## TECHNICAL ADVANCE

## Open Access

# Coronary angiography video segmentation method for assisting cardiovascular disease interventional treatment

Dongxue Liang[1]* , Jing Qiu[1], Lu Wang[1], Xiaolei Yin[1], Junhui Xing[3], Zhiyun Yang[2], Jiangzeng Dong[2,3] and Zhaoyuan Ma[1]

## Abstract

**Background:** Coronary heart disease is one of the diseases with the highest mortality rate. Due to the important position of cardiovascular disease prevention and diagnosis in the medical field, the segmentation of cardiovascular images has gradually become a research hotspot. How to segment accurate blood vessels from coronary angiography videos to assist doctors in making accurate analysis has become the goal of our research.

**Method:** Based on the U-net architecture, we use a context-based convolutional network for capturing more information of the vessel in the video. The proposed method includes three modules: the sequence encoder module, the sequence decoder module, and the sequence filter module. The high-level information of the feature is extracted in the encoder module. Multi-kernel pooling layers suitable for the extraction of blood vessels are added before the decoder module. In the filter block, we add a simple temporal filter to reducing inter-frame flickers.

**Results:** The performance comparison with other method shows that our work can achieve 0.8739 in *Sen*, 0.9895 in *Acc*. From the performance of the results, the accuracy of our method is significantly improved. The performance benefit from the algorithm architecture and our enlarged dataset.

**Conclusion:** Compared with previous methods that only focus on single image analysis, our method can obtain more coronary information through image sequences. In future work, we will extend the network to 3D networks.

**Keywords:** Coronary angiography, Medical assistance, Video segmentation, Medical imaging, Medical assistance

## Background

Cardiovascular disease, especially coronary sclerotic heart disease, also known as coronary heart disease, is one of the diseases with the highest mortality rate. Traditionally, doctors diagnose these cardiovascular diseases by directly observing angiographic images with their eyes. According to experience, they make a qualitative judgment on the patient's condition. This diagnostic method is greatly affected by human factors and lacks accuracy and objectivity. With the widespread application of image segmentation technology in the field of medical image analysis and processing, and the increasing emphasis on the prevention, diagnosis, and treatment of cardiovascular disease, the segmentation of coronary angiography images and videos has gradually become a research hotspot. The segmentation and extraction of coronary blood vessels are also gradually being applied, in practice such as assisted diagnose of the disease, precise location of lesions, quantitative analysis of vascular tissue, and research on three-dimensional reconstruction of coronary arteries.

*Correspondence: liang.laurel@hotmail.com
[1]The Future Laboratory, Tsinghua University, Chengfu Road, Beijing, China
Full list of author information is available at the end of the article

Due to the particularity of medical images, there have been some difficulties in the processing and analysis of the images. The analysis of angiograms has the following difficulties: first, the shape of the blood vessels in the video is complex and easily deformed. Blood vessels have a tubular curved structure, and some blood vessels can block, cover or entangle with each other, which brings some difficulties to image processing. Second, the density and diameter of blood vessels vary. With the extension of the blood vessels, the blood vessels gradually become thinner, and there is vascular stenosis caused by the blockage. This causes the contrast and resolution of the small blood vessels in the contrast image to be very low and difficult to process. Third, the background noise in the image is relatively high. In coronary angiography images, besides blood vessels, there are different tissues such as the chest, lungs, and ribs. The shape or grayness of some tissues is similar to that of the blood vessels, which makes it difficult to extract blood vessels.

So far, many methods have been used to segment blood vessels. These methods maybe classified as follows: pattern recognition, model-based tracking, propagation, and artificial intelligence-based methods [1–5]. Most of the vessel detection methods are performed in the spatial domain of the original image such as single-scale top-hat operator [6], hit-or-miss transform [7], and Gaussian matched filter [8]. Poli [9] also proposed an efficient method based on linear filters, which used a shifted Gaussian kernel as it is more sensitive to vessels with different orientations and radius. An adaptive tracking method was also presented to extract vessels from the X-ray angiograms [10]. Model-based approaches for medical image segmentation were also applied. In [11], the author proposed a snake model to extract the vessels by deforming the spline to minimize energy functions. All these previous methods for vessel segmentation are limited by at least one of the following disadvantages: they either were unable to suppress sudden noise, could not detect vessels in a wide range from a fixed scale image, or needed heavy computation for the vessel segmentation.

