Comparative Study of Vessel Detection Methods for Contrast Enhanced Computed Tomography: Effects of Convolutional Neural Network Architecture and Patch Size

Yuki SUZUKI,* Masatoshi HORI,** Shoji KIDO,* Yoshito OTAKE,*** Mariko ONO,*** Noriyuki TOMIYAMA,† Yoshinobu SATO***, #

Abstract Segmenting blood vessels is an important step in a wide variety of tasks in medical image analysis. Patch-based convolutional neural networks (CNNs) are often used for vascular detection, but the impact of patch size and choice of CNN architecture have not been addressed in detail in previous studies. In this study, we aim to investigate the impact of patch size and CNN architecture on the accuracy of vascular detection from contrast enhanced computed tomography (CT). We targeted the renal arteries as the primary focus of detection. We conducted experiments using contrast enhanced abdominal CT data of 30 cases. For the experiments, arteries in pre-defined regions of interest were manually labeled to build a dataset of input CT images and ground truth labels. We repeated the experiments with four patch sizes and two patch-based 3D CNN architectures (U-Net-like model and a simple sequential model) and evaluated the differences. Moreover, a Hessian-based line enhancing method was included in the evaluation to compare this non-deep learning method with the CNNs. The experimental results showed that patch size had a significant impact on detection accuracy. U-Net-like model showed peak accuracy at a certain patch size, unlike the sequential model that plateaued at large patch sizes. Although both CNNs outperformed Hessian-based line enhancement by a large margin, Hessian-based line enhancement achieved good recall when enhancing vessel structures not included in the CNN training. Our experiments show that different network architectures have different characteristics regarding their response to various patch sizes and vessel structures unseen during training.

Keywords: segmentation, blood vessel, renal artery, abdominal CT, deep learning, CNN.

Adv Biomed Eng. 10: pp. 138–149, 2021.

1. Introduction

Blood vessels play a vital role in a wide variety of medical tasks such as surgical planning and diagnosis. For example, in organ transplants, a patient’s vessel structure is one of the factors that determines the operability of the patient. In a partial resection operation (such as partial hepatectomy and nephrectomy), optimal resection is designed by estimating blood vessel-dominant regions based on the vascular structure.

Understanding vascular structure is a challenging task because the structure has an elongated tubular structure with curves and branchings. Therefore, assisting humans with vessel recognition has been a main topic of the studies in the medical image analysis field. Common approaches in these studies are vessel enhancement [1, 2] and automated vessel segmentation [3–5]. Vessel structures are extracted from images in a broad range of image protocols and body parts, such as fundus images for diabetic retinopathy, cardiac CT for coronary stenosis, and head CT for aneurysms [6].

Convolutional neural networks (CNNs) have been widely used in semantic segmentation since AlexNet [7] won an image classification contest by a large margin, and a large number of methods and tools have been developed. In medical image analysis, researchers have be-
gun to use CNNs for a variety of tasks such as classification and detection. A fully convolutional network [8] is typically used for segmentation tasks. The semantic segmentation of blood vessels is no exception to this trend, and a number of new techniques specifically for segmenting tubular structures have been proposed [9–13].

Patch-based networks are commonly used in the semantic segmentation of medical images. This is because the target object is often localized in a limited region of interest (ROI) and because 3D volumetric images are too big to fit in GPU memory, which is commonly used in the training of CNNs. Although a CNN is capable of learning its weight parameters through training, there are hyper-parameters for the network that are not optimized by the training process. In patch-based networks, the patch size is one hyper-parameter that is often heuristically determined. Whereas patch-based CNN is often used for vascular segmentation, the impact of patch size has not been addressed in detail so far. In many previous studies, a fixed patch size was used ([14–16] to name a few). Although a couple of papers [17, 18] have mentioned patch size in their studies, more detailed study on the impact of patch size is needed to achieve better understanding of vascular segmentation using patch-based CNN.

Oda et al. [17] proposed a method that segments abdominal arteries in contrast enhanced CT and reported the segmentation accuracies for different patch sizes. A limitation in their work is that they did not test a 3D network. They used 2.5D network that incorporates three planes; namely the axial, sagittal, and coronal planes. In their work, the 2.5D CNN performed better than a simple 2D CNN. However, the development of a 3D CNN that is capable of capturing a 3D structure remained unaddressed.

