A Deep Learning-based Method to Extract Lumen and Media-Adventitia in Intravascular Ultrasound Images

Fubao Zhu1, Zhengyuan Gao1, Chen Zhao2, Hanlei Zhu1, Jiaofen Nan1, Yanhui Tian1, Yong Dong3, Jingfeng Jiang4, Xiaohong Feng5, Neng Dai6,7, and Weihua Zhou2,8

Abstract
Intravascular ultrasound (IVUS) imaging allows direct visualization of the coronary vessel wall and is suitable for assessing atherosclerosis and the degree of stenosis. Accurate segmentation and lumen and median-adventitia (MA) measurements from IVUS are essential for such a successful clinical evaluation. However, current automated segmentation by commercial software relies on manual corrections, which is time-consuming and user-dependent. We aim to develop a deep learning-based method using an encoder-decoder deep architecture to automatically and accurately extract both lumen and MA border. Inspired by the dual-path design of the state-of-the-art model IVUS-Net, our method named IVUS-U-Net++ achieved an extension of the U-Net++ model. More specifically, a feature pyramid network was added to the U-Net++ model, enabling the utilization of feature maps at different scales. Following the segmentation, the Pearson correlation and Bland-Altman analyses were performed to evaluate the correlations of 12 clinical parameters measured from our segmentation results and the ground truth. A dataset with 1746 IVUS images from 18 patients was used for training and testing. Our segmentation model at the patient level achieved a Jaccard measure (JM) of 0.9080 ± 0.0321 and a Hausdorff distance (HD) of 0.1484 ± 0.1584 mm for the lumen border; it achieved a JM of 0.9199 ± 0.0370 and an HD of 0.1781 ± 0.1906 mm for the MA border. The 12 clinical parameters measured from our segmentation results agreed well with those from the ground truth (all p-values are smaller than .01). Our proposed method shows great promise for its clinical use in IVUS segmentation.

Keywords
intravascular ultrasound, deep learning, U-Net++, feature pyramid, segmentation

Introduction
Atherosclerosis is a vessel wall disease and is responsible for many fatal cardiovascular diseases.1 Compared with the in vitro screening, the widespread application of the intravascular ultrasound (IVUS) technique relies on its capability to visualize the inner structure and the blood flow in real-time to diagnose the atherosclerotic disease of the coronary artery.2 It can assess quantitative clinical measurements, such as the lumen cross-sectional area (CSA), external elastic membrane (EEM) CSA, plaque plus media CSA, and other relevant structural information of the coronary artery.3 However, accurate extraction of the lumen and median-adventitia (MA) border is essential for assessing plaque volume and stenosis degree. The current clinical practice relies on automated segmentation in the IVUS frames by commercial software. However, the accuracy is limited, and manual corrections are frequently needed, which is time-consuming and user-dependent.

Traditional image processing techniques and deep learning approaches have been investigated to extract the lumen and MA border automatically. Unal et al.4 explored the method based on the use of shape and intensity priors. Zhu et al.5 exploited the gradient vector flow in a nonparametric energy function to detect the target’s border. Considering different characteristics of the imaging caused by radio frequency signals, Mendizabal-Ruiz and Kakadiaris6 proposed a physics-based IVUS image reconstruction method for lumen segmentation.

Recently, several techniques based on deep learning have been proposed for IVUS segmentation. Su et al.7 proposed a coding method that used an artificial neural network (ANN) to determine if the pixel was located in the border. Based on the U-Net,8 Yang et al.9 designed an IVUS-Net model and a Dual-Path U-Net model10 for lumen and MA segmentation. However, all of the existing approaches based on U-Net all adopted a symmetric network structure with the skip-connection between
two blocks in the same deep layers with the same size. Hence, feature information at different spatial scales cannot be effectively fused and propagated over the entire network.

To incorporate the multiscale feature information, in this paper, we proposed a novel method (IVUS-U-Net++) that used nested and dense connections\(^\text{11}\) to extract the lumen and MA borders with the feature pyramid network.\(^\text{12}\) The proposed IVUS-U-Net++ model was evaluated with patient data from our internal imaging database. Three state-of-the-art models, including U-Net, IVUS-Net and U-Net++ were selected for the comparisons.