Nowadays, the learning-based methods were very popular in segmenting medical images. An unsupervised learning-based method was proposed by Aganj [12] to segment the X-ray and MRI images. Tong [13] also proposed a multi-organ segmentation method by combing dictionary learning and sparse coding. There was also some work [14, 15] using the pixel-level classification method to get the ROI of the medical image by using the pre-trained data. The main drawback of these methods is that it is difficult to design the representative features for different applications, and if we change the kind of the input image, it is hard to capture the features. With the development of the deep learning method, the convolutional neural network played an important role in medical image analysis, such as [16, 17]. Different from the traditional classification method, deep learning methods learn the features automatically. For the medical image segmentation, most of the earlier deep learning-based methods use the image patches and sliding window block like [18]. But this kind of method will have a huge amount of computation caused by sliding the window block and ignore the global feature at the same time. In 2015, U-net [19] was proposed for medical image segmentation and achieve a good result. After that many different methods based on U-net architecture had been proposed. M-net [20] added a multi-scale input image and deep supervision to the original U-net architecture. Also, some new modules were proposed to replace some blocks in the U-net architecture to enhance the feature learning ability. Gibson et al.[21] proposed a dense connection in the encoder block to do the organ segmentation task on CT images. Also, to improve the segmentation performance, Zhao et al.[22] introduce a modified U-net by adding a spacial pyramid pooling.

The U-net and its modification methods have a common drawback which is the continuous pooling and convolution striding will lose the feature resolution, and this is a big limitation for segmenting the vessels witch require small spatial information. Especially, in medical video analysis, the continuous segmentation between frames needs more detail information of the segmentation results. With the previous discussion and reference the work of Gu et al.[23], in this paper, we focus on a new application of the segmentation method used for coronary angiography video. Similar to [23], the method combines with three main modules: the video encoder module, the video decoder module, and the video filter module. In the encoder module, to achieve more high-level semantic features, a dense atrous convolution (DAC) block and a residual multi-kernel pooling (RMP) block are added. After the normal decoder module, we add a video filter block to get the smooth video sequences.

The main contributions of our work are summarized as follows:

– A novel method to extract vessel from the coronary video.

– A new application of the segmentation algorithm used in the medical assistance field.

– A new dataset used for coronary angiography video segmentation.

## Methods
Figure 1 shows an overview of the proposed method. The method uses the coronary angiography video sequences as input, and outputs are the segmented sequences. The proposed method consists of three main parts: the sequence encoder module, the sequence decoder module, and the sequences filter module.

**Fig. 1** Framework for video segementation
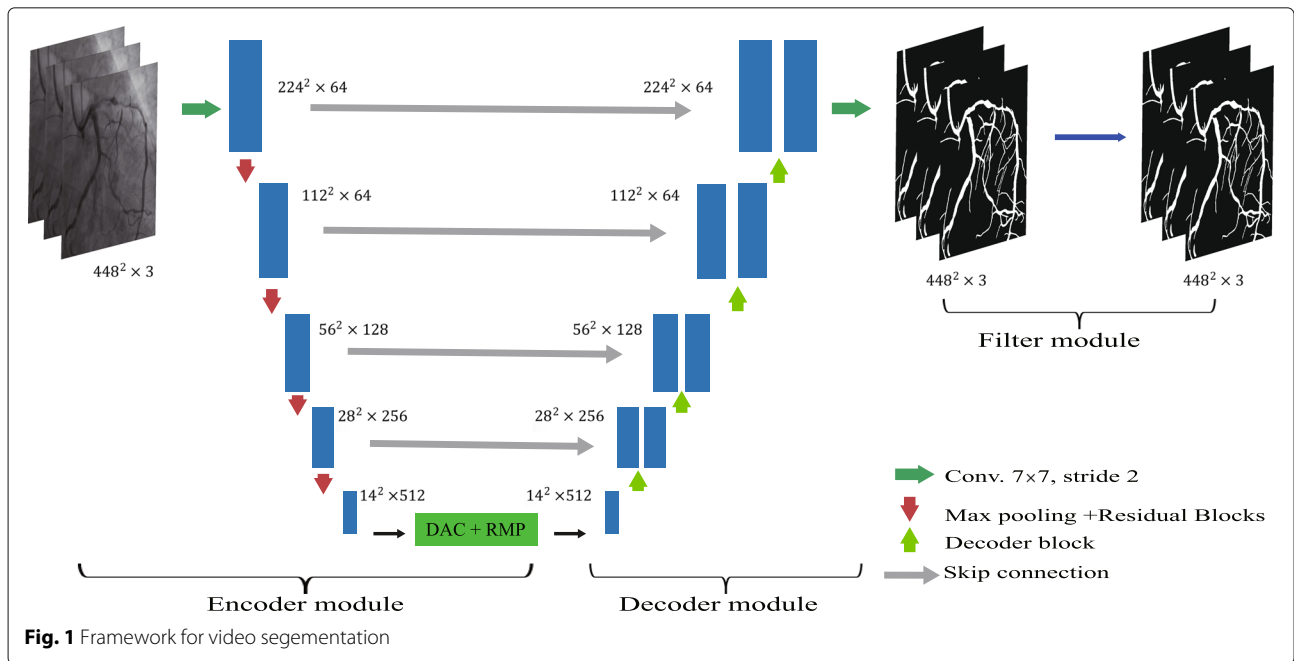
### Encoder module

We use the U-net as the basic architecture for the proposed method. In this architecture, the original encoder replaced with the pre-trained ResNet-34, as ResNet can avoid gradient disappearance and accelerate network convergence. In the ResNet-34, we keep the first four feature extracting blocks only. We use the modified U-net with pre-trained ResNet as the backbone method.