Yang et al. [18] proposed a method that segments catheters in 3D ultrasound images using techniques such as focal loss and dense sampling. In their experiments, a 3D U-Net was used, and they reported that changing patch size changed the segmentation accuracy. The limitation of in their work regarding patch size in the context of vascular segmentation is as follows. First, their target is a catheter, which has a tubular structure similar to that of blood vessels. However, unlike vessels, a catheter does not have branches nor large changes in diameter. Second, because they used 3D ultrasound, image characteristics such as field of view and background organs are different from those in other image modalities such as CT and MR.

The task of this work was to detect abdominal arteries in contrast enhanced 3D CT images using patch-based 3D CNN. Our contribution was three-fold. First, we studied the impact of the patch size and network architecture in the task of abdominal artery detection in contrast enhanced 3D CT through experiments. We determined the best-performing CNN architecture and its optimal patch size as well as how different CNN architectures were affected by patch size. Second, we studied the behavior of CNN when detecting categories of vessels that were not included in the training data. Third, we compared a typical non-deep learning method; namely the Hessian-based multiscale line enhancement method [1], with CNN-based methods.

2. Materials and Methods

2.1 Materials

In this work, we used 30 cases of contrast enhanced abdominal CT scanned at Osaka University Hospital. The average voxel size of the images was approximately $0.68 \times 0.68 \times 0.63$ mm$^3$. The ground truth data of vessel regions were prepared for three types of ROIs, which were defined in individual cases to observe the differences among methods in different regions. The three types of ROIs, as illustrated in Fig. 1, were as follows.

- Renal ROI (red rectangles in Fig. 1) was a bounding box that circumscribed one kidney and the arterial tree stemming from the aorta. Two (left and right) renal ROIs were defined in each patient. The average size of each renal ROI was $142 \times 121 \times 171$ voxels and the standard deviation was $18.2 \times 7.41 \times 20.1$

![Fig. 1](image)

Fig. 1 Placement of the three types of ROIs. Colored rectangles indicate ROIs (red: renal ROI; green: lung ROI; blue: spine ROI). (a): Maximum intensity projection (MIP) in the coronal plane. (b), (c), and (d): Image slices with renal, lung, and spine ROIs, respectively.
Lung ROI was a cube of $32 \times 32 \times 32$ voxels placed at the bottom of the right lung. No organs except for lung tissues and vessels were in the ROI.

Spine ROI was a cube of $32 \times 32 \times 32$ voxels placed next to the 12th thoracic vertebra with an intercostal artery at the center of the ROI. A portion of spinal bone tissue was included in the ROI.

The main ROI type was the renal ROI, which was used for both training of the CNNs and evaluation using cross-validation. The other two types of ROIs were used only for evaluation. The CT images from our contrast enhanced abdominal CT scan have four contrast phases (non-contrasted, early arterial, late arterial, and venous phases). The early arterial phase has the best contrast between arteries and other non-arterial regions such as veins in the renal ROI. Additionally, vessels in the lung and spine ROIs are not very sensitive to the contrast phase because non-vascular regions in these ROIs are primarily air and bone, respectively. Therefore we used early arterial phase for all three types of ROIs. All arterial regions in the renal and spine ROIs and vessels in the lung ROIs were manually labeled by medical image researchers under supervision of an experienced radiologist.

2.2 Methods
In this work, the following three methods were evaluated.

1. A U-Net [19]-like network (UN): A CNN that uses a U-Net-like architecture, which is a network architecture commonly used in semantic segmentation of medical images.

2. Sequential network (SN): A CNN that uses a simple sequential network without skip connections.

3. Hessian-based line enhancement method (HM) [1].

The UN architecture used in the experiments is shown in Fig. 2(a). Our UN architecture is a 3D network using 3D convolution layers and 3D max pooling layers. The network has two max pooling layers for downsampling in the encoding part and two upsampling layers in the decoder part. The number of convolutional kernels are doubled at each convolutional layer before max pooling to avoid bottlenecks [20]. Batch normalization layers (BNs)

![Network Architectures](image-url)
are inserted between the convolutional layers and their activations.

The SN is a simple network shown in Fig. 2(b) that only consists of convolutional layers. Unlike UN, there are no skip-connections or downsampling/upsampling layers. The SN has the same number of convolutional layers as the UN.

