

A Two-step Surface-based 3D Deep Learning Pipeline for Segmentation of Intracranial Aneurysms

Xi Yang^{1,2} ✉, Ding Xia¹, Taichi Kin¹, Takeo Igarashi¹

© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract The exact shape of intracranial aneurysms is critical in medical diagnosis and surgical planning. While voxel-based deep learning frameworks have been proposed for this segmentation task, their performance remains limited. In this study, we offer a two-step surface-based deep learning pipeline that achieves significantly better results. Our proposed model takes a surface model of an entire set of principal brain arteries containing aneurysms as input and returns aneurysm surfaces as output. A user first generates a surface model by manually specifying multiple thresholds for time-of-flight magnetic resonance angiography images. The system then samples small surface fragments from the entire set of brain arteries and classifies the surface fragments according to whether aneurysms are present using a point-based deep learning network (PointNet++). Finally, the system applies surface segmentation (SO-Net) to surface fragments containing aneurysms. We conduct a direct comparison of the segmentation performance of our proposed surface-based framework and an existing voxel-based method by counting voxels: our framework achieves a much higher Dice similarity (72%) than the prior approach (46%).

Keywords Intracranial aneurysm segmentation, Point-based 3D deep learning, Medical image segmentation.

¹ The University of Tokyo, 7 Chome-3-1 Hongo, Bunkyo City, Tokyo, 113-8654, Japan. E-mail: X. Yang, earthyangxi@gmail.com, D. Xia, dingxia1995@gmail.com, T. Kin, tkin-ty@umin.ac.jp, T. Igarashi, takeo@acm.org.

² Jilin University, No.2699, Qianjin Road, Changchun City, Jilin Province, China. E-mail: earthyangxi@gmail.com.

Manuscript received: 2014-12-31; accepted: 2015-01-30.

Tab. 1 Comparison of related works and our method.

Method	Entire image (Practicality)	Surface-based (Effectiveness)
Park et al. [18]	✓	✗
Sichterman et al. [22]	✓	✗
Yang et al. [28]	✗	✓
Bizjak et al. [3]	✗	✓
Our method	✓	✓

1 Introduction

An intracranial aneurysm (IA) is a weakened or thinned portion of a blood vessel in the brain that bulges dangerously and fills up with blood. Bloated aneurysms compress the surrounding nerves and brain tissue, and have a high risk of rupture, which results in subarachnoid hemorrhage (SAH). The risk of such rupture is related to the size and form of the IA [7]. The practical surgical approach to prevent rupture is to clip the neck of the aneurism. Therefore, extracting the shape of aneurysms is a crucial aspect not only of IA diagnosis but also of preoperative examination to determine the position and posture of the necessary clips [1]. In current practice, this process requires manual identification by medical experts, taking several minutes per case. Clearly, automating this process is a very worthwhile venture. Furthermore, employing automation, we can also obtain large segmented datasets, which can open up new avenues for research aimed at gaining further insights into IA through statistical analysis.

Over the last decade, many extraction algorithms have been designed by calculating local geometric features [13, 17]; however, rule-based methods have high computational costs and time requirements, and their performance is limited because of the wide variety of aneurysm shapes. Meanwhile, deep learning techniques are becoming increasingly popular

in medical image processing; however, they are mostly used for classification and detection. Few prior research works have explored the application of deep learning methods to the segmentation of IAs, and their performance remains limited [17] (see Section 2).

This study builds on Yang et al.’s work the IntraA dataset [28], which was created for surface-based classification and segmentation of IAs, and reported the performance of existing neural network models on both tasks. However, in their work, the dataset and execution process were fully separated for both classification and segmentation tasks. Segmentation was evaluated only on manually sampled surface fragments containing aneurysms. This process is unrealistic in clinical practice. In addition, the per-fragment segmentation results were not integrated. Therefore, in this study, we present a complete processing pipeline for segmenting IAs, shown in Figure 1, which integrates deep learning and geometry processing techniques to achieve better results. Our proposed pipeline takes an entire intracranial vessel network model as input and returns IA fragments as output.

The main contributions of this study are as follows:

1. A complete pipeline using point-based 3D deep neural networks for aneurysm segmentation from entire medical images. The proposed pipeline uses automatic sampling and achieves state-of-the-art results comparable to segmentation based on manual sampling [28].
2. Adjusted algorithms for fragment sampling, refinement, and integration by geometry processing techniques to support the proposed pipeline.
3. A demonstration of the advantage of our two-step pipeline combining a classification step and a segmentation step by comparing it to a segmentation-only pipeline, and a direct comparison of our surface-based framework to a state-of-the-art voxel-based method, showing the superiority of our proposed framework.

In the field of computer vision in general, detection tasks usually involve determining the rough location and size of target objects, for example, a boundary box, whereas, in medical imaging, detection may indicate only the presence of target objects. The difficulties of both are lower than those of the segmentation task, which requires algorithms to predict the precise shape of the target objects.

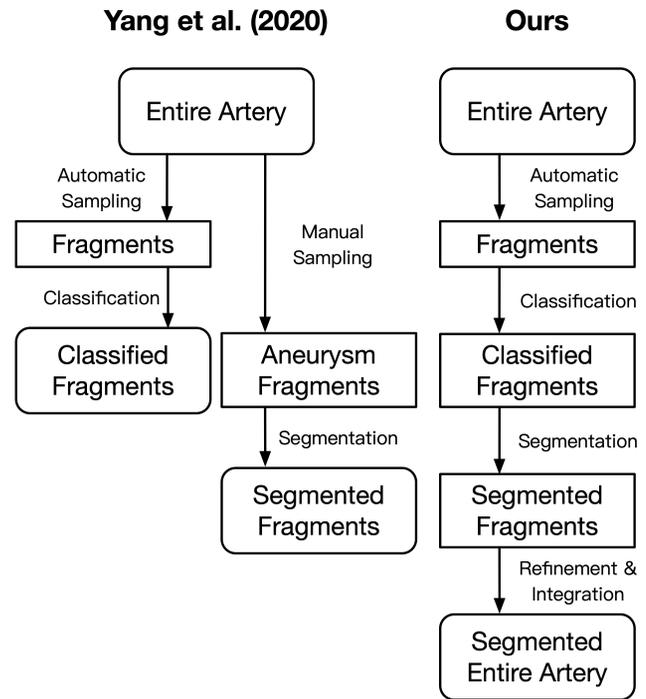


Fig. 1 Comparison of the pipelines of Yang et al. [28] and the proposed method.