The context extractor consists of the DAC (Dense Atrous Convolution) block and the RMP (Residual Multi-kernel pooling) block [23], which can extract more high-level features. Based on the Inception-ResNet-V2 block

[24] and atrous convolution [25], high-level features can be encoded by the dense atrous convolution (DAC) block. As shown in Fig. 2, we use the same four cascade branches with the gradual increment of the number of atrous convolution in the DAC block, this block employs different receptive fields. The convolution of a large reception field could extract and generate more abstract features of large objects, while the convolution of a small reception field is better for small objects. By combining the atrous convolution of different atrous rates, the DAC block can extract features for objects with various sizes. A residual multi-kernel pooling (RMP) is also adopted in this method, which works on a large variety of object sizes in a medical
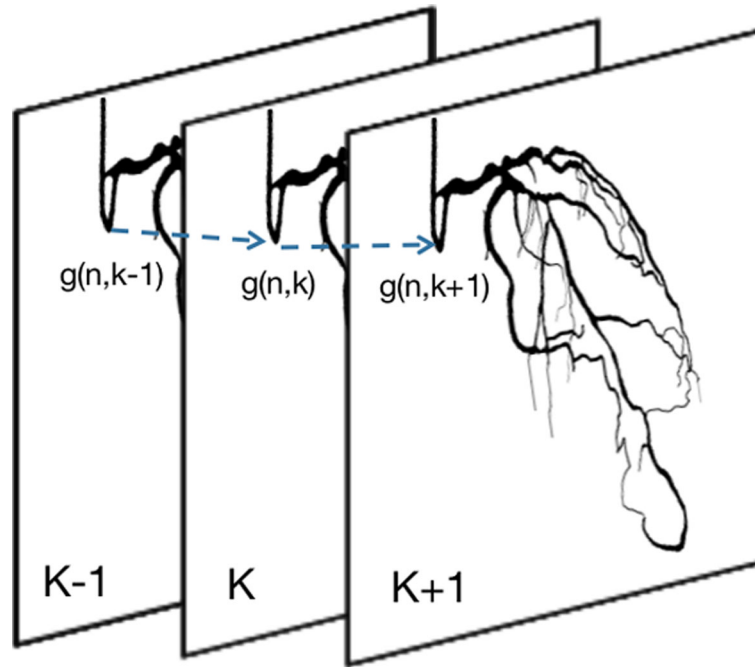


**Fig. 2** Dense atrous convolution module
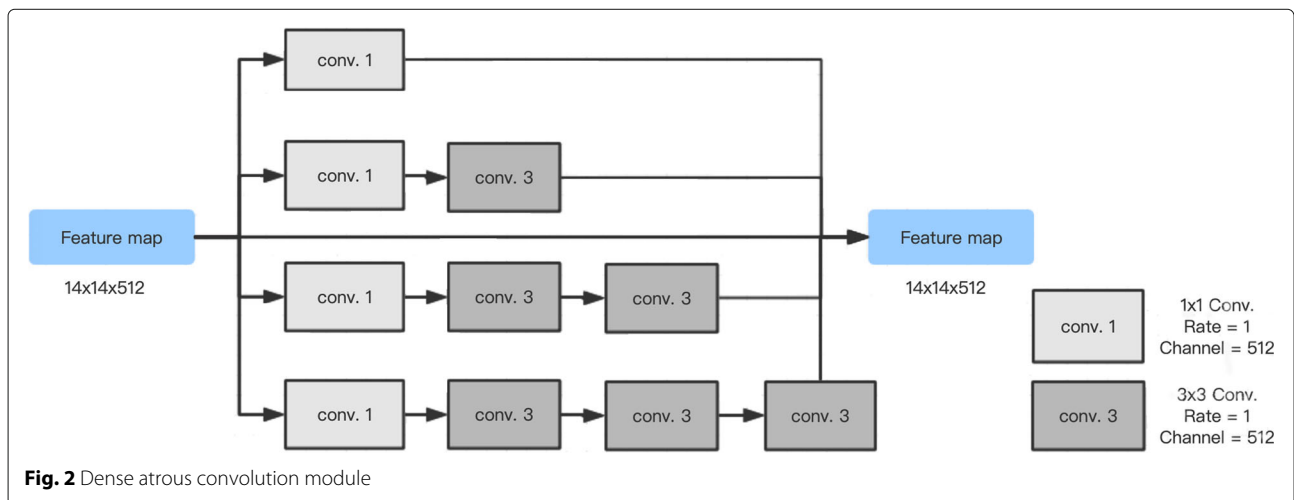
**Fig. 3** Video filter of the image element (**n**,k)

image, like coronary angiography with different sizes of vessels. The RMP extracts the feature with four different sizes of poolings: 2×2, 3×3, 5×5 and 6 ×6. After every pooling step, a 1 ×1 convolution is used to reduce the dimension of the computational cost. In the end, by bilinear interpolation, we do the upsampling to get the final features.