**Methodology**

*Image Acquisition and Description*

In this study, all IVUS data, including 1746 cross-sectional images from 18 patients, were collected from the 7th People’s Hospital of Zhengzhou between February 2020 and March 2020. Each patient was scanned using a commercial IVUS machine equipped with a single element transducer pulsed at 40 MHz (EagleEye, Volcano Corporation, Cordova, CA, USA). All scans were performed from the target lesion’s distal end and pulled back to the proximal end at a fixed speed (0.5 mm per second). An experienced IVUS analyst selected the range of slices containing the target lesion and then manually annotated the lumen and MA borders on these images. As shown in Figure 1, the yellow and red contours represent the lumen and MA borders. The plaque is located between the lumen and MA. Of note, the study was approved by the Ethics of 7th People’s Hospital of Zhengzhou. Subjects provided written informed consent for scan and participation in anonymized analyses.

*Segmentation Network*

Figure 2 shows the classical symmetrical neural network architecture with the skip-connection of the U-Net model. In the down-sampling path, the features are captured from the images through the convolutional operators automatically. With the transpose convolution adopted in the up-sampling path, the spatial structure of the feature maps is restored to the original scale gradually. In the IVUS-Net,\(^\text{9}\) the dual-path convolution blocks, which are similar to Google’s inception architecture,\(^\text{13}\) are employed as the convolutional blocks. This dual-path block combined with two different scale information is shown in Figure 3. As a supplement feature, the refined feature map was concatenated to the main branch, which led to a better extraction than the normal deep learning method.

Although the multiscale feature representation and fusion capability have been enhanced to a certain extent through the dual-path design from IVUS-Net, a heterogeneous combination of feature maps at different spatial scales is designed to better integrate the multiscale information. As shown in Figure 4, the architecture of our deep neural network IVUS-U-Net++ incorporates two parts. One is the U-net++ network, which leads to a highly flexible feature fusion scheme to integrate different-scale features from different-depth encoder layers. The other is the feature pyramid network.

Figure 1. An example showing annotated boundaries of the lumen (yellow) and MA (red) by a radiologist.

---

1School of Computer and Communication Engineering, Zhengzhou University of Light Industry, Zhengzhou, Henan, China
2Department of Applied Computing, Michigan Technological University, Houghton, MI, USA
3Department of Cardiology, The 7th People’s Hospital of Zhengzhou, Zhengzhou, Henan, China
4Department of Biomedical Engineering, Michigan Technological University, Houghton, MI, USA
5Department of Pediatrics, Yicheng Maternity and Child Health Care Hospital, Yicheng, Hubei, China
6Department of Cardiology, Zhongshan Hospital, Fudan University, Shanghai Institute of Cardiovascular Diseases, Shanghai, China
7National Clinical Research Center for Interventional Medicine, Shanghai, China
8Center of Biocomputing and Digital Health, Institute of Computing and Cybersystems, and Health Research Institute, Michigan Technological University, Houghton, MI, USA

**Corresponding Authors:**
Neng Dai, MD, PhD, Department of Cardiology, Zhongshan Hospital, Fudan University, Shanghai Institute of Cardiovascular Diseases, National Clinical Research Center for Interventional Medicine, 180 Fenglin Road, Xuhui District, Shanghai 200032, China.
Email: niceday1987@hotmail.com

Weihua Zhou, PhD, Department of Applied Computing, Michigan Technological University, 1400 Townsend Dr, Houghton, MI 49931, USA.
Email: whzhou@mtu.edu
which is able to synthesize multiscale feature maps to the final output from different-depth decoder layers. Of note, upscale operators are adopted to ensure size compatibility of the feature maps during the fusion of multiscale features. Our goal is to enhance the multiscale feature representation and fusion. For example, the feature map from the convolution block (0,3) (see Figure 4) is shown in equation (1).

\[
F(i = 0, j = 3) = F(0,0) + F(0,1) + F(0,2) + \text{Tr}(F(1,2))
\]

(1)

where the \(F(i, j)\) represents the feature map of the convolution block \((i, j)\) and the \(\text{Tr}\) is the convolutional transpose operator.