2 Related work

2.1 Detection

Deep learning methods have been widely used as diagnosis aids to detect IAs. Nakao et al. [15] detected intracranial aneurysms in MRA images using a basic deep convolutional neural network. Ueda et al. [24] used ResNet-18 for an automated diagnosis of cerebral aneurysms from TOF MR angiography image data from several sources. Zhou et al. [29] proposed a transferable multi-model ensemble (MMEN) architecture to predict the possibility of aneurysms using a mesh model. This approach used 3D objects as input, but also still used 2D neural networks by conformal mapping.

2.2 Segmentation

Segmentation of IAs requires obtaining detailed location and shape information for aneurysms. Conventional approaches have used rule-based 2D or 3D shape analyses. For example, Nikravanshalmani et al. [16, 17] used a level set algorithm and a region growing based approach for semi-automatic segmentation of cerebral aneurysms from CTA images. Law et al. [12, 13] proposed an intensity-based algorithm to segment intracranial vessels and embedded aneurysms using multirange filters and local variances. Wang et al. [26] presented a multilevel segmentation

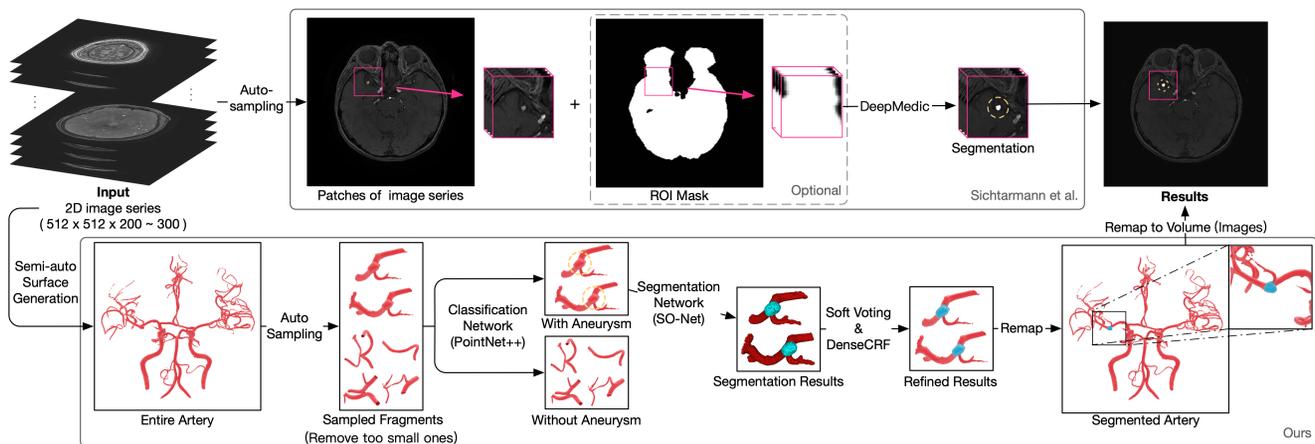


Fig. 2 Comparison of our proposed pipeline and a voxel-based method.

method based on the lattice Boltzmann method (LBM) and level sets with ellipses for accurate segmentation of intracranial aneurysms. Sulayman et al. [23] proposed a scheme for semi-automatic detection and segmentation of intracranial aneurysms. Dakua et al. [4] presented a PCA-based approach to segmenting the brain vasculature in low contrast cerebral blood vessels. Jerman et al. [8] proposed automated cutting plane (ACP) positioning, based on the detection of specific geometric descriptors of an aneurysm and its parent vasculature.

Recently, learning-based methods have become increasingly popular alongside the development of deep learning. However, few studies have focused on segmentation of IAs. Podgorsak et al. [19] claimed that segmenting IAs and the surrounding vasculature from digital subtraction angiography (DSA) images using a convolutional neural network was not inferior to manually identifying the contours of aneurysms. However, they extracted only 2D contours of the IAs. Park et al. [18] developed a neural network segmentation model called HeadXNet to generate voxel-by-voxel predictions of intracranial aneurysms on tomographic angiography (CTA) imaging to augment the performance of clinical intracranial aneurysm diagnosis. However, they evaluated their model based on the sensitivity, specificity, and accuracy of the entire image; these metrics cannot reflect actual segmentation performance in practice. Sichtermann et al. [22] applied a popular software-based method using a volume-based neural network, called DeepMedic [9], to segment IAs from MRA images. However, the quality of their results was poor (46% in DSC). Importantly, Yang et al. [28] and Bizjak et al. [3] made useful attempts to apply point-based networks to segment IAs. However,

segmentation was only performed for surface fragments that were manually labeled as containing aneurysms. This approach is unrealistic in clinical practice. In addition, Bizjak et al. [3] employed only the sensitivity of the entire input as an evaluation metric.

2.3 3D deep learning

3D surface models have several representations, including projected views, voxels or pixels, point clouds, and meshes. Voxel-based deep learning approaches are easy to implement using networks developed for 2D image tasks. However, point-based methods have shown great promise and improved performance compared to previous voxel-based methods in deep learning 3D shape analysis [5, 20, 25, 27]. In addition, using point-based rather than mesh-based methods [6] avoids arduous pre-processing steps, including model cleaning and manifold mesh construction. A problem with point-based methods is that they require surface models and cannot be directly applied to medical images. Therefore, we introduce an interactive surface reconstruction process before applying point-based classification and segmentation. We leverage Yang et al. [28]’s surface model data set for training and evaluation.

3 Proposed Pipeline

3.1 Background

Figure 2 compares our surface-based pipeline with a voxel-based method [22]. In the latter, the medical image is directly fed to a neural network, which affixes a label to each voxel indicating whether it is part of an aneurysm or not.

Our surface-based pipeline first interactively reconstructs a surface model of the entire set of

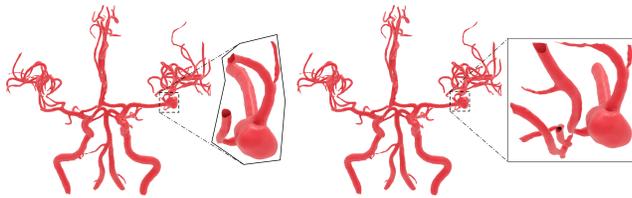


Fig. 3 In contrast to grid sampling, our sampling based on geodesic distance avoids involving noisy blood vessels.

principal brain arteries using a multiple threshold method. We then generate small samples along vessels within the entire model and perform surface-based classification (using PointNet++) on them. Finally, we perform surface-based segmentation (using SO-Net) on samples classified as containing aneurysms. The deep learning models were selected following [28]. To compare our results to those obtained by voxel-based methods, we voxelized the surface model of the segmented aneurysms into volumes using winding number [2].