### Decoder module

Using the decoder module, we can recover the high-level features extracted from the feature extraction encoder module. Just like U-net, as normal striding convolution and continuous pooling operations will lead to loss of information, we can skip the connection and transfer some detailed information directly from the encoder to the decoder. The upscaling and deconvolution are two common operations of the decoder in U-net. The upscaling operation increases the image size by linear interpolation, while the deconvolution uses a convolution operation to enlarge the image. Here, it mainly includes 1×1 convolution, 3×3 transposed convolution, and 1×1 convolution consecutively. The deconvolution operation can adaptively learn the mapping to recover features using more detailed information. Therefore, in our method, deconvolution was selected to recover the high-resolution features in the decoder. The output of the final feature recovery decoder module is a mask of the same size as the original input image.

### Video filter module

We observe that there exists some confusion between frames in the output mask from the decoder. Since we only do the segmentation in a one-shot manner without considering temporal information in other frames. To
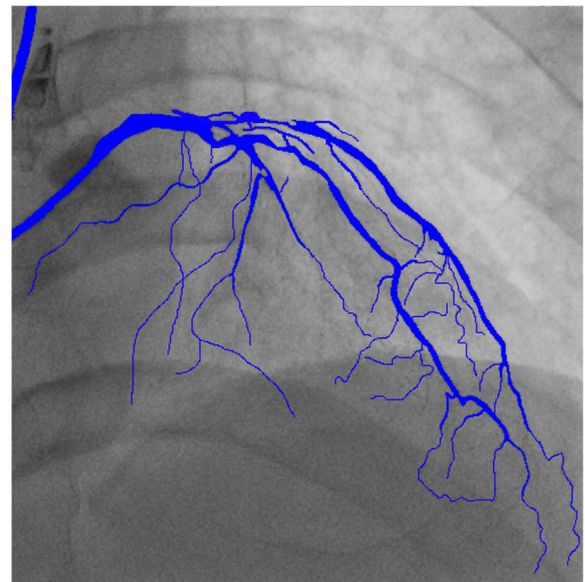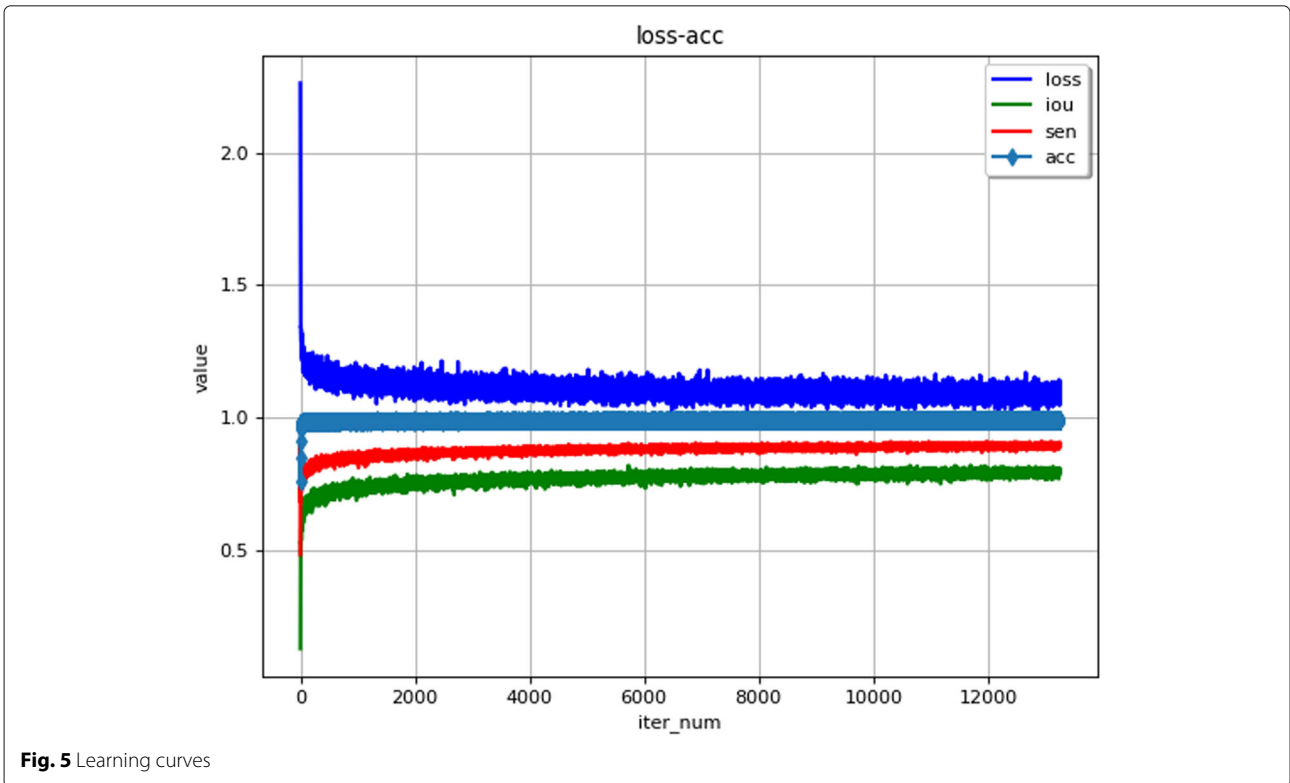