We used a patch-based method to detect arteries with 3D CNNs, in which each patch is cubic. This is because the entire 3D volumetric images are too large to fit in the memory available in a GPU, which is essential for training the CNNs. Furthermore, arterial regions and non-arterial regions are imbalanced in the entire volume and it is easier to rectify this volumetric imbalance using a patch-based method by extracting more patches that contain arteries from the training images than patches that do not contain arteries. With a patch-based method, an input volume is divided into 3D patches to be fed to the network and the output patches are stitched together to reconstruct a full-sized output volume.

In our experiments, 3D patches were sampled using a sliding window with no overlap between output patches. Therefore, the step size of the sliding window is simply $s_n$, where $s_n$ is the output patch size. In addition to the patches sampled by the sliding window, patches that contained the arterial center line were also sampled to mitigate the volumetric imbalance. A 3D patch was sampled at every voxel of the arterial center lines. Only patches sampled from renal ROIs were used for training.

To study the impact of patch size, where the patch size denotes the edge length of each cubic patch, we repeated the experiment using four different patch sizes with the other parameters fixed. In our experiments, convolutional layers were applied without padding to avoid introducing false signals in the perimeter of the input; hence convolutional layers reduced the output sizes. The minimum patch size applicable in the network was determined by the number of convolutional and pooling layers, and it was 48 voxels in our experiments. The maximum patch size was limited by the amount of GPU memory, and it was 96 voxels in our experimental environment, which is described later. Therefore, patch sizes used in our experiments ranged from 48 voxels to 96 voxels with a step size of 16, where the step size was heuristically determined to be small enough to study the impact of patch size.

In addition to the UN shown in Fig. 2, shallower and deeper UNs were tested to observe the impact of the depth of the network. While the UN has three levels of resolution, shallower and deeper UNs have two and four levels of resolution, respectively. In other words, compared to the UN, shallower UN has one less pair of encoder and decoder paths and deeper UN has one more pair of encoder and decoder paths.

Early stopping was used to dynamically determine the number of training epochs. The validation set was extracted from the training set to calculate the validation loss. The early stopping algorithm monitored the validation loss and stopped the training process at the end of the epoch in which the validation loss started to increase. After early stopping, the best weight parameters (those that minimized the validation loss) were used for evaluation.

The other hyper-parameters used in the experiments were the optimization algorithm Adam [21] with learning rate $\eta = 0.001$ (default); the loss function, which was binary cross-entropy; and batch size of 32. The segmentation accuracy of the CNN methods was evaluated using experiments with 3-fold cross-validation. The training dataset was split at patient level so that the patches of one patient did not split over different sets. Training and evaluation were conducted using a workstation with single NVIDIA TITAN RTX with 24 GB of GPU memory. Training one UN required 461 minutes and training one SN 352 minutes on average.

In addition to the two CNNs, the HM [1] was also evaluated to compare the CNNs with a conventional method that does not require training, thus does not require a labeled dataset. The gaussian standard deviation $\sigma$ was adjusted for thin arteries in the HM to the values $\sigma = 1, \sqrt{2}, 2$ voxels.

Instead of volumetric evaluation metrics such as Dice similarity coefficient [22] or Jaccard index [23], we chose metrics based on the similarity of the center lines of arteries for the following two reasons.

1. Because blood vessels are small in volume, volumetric similarity metrics are too sensitive to slight differences in the boundaries between manual traces and automatically extracted vessel regions.

2. When understanding vessel structure is the main purpose of detection, center line-based metrics are more appropriate than volumetric based metrics.

The output images of the methods were binarized by thresholding and a binary thinning algorithm [24] was applied to obtain the center lines of the extracted arteries. Likewise, the ground truth of the vessel center lines was generated by applying a binary thinning algorithm to the manually labeled vessels.

The area under the precision-recall curve [25] (AUPRC) was used to evaluate recall and precision of the results. The evaluation metrics for the accuracy were defined as follows.

- The number of true positives (TPs) was defined as the number of extracted center line voxels that were within two voxels of the center line of the ground truth.
• The number of false negatives (FNs) was defined as the number of ground truth voxels that did not have an extracted centerline within a two-voxel proximity.
• The number of false positives (FPs) was defined as the number of extracted centerline voxels that were not within two voxels of the centerline of the ground truth.
• Recall = TPs/(TPs + FNs)
• Precision = TPs/(TPs + FPs)

Although the arteries in the lung and spine ROIs were not used for training, they were treated as true arteries in the evaluation. Wilcoxon signed-rank test was used to analyze statistical significance.