3.2 Interactive reconstruction of surface models

We first obtain surface models of the principal brain arteries of patients using TOF-MRA image sets. We perform this semi-automatically using a software package (Amira 2019 by Thermo Fisher Scientific, MA, USA) based on a multi-threshold method [10]. Importantly, we focus on dealing with the brain regions surrounding aneurysms to ensure that the complete shape of the aneurysms was exhibited in the extracted 3D surface model, compared to the data in Intra [28]. In future, we envision that this process can be mostly or fully automated using a reconstruction network specifically designed for brain arteries.

3.3 Fragment sampling

Segmentation network does not work well if the entire model is input directly because aneurysms are a tiny portion of the entire model. Therefore, we first sample small fragments from the entire model. A typical method used to obtain small patches in medical images involves systematically sampling rectilinear boxes; however, it is difficult to obtain clean artery surfaces because surface boundaries are not aligned with canonical axes. Thus, we have designed an algorithm to sample along a surface, as shown in Figure 3. We set the size of the fragments to roughly cover an aneurysm of typical size in medical experts' experience. Then, we divide the 3D space into regular grid cells. From the center of each grid cell, the nearest

point on the surface model closer than a threshold (α) is selected as a starting point, while grid cells that do not have nearby surface model points are ignored. Finally, we collect the surface points around the starting points whose geodesic distance is less than a threshold (β). Note that this sampling is designed to cover the model with some overlap; uniform sampling is not a vital requirement of our proposed method.

3.4 Classification step

We use PointNet++ [21] to classify the fragments into two classes, those with and without aneurysms. Fragments with few points are discarded before classification. However, the number of fragments with aneurysms is still significantly fewer than those without aneurysms. Therefore, we use a weighted soft-max cross-entropy loss function to train the classification network to deal with the imbalance between the two classes. The purpose of the classification step is to reduce the number of candidate fragments fed to the segmentation network and improve its accuracy.

By design, we sacrifice some classification performance to obtain a better segmentation result. See Section 4.5 for a detailed discussion. The evaluation of our classification results does not match the accuracy of the detection task. Our sampling method allows one IA to be sampled in several fragments, and fragments with a tiny portion of IA may be misclassified, but the same IA may be detected from other fragments. Therefore, the real-world detection performance results are much better than the performance of the classifier itself.

We expected the classification step to reduce training and prediction time as well as data noise in the final segmentation result. To demonstrate this, we conducted an experiment comparing results between the proposed two-step pipeline and a pipeline without the classification step (see Section 4.6.2).

3.5 Segmentation step

We next feed fragments with aneurysms into the segmentation network, SO-Net [14]. Only a fraction of the original points are classified after segmentation because the point-based network uses random sampling to deal with input models with varying numbers of points. Thus, we perform segmentation with random sampling multiple times and assign labels to all points based on a voting criterion to enrich segmentation details. There is the possibility that a small number of points may fail to be sampled; we label them as arteries rather than aneurysms. However, these are few,

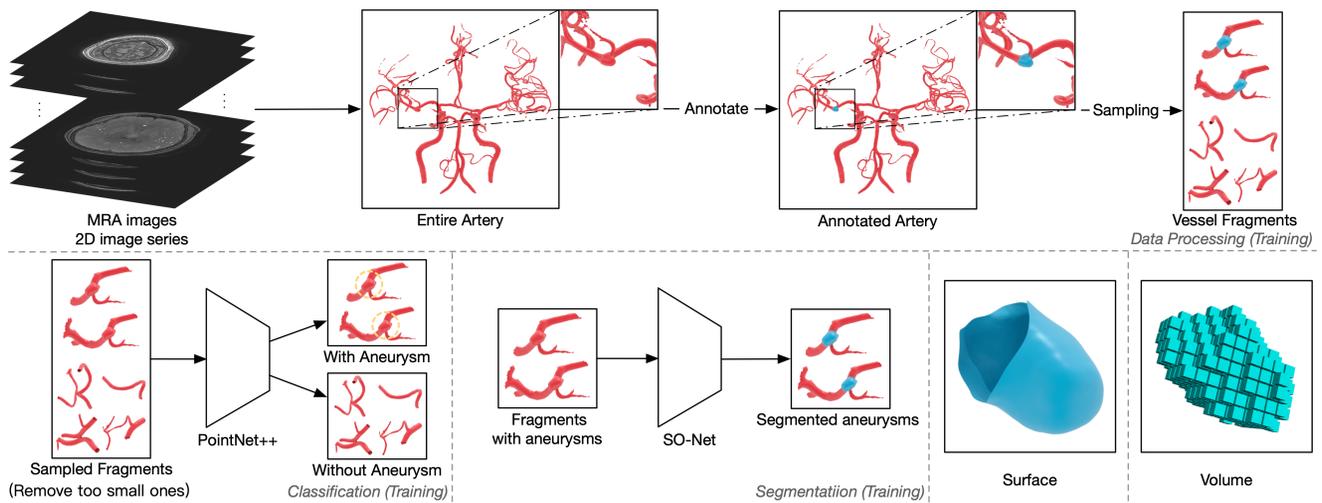


Fig. 4 Data-processing pipeline and algorithm.

and do not significantly affect the segmentation results. Next, we use a conditional random field (CRF) to refine each voting result, specifically DenseCRF [11]. Finally, the segmentation results of the individual fragments are remapped into the original entire model using a global ID for each point to obtain a complete segmentation result over the entire surface. Points in overlapping parts are commonly sampled twice, so we do not use majority voting. Points with multiple labels are marked as aneurysms if they had an aneurysm label.

3.6 Voxelization

We convert the results of our surface-based segmentation to a volume to perform a direct comparison with voxel-based methods. An example is shown in Figure 4. We first obtain a set of query points through uniform sampling using the same interval as for the MRA images. We then compute the winding number of each query point using the fast winding number method [2] to determine whether a given point is inside or outside an aneurysm. We set the winding number threshold to 0.5, as suggested in their study. This step is not necessary for clinical practice if segmentation results are required only on a surface model.

4 Experiments

4.1 Imaging

Various medical imaging techniques, such as computed tomography angiography (CTA), magnetic resonance angiography (MRA), and digital subtraction angiography (DSA) can be used to image the brain. DSA is the most sensitive method for diagnosing

intracranial IAs; however, it is invasive and time-consuming. Although CTA scans are efficient, distinguishing details of vessels and aneurysms using CTA remains difficult. TOF-MRA is less invasive and has high sensitivity for diagnosing IAs. Therefore, we decided upon TOF-MRA as a suitable technique for preoperative examination. However, our proposed pipeline is not affected by the type of medical image, as it is based on reconstructed surface models.