**Fig. 4** Annotated image

**Fig. 5** Learning curves

improve the segmentation results and reduce computational complexity, we use a simple temporal filter to carry out a weighted averaging of successive frames [26]. As shown in Eq.(1),

$$\hat{f}(\mathbf{n}, k) = \sum_{l=-K}^{K} h(l)g(\mathbf{n}, k - l) \qquad (1)$$

where $g(\mathbf{n}, k)$ is the recorded image sequence, $k$ is the number of the sequence. $\mathbf{n}=(n1, n2)$ refers to the spatial coordinates. $h(l)$ are the temporal filter coefficients used to weight $2K+1$ consecutive frames. In case the frames are considered equally important we have $h(l) = 1/(2K + 1)$. The motion artifacts can greatly be reduced by operating the filter along with the image elements that lie on the same motion trajectory [27], as shown in Fig 3.
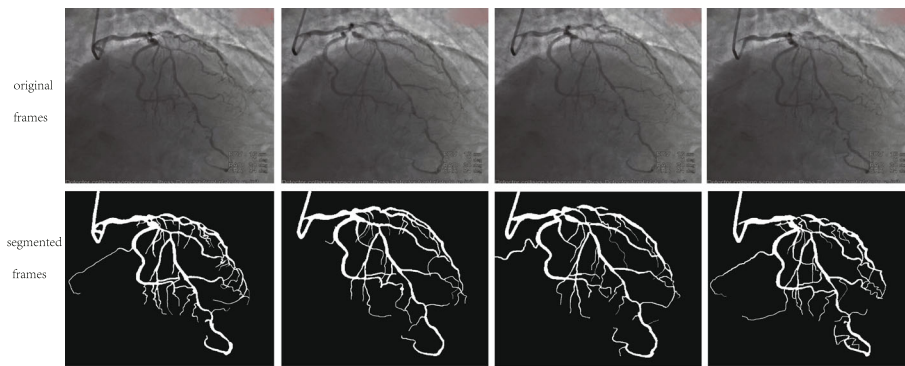


**Fig. 6** Result of angiography image segmentation
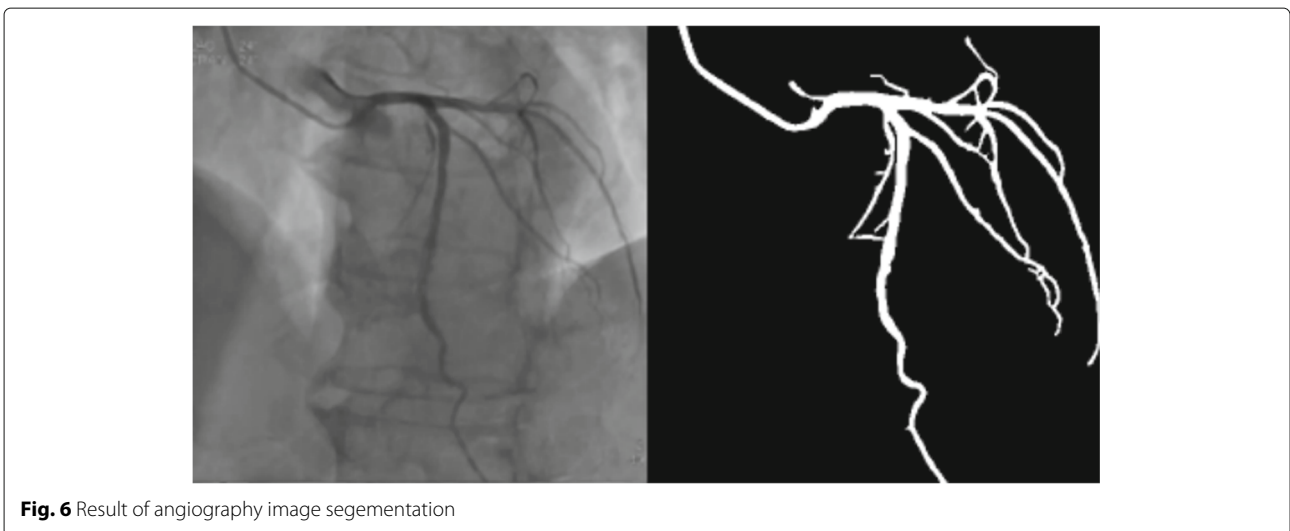
**Fig. 7** Video sequences segmentation results. For frame 3060, 3065, 3070, 3075

## Results

### Data

In our work, we choose two parts of the dataset for training, one is the public DRIVE [28] dataset which is widely used for vessel detection, and the other dataset comes from the angiography video of coronary interventional surgery. 170 video clips contain clear blood vessels with contrast agents were annotated manually by medical students. We record hundreds of videos of interventional surgeries, clip the videos which have contrast agents, then do interframe sampling about every three frames, finally we get 4904 annotated angiography images, as shown in Fig. 4. The annotated images are divided into two sets by selecting 4354 images for the train set and 550 images for the validation set during the training.