3. Results

In this section, the results of the three methods for the three ROIs are shown. Hereafter, the UN and SN may have patch size indicated in parentheses such as SN(64) and UN(80).

3.1 Evaluation of the renal ROI

Figure 3 shows box-and-whisker plots of the AUPRC comparing results of different patch sizes. As shown in Fig. 3(a), the SN plateaued at the patch size of 64. There was no significant difference among the models with patch sizes of 64, 80, and 96, and larger patch sizes only slightly increased average accuracy.

In the UNs, the model with a patch size of 80 had the best accuracy with statistical significance (Fig. 3(b)). Unlike the results of the SNs, the model with the largest patch size of 96 had significantly worse accuracy than the model with a smaller patch size of 80.

Figure 4 compares the results of the three methods. For each of the CNN methods, the patch size with the best result was chosen for the comparison. As shown in Fig. 4(a), the UN yielded in the best accuracy of the three methods.

In the precision-recall curve shown in Fig. 4(b), the UN had better recall and precision. Although the HM had much lower precision, the best recall was better than those of the CNN-based methods.

Figure 5 shows a qualitative comparison of the methods for the renal ROI results. The CNNs were able to enhance the arterial regions while suppressing the non-arterial regions such as kidneys and veins. No major difference in results was observed between the UN and SN results. However, the UN results tended to show better accuracy for the thinner arteries as indicated by the red arrows. Some veins were mistakenly enhanced in the CNN models (cyan arrows). HM enhanced not only arteries but also veins (cyan arrow) and the boundaries of renal cortex and medulla (green arrow).

![Fig. 3](image-url) Box-and-whisker plots of the results for the renal ROI. The mean (standard deviation) of each method is indicated below each method. Statistical significance is indicated by **(p < 0.01.) (a) SN results. (b) UN results. In (b), statistical significance is detected only for the patch size of 80.
3.2 Evaluation of the lung ROI

As shown in Fig. 6(a), the HM had the best score in the lung ROI. As shown in Fig. 6(b), HM achieved better results both in recall and precision. For the lung ROI, UN yielded the worst AUPRC, and its recall score was much lower than that of other two methods.

In this evaluation, the lung vessels were defined as targets for detection. Therefore, detected lung vessels were counted as TPs even though lung vessels are not renal artery or any of the abdominal arteries.

Figure 7 shows the qualitative comparisons of the lung ROI results. Consistent with the quantitative results, the UN removed most of the lung vessels whereas the SN retained some of the vessels. As noted in the quantitative results, if lung vessels were not targets for detection, the UN would have had the best result because there was almost nothing to detect in the lung ROIs.
3.3 Evaluation of the spine ROI

As was the case with the evaluation of the lung ROI, arteries in the spine ROI were considered targets for detection. A notable difference between lung ROI and spine ROI was that the arteries (intercostal arteries) were more similar to renal arteries.

Figure 8 shows the quantitative comparisons for the spine ROI results. For the spine ROI, the SN and UN achieved equally good results [Fig. 8(a)]. The HM yielded low precision and there were too many false-positive enhancements.

Figure 9 shows a quantitative comparison of the spine ROI results. In this example, CNNs achieved good accuracy on the intercostal artery with no false positive enhancement of the spinal bone. The HM had a strong response at the edge of the spinal bone. This is because cortical bones have an intensity distribution similar to the distribution that the HM is mathematically designed for.

3.4 Overall image evaluations

Figure 10 shows a qualitative comparison of the results of all the image. The differences between the UN and the SN results are as follows.

1. In the lung region, the SN enhanced more vessels than the UN.
2. Thin vessels such as the intercostal artery were better enhanced by the UN.

Among the three methods, the HM had the best result for the lung vessels, although many artifacts originating from bones and other non-arterial organs were included. As shown in the images in the right column of Fig. 10,
the HM results were good when artifacts due to the bones were removed. However, abdominal arteries were not selectively enhanced; instead, abdominal veins and lung vessels were also enhanced.

3.5 U-Net depth

Figure 11 shows the results of three UNs [UN (default setting), shallower UN, and deeper UN] in the renal ROI. Note that we were able to run deeper UN only with the patch size 96, because deeper UN had too many convolution and max pooling layers for smaller patch sizes. As shown in Fig. 11, the UN (default setting) was better than shallower and deeper UNs in all patch sizes, and both shallower UN and the UN (default setting) had peak accuracy with patch size 80.