4.2 Dataset

We collected TOF-MRA image sets from 103 patients with 114 aneurysms. Each set contains at least one IA, and 180–300 2D images of resolution 512×512 sliced at 0.496 mm. Our dataset does not include small aneurysms (<3.00 mm), as our objective is to segment aneurysms requiring surgery. We calculated the size of each aneurysm based on maximum diameter. Figure 5 shows the distribution (Mean: 7.49 mm, SD: 2.72 mm; Range: 3.48–18.66 mm) of aneurysm sizes on our dataset. In terms of IA type, most were saccular aneurysms, and one fusiform aneurysm was included, but no dissecting aneurysm. Another special case was that we treated two aneurysms very close together as being one. We annotated the aneurysm portions on both the entire surface models of the brain arteries and on TOF-MRA images to generate ground truth for classification and segmentation for training neural networks. It took a total of three experts 21 working days to perform this task. We used five-fold cross-validation to conduct our experiments. A total of 103 sets were shuffled and divided into five subsets, of which four were used as training data, and one was used as testing data. This study design was approved by an

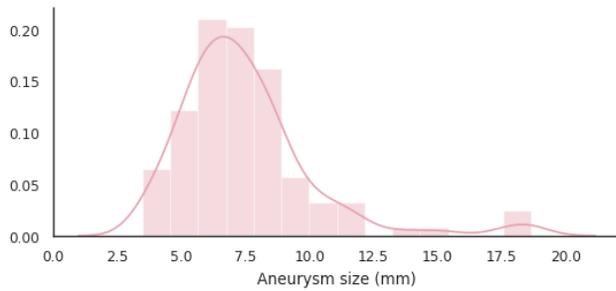


Fig. 5 Distribution of aneurysm sizes on our dataset. Mean: 7.49 mm, SD: 2.72 mm; Range: 3.48–18.66 mm.

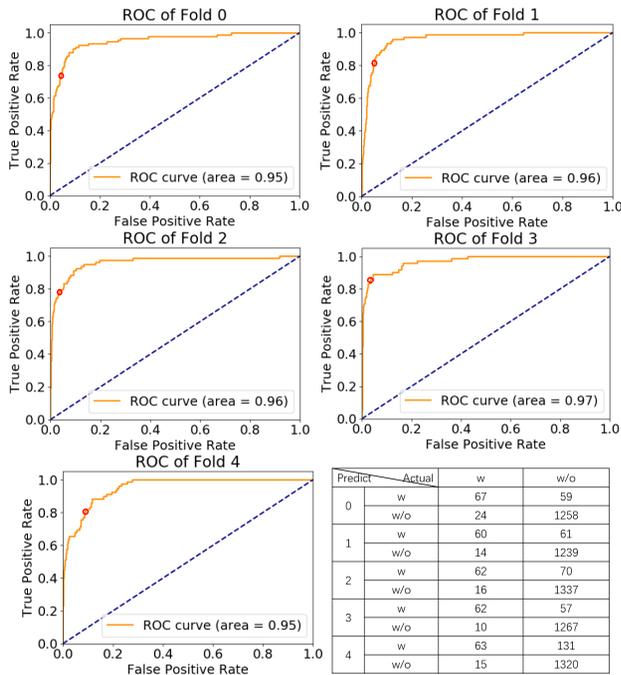


Fig. 6 ROC curves and confusion matrices of five trained classification networks.

appropriate ethics review board.

4.3 Evaluation metrics

Several evaluation metrics were employed to assess the models in different tasks. Accuracy, recall, and sensitivity are typically used to evaluate performance on classification tasks. For segmentation task, Dice similarity coefficient (DSC) or Intersection over Union (IoU) is employed to indicate the prediction of the target region. Sensitivity for the entire input can easily be high because the target region may be tiny. Moreover, in this situation, the overall statistics may conflict with the part-wise statistics due to Simpson's paradox.

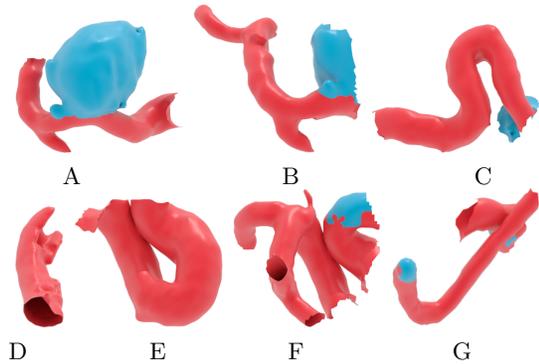


Fig. 7 Fragment examples. A has a complete aneurysm. B and C partly overlap with A, but only include a part of the aneurysm. D, E, F, and G lack aneurysms, but are misclassified. The original data for D is noisy, while E, F, and G have shapes very similar to a part of the aneurysm.

4.4 Implementation details

The experiments were performed on a PC with a GeForce RTX 2080Ti GPU. During data preprocessing, the normal vector of each point was estimated using the original surface model. We also recorded the point index of the entire model as a global ID on the sampled fragments for each point to improve the efficiency of voting. We set the sampling thresholds to $\alpha = 15$, $\beta = 1.5\alpha$, and samples with fewer than 500 points were removed. We automatically generated 7192 vessel fragments from the 103 entire models; 392 fragments contained aneurysms.

The training hyper-parameters were set as follows. For the classification network, the number of sample points for each fragment was 1024. The weights of the loss function were determined according to the number of fragments. We trained the network using 251 epochs and a batch size of 8. The classification results were predicted by setting a discrimination threshold of 0.23. For the segmentation network, the number of sample points was 2048. We trained the network for up to 401 epochs, with a batch size of 12. For each network, we used the Adam optimizer with a learning rate of 10^{-3} .

4.5 Results

The receiver operating characteristic (ROC) curve and confusion matrix of each classification network are shown in Figure 6. It can be observed that all areas of the ROC curves are higher than 0.95, demonstrating that the trained classification networks generalize well. The sensitivities to the aneurysm class of the five networks were 73.63%, 81.08%, 79.49%, 86.11%, and 80.77%, respectively. This shows that we can precisely detect fragments with IA portions. By analyzing the confusion matrices, we observed that only a few

