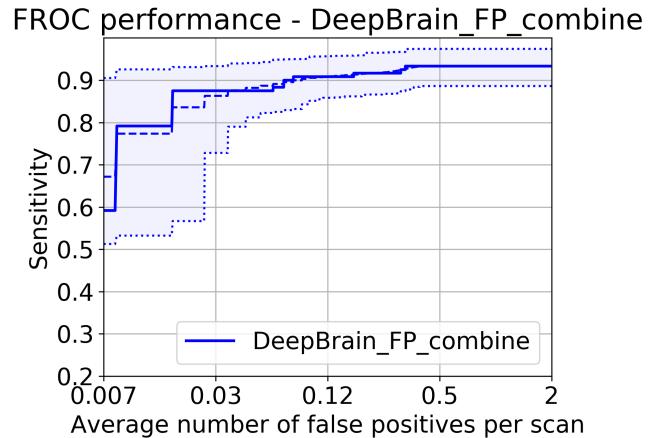
**Figure 9: FROC on the test set.**

The detection network and FP reduction network Archi-2 are used.

(FROC) as shown in Figure 11. DeepBrain achieved a sensitivity of 93.3% with an average of 0.3 false positives per patient. Table 1 gives a summary of the performance of the proposed method compared to published results. Most of the work is done on MR, the noise of MR images is lower than that of CTA, and it is not disturbed by cortical bone. There is also software for locating aneurysm candidates in MR. The latest work of Park et al. Used a UNet-like network for aneurysm detection on CTA, but they excluded all SAH patients, which is not applicable to the actual situation. The result of only passing the detection network is shown in Fig. 4. By choosing an appropriate threshold, we can achieve a sensitivity of 80%, with one false positive per scan (threshold 0.95). Or to maintain 90% sensitivity, each scan brings two false positives (threshold 0.85). As shown in Figure 5, we selected 34 cerebral aneurysms with a radius

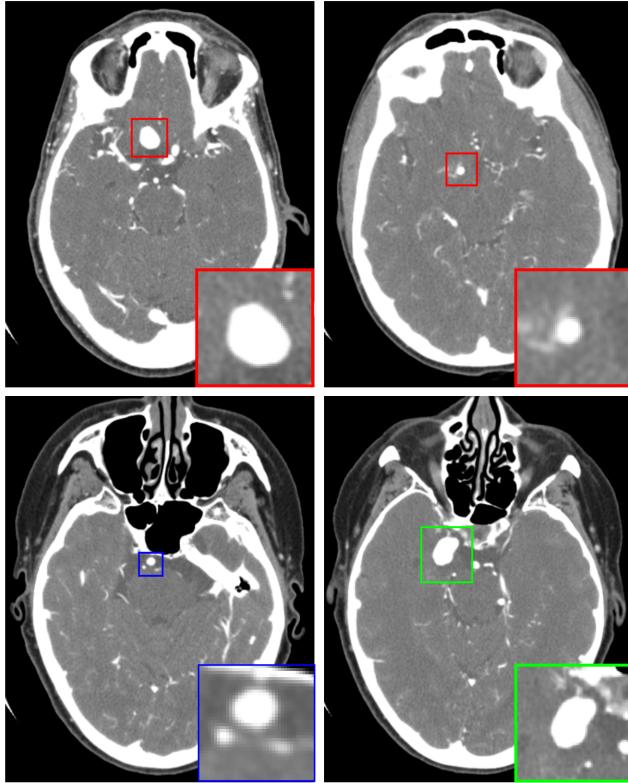
**Figure 10: FROC on the test set.**

The detection network and FP reduction network Archi-3 are used.

**Figure 11: FROC on the test set.**

The detection network and all three FP reduction networks are used.

of 3 to 5 mm for testing. In this case, 30 were detected. In this case, we can see that it is difficult to detect very small aneurysms, and its performance is slightly worse than average. But even then, we can still guarantee a sensitivity of 75%, with two false positives per scan, or a sensitivity of 88%, with four false positives at the same time. As shown in Figure 6, we selected 66 cerebral hemangiomas with a radius of 5 to 10 mm. In this case, 64 were detected. In this case, we can see that our model shows good performance. We can guarantee a sensitivity of 80% with one false positive per scan, or a sensitivity of 90% with two false positives per scan. As shown in Figure 7, we selected 18 cerebral aneurysms with a radius more than 10 mm. In this case, 17 were detected. In this case, we can see that our model can guarantee 94% sensitivity, and each scan brings two false positives. Figures 8, 9, and 10 show the FROC after



**Figure 12: Examples of aneurysm detection results. The red boxes are true positive cases, the blue box is a false positive case and the green box is a false negative case.**

**Table 1: Comparison with the published results. [6] reported sensitivity and specificity on patient level instead of per-aneurysm level.**

Method	Sensitivity	FPs/scan	#scans	Modality
Arimura et al. [1]	100%	2.40	29	MRA
Yang et al. [10]	95.0%	9.00	287	MRA
Nakao et al.[5]	94.2%	2.90	450	TOF MRA
Ueda et al. [9]	92.5%	9.00	683	TOF MRA
Park et al. [6]	94.9%	N/A	818	CTA
DeepBrain	93.3%	0.33	292	CTA

reducing the Archi-1, Archi-2, and Archi-3 of the false positive model. It can be seen that even if only a single 3D CNN network is passed, the performance obtained is far better than only passing the detection network. After synthesizing the three networks, we get the final result, as shown in Figure 11, showing the most superior performance. Figure 16 shows examples of true positives (TP), false positives (FP) and false negatives (FN). FP cases are blood vessels with an axial morphology similar to an aneurysm. FN cases showed a large aneurysm, which may be caused by the lack of a large aneurysm in the training patient.