Since the encoder and decoder network modules and the feature pyramid network are used in this study, it is necessary to optimize the network configuration, including depths and fusion strategy. Based on our network structure, we tested the U-net++ with three different depths for as the backbone and four fusion strategies (as shown in Figure 5): fusion by Concatenate (Con), fusion by Add, fusion by Layer-Concatenate (LyCon) and fusion by Layer-Add (LyAdd). After the fusion, we used the final convolution and softmax classifier to obtain the model output.

Several strategies are used to accelerate the training process. First, to overcome the gradient vanishing problem, pre-trained weights are used to initialize the weights of the

Figure 2. A schematic diagram showing a symmetric network model with the skip-connection.

Figure 3. A diagram illustrating convolution blocks used in the IVUS-Net. This dual-path avoids information loss due to the pooling operations and produces a refined feature map to the main branch. PReLU = parametric rectified linear unit; BN = batch normalization.
backbone. Second, batch normalization (BN) and ReLU activation function are applied after each convolution layer to accelerate network training and guarantee the non-linearity of the network. Third, to fully utilize the information extracted from different scales, five resized feature maps are concatenated as the feature map of the convolution block (0,5) and employed for supervised learning. By using this feature pyramid network, the final probability map is generated by a voting mechanism with this parallel connection.

**Post-processing**

The proposed IVUS-U-Net++ model is trained to solve a binary classification problem. In the prediction stage, the
output from the proposed IVUS-U-Net++ model is a probability map ranging from 0 to 1. Using a self-adapting threshold, OSTU\(^{16}\) as equation (2), the pixels with high probability in the probability map are converted to the foreground, and the pixels with low probability indicate the background. Thus, a binary map is obtained. As the anatomical target only occupies one connected region, only the largest connected region from the binary map is retained, and all other regions are removed\(^{17}\) after the flood fill technique.

\[
\text{dst}(i, j) = \begin{cases} 
1 & \text{if src}(i, j) > \text{thresh} \\
0 & \text{otherwise}
\end{cases}
\]  \hspace{1cm} (2)

**Model Training**

A loss function based on the Dice similarity coefficient (DSC) is used as our optimization function. It penalizes the difference between the prediction output and the ground truth. The definition is shown in equation (3).

\[
\text{DSC-Loss}(R, R') = 1 - \frac{2|R_y \cap R'_y|}{|R_y| + |R'_y|} \times 100\%
\]  \hspace{1cm} (3)

where \(R'_y\) represents the ground truth in pixel \((i, j)\). \(R_y\) indicates the prediction output of the network corresponding to the \(R'_y\).

We trained the models on a Linux machine equipped with a single Tesla P100 GPU with 16 GB GPU memory. Our method was implemented by Keras 2.2.24 with Cuda 9.0. In addition, in the training stage, an RMSprop optimizer\(^{18}\) was used to optimize the model weights with the learning rate of 0.0001 and the number of epochs set to 201. To accelerate the convergence, the pre-trained weights in ImageNet\(^{19}\) were employed as the backbone of the encoder. Of note, as there were no pre-trained weights for IVUS-Net according to the paper,\(^8\) the IVUS-Net was trained from scratch.

Restricted by the number of patients, the entire dataset containing 1746 images from 18 patients was randomly split into the training data with 1572 images for the 10-fold cross-validation and the test data with 174 images for evaluating the performance of models on the slice level. In addition, a leave-one-out cross-validation on the patient level was employed to further verify our method’s robustness. The data augmentation technique, including rotation, scale transformation, and elastic transformation, was also consistently applied to increase the sample size of the training dataset and mitigate the overfitting problem.

**Evaluation Metrics**

To evaluate the effectiveness of the model, the Jaccard Measure (JM) is used to measure the consistency between the predicted binary map and the ground truth. The definition of JM is shown as:

\[
\text{JM}(R, R') = \frac{|R \cap R'|}{|R \cup R'|} \times 100\%
\]  \hspace{1cm} (4)

where the \(R\) is the predicted binary map after post-processing, and the \(R'\) is the ground truth.