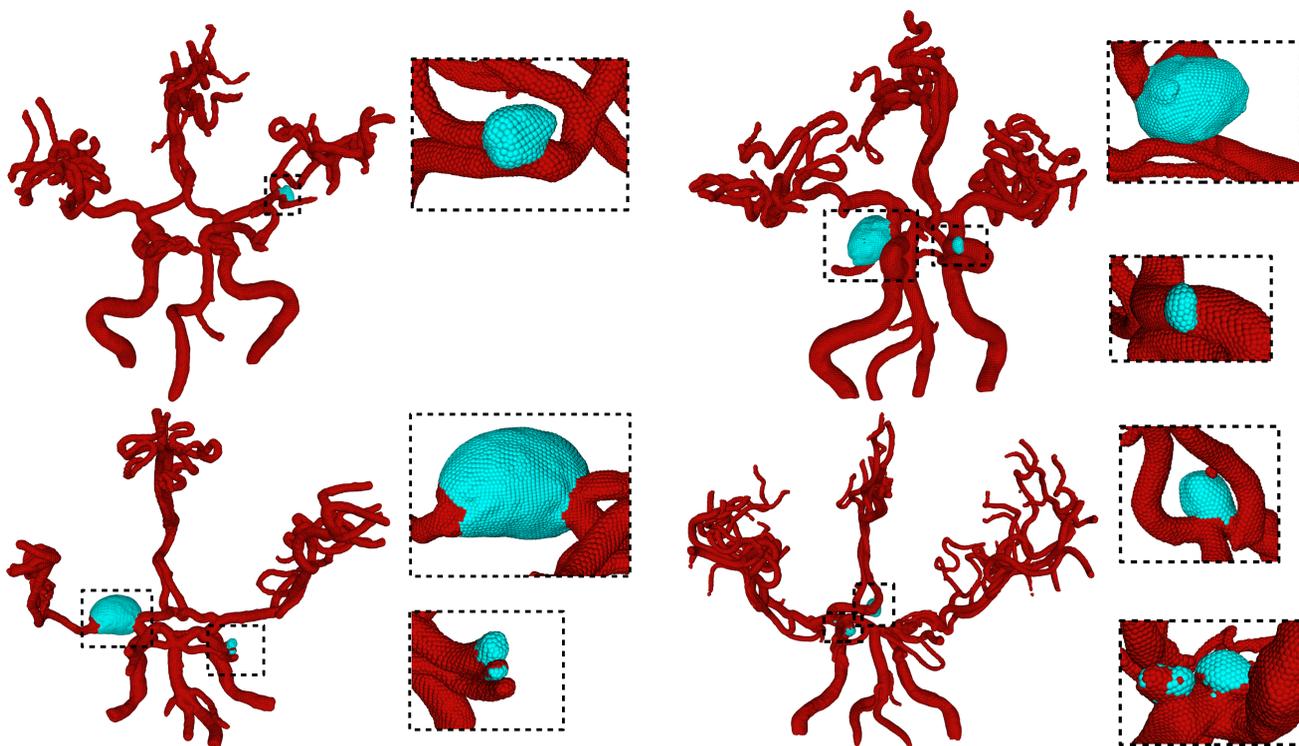


Fig. 8 Examples of final segmentation results (rendered as points); enlarged figures may be captured from a different viewpoint to show the aneurysm shape. Above left: saccular aneurysms. The proposed pipeline obtained a perfect segmentation. Above right: saccular aneurysms. The aneurysm was segmented clearly without the impact of close blood vessels (see top enlargement). A potential aneurysm was also found, which was not annotated by experts (see bottom enlargement). Below left: fusiform aneurysms. The aneurysm was well segmented (see top enlargement). The ends of normal blood vessels were segmented incorrectly into aneurysms (see bottom enlargement). Below right: double saccular aneurysm: multiple aneurysms in one case. Two close aneurysms were annotated as one, and our segmentation did not achieve the same result as Yang et al. [28] as our training data included much more complex shapes.

fragments with aneurysms were misclassified because they had tiny aneurysms or contained only a small part of the aneurysm. However, 100% sensitivity is not necessary for our classification network because the sampled fragments overlap, as shown in Figure 7. Using our sampling algorithm, 80% sensitivity of the classifier does not mean that 1/5 of the aneurysms are already missed before the segmentation step. In fact, in this experiment, only 5 out of 114 IAs were missing. The real sensitivity to IAs is satisfactory, as shown in Table 2. However, some fragments lacking IAs were not classified correctly because the original data contained significant noise and the fragments had a very similar shape to a small part of the full aneurysms. These misclassified cases were also sometimes difficult for segmentation networks; they did not have a significant impact on the final segmentation results.

We added the classification step before segmentation to filter out the majority of fragments that did not contain aneurysms. This helped the segmentation

Tab. 2 Sensitivities for aneurysm fragments and whole IAs (%).

Fold	0	1	2	3	4
Fragments	73.63	81.08	79.49	86.11	80.77
Aneurysms	95.24	100.00	100.00	95.00	85.71

network to avoid predicting false positive results on non-aneurysm regions, as well as improving the balance between fragments with and without aneurysms, leading to better final segmentation accuracy. The benefit of the classification process is shown in Section 4.6.2, in which we compare the proposed two-step pipeline and a segmentation-only pipeline.

Four examples of the final segmentation results are presented in Figure 8. In this figure, we show the entire 3D surface models and enlarged important parts marked by black dotted boxes. Segmented aneurysms are shown in cyan, and other normal blood vessels are red. We can see that our proposed pipeline obtained satisfactory segmentation results for various

Tab. 3 Comparison of segmentation results between segmentation only, and our two-step design, on surface DSCs (%).

	Segmentation only		Two-step	
	Mean	STD	Mean	STD
Overall	31.43	16.92	74.74	26.47
Fold 0	36.10	15.63	76.73	25.83
Fold 1	38.21	19.28	80.18	17.75
Fold 2	33.74	18.81	80.66	20.33
Fold 3	26.67	11.78	73.78	25.67
Fold 4	22.52	14.00	62.54	36.62

shapes and sizes of saccular aneurysms. We also found that unannotated potential aneurysms could also be segmented. However, a few normal vessel ends were segmented as aneurysms because their shapes were extremely similar to IAs. In addition, our networks predicted a suitable segmentation result for the fusiform aneurysm, even though they were trained only on saccular aneurysms. This demonstrates the excellent generalizability of point-based deep learning models. Furthermore, our proposed pipeline obtained superb segmentation of multiple aneurysms in one case. A more detailed statistical analysis of our final segmentation results is provided in the comparison experiments 4.6.

4.6 Comparisons

4.6.1 3D U-Net

We first applied the original 3D U-Net to our data; however, the network did not predict any segmentation result because the aneurysms were too small compared to the entire image. This demonstrates the difficulty of this segmentation task.

4.6.2 Segmentation only

To indicate the utility of the two-step design, we performed an ablation study by removing the classification step from our proposed pipeline: we fed all fragments to the pre-trained segmentation network to segment the aneurysm regions in the fragments. We compared the final results of this segmentation-only pipeline with those of the two-step pipeline on the entire artery surface models, as shown in Table 3. The results demonstrate that the classification step greatly reduces noise and improves the final segmentation result.