4. Discussion and Conclusion

4.1 Discussion

In this study, we experimentally investigated the impact of patch size and network architecture on vessel detection and discrimination accuracy in 3D CT data. Two patch-based 3D CNNs; that is, a U-Net-like architecture and a sequential architecture, were investigated. In addition to the CNNs, a conventional HM was also tested.

The characteristics of the two CNNs and HM are summarized below. First, the UN had the characteristics of high selectivity when extracting specific vessels such as abdominal and intercostal arteries, which have similar features to the renal arteries in the training data. Therefore, the UN showed the best results on abdominal artery extraction when trained on renal artery data, and it will
be suitable for selective extraction of vessels similar to those of the training data. The patch size needed to be fine-tuned when using UN. Second, the SN had similar characteristics to those of the UN, but with less strong selectivity to specific vessels. The patch size needed to be tuned, but the impact on accuracy was less sensitive than in the UN. Third, HM produced many false positives in renal and spine ROIs, but it should be noted that HM had the best performance in lung ROI, and recall (the true positive rate) was better than or equivalent to the CNN methods in all three ROIs in our experiments. The advantages of the HM is that it does not require training data and it responds to vessels generally (while producing false positives); therefore it will be still useful as a general purpose method when training dataset is not available.

As shown in Fig. 3, the patch size had a significant influence on the detection accuracy of patch-based 3D CNNs. The optimal patch size depended on the CNN architecture. Therefore it is necessary to adjust the patch size to achieve better accuracy in vascular detection.

The UN and SN had different responses to the input patch size. The UN was sensitive to patch size because it had a peak accuracy at the patch size of 80, whereas the SN was less sensitive to patch size because accuracy plateaued. In previous papers, most studies [14–16] used a fixed patch size in a patch-based CNN, and the impact of patch size was not investigated in detail. Some papers [12, 18, 26, 27] addressed the impact of patch size, two of which [26, 27] showed that large patch sizes always improved accuracy. In these papers, however, the patch was not 3D; instead, patch-based 2D or 2.5D CNNs were used. Yang et al. [18] used a 3D CNN and showed partly similar results to those of our work. However, they were not able to show optimal patch size by demonstrating a peak or plateau in accuracy due to the hardware restriction.

Fig. 10 Qualitative illustration of all the images. 3D volumetric images were reduced to 2D images using MIP in the coronal plane. Examples 1, 2 and 3 correspond to the symbols in Fig. 4. For better visualization, MIP image with bones masked out are also shown (Input w/o bones and HM w/o bones). Note that the bones were only masked for visualization; the original images were used as the input for each method.

Fig. 11 Average AUPRCs of UN variants with different depths and patch sizes.
As shown in Fig. 4, network architecture significantly changes the detection accuracy. Although the UN had the best results for the renal ROI, other methods had better or equivalent result in other ROIs. Because the SN had a better generalization ability, it can be used in general vessel extraction when only limited training data are available. The HM responds to any vessel, but the responses largely depend on image contrast.

Although there are numerous new network architectures designed for semantic segmentation in medical images [28–32], we focused on two of the most fundamental network architectures in this study. We believe that it is still valuable to test these two architectures because testing fundamental architectures gives us insights to design or assess new architectures, which are typically derived from these fundamental architectures. Future work will include further comparative study of the recent architectures.

As shown in Fig. 11, deeper UN performed significantly worse compared to other two UNs in our experiments. Before conducting the experiments, we expected that deeper UN would perform better because deeper networks are able to capture larger context, which usually helps distinguish objects. One potential explanation why deeper UN failed is that it is harder to train deeper UN because deeper UN has more layers and more trainable weight parameters. Another potential reason is that large context is not always as important as local appearance when it comes to recognizing arteries, which typically have elongated structure but are small in diameter. As shown in the results, the depth has greater impact on accuracy than the patch size. Therefore, in hyper-parameter tuning, we recommend tuning the depth first and then tuning other hyperparameters such as patch size for the U-Net-like architectures.

Because of the limited amount of GPU memory available in our experimental environment, we were not able to conduct experiments with patch sizes larger than 96 voxels. However, as shown in the results, we expect that larger patch sizes will result in similar or worse accuracy compared with that achieved by the optimal patch sizes in the experiments.