### Train

**loss function** This paper uses the end-to-end learning process based on U-net. Blood vessel segmentation is a pixel-level classification task. Usually, the cross-entropy loss function is used for this kind of segmentation task. But, in the coronary segmentation task, the background occupies a large area of the image, and the blood vessels are unevenly distributed in the image, so dice loss [29] is selected to replace the cross-entropy loss function. If the dice coefficient is higher, the similarity between the predicted result and the ground truth is higher. Also, it is more feasible to train for minimizing the loss value. The loss function is defined as:

$$L_{dice} = 1 - \sum_m^M \frac{2w_m \sum_i^N p_{(m,i)} g_{(m,i)}}{\sum_i^N p_{(m,i)}^2 + \sum_i^N g_{(m,i)}^2} \qquad (2)$$

where $M$ is the number of the pixel, the $p_{(m,i)}$ and $g_{(m,i)}$ are predicted probability and ground truth for class $m$. and $\sum_m w_m = 1$ are the class weights, here $w_m = 1/M$. The final loss function is defined as:

$$L_{loss} = L_{dice} + L_{reg} \qquad (3)$$

here $L_{reg}$ represents the regularization loss [30]. The loss curve is shown in Fig. 5.

### Implementation environment

The model in this method is implemented using PyTorch based on GeForce GTX 1070 GPU. The system is Ubuntu 16.04, and the CPU is intel i5, the RAM is 16G.

### Method results and performance

In this section, we show the results and evaluate the performance of the proposed method on the tasks of coronary angiography video segmentation.

Figure 6 shows the segmentation result of one image. Figure 7 shows the results of a video sequence. From the results we can see the detail of the coronary arteries. To evaluate the performance of the vessel segmentation, we compute the sensitivity and the accuracy, which are also calculated in [31].

$$Sen = \frac{TP}{TP + FN} \qquad (4)$$

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \qquad (5)$$

where $TP$ is the number of true positives, $TN$ is the number of true negatives, $FP$ and $FN$ represent the number of false positives and false negatives. In addition, we also compute the IoU overlap to measure segmentation performance as shown in Fig. 5. Table 1 shows the performance comparison with other method. What needs to be explained here is that we use not only the public data set DRIVE [28] but also our own data set. Our work achieves

**Table 1** Performace comparison of vessel segmentation

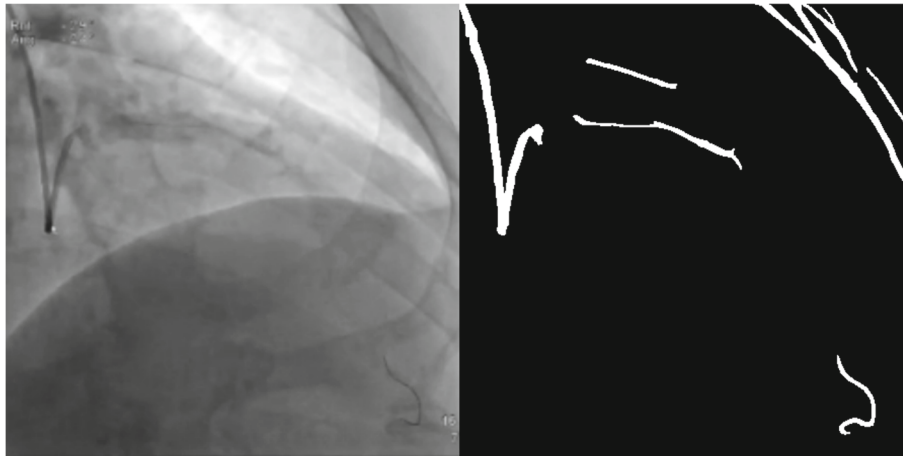| Method | *Sen* | *Acc* |
|---|---|---|
| U-Net[19] | 0.7537 | 0.9531 |
| DeepVessel[31] | 0.7603 | 0.9523 |
| Backbone | 0.7781 | 0.9477 |
| Ours | **0.8204** | **0.9867** |

**Fig. 8** Fail sample

0.8739 in *Sen*, 0.9895 in *Acc*, the average values of *TP*, *TN*, *FP*, and *FN* are 3929912.0, 115029685.0, 747732.0, and 860167 for all the test frames. From the performance of the results, the accuracy of our method is significantly improved. The performance benefit from the algorithom architecture in [23] and the enlarged dataset.

## Discussion

As shown in Fig. 8, two situations can cause segmentation results to fail. First, there are some tissue structures in the video image which have close grayscale value and similar shape with the angiographic vessels, which will cause the tissue to be segmented together. The second is that when the contrast agent is very thin or uneven, the coronary blood vessels will not be segmented. Considering these two situations, we need to make the dataset bigger. At the same time, we need also to consider the differences between the imaging effects of different contrast devices. The time-consuming of this method is about 59ms per frame on a GeForce GTX 1070 GPU, it is almost real-time if the rate set by the coronary angiography system is 15f/s. The algorithm still has room for continuous optimization and will continue to be improved in the future.