4.6.3 DeepMedic

We also applied the method described in [22] to our dataset for comparison. Four preprocessing approaches *A*, *B*, *C*, and *D* were applied in their study. *A* was only been applied as a necessary step in DeepMedic,

while *B*, *C*, and *D*, were used as additional masks for the skull-stripping of the TOF-MRA images. *B* generated masks with a fixed threshold, *C* used a manual threshold for skull-stripping of each sample, and *D* added N4 bias correction to the result of *C*. By analyzing the segmentation results, we found that skull-stripping could improve performance; however, there was not much difference between the results of *B*, *C*, and *D*. Therefore, we compared our method with *B*, which has the highest reproducibility. We used BET2 to obtain masks for skull-stripping using a fixed threshold of 0.2. The input of the TOF-MRA images was resized to 256×256 by down-sampling, to meet the requirements of DeepMedic.

The DSC of the aneurysm parts was employed to evaluate the segmentation results. A comparison of the final segmentation results is shown in Figure 9 and Table 4. The performance of the voxel-based method was comparable to that reported in the original paper ([22]). Our surface-based method obtained much better segmentation results than the voxel-based method on most of the data. However, a few samples with tiny aneurysms were challenging both for the voxel-based method and for ours.

4.6.4 DeepMedic with surface mask (DeepMedic S)

To directly compare performance between voxel-based and points-based networks, we generated artery region masks by converting our entire surface models into solid models and then mapping them back to the original MRA images, as shown in Figure 10. Having done so, the voxel-based network obtained the same region of interest (ROI) as point-based models. We can see that the segmentation results were improved compared to the model trained with skull-stripping masks. However, the performance was still worse than for our surface-based method. This experiment shows that point-based networks can learn more accurate topological and geometric shape information than voxel-based models.

5 Discussion

5.1 Impact of classification accuracy

To quantify the impact of the classification accuracy on the segmentation results, we show a comparison for different discrimination thresholds for Fold 0 in Figure 11. Even for an extremely small threshold, the classifier performs superbly compared to the method of segmentation only. When adjusting the discrimination threshold from small to large, the classification model

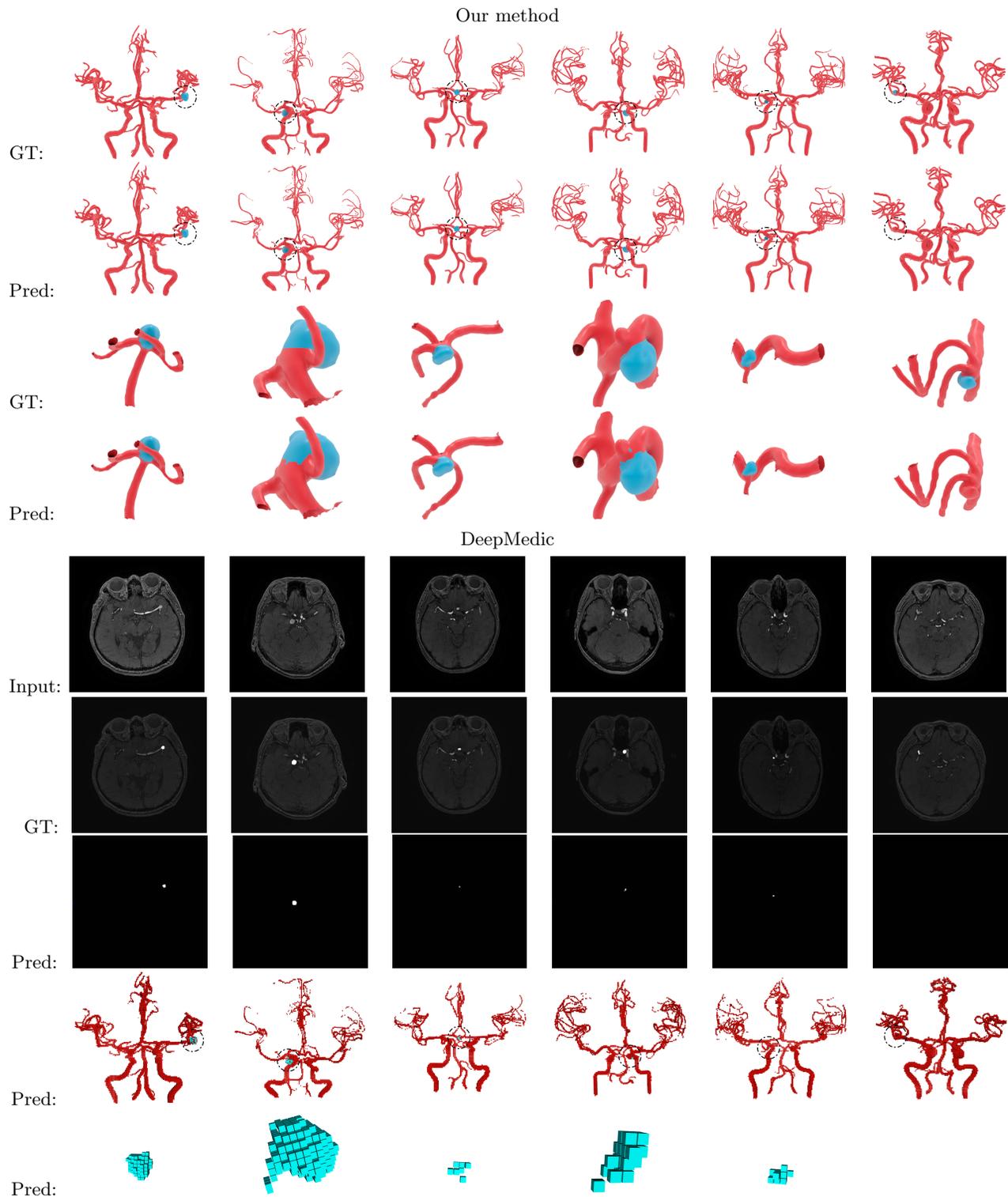


Fig. 9 Comparison of segmentation results. Left to right: both our method and DeepMedic yielded high segmentation accuracy on the first two examples. Our method yielded significantly better results than DeepMedic on the next three examples. Both our method and DeepMedic failed to obtain the parts with IA in the final example, as the fragment with the IA was filtered out by our classification network. We also show the results predicted by DeepMedic as volumes (last two rows).

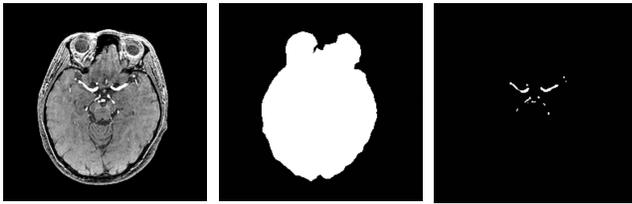


Fig. 10 Left: an original MRA image slice. Center: the corresponding ROI mask for skull-stripping generated by BET2. Right: the corresponding ROI surface mask generated by our surface model to provide the same input region for the DeepMedic model.