4.2 Conclusions and future work
The results of this study show that it is necessary to adjust the patch size to achieve better accuracy in vascular detection in 3D data, especially when patch-based 3D UN is used. Although we cannot determine a conclusive procedure about the hyper-parameter tuning, we recommend tuning the depth of the network first and then tuning patch size, because the depth has greater impact on accuracy and depth tuning seems to be more insensitive to patch size variation. Regarding comparison among different methods, UN provides the best result in the experiments when the task is to selectively extract specific vessels with features particularly matching those of the training data. However, the other methods (SN and HM) will be more suitable when the task is to extract vessels more generally.

In this study, we performed experiments on abdominal arteries. Future work will include applications to arteries in other domains and veins, which may have different properties and characteristics such as thickness and curvature.

Acknowledgement
This work was supported in part by MEXT/JSPS KAKENHI Grant No. 17K10403.

Conflicts of interest
The authors declare no conflicts of interest with any companies or commercial organizations per the definition of Japanese Society for Medical and Biological Engineering.

Ethics declaration
The study received Institutional Review Board approval (08162) at Osaka University Hospital. Requirement of informed consent was waived due to the retrospective nature of the study.

References
1. Sato Y, Nakajima S, Shiraga N, Atsumi H, Yoshida S, Koller T, Gerig G, Kikinis R: Three-dimensional multi-scale line filter for segmentation and visualization of curvilinear structures in medical images. Med Image Anal. 2(2), 143–168, 1998.
2. Frangi AF, Niessen WJ, Vincen KL, Viergever MA: Multiscale vessel enhancement filtering. Med Image Comput Assist Interv, 130–137, 1998.
3. Bibiloni P, Gonzalez-Hidalgo M, Massanet S: A survey on curvilinear object segmentation in multiple applications. Pattern Recognit. 60, 949–970, 2016.
4. Moccia S, De Momi E, El Hadji S, Mattos LS: Blood vessel segmentation algorithms review of methods, datasets and evaluation metrics. Comput Methods Programs Biomed. 158, 71–91, 2018.
5. Zhao F, Chen Y, Hou Y, He X: Segmentation of blood vessels using rule-based and machine-learning-based methods: A review. Multimed Syst. 25(2), 109–118, 2019.
6. Hanaoka S, Nomura Y, Takenaga T, Murata M, Nakao T, Miki S, Yoshikawa T, Hayashi N, Abe O, Shimizu A: HoTPiG: A novel graph-based 3-D image feature set and its applications to computer-assisted detection of cerebral aneurysms and lung nodules. Int J CARS. 14(12), 2095–2107, 2019.
7. Krizhevsky A, Sutskever I, Hinton GE: Imagenet classification with deep convolutional neural networks. Neural Inf Process Syst, 1097–1105, 2012.
8. Long J, Shelhamer E, Darrell T: Fully convolutional networks for semantic segmentation. Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit, 3431–3440, 2015.
9. Wang C, Hayashi Y, Oda M, Itoh H, Kitasaka T, Frangi AF, Mori K: Tubular Structure Segmentation Using Spatial Fully Connected Network with Radial Distance Loss for 3D Medical Images. Med Image Comput Comput Assist Interv, 348–356, 2019.

10. Taha A, Lo P, Li J, Zhao T: Kid-net: Convolution networks for kidney vessels segmentation from CT-volumes. Med Image Comput Comput Assist Interv, 463–471, 2018.

11. Keshwani D, Kitamura Y, Ihara S, Iizuka S, Simo-Serra E: TopNet: Topology Preserving Metric Learning for Vessel Tree Reconstruction and Labelling. Med Image Comput Comput Assist Interv, 14–23, 2020.

12. Patel TR, Paliwal N, Jaiswal P, Waqas M, Mokin M, Siddiqui AH, Meng H, Rai R, Tutino V: Multi-resolution CNN for brain vessel segmentation from cerebrovascular images of intracranial aneurysm: A comparison of U-Net and DeepMedic. Medical Imaging 2020: Computer-Aided Diagnosis. 11314, 113142W, 2020.

13. Yu F, Zhao J, Gong Y, Wang Z, Li Y, Yang F, Dong B, Li Q, Zhang L: Annotation-free cardiac vessel segmentation via knowledge transfer from retinal images. Med Image Comput Comput Assist Interv, 714–722, 2019.

14. Alom MZ, Hasan M, Yakopcic C, Taha TM, Asari VK: Recurrent residual convolutional neural network based on u-net (R2u-net) for medical image segmentation. arXiv preprint arXiv: 1802.06955, 2018.