The Hausdorff Distance (HD) is used as an evaluation metric as well. HD measures the largest distance between all points belonging to the border of the predicted binary map after post-processing \((CR)\) to the closest point in the corresponding border of the ground truth \((CR')\). And the \(ps\) indicates the pixel spacing of the image. HD is defined as in equation (5):

\[
\text{HD}(CR, CR') = \text{MAX} \{d(CR, CR'), d(CR', CR)\} \times ps
\]  \hspace{1cm} (5)

In addition, to further evaluate the clinical accuracy, 12 clinical parameters were calculated according to guidelines from the American College of Cardiology consensus statement on IVUS.\(^{20}\) The detailed clinical parameters are as follows.

- Maximum EEM diameter: The longest diameter through the center of the mass of the MA.
- Minimum EEM diameter: The shortest diameter through the center of the mass of the MA.
- EEM CSA: The area bounded by the MA border.
- Maximum lumen diameter: The largest diameter through the center of the mass of the lumen.
- Minimum lumen diameter: The shortest diameter through the center of the mass of the lumen.
- Lumen CSA: The area bounded by the lumen border.
- Lumen eccentricity: \([(\text{maximum lumen diameter} - \text{minimum lumen diameter}) / \text{maximum lumen diameter}]\).
- Maximum plaque plus media thickness: The largest distance from the intimal leading edge to the EEM along with any line passing through the luminal center of mass.
- Minimum plaque plus media thickness: The shortest distance from the intimal leading edge to the EEM along any line passing through the center of mass of the lumen.
- Plaque plus media CSA: (The EEM CSA – the lumen CSA.)
- Plaque plus media eccentricity: (Maximum plaque plus media thickness – minimum plaque plus media thickness)/maximum plaque plus media thickness.
- Plaque burden: Plaque plus media CSA/EEM CSA.

The longest or shortest diameter is obtained by searching all lines through the center point of mass and the area bounded by the border is calculated by counting the pixels inside the border.
Results

Optimization of Model Architecture

As shown in Figure 6, the fusion strategy with a Concatenate operation obtained a better result than other fusion strategies, and the optimal depth of the backbone is 5. As the depth increases, the capability of deep feature extraction is enhanced; however, when the depth of the model is too large, the gradient vanishing occurred and the model performance is decreased. Compared to the concatenate operator, the computation compression and spatial dimension by the add operator may lead to the information loss and the noise generation, thus influencing the performance.21

Extraction and Quantitation

To determine the optimal model and verify the robustness of the models, all the models from the 10-fold cross-validation were evaluated with the test dataset. The results per the slice level can be seen in Table 1. The model at the slice level achieved a JM of 0.9355 ± 0.0035 and an HD of 0.2720 ± 0.3115 for the lumen border; it achieved a JM of 0.9443 ± 0.0053 and an HD of 0.1289 ± 0.0311 for the MA border. Table 2 demonstrates the experiment results of the leave-one-out cross-validation at the patient level. The model per the patient level achieved a JM of 0.9080 ± 0.3021, an HD of 0.1484 ± 0.1584 mm for the lumen border, and a JM of 0.9199 ± 0.0370, an HD of 0.1781 ± 0.1906 mm for the MA border, respectively. Since the adjacent slices had high similarity in the same patient, better results were seen per the slice level than per the patient level.

The U-Net, IVUS-Net, and U-Net++ were used as the benchmark models. Table 3 compares the number of weights, time cost per slice, JMs, and HDs by different segmentation models for the test dataset. For both the lumen and the MA border, IVUS-U-Net++ achieved the best JM and the best HD compared with U-net, IVUS-Net, and Unet++. It took 0.091 ± 0.0578 (min 0.014, max 0.314) seconds for our IVUS-U-Net++ model to predict the segmentation result for each slice. The use of pre-trained weights led to better performance for U-Net, U-Net++ and IVUS-U-Net++ compared to IVUS-Net.