Tab. 4 Comparison of segmentation results on DSCs (%). U-Net failed to provide segmentation results.

	Ours (Surface-based)		DeepMedic S (Voxel-based)		DeepMedic B (Voxel-based)	
	Mean	STD	Mean	STD	Mean	STD
Overall	71.79	29.91	52.55	31.37	45.90	31.00
Fold 0	73.10	32.55	59.38	28.20	56.56	28.77
Fold 1	75.72	27.23	61.70	30.73	47.40	30.25
Fold 2	73.81	30.43	47.96	32.34	38.86	31.57
Fold 3	72.37	30.50	53.97	27.58	45.51	30.88
Fold 4	64.17	34.85	40.26	35.27	41.20	33.22

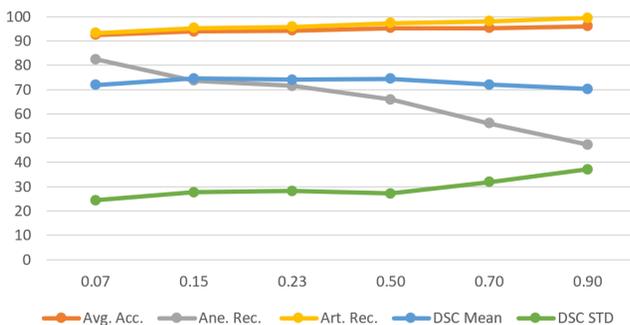


Fig. 11 Comparison with different discrimination thresholds for Fold 0. (Due to the random sampling of the input points, the results may fluctuate slightly each time.)

tends to classify fragments as arteries. The average accuracy improves as the number of artery fragments is much higher than of aneurysm fragments, while the recall of aneurysms becomes worse. However, the final results of segmentation perform stably due to the designs of the methods of fragments sampling, refinement, and integration. It shows that our proposed pipeline is convenient for use in clinical practice.

5.2 Limitations

Figure 12 shows two examples comparing 3D surface models of aneurysms created automatically and manually. Our current pipeline requires manual effort by medical experts to obtain surface models of

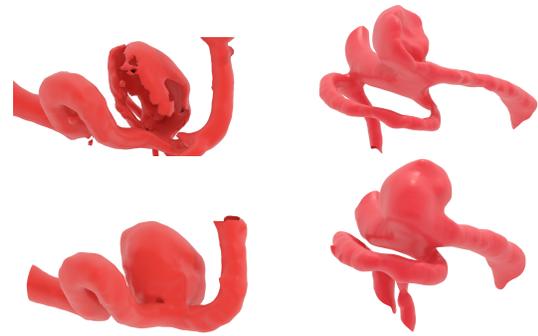


Fig. 12 Above: complete 3D surface models of aneurysms are difficult to reconstructed automatically. Below: corresponding models manually created by experts.

intracranial artery networks. Thus, a possible criticism of our method is that this process severely limits its practical value. There are three reasons why we still believe that our method has significant practical value. Firstly, neurosurgeons are presently already constructing surface models regularly in practice for preoperative examinations. Thus, in this context, we can assume that the surface model is already available. Secondly, construction of the surface model is mostly performed through simple thresholding [10]. An expert manually sets a threshold, and voxels with intensities higher than the threshold are automatically extracted. In this process, the expert does not pay attention to the details of individual aneurysms. Aneurysms need to be carefully segmented manually using surface editing tools in current practice, and we expect automation of this process to be highly appreciated. Finally, we expect that, with advances in deep learning methods, surface extraction will become largely or even fully automatic in future. Consequently, the entire process may be fully automated, which has the potential to significantly impact the field.

Our experimental results show a baseline performance for the proposed framework. We believe that our results can be further improved significantly by adjusting hyper-parameters.

5.3 Benefits to clinical practice

Clipping surgery of intracranial aneurysms uses clips at the boundary between aneurysm and parent artery. The number of clips and their position and posture determine the artery deformation and blood flow, impacting the effect of the surgery. Discussion and simulation of these issues are critical in both surgery planning and the education of students. 3D visualization provides an intuitive way of observing and simulating the real brain of the patient. The exact

aneurysm boundary on 3D surface models of brain arteries is of great help in the automatic computation of positions and postures of clips and the simulation of blood flow.

6 Conclusions

In this study, we proposed a new surface-based framework for the segmentation of intracranial aneurysms from TOF-MRA images. Our framework applied a two-step design, based on classification then segmentation, using state-of-the-art point-based deep learning networks. We also designed sampling and refinement methods for the IA segmentation task. The segmentation results show that our framework significantly outperforms an existing voxel-based method. Surface-based methods are as yet not prevalent in medical diagnosis and surgical planning. Our results show that surface-based methods can be a reliable alternative to popular voxel-based methods, and we hope this work will inspire further research efforts in this direction in other medical application domains.

Statement

Conflict of Interest: The authors declare that they have no conflict of interest.

