Automatic Segmentation of Coronary Lumen and External Elastic Membrane in IVUS Images Using 8-layer U-Net

Liang Dong  
The Second Affiliated Hospital of Zhejiang University

Wei Lu  
ArteryFlow Technology Co., Ltd.

Jun Jiang  
Second Affiliated Hospital, School of Medicine, Zhejiang University

Ya Zhao  
ArteryFlow Technology Co., Ltd.

Xiangfen Song  
ArteryFlow Technology Co., Ltd.

Xiaochang Leng  
ArteryFlow Technology Co., Ltd.

Hang Zhao  
Arteryflow Technology Co., Ltd.

Jian'an Wang  
The Second Affiliated Hospital of Zhejiang University

Changling Li  
The Second Affiliated Hospital of Zhejiang University

Jianping Xiang (jianping.xiang@arteryflow.com)  
ArteryFlow Technology Co., Ltd.

Research

Keywords: IVUS, Coronary, Segmentation, Lumen, EEM, MeshGrid, U-Net

DOI: https://doi.org/10.21203/rs.3.rs-42239/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Background: Intravascular ultrasound (IVUS) is the golden standard in accessing the coronary lesions, stenosis, and atherosclerosis plaques. In this paper, a fully-automatic approach by an 8-layer U-Net is developed to segment the coronary artery lumen and the area bounded by external elastic membrane, i.e. EEM cross section area (EEM-CSA). The database comprises of single-vendor and single-frequency IVUS data. Particularly, the proposed data augmentation of MeshGrid combined with flip and rotation operations is implemented, improving the model performance without pre- or post-processing of the raw IVUS images.

Results: The mean intersection of union (mIoU) of 0.941 and 0.750 for the lumen and EEM-CSA respectively were achieved, which exceeded the manual labeling accuracy of the clinician.

Conclusion: The accuracy shown by the proposed method is sufficient for subsequent reconstruction of 3D IVUS images, which is essential for doctors’ diagnosis in the tissue characterization of coronary artery walls and plaque compositions, qualitatively and quantitatively.

Background

Coronary heart disease has been the leading cause of death worldwide [1], and the coronary atherosclerosis is the dominant cause of coronary heart disease. In early atherosclerosis, coronary artery remodeling slows down the progression of vascular stenosis with the accumulation of coronary plaques. Intravascular ultrasound (IVUS) is one of the most effective real-time medical imaging techniques, which play critical roles in the diagnosis and treatment of coronary heart disease.

2D IVUS images are acquired serially by an IVUS catheter pulling back through the coronary artery and can evaluate arterial distensibility caused by atherosclerotic plaque. The accurate segmentation of lumen and external elastic membrane cross section area (EEM-CSA) from 2D coronary IVUS images that contributes to assess the atherosclerosis plaque and its vulnerability by measuring lumen diameter, plaque eccentricity, plaque burden, etc. has crucial clinical significance. However, it is time-consuming and experience-dependent for doctors to manually delineate the lumen and EEM contours on the 2D IVUS images. A typical IVUS pullback contains more than 3000 images, so an accurate, fast, and fully automatic segmentation of lumen and EEM-CSA is highly desirable but remains a challenging task due to the relative complexity of IVUS images.

Several segmentation techniques and methods in image processing and computer vision have been performed for coronary IVUS images [3–8]. Traditional image processing methods, including graph search, active surfaces, and active contours, were applied to segment IVUS images based on local image properties or global gray level properties. 3D fast marching method [7, 8], incorporating the texture gradient and the gray level gradient, was applied to segment the walls of the coronary artery with an interactive initialization on EEM borders. In recent years, deep learning has been widely applied in the
medical imaging analysis and achieved remarkable results [13–15]. It has been done to detect the lumen and media-adventitia borders in IVUS due to its capabilities in automatic feature extraction [16–18].

In this paper, we describe the development and evaluation of a U-Net [19] based pipeline that automatically segments the lumen and EEM-CSA of 2D IVUS images. The pipeline has two major steps: first, the data augmentation of MeshGrid combined with flip and rotation operations (MeshGrid + Flip + Rotate) is performed on raw IVUS images; second, an 8-layer deep U-Net is used for pixel-level prediction.

Results

Experiments were carried out for segmenting the lumen and EEM-CSA with four augmentation strategies of No Augmentation, Flip + Rotate, MeshGrid, MeshGrid + Flip + Rotate. The contours predicted by the method with MeshGrid + Flip + Rotate augmentation (3rd row of Fig. 3) were in higher agreement with the ground truth (2nd row of Fig. 3) for a range of morphologies, in comparison with those with No Augmentation (4th row of Fig. 3), MeshGrid (5th row of Fig. 3), and Flip + Rotate (6th row of Fig. 3). Particularly obvious difference was shown in the 2nd column of Fig. 3. Table 1 quantifies the segmentation results from the four data augmentation strategies. Compared with the other three data augmentation strategies, the way of MeshGrid + Flip + Rotate presented better segmentation performance in both lumen and EEM-CSA, with MIoU of 0.941 and 0.750, respectively.