Figures 7 and 8 compare segmentation results from the above-mentioned four segmentation methods/groups: U-Net,
Table 2. Performance of the IVUS-U-Net++ Models Generated From the Leave-One-Out Cross-Validation Over the Test Dataset at the Patient Level.

| Fold | Lumen JM | Lumen HD (mm) | Media-adventitia JM | Media-adventitia HD (mm) |
|------|----------|---------------|---------------------|-------------------------|
| 1    | 0.9140   | 0.2006        | 0.9407              | 0.1019                  |
| 2    | 0.9136   | 0.0777        | 0.9333              | 0.1335                  |
| 3    | 0.9411   | 0.2110        | 0.8961              | 0.1804                  |
| 4    | 0.8826   | 0.0909        | 0.8868              | 0.2969                  |
| 5    | 0.9391   | 0.9494        | 0.9672              | 0.0731                  |
| 6    | 0.9333   | 0.1189        | 0.8569              | 0.1788                  |
| 7    | 0.9356   | 0.7113        | 0.9515              | 0.1220                  |
| 8    | 0.9176   | 0.8801        | 0.9250              | 0.1082                  |
| 9    | 0.9346   | 0.9356        | 0.8868              | 0.1505                  |
| 10   | 0.9373   | 0.9373        | 0.9437              | 0.0845                  |
| 11   | 0.9285   | 0.1182        | 0.9176              | 0.1121                  |
| 12   | 0.8992   | 0.0881        | 0.9424              | 0.8632                  |
| 13   | 0.9284   | 0.0710        | 0.9468              | 0.0803                  |
| 14   | 0.8856   | 0.1611        | 0.9356              | 0.1121                  |
| 15   | 0.8303   | 0.1682        | 0.8371              | 0.1546                  |
| 16   | 0.8620   | 0.0871        | 0.9513              | 0.0977                  |
| Mean | 0.9080   | 0.1484        | 0.9199              | 0.1781                  |
| Standard deviation | 0.0321 | 0.1584 | 0.0370 | 0.1906 |

The quantitative results are evaluated by Jaccard Measure (JM).

Table 3. Comparison of Different Segmentation Models Generated From the 10-Fold Cross-Validation Over the Test Dataset at the Slice Level.

| U-Net | IVUS-Net | U-Net++ | IVUS-U-Net++ |
|-------|----------|---------|--------------|
| Weights (M) | 11.6 | 66.2 | 255.9 | 255.9 |
| Prediction time per slice (s) | 0.091 ± 0.0742 | 0.117 ± 0.500 | 0.091 ± 0.0575 | 0.091 ± 0.0578 |
| Lumen JM | 0.9252 ± 0.0067 | 0.9187 ± 0.0116 | 0.9294 ± 0.0076 | 0.9355 ± 0.0035 |
| Lumen HD (mm) | 0.4368 ± 0.4826 | 1.6427 ± 0.7638 | 0.6726 ± 0.4807 | 0.2720 ± 0.3115 |
| Media-adventitia JM | 0.9338 ± 0.0072 | 0.9272 ± 0.0064 | 0.9351 ± 0.0084 | 0.9443 ± 0.0053 |
| Media-adventitia HD (mm) | 0.6003 ± 0.3844 | 1.8366 ± 0.6791 | 0.3856 ± 0.0763 | 0.1289 ± 0.0311 |

The evaluation measures are the number of weights, prediction time per slice, Jaccard Measure (JM), and Hausdorff Distance (HD). Data are shown as mean ± standard deviation.

IVUS-Net, U-Net++ and IVUS-U-Net++, respectively. The visual comparisons are listed in terms of the lumen and MA segmentation results using the same image slice.

The consistency and discrepancy between the commonly used clinical parameters measured from the ground truth and those from our IVUS-U-Net++ model were further evaluated to validate the accuracy of our model. Table 4 shows $R$ (measured by the Pearson correlation analysis with IBM SPSS Statistics 20), mean absolute error, root mean square error and relative error. It can be observed that there is a high correlation with statistical significance for all the 12 clinical parameters. The scatter plots (Figure 9) and Bland-Altman analysis (Figure 10) further visualize the linear relationship and the consistency for evaluating those clinical parameters. Figure 9 reflects the excellent linear correlations between them from the prediction and the ground truth. Figure 10 shows that with the 95% limits of agreement, most of the points are within the upper and lower consistency interval.