## Conclusion

Image segmentation is important in the field of medical image analysis, and segmentation of medical image sequences can obtain more motion information than single-image segmentation, which also has a positive significance for assisting doctors in the diagnosis and treatment. This article introduces the application and implementation of a segmentation method for medical image sequences in coronary angiography video. Compared with previous methods that only focus on single image analysis, we can obtain more coronary information through image sequences. At the same time, we annotated our coronary sequence data set. Experimental results show that this method can segment coronary angiography image sequences with high accuracy. In future work, we will pay more attention to the continuity of information between frames, so that the results look more natural. Our method is now validated on 2D networks and will extend to 3D networks in future work.

**Authors' contributions**
DXL proposed the method and analyzed data and wrote the manuscript. JQ participated in manuscript revisions, LW and XLY conducted experiments,JHX and ZYY helped to analyze the medical data, JZD and ZYM provided critical review that helped in improving the manuscript. All authors read and approved the final manuscript.

**Availability of data and materials**
Data related to the current study are available from the corresponding author on reasonable request.

**Ethics approval and consent to participate**
All enrolled patients provided written informed consent after receiving detailed explanations. The research protocol was supervised and authorizedby the Capital Medical University Ethics(No.2006003X).

**Consent for publication**
Not applicable.

Liang *et al. BMC Medical Imaging*          (2020) 20:65

Page 8 of 8

**Author details**
[1]The Future Laboratory, Tsinghua University, Chengfu Road, Beijing, China.
[2]Center for Cardiology, Beijing Anzhen Hospital, Capital Medical University,
Anzhen Road, 100029 Beijing, China. [3]The First Affiliated Hospital of
Zhengzhou University, Zhengzhou, 450052, China, 1 Jianshe East Road, Erqi
District, 450052 Zhengzhou, China.