Table 1.
The quantitative performance for the four different data augmentation strategies

| Augmentation Strategies    | MIoU (Lumen) | MIoU (EEM-CSA) |
|----------------------------|--------------|----------------|
| No Augmentation            | 0.883        | 0.639          |
| Flip + Rotate              | 0.927        | 0.704          |
| MeshGrid                   | 0.905        | 0.688          |
| MeshGrid + Flip + Rotate   | 0.941        | 0.750          |

Discussion

The IVUS images varied significantly from the intensity gradient of edge of lumen to the contour curvature of plaques. The current dataset was limited to the single-vendor and single-brand. However, the proposed method provided acceptable segmentation results for both lumen and EEM-CSA from the frames in the dataset. On the visual comparison, in the case of the EEM-CSA segmentation, the performance was lower for complex frames while it was comparable good for simple frames; excellent results were seen for lumen segmentation for all cases. The segmentation of bifurcation images might be difficult due to the ambiguous vessel definition. Shadows at the back of calcified plaques also might
cause troubles to train the model. The dataset needs to be enriched in future to test the robustness of the proposed method.

The data augmentation of MeshGrid + Flip + Rotate helped improve the segmentation performance. It was capable to generalize well to eliminate the outliers (2nd column of Fig. 4). Neither pre-processing steps nor post-processing steps were necessary. The model was trained well on the current dataset, which provided MIoU of 0.941 for lumen predictions. However, when the testing set deviated far from the training set, such as serious artifacts, mixture plaques and branch vessels, the accuracy for EEM-CSA became relatively low (MIoU of 0.750). It can be improved largely when more coronary IVUS data of different categories are collected for training in the future.

From clinical prospective, a clinical threshold to assess the quality of the method should be provided by expert physicians to interpret the segmentation feature with complex or simple frames. For example, a fast pullback through the calcified lesion may result in loss of image features, an increase of catheter artifacts and calcified shadows from the echogenicity of the lumen and plaque textures, which is tougher to be annotated even for experienced physicians. The cardiac cycle motion and coronary vessel pulsation due to the variability or arrhythmia of heart rate might push the catheter to touch the vessel and plaque boundaries, which increase the artifacts and motion jitters in the IVUS images.

Our future research is to extend the current dataset to enhance the robustness and generalization of our method presented in this paper. The heterogeneous dataset of IVUS images shall cover different medical centers, different probe frequencies from different vendors. More IVUS image categories from different artery pullback sections and different characteristics should be considered: plaque, bifurcations, branches, shadow artifact, stent, and catheter artifact, etc. Each frame shall be cross-labeled by three expert physicians according to the respective categories to assess the method, which will make it more convincing.

**Conclusion**

In this paper, an 8-layer U-Net is proposed with the data augmentation of MeshGrid + Flip + Rotate, which is specifically fits for the coronary IVUS lumen and EEM-CSA segmentation task. The experimental results show its superiority in segmentation accuracy and efficiency. Furthermore, it provides a good start for the image-based gating to implement 3D IVUS reconstruction when fused with X-ray projections, which allows fluid and dynamic analysis on plaques and vascular walls of coronary arteries.

**Method**

In this section, we first introduce the coronary IVUS dataset used for training and testing. Then, the 8-layer deep U-Net architecture that predicts the masks for the lumen and the EEM-CSA of IVUS images is presented. The training details are described, and the metric for evaluating the proposed method is illustrated.
Dataset and Augmentation

We use the coronary IVUS dataset from The Second Affiliated Hospital of Zhejiang University School of Medicine. It consists of in vivo pullback of coronary artery acquired by the iLab IVUS from Boston Scientific Corporation equipped with the 40-MHz OptiCross catheter. It contains IVUS frames from 30 patients, which are chosen at the end-diastolic cardiac phase in DICOM formats, with the resolution for each frame is 512×512. The dataset is split into two parts, 567 frames for training and 108 frames for testing, respectively. The training set is used for building the deep learning model and the testing set is used to evaluate the model performance.

IVUS images contain catheter, lumen, endothelium, intima, media, external elastic membrane, adventitia, atherosclerosis plaque. The external elastic membrane is usually treated as the borders of media and adventitia. The media is gray or dark as it contains dense smooth muscle. The adventitia is similar to external tissues surrounding the vascular walls. The endothelium and intima are thinner than the lumen and media. Thus, the lumen and EEM-CSA can be manually annotated by experienced physicians as the ground truth for metric evaluation. Each IVUS frame has been manually annotated for the lumen and EEM-CSA in the short-axis view by three clinical experts, daily working with the specific IVUS brand from the Cardiology Department, shown in Figure 1. Each expert is blinded to the other two experts’ annotations and every frame is repeatedly labeled by each of the three experts to ensure the correctness and blindness of the annotations.