**Discussion**

In the current study, a novel segmentation method has been proposed to extract the lumen and MA border in IVUS images. Pyramid feature extraction is added to the known U-Net++ model. The overall JMs and HDs at the patient level were 0.9080 ± 0.0321 and 0.1484 ± 0.1584 mm for the lumen border segmentation, and 0.9199 ± 0.0370 and 0.1781 ± 0.1906 mm for the MA border segmentation, respectively. There was an excellent agreement between the clinical parameters from our prediction results by
When compared with the other three state-of-the-art models at the slice level, IVUS-U-Net++ achieved the best JM and the best HD for both the lumen and MA border, as shown in Table 5; however, it only provides a limited comparison due to the dataset difference. In addition, our model could predict the segmentation result in \(0.1203 \pm 0.0046\) seconds for a single slice. It significantly reduces the annotation burden and has excellent potential for clinical use.

IVUS-U-Net++ shows significant advantages in retaining the global contours because of the multiscale feature incorporation. It can be seen that, in Figures 7(d), (h), 8(d), (h), and (l), there are significant differences in the global contour shapes of lumen and MA between U-Net/ IVUS-Net/ U-Net++ and the ground truth. In contrast, IVUS-U-Net++ captures the global shapes with a consistent accuracy as shown in both Figures 7 and 8. This is applicable even for the slices with the lowest JMs generated by IVUS-U-Net++ (Figures 7(p) and 8(p)). In addition, by incorporating the multiscale features, U-Net++ and IVUS-U-Net++ generated smoother borders than those of U-Net and IVUS-Net. This illustrates that by applying the pyramid feature-related

---

**Figure 7.** Segmentation results for the lumen border in the test dataset. The red contours represent the ground truth and the green ones represent the segmentation results from U-Net (a, e, i, m), IVUS-Net (b, f, j, n), U-Net++ (c, j, k, o), or IVUS-U-Net++ (d, h, l, p). Figures (a, f, k, and p) (i.e., the diagonal direction) show the slices with the lowest JMs generated by U-Net, IVUS-Net, U-Net++ and IVUS-U-Net++, respectively.
Figure 8. Segmentation results for the MA border in the test dataset. The red contours represent the ground truth and the green ones represent the segmentation results from U-Net (a, e, i, m), IVUS-Net (b, f, j, n), U-Net++ (c, j, k, o), or IVUS-U-Net++ (d, h, l, p). Figures (a, f, k, and p) (i.e., the diagonal direction) show the slices with the lowest JMs generated by U-Net, IVUS-Net, U-Net++ and IVUS-U-Net++, respectively.

Table 4. Correlation and Discrepancy Between the Clinical Parameters Measured From the Ground Truth and Those From Our IVUS-U-Net++ Model.

| Target                | Metric       | R       | MAE       | RMSE       | RE (%) (min, max) |
|-----------------------|--------------|---------|-----------|------------|-------------------|
| Maximum EEM diameter  | .984         | 0.0806 (mm) | 0.1131 (mm) | (−0.0713, 0.1055) |
| Minimum EEM diameter  | .991         | 0.0737 (mm) | 0.0954 (mm) | (−0.0597, 0.1107) |
| EEM CSA               | .992         | 0.4000 (mm²) | 0.5473 (mm²) | (−0.1573, 0.1362) |
| Maximum lumen diameter| .986         | 0.0682 (mm) | 0.1002 (mm) | (−0.1166, 0.0670) |
| Minimum lumen diameter| .985         | 0.0779 (mm) | 0.1099 (mm) | (−0.0850, 0.2660) |
| Lumen CSA             | .995         | 0.1845 (mm²) | 0.2482 (mm²) | (−0.1224, 0.1965) |
| Lumen eccentricity    | .858         | 0.0347 | 0.0466 | (−0.6270, 1.1489) |

(continued)
Table 4. (continued)

| Target                                | R    | MAE     | RMSE      | RE (%) (min, max) |
|----------------------------------------|------|---------|-----------|-------------------|
| Maximum plaque plus media thickness    | .975 | 0.0805 (mm) | 0.1074 (mm) | (−0.1507, 0.3494) |
| Minimum plaque plus media thickness    | .948 | 0.0514 (mm) | 0.0698 (mm) | (−0.7764, 7.0623) |
| Plaque plus media CSA                  | .981 | 0.4412 (mm²) | 0.5831 (mm²) | (−0.3145, 0.1818) |
| Plaque plus media eccentricity         | .955 | 0.0420   | 0.058     | (−0.3880, 0.3105) |
| Plaque burden                          | .976 | 0.0163   | 0.0221    | (−0.1865, 0.1108) |
| Mean                                   | .969 | 0.1292   | 0.1744    | -                 |

MAE = mean absolute error; RMSE = root mean square! error; RE = relative error. For $R$ measures, all $p < .01$. $p < .001$ for the $R$-value in all the listed clinical parameters.