## 5 CONCLUSION AND FUTURE DIRECTIONS

In this paper we proposed DeepBrain, an automatic aneurysm detection algorithm for CTA scans. It consisted of a 3D Faster R-CNN for aneurysm detection and a 3D multi-scale CNN for FP reduction. The two-step design granted the method with robustness and ease-to-train. A novel hierarchical 3D NMS algorithm was also proposed which significantly accelerated the computational time. Experimental results on 550 CTA scans demonstrated comparable performance to reported data on MR. One of the major advantages of the proposed method is that it needs only the location and diameter of aneurysms in the annotation instead of a full segmentation. This made DeepBrain easy to be transferred to other institutions or new scanners through transfer learning, where around a hundred annotations will be needed to fine-tune the model for the new protocol. The performance of DeepBrain can be further improved from several aspects. The sample size of our dataset can be further enlarged to include more larger aneurysms in the training dataset, which could reduce some of the false negatives as shown in Figure 12. We will also incorporate focal loss [4] in our training process to further reduce FP rate. Furthermore, the structure and training parameter of networks could also be improved and optimized.

## ACKNOWLEDGEMENTS

This work was completed at Massachusetts General Hospital and Harvard Medical School, Boston MA 02114, United States and was supported in part by National Institute of Health under Grant 1RF1AG052653 and 5P41EB022544.

## REFERENCES

- [1] Hidekata Arimura, Qiang Li, Yukunori Korogi, Toshinori Hirai, Hiroyuki Abe, Yasuyuki Yamashita, Shigehiko Katsuragawa, Ryuji Ikeda, and Kunio Doi. Automated computerized scheme for detection of unruptured intracranial aneurysms in three-dimensional magnetic resonance angiography1. *Academic radiology*, 11(10):1093–1104, 2004.
- [2] Qi Dou, Hao Chen, Lequan Yu, Jing Qin, and Pheng-Ann Heng. Multilevel contextual 3-d cnns for false positive reduction in pulmonary nodule detection. *IEEE Transactions on Biomedical Engineering*, 64(7):1558–1567, 2016.
- [3] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [4] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision*, pages 2980–2988, 2017.
- [5] Takahiro Nakao, Shouhei Hanaoka, Yukihiko Nomura, Issei Sato, Mitsutaka Nemoto, Soichiro Miki, Eriko Maeda, Takeharu Yoshikawa, Naoto Hayashi, and Osamu Abe. Deep neural network-based computer-assisted detection of cerebral aneurysms in mr angiography. *Journal of Magnetic Resonance Imaging*, 47(4):948–953, 2018.
- [6] Allison Park, Chris Chute, Pranav Rajpurkar, Joe Lou, Robyn L Ball, Katie Shpanskaya, Rashad Jabarkheil, Lily H Kim, Emily McKenna, Joe Tseng, et al. Deep learning-assisted diagnosis of cerebral aneurysms using the headnet model. *JAMA network open*, 2(6):e195600–e195600, 2019.
- [7] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. In *Advances in neural information processing systems*, pages 91–99, 2015.
- [8] Jose I Suarez, Robert W Tarr, and Warren R Selman. Aneurysmal subarachnoid hemorrhage. *New England Journal of Medicine*, 354(4):387–396, 2006.
- [9] Daiju Ueda, Akira Yamamoto, Masataka Nishimori, Taro Shimono, Satoshi Doishita, Akitoshi Shimazaki, Yutaka Katayama, Shinya Fukumoto, Antoine Choppin, Yuki Shimahara, et al. Deep learning for mr angiography: automated detection of cerebral aneurysms. *Radiology*, 290(1):187–194, 2019.
- [10] Xiaojiang Yang, Daniel J Blezek, Lionel TE Cheng, William J Ryan, David F Kallmes, and Bradley J Erickson. Computer-aided detection of intracranial aneurysms in mr angiography. *Journal of digital imaging*, 24(1):86–95, 2011.
- [11] Wentao Zhu, Chaochun Liu, Wei Fan, and Xiaohui Xie. Deeplung: Deep 3d dual path nets for automated pulmonary nodule detection and classification. In *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages

, ,

Ziheng Duan, Daniel Montes, Javier M. Romero, Ramon Gilberto Gonzalez, Yangsibo Huang, Dufan Wu, and Quanzheng Li

673–681. IEEE, 2018.