The training set comprises of 567 frames, which is not large enough for training a CNN model from scratch. Data augmentation is essential for better performance. The augmentation is twofold and performed online. First, the coronary IVUS raw images and the corresponding ground truth are randomly (1) rotated at angles: 90°, 180° or 270°; (2) flipped up-down or left-right. Secondly, the MeshGrid is added to the raw image at pixel-level, providing the relative location information. Due to the relatively fixed position like intima and adventitia in IVUS images, MeshGrid could play a good guiding role in training process, which guides the neural network where to look. The ground truth makes no changes.

Model Architectures

The U-Net is one type of the fully convolutional network (FCN) [20] and is the most common convolutional network architecture for biomedical image segmentation. It consists of encoder and decoder parts and predicts segmentation mask at pixel-level instead of image-level classification. The encoder part is used for down-sampling and extracts higher-level features. The decoder part is used for up-sampling the output from the encoder part and concatenates the feature maps of the corresponding layer by skip-connection. The skip-connection is to relieve the gradient diffusion problem due to deep layers. The final decoder layer is activated by softmax to produce the class probability map to recover the segment predictions.

The encoder part has 9 blocks and each incorporates two repeated operations of 3×3 convolution, batch normalization (BN) and LeakyReLU activation. The down-sampling operation of 3×3 convolution with
stride 2×2 reduces feature maps by half. The size of the 8\textsuperscript{th} block is 2×2 to capture the deeper abstract information. The decoder part has 8 blocks to restore the image dimension. Each up-sampling operation contains a 5×5 deconvolution with stride 2ni. The skip-connection concatenates the corresponding feature maps. The last convolution of 1co outputs the probability map of mask class prediction by softmax activation. The entire architecture is shown in Figure 2.

**Implementation Details**

The model was trained and evaluated on Dell PowerEdge T640 server with Xeon Silver 4114 processor, 128GB of RAM, and four Nvidia GTX 1080Ti graphics cards. It took less than 90 minutes for training and 10ms per image for inference.

The deep learning framework used in this study was Tensorflow 1.13. The optimizer was Adam \[21\], which was fast and robust. The weights were initialized randomly and the batch size was set to 16. The initial learning rate was 0.001 with the decay of 0.1 every 2000 iterations. A total of 8000~10000 iterations were done for training. Lumen and EEM-CSA were trained and predicted at one shot with the softmax function as the output activation, which gave each pixel its class probability. The loss function was the sparse softmax cross entropy:

\[
L(p_{\hat{j}}, p) = -\sum_{j=1}^{K} p_{\hat{j}j} \log(p_j)
\]

\[\text{(1)}\]

\[
p_j = \text{softmax}(x_j) = \frac{e^{x_j}}{\sum_{k=1}^{K} e^{x_k}}
\]

\[\text{(2)}\]

with $K$ being the number of classes, $p_{\hat{j}j}$ being the predicted probability belonging to class $j$, $p_{yj}$ being the true probability.

**Evaluation Criteria**

In semantic segmentation, the mean intersection over union (MIoU) is the standard metric to evaluate the model, alternatively called Jaccard measure (JM). We compute the MIoU score between the ground truth and the predicted masks:

\[
\text{MIoU} = \frac{1}{k+1} \sum_{i=0}^{k} \frac{p_{ii}}{\sum_{j=0}^{k} p_{ij} + \sum_{j=0}^{k} p_{ji} - p_{ii}}
\]

\[\text{(3)}\]
with \( k \) being the number of classes excluding background, \( p_{ij} \) being the number of pixels of class \( i \) predicted to class \( j \).

**Abbreviations**

**3D**: Three-dimensional  
**IVUS**: Intravascular ultrasound  
**EEM-CSA**: External Elastic Membrane Cross Section Area  
**mIoU**: mean Intersection of Union  
**FCN**: fully convolutional network  
**JM**: Jaccard Measure

**Declarations**

**Acknowledgements**

The authors thanks the Second Affiliated Hospital of Zhejiang University for providing the IVUS datasets.

**Funding**

This study was supported by Zhejiang Provincial Public Welfare Technology Research Project (NO. LGF20H020012).

**Ethics approval and consent to participate**

This study protocol was approved by the Second Affiliated Hospital of Zhejiang University. We followed CONSORT guideline to perform this study.

**Consent for publication**

Not applicable.

**Competing interests**

No benefits in any form have been received or will be received from a commercial party related directly or indirectly to the subject of this article.