15. Huang Q, Sun J, Ding H, Wang X, Wang G: Robust liver vessel extraction using 3D U-Net with variant dice loss function. Computers in Biology and Medicine, 101, 153–162, 2018.

16. Wang C, Zhao Z, Ren Q, Xu Y, Yu Y: Dense U-net based on patch-based learning for retinal vessel segmentation. Entropy. 21(2), 168, 2019.

17. Oda M, Roth HR, Kitasaka T, Misawa K, Fujiwara M, Mori K: Abdominal artery segmentation method from CT volumes using fully convolutional neural network. Int J CARS. 14(12), 2069–2081, 2019.

18. Yang H, Shan C, Kolen AF, de With PH: Automated catheter localization in volumetric ultrasound using 3D patch-wise U-net with focal loss. 2019 IEEE International Conference on Image Processing (ICIP), 1346–1350, 2019.

19. Ronneberger O, Fischer P, Brox T: U-net: Convolutional networks for biomedical image segmentation. Med Image Comput Comput Assist Interv, 234–241, 2015.

20. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O: 3D U-Net: Learning dense volumetric segmentation from sparse annotation. Med Image Comput Comput Assist Interv, 424–432, 2016.

21. Kingma DP, Ba J: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980, 2014.

22. Dice LR: Measures of the amount of ecologic association between species. Ecology. 26(3), 297–302, 1945.

23. Jaccard P: The distribution of the flora in the Alpine zone. New Phytologist. 11(2), 37–50, 1912.

24. Lee TC, Kashyap RL, Chu CN: Building skeleton models via 3-D medial surface axis thinning algorithms. CVGIP: Graphical Models and Image Processing. 56(6), 462–478, 1994.

25. Davis J, Goadrich M: The relationship between Precision-Recall and ROC curves. Proceedings of the 23rd International Conference on Machine Learning, 233–240, 2006.

26. Hamwood J, Alonso-Caneiro D, Read SA, Vincent SJ, Collins MJ: Effect of patch size and network architecture on a convolutional neural network approach for automatic segmentation of OCT retinal layers. Biomedical Optics Express. 9(7), 3049–3066, 2018.

27. Livne M, Rieger J, Aydin OU, Taha AA, Akay EM, Kossen T, Sobesky J, Kelleher JD, Hildebrandt K, Frey D, Madai VI: A U-Net Deep Learning Framework for High Performance Vessel Segmentation in Patients With Cerebrovascular Disease. Front Neurosci. 13, 2019.

28. Roth HR, Shen C, Oda H, Sugino T, Oda M, Hayashi Y, Misawa K, Mori K: A multi-scale pyramid of 3D fully convolutional networks for abdominal multi-organ segmentation. Medical Image Computing and Computer Assisted Intervention MICCAI, 2018, 417–425, 2018.

29. Chen W, Liu B, Peng S, Sun J, Qiao X: 3D-UNet: Separable 3D U-Net for brain tumor segmentation. Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, 358–368, 2019.

30. Li X, Chen H, Qi X, Dou Q, Fu C-W, Heng P-A: H-DenseUNet: Hybrid densely connected U-Net for liver and tumor segmentation from CT volumes. IEEE Transactions on Medical Imaging. 37(12), 2663–2674, 2018.

31. He Y, Yang G, Yang J, Chen Y, Kong Y, Wu J, Tang L, Zhu X, Dillenseger J-L, Shao P, Zhang S, Shu H, Coatrieux J-L, Li S: Dense biased networks with deep priori anatomy and hard region adaptation: Semi-supervised learning for fine renal artery segmentation. Med Image Anal. 63, 101722, 2020.

32. Wang C, Oda M, Hayashi Y, Yoshino Y, Yamamoto T, Frangi AF, Mori K: Tensor-cut: A tensor-based graph-cut blood vessel segmentation method and its application to renal artery segmentation. Med Image Anal. 60, 101623, 2020.