## References

1. Klein AK,  Lee F,  Amini AA. Quantitative coronary angiography with deformable spline models. IEEE Trans Med Imaging. 1997;16(5):468–82. https://doi.org/10.1109/42.640737.
2. Liu I,  Sun Y. Recursive tracking of vascular networking angiograms based on the detection-deletion scheme. IEEE Trans Med Imaging. 1993;12(2): 334–41. https://doi.org/10.1109/42.232264.
3. Quek FKH,  Kirbas C. Vessel extraction in medical images by wave-propagation and traceback. IEEE Trans Med Imaging. 2001;20(2): 117–31. https://doi.org/10.1109/42.913178.
4. Kobashi S,  Kamiura N,  Hata Y,  Miyawak F. Fuzzy information granulation on blood vessel extraction from 3d, tof mra image. Int J Pattern Recognit Artif Intell. 2000;14(4):409–25. https://doi.org/10.1142/S0218001400000271.
5. Kayikcioglu T,  Ganal A,  Turhal M, et al. A surface based method for detection of coronary boundaries in poor quality x-ray angiogram images. Pattern Recogn Lett. 2002;23(7):783–802. https://doi.org/10.1016/s0167-8655(01)00156-8.
6. Eiho S,  Qian Y. Detection of coronary artery tree using morphological operator. Comput Cardiol. 1997;24:525–8. https://doi.org/10.1109/CIC.1997.647950.
7. Bouraoui B,  Ronse C,  Baruthio J,  Passat N,  Germain P. Fully automatic 3d segmentation of coronary arteries based on mathematical morphology. In: International Symposium on Biomedical Imaging: from nano to macro. IEEE; 2008.  p. 1059–62. https://doi.org/10.1109/ISBI.2008.4541182.
8. Chaudhuri S,  Chatterjee S,  Katz N,  Nelson M,  Goldbaum M. Detection of blood vessels in retinal images using two-dimensional matched filters. IEEE Trans Med Imaging. 1989;8(3):263–9. https://doi.org/10.1109/42.34715.
9. Poli R,  Valli G. An algorithm for real-time vessel enhancement and detection. Comput Methods Programs Biomed. 1996;52(1):1–22.
10. Chen W,  Smith R,  Ji SY,  Ward KR,  Najarian K. Automated ventricular systems segmentation in brain ct images by combining low-level segmentation and high-level template matching. Bmc Med Inform Decis Making. 2009;9(1):4. https://doi.org/10.1186/1472-6947-9-S1-S4.
11. Wilkinson MHF,  Westenberg MA. Shape preserving filament enhancement filtering. Int Conf Med Image Comput Comput Assist Interv. 2001;2208:770–7. https://doi.org/10.1007/3-540-45468-3_92.
12. Aganj I,  Harisinghani MG,  Weissleder R,  Fischl B. Unsupervised medical image segmentation based on the local center of mass. Sci Rep. 2018;8(1):13012. https://doi.org/10.1038/s41598-018-31333-5.
13. Tong T,  Wolz R,  Wang Z,  Gao Q,  Misawa K,  Fujiwara M,  Mori K,  Hajnal JV,  Rueckert D. Discriminative dictionary learning for abdominal multi-organ segmentation. Med Image Anal. 2015;23(1):92–104. https://doi.org/10.1016/j.media.2015.04.015.
14. Cheng J,  Liu J,  Xu Y,  Yin F,  Wong TY. Superpixel classification based optic disc and optic cup segmentation for glaucoma screening. IEEE Trans Med Imaging. 2013;32(6):1019–32. https://doi.org/10.1109/TMI.2013.2247770.
15. Abra'moff MD,  Alward WLM,  Greenlee EC,  Shuba L,  Kim CY,  Fingert JH,  Kwon YH. Automated segmentation of the optic disc from stereo color photographs using physiologically plausible features. Investig Ophthalmol Vis Sci. 2007;48(4):1665–73. https://doi.org/10.1167/iovs.06-1081.
16. Krizhevsky A,  Sutskever I,  Hinton G. Imagenet classification with deep convolutional neural networks. Adv Neural Inf Process Syst. 2012;25(2): 1097–105. https://doi.org/10.1145/3065386.
17. Wang Z,  Yin Y,  Shi J,  Fang W,  Li H,  Wang X. Zoom-in-net: Deep mining lesions for diabetic retinopathy detection. In: International Conference on Medical Image Computing and Computer Assisted Intervention. Springer; 2017.  p. 267–75. https://doi.org/10.1007/978-3-319-66179-7_31.
18. Ciresan D,  Giusti A,  Gambardella LM, Schmidhuber. Deep neural networks segment neuronal membranes in electron microscopy images. Adv Neural Inf Process Syst. 2012;25:2843–51.
19. Ronneberger O,  Fischer P,  Brox T. U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer; 2015. p. 234–41. https://doi.org/10.1007/978-3-319-24574-4_28.
20. Fu H,  Cheng J,  Xu Y,  Wong DWK,  Liu J,  Cao X. Joint optic disc and cup segmentation based on multi-label deep network and polar transformation. Trans Med Imaging IEEE. 20181–1. https://doi.org/10.1109/tmi.2018.2791488.
21. Gibson E,  Giganti F,  Hu Y,  Bonmati E,  Bandula S,  Gurusamy K,  Davidson B,  Pereira SP,  Clarkson MJ,  Barratt DC. Automatic multi-organ segmentation on abdominal ct with dense v-networks. IEEE Trans Med Imaging. 2018;37(8):1. https://doi.org/10.1109/TMI.2018.2806309.
22. Zhao H,  Shi J,  Qi X,  Wang X,  Jia J. Pyramid scene parsing network. 2017;1:2881–90. https://doi.org/10.1109/CVPR.2017.660.
23. Gu Z,  Cheng J,  Fu H,  Zhou K,  Hao H,  Zhao Y,  Zhang T,  Gao S,  Liu J. Ce-net: Context encoder network for 2d medical image segmentation. IEEE Trans Med Imaging. 2019;38(10):2281–92. https://doi.org/10.1109/TMI.2019.2903562.
24. Szegedy C,  Ioffe S,  Vanhoucke V,  Alemi A. Inception-v4,inception-resnet and the impact of residual connections on learning. AAAI. 2016;4(12):12.
25. Chen LC,  Papandreou G,  Kokkinos I,  Murphy K,  Yuille AL. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. IEEE Trans Pattern Anal Mach Intell. 2016;40(4):834–48. https://doi.org/10.1109/TPAMI.2017.2699184.
26. Huang TS. Image sequence analysis, vol. 5. Berlin: Springer Verlag; 1981. https://doi.org/10.1007/978-3-642-87037-8.
27. Dubois E,  Sabri S. Noise reduction in image sequences using motion-compensated temporal filtering. IEEE Trans Comm. 1984;32(7): 826–31. https://doi.org/10.1109/TCOM.1984.1096143.
28. Staal J,  Abramoff MD,  Niemeijer M,  Viergever MA,  van Ginneken B. Ridge-based vessel segmentation in color images of the retina. IEEE Trans Med Imaging. 2004;23(4):501–9. https://doi.org/10.1109/tmi.2004.825627.
29. Crum WR,  Camara O,  Hill DL. Generalized overlap measures for evaluation and validation in medical image analysis. IEEE Trans Med Imaging. 2006;25(11):1451–61. https://doi.org/10.1109/TMI.2006.880587.
30. Hoerl AE,  Kennard RW. Ridge regression: Biased estimation for nonorthogonal problems. Technometrics. 1970;12(1):55–67. https://doi.org/10.1080/00401706.1970.10488634.
31. Fu H,  Xu Y,  Lin S,  Wong DWK,  Jiang L. Deepvessel: Retinal vessel segmentation via deep learning and conditional random field. Springer; 2016. p. 132–9. https://doi.org/10.1007/978-3-319-46723-8_16.

## Publisher's Note