**Authors' contributions**

Liang Dong: Project administration; Wei Lu: Original draft writing; Jun Jiang: Data curation; Ya Zhao: Original draft writing; Xiangfen Song: Formal analysis and paper revise; Xiaochang Leng: Data
Availability of data and materials

The datasets analyzed during the current study are available from the corresponding author on reasonable request.

References

1. T. Vos, A. A. Abajobir, K. H. Abate, C. Abbafati, K. M. Abbas, F. Abd-Allah, et al., "Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016," The Lancet, vol. 390, pp. 1211-1259, 2017.

2. J.-Y. Hahn, Y. Honda, and P. J. Fitzgerald, "Intravascular Ultrasound," in Catheter-Based Cardiovascular Interventions, ed: Springer, 2013, pp. 325-348.

3. M. Sonka, X. Zhang, M. Siebes, M. S. Bissing, S. C. DeJong, S. M. Collins, et al., "Segmentation of intravascular ultrasound images: A knowledge-based approach," IEEE Transactions on Medical Imaging, vol. 14, pp. 719-732, 1995.

4. R. Shekhar, R. Cothren, D. Vince, S. Chandra, J. Thomas, and J. Cornhill, "Three-dimensional segmentation of luminal and adventitial borders in serial intravascular ultrasound images," Computerized Medical Imaging and Graphics, vol. 23, pp. 299-309, 1999.

5. M. E. Plissiti, D. I. Fotiadis, L. K. Michalis, and G. E. Bozios, "An automated method for lumen and media-adventitia border detection in a sequence of IVUS frames," IEEE Transactions on Information Technology in Biomedicine, vol. 8, pp. 131-141, 2004.

6. R. Downe, A. Wahle, T. Kovarnik, H. Skalicka, J. Lopez, J. Horak, et al., "Segmentation of intravascular ultrasound images using graph search and a novel cost function," in Proc. 2nd MICCAI workshop on computer vision for intravascular and intracardiac imaging, 2008, pp. 71-79.

7. M.-H. Cardinal, J. Meunier, G. Soulez, R. L. Maurice, É. Therasse, and G. Cloutier, "Intravascular ultrasound image segmentation: a three-dimensional fast-marching method based on gray level distributions," IEEE transactions on medical imaging, vol. 25, pp. 590-601, 2006.

8. M. H. R. Cardinal, G. Soulez, J. C. Tardif, J. Meunier, and G. Cloutier, "Fast, J. C. Tardif, J. Me Meation: a three-dimensional fast-marching method based on gray level distributions," 1990 Medical physics, vol. 37, pp. 3633-3647, 2010.

9. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in neural information processing systems, 2012, pp. 1097-1105.

10. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014.
11. M. E. Olszewski, A. Wahle, S. C. Mitchell, and M. Sonka, "Segmentation of intravascular ultrasound images: a machine learning approach mimicking human vision," in *International congress series*, 2004, pp. 1045-1049.

12. A. Katouzian, B. Baseri, E. E. Konofagou, and A. F. Laine, "Automatic detection of blood versus non-blood regions on intravascular ultrasound (IVUS) images using wavelet packet signatures," in *Medical imaging 2008: Ultrasonic imaging and signal processing*, 2008, p. 69200H.

13. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, pp. 436-444, 2015.

14. D. Shen, G. Wu, and H.-I. Suk, "Deep learning in medical image analysis," *Annual review of biomedical engineering*, vol. 19, pp. 221-248, 2017.

15. G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, et al., "A survey on deep learning in medical image analysis," *Medical image analysis*, vol. 42, pp. 60-88, 2017.

16. S. Su, Z. Gao, H. Zhang, Q. Lin, W. K. Hau, and S. Li, "Detection of lumen and media-adventitia borders in IVUS images using sparse auto-encoder neural network," in *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, 2017, pp. 1120-1124.

17. S. Su, Z. Hu, Q. Lin, W. K. Hau, Z. Gao, and H. Zhang, "An artificial neural network method for lumen and media-adventitia border detection in IVUS," *Computerized Medical Imaging and Graphics*, vol. 57, pp. 29-39, 2017.

18. J. Yang, L. Tong, M. Faraji, and A. Basu, "IVUS-Net: an intravascular ultrasound segmentation network," in *International Conference on Smart Multimedia*, 2018, pp. 367-377.

19. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *International Conference on Medical Image Computing & Computer-assisted Intervention*, 2015.

20. J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 39, pp. 640-651, 2014.

21. D. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," *Computer Science*, 2014.

**Figures**
Figure 1

Ground truth labeling for lumen and EEM-CSA

Figure 2

The proposed U-Net Architecture with eight layers
Figure 3

Result comparison for four data augmentation strategies.