Acknowledgements

This research was supported by AMED under Grant Number JP18he1602001.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- [1] A. Alaraj, C. J. Luciano, D. P. Bailey, A. Elsenousi, B. Z. Roitberg, A. Bernardo, P. P. Banerjee, and F. T. Charbel. Virtual reality cerebral aneurysm clipping simulation with real-time haptic feedback. *Operative Neurosurgery*, 11(1):52–58, 2015.
- [2] G. Barill, N. G. Dickson, R. Schmidt, D. I. Levin, and A. Jacobson. Fast winding numbers for soups and clouds. *ACM Transactions on Graphics (TOG)*, 37(4):1–12, 2018.
- [3] Ž. Bizjak, B. Likar, F. Pernuš, and Ž. Špiclin. Vascular surface segmentation for intracranial aneurysm isolation and quantification. *arXiv preprint arXiv:2005.14449*, 2020.
- [4] S. P. Daku, J. Abinshed, and A. Al-Ansari. A pca-based approach for brain aneurysm segmentation. *Multidimensional Systems and Signal Processing*, 29(1):257–277, 2018.
- [5] M.-H. Guo, J.-X. Cai, Z.-N. Liu, T.-J. Mu, R. R. Martin, and S.-M. Hu. Pct: Point cloud transformer. *Computational Visual Media*, 7(2):187–199, 2021.
- [6] R. Hanocka, A. Hertz, N. Fish, R. Giryes, S. Fleishman, and D. Cohen-Or. Meshcnn: a network with an edge. *ACM Transactions on Graphics (TOG)*, 38(4):1–12, 2019.
- [7] U. J. Investigators. The natural course of unruptured cerebral aneurysms in a japanese cohort. *New England Journal of Medicine*, 366(26):2474–2482, 2012.
- [8] T. Jerman, A. Chien, F. Pernuš, B. Likar, and Ž. Špiclin. Automated cutting plane positioning for intracranial aneurysm quantification. *IEEE Transactions on Biomedical Engineering*, 67(2):577–587, 2019.
- [9] 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.
- [10] T. Kin, H. Nakatomi, M. Shojima, M. Tanaka, K. Ino, H. Mori, A. Kunimatsu, H. Oyama, and N. Saito. A new strategic neurosurgical planning tool for brainstem cavernous malformations using interactive computer graphics with multimodal fusion images. *Journal of neurosurgery*, 117(1):78–88, 2012.
- [11] P. Krähenbühl and V. Koltun. Efficient inference in fully connected crfs with gaussian edge potentials. In *Advances in neural information processing systems*, pages 109–117, 2011.
- [12] M. W. Law and A. C. Chung. Vessel and intracranial aneurysm segmentation using multi-range filters and local variances. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 866–874. Springer, 2007.
- [13] M. W. Law and A. C. Chung. Segmentation of intracranial vessels and aneurysms in phase contrast magnetic resonance angiography using multirange filters and local variances. *IEEE Transactions on image processing*, 22(3):845–859, 2012.
- [14] J. Li, B. M. Chen, and G. Hee Lee. So-net: Self-organizing network for point cloud analysis. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 9397–9406, 2018.
- [15] T. Nakao, S. Hanaoka, Y. Nomura, I. Sato, M. Nemoto, S. Miki, E. Maeda, T. Yoshikawa, N. Hayashi, and O. Abe. Deep neural network-based computer-assisted detection of cerebral aneurysms in mr angiography. *Journal of Magnetic Resonance Imaging*, 47(4):948–953, 2018.
- [16] A. Nikravanshalmani, M. Karamimohammdi, and

- J. Dehmshki. Segmentation and separation of cerebral aneurysms: A multi-phase approach. In *2013 8th International Symposium on Image and Signal Processing and Analysis (ISPA)*, pages 505–510. IEEE, 2013.
- [17] A. Nikravanshalmani, S. D. Qanadli, T. J. Ellis, M. Crocker, Y. Ebrahimdoost, M. Karamimohammadi, and J. Dehmshki. Three-dimensional semi-automatic segmentation of intracranial aneurysms in cta. In *Proceedings of the 10th IEEE International Conference on Information Technology and Applications in Biomedicine*, pages 1–4. IEEE, 2010.
- [18] A. Park, C. Chute, P. Rajpurkar, J. Lou, R. L. Ball, K. Shpanskaya, R. Jabarkheel, L. H. Kim, E. McKenna, J. Tseng, et al. Deep learning–assisted diagnosis of cerebral aneurysms using the headxnet model. *JAMA network open*, 2(6):e195600–e195600, 2019.
- [19] A. R. Podgorsak, R. A. Rava, M. M. S. Bhurwani, A. R. Chandra, J. M. Davies, A. H. Siddiqui, and C. N. Ionita. Automatic radiomic feature extraction using deep learning for angiographic parametric imaging of intracranial aneurysms. *Journal of neurointerventional surgery*, 2019.
- [20] C. R. Qi, H. Su, K. Mo, and L. J. Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 652–660, 2017.
- [21] C. R. Qi, L. Yi, H. Su, and L. J. Guibas. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. In *Advances in neural information processing systems*, pages 5099–5108, 2017.
- [22] T. Sichter, A. Faron, R. Sijben, N. Teichert, J. Freiherr, and M. Wiesmann. Deep learning–based detection of intracranial aneurysms in 3d tof-mra. *American Journal of Neuroradiology*, 40(1):25–32, 2019.
- [23] N. Sulayman, M. Al-Mawaldi, and Q. Kanafani. Semi-automatic detection and segmentation algorithm of saccular aneurysms in 2d cerebral dsa images. *The Egyptian Journal of Radiology and Nuclear Medicine*, 47(3):859–865, 2016.
- [24] D. Ueda, A. Yamamoto, M. Nishimori, T. Shimono, S. Doishita, A. Shimazaki, Y. Katayama, S. Fukumoto, A. Choppin, Y. Shimahara, et al. Deep learning for mr angiography: automated detection of cerebral aneurysms. *Radiology*, 290(1):187–194, 2019.
- [25] Y. Wang, Y. Sun, Z. Liu, S. E. Sarma, M. M. Bronstein, and J. M. Solomon. Dynamic graph cnn for learning on point clouds. *Acm Transactions On Graphics (tog)*, 38(5):1–12, 2019.
- [26] Y. Wang, Y. Zhang, L. Navarro, O. F. Eker, R. A. Corredor Jerez, Y. Chen, Y. Zhu, and G. Courbebaisse. Multilevel segmentation of intracranial aneurysms in ct angiography images. *Medical physics*, 43(4):1777–1786, 2016.
- [27] W. Wu, Z. Qi, and L. Fuxin. Pointconv: Deep convolutional networks on 3d point clouds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9621–9630, 2019.
- [28] X. Yang, D. Xia, T. Kin, and T. Igarashi. Intra: 3d intracranial aneurysm dataset for deep learning. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.
- [29] M. Zhou, X. Wang, Z. Wu, J. M. Pozo, and A. F. Frangi. Intracranial aneurysm detection from 3d vascular mesh models with ensemble deep learning. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 243–252. Springer, 2019.



Xi Yang is currently an associate professor in the School of Artificial Intelligence at Jilin University. He was a project assistant professor in the Graduate School of Information Science and Technology at The University of Tokyo. He received a Ph.D. degree in the Graduate School of Engineering at Iwate University. His research interests include geometric processing, visualization, deep learning, and user interfaces.



Ding Xia is a research assistant in the Graduate School of Information Science and Technology at The University of Tokyo. He received his master’s degree from the School of Automation Science and Engineering in South China University of Engineering. His research interests are in computer vision, visualization and medical analysis.



Taichi Kin is an assistant professor in Department of Neurosurgery at The University of Tokyo Hospital. He received his M.D. and Ph.D. degrees from the Department of Neurosurgery at The University of Tokyo Graduate School of Medicine in 2011. His research interests are in medical 3DCG and virtual reality surgical simulation.



Takeo Igarashi is a professor in the Computer Science Department at the University of Tokyo. He received his Ph.D. degree from the Information Engineering Department at the University of Tokyo in 2000. His research interest is in user interfaces in general and his current focus is on interaction techniques for 3D graphics.