Yuki Suzuki

Yuki Suzuki received his B.S. degree in Engineering and M.S. degree in Information Science from Osaka University in 2012 and 2014 respectively. From 2014 to 2019, he worked as a research assistant in Nara Institute of Science and Technology. He is currently a Specially Appointed Assistant Professor in the Department of Artificial Intelligence Diagnostic Radiology, Osaka University Graduate School of Medicine. His research interests include medical image analysis, computational anatomy, and image processing.
Masatoshi Hori

Masatoshi Hori received the B.S. and M.S. degrees in electrical engineering from Osaka University, Osaka, Japan in 1986 and 1988, respectively. Then he received the M.D. degree and the Ph.D. degree in medicine from Osaka University in 1993 and 2001, respectively. He is currently a Professor in the Department of Radiology, Kobe University Graduate School of Medicine. His research interests include diagnostic radiology, image analysis, and computer aided diagnosis. Dr. Hori is a member of Radiological Society of North America, Fleischner Society, Japanese Society of Medical Imaging Technology, International Society of Magnetic Resonance in Medicine, and Radiological Society of North America.

Shoji Kido

Shoji Kido received his M.D. degree from Osaka University in 1988. He received his Ph.D. degrees in Medicine and Information Science from Osaka University in 1992 and 1999, respectively. He was a professor in the Department of Computer Science and Systems Engineering at Yamaguchi University from 1999 to 2006, in the Department of Applied Medical Engineering, Graduate School of Medicine, Yamaguchi University from 2006 to 2016, and in the Division of Information Sciences and Engineering, Graduate School of Sciences and Technology for Innovation, Yamaguchi University from 2016 to 2019. He is currently a professor in the Department of Artificial Intelligence Diagnostic Radiology, Osaka University Graduate School of Medicine. His research interests include chest radiology, and computer-aided diagnosis based on artificial intelligence. He is a member of Japan Radiological Society, Radiological Society of North America, Japanese Society for Medical and Biological Engineering, Japanese Society of Medical Imaging Technology, Institute of Electronics, Information and Communication Engineers Medical imaging.

Yoshito Otake

Yoshito Otake received the B.S., M.S. and Ph.D. degrees from the Department of Electronics Information and Communication Engineering, Waseda University, in 2000, 2002, and 2007. From 2002 to 2008, he worked as an Assistant Professor at the Institute for High Dimensional Medical Imaging in Jikei University School of Medicine. From 2008 to 2014, he was a research scientist at the Departments of Computer Science and Biomedical Engineering at Johns Hopkins University. Since 2014, he has been an Associate Professor at Nara Institute of Science and Technology. His research interests include medical image analysis, computer-assisted surgery, and computational anatomy.

Mariko Ono

Mariko Ono received her M.S. degree in Engineering from Nara Institute of Science and Technology in 2018. In 2018, she joined Accenture plc. Her research interests include medical image analysis.

Noriyuki Tomiyama

Noriyuki Tomiyama received his degrees of M.D. and Ph.D. in Medicine from Osaka University in 1987 and 1993 respectively. In 1987, he started his carrier as a medical doctor in Osaka University Hospital. After he graduated from Osaka University Graduate School of Medicine, he held the various posts such as Assistant professor and Associate professor in the Department of Radiology of the school. He worked as a research fellow in Vancouver General Hospital, Canada in 1999, and as a staff in Ehime National Hospital from 2000 to 2002. From 2010, he is the Professor and Chairman of the Department of Radiology, Osaka University Graduate School of Medicine. Dr. Tomiyama is a Vice President of Japan Radiological Society and the Presidential Elect of Asian Oceanian Society of Radiology. In addition, he is a board member of Japanese Society of Thoracic Radiology and Japan Radiology Congress and a member of Radiological Society of North America, Fleischner Society, International Society for Strategic Studies in Radiology, and Japanese Medical Specialty Board.

Yoshinobu Sato

Yoshinobu Sato received his B.S., M.S. and Ph.D degrees in Information and Computer Sciences from Osaka University, Japan in 1982, 1984, 1988, respectively. From 1988 to 1992, he worked at the NTT Human Interface Laboratories. In 1992, he joined Osaka University Medical School. From 1996 to 1997, he was a Visiting Research Fellow at the Surgical Planning Laboratory, Harvard Medical School and Brigham and Women’s Hospital. In 1999, he became an Associate Professor at Osaka University Graduate School of Medicine. From 2014, he is a Professor at Nara Institute of Science and Technology (NAIST). His research interests include medical image analysis, computer assisted surgery, and computational anatomy. Dr. Sato is an Editorial Board member of Medical Image Analysis journal and International Journal of Computer Assisted Radiology and Surgery (IJACRS) journal. He was Program Chair of the 16th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI 2013).