![Figure 9. The scatter plots to visualize the linear correlations between clinical parameters measured from the ground truth (horizontal axis) and those from the predicted results by IVUS-U-Net++ (vertical axis).](image-url)
Table 5. Comparison of the Existing Methods With The Proposed Approach for Lumen and MA Segmentation.

| Data Level | JM       | HD (mm) | JM       | HD (mm) |
|------------|----------|---------|----------|---------|
| Faraji et al. | 0.87 ± 0.06 | 0.30 ± 0.20 | 0.77 ± 0.17 | 0.67 ± 0.54 |
| Downe et al. | 0.77 ± 0.09 | 0.47 ± 0.22 | 0.74 ± 0.17 | 0.76 ± 0.48 |
| Exarchos et al. | 0.81 ± 0.99 | 0.42 ± 0.22 | 0.79 ± 0.11 | 0.60 ± 0.28 |
| Yi et al. | 0.77 ± 0.09 | 1.21 ± 0.68 | 0.83 ± 0.08 | 1.13 ± 0.69 |
| Su et al. | 0.92 | 0.22 | 0.91 | 0.32 |
| Proposed | 0.91 ± 0.03 | 0.15 ± 0.16 | 0.92 ± 0.04 | 0.18 ± 0.19 |

The evaluation measures are Jaccard Measure (JM) and Hausdorff Distance (HD). Data are shown as mean ± standard deviation.

Figure 10. Bland-Altman plots (difference against average) between clinical parameters measured from the ground truth and those from the predicted results by IVUS-U-Net++. (a) Maximum EEM diameter. (b) Minimum EEM diameter. (c) EEM CSA. (d) Maximum lumen diameter. (e) Minimum lumen diameter. (f) Lumen CSA. (g) Lumen eccentricity. (h) Maximum plaque plus media thickness. (i) Minimum plaque plus media thickness. (j) Plaque plus media CSA. (k) Plaque plus media eccentricity. (l) Plaque burden.

techniques, the receptive field was significantly enlarged. However, it is noteworthy that with the extension of U-Net++, IVUS-U-Net++ not only produced smooth borders which were similar to the radiologist’s annotation but also achieved the highest segmentation accuracy.

The accuracy of our IVUS-U-Net++ model was further validated by the measurement of clinical parameters. There is a statistical significance (p-values < .01) in all the 12 clinical parameters measured from the segmentation results from our model compared with the ground truth. A correlation of
Conclusion

A new method based on deep learning has been proposed to extract the lumen and MA borders from IVUS images. The quantitative evaluations in this feasibility study demonstrate that our method has high accuracy and good robustness. It produces smooth contours with high computational efficiency.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by grants from Henan Science and Technology Development Plan 2020 (Project Number: 202102210384), 2019 Maker Space Incubation Project of Zhengzhou University of Light Industry (Project Number: 2019ZCKJ228), and Key Scientific Research Projects of Colleges and Universities in Henan Province (Project Number: 20A510014). This research was also supported in part by grants from the National Natural Science Foundation of China (Project Number: 81600279), from China Postdoctoral Science Foundation (Project Number: 2019M650075), and Pilot Project for Disruptive Technology (Biotechnology) from China National Center for Biotechnology Development (Project Number: 2020YFC1316700).

ORCID iDs

Jingfeng Jiang https://orcid.org/0000-0001-8812-6246
Weihua Zhou https://orcid.org/0000-0002-6039-959X

References

1. Mahtta D, Mahmoud AN, Mojadidi MK, Elgendy IY. Intravascular ultrasound-guided percutaneous coronary intervention: an updated review. Cardiovasc Innov Appl. 2018;3(2):127-36.
2. Wali E, Nathan S. What is the clinical utility of intravascular ultrasound? Curr Cardiol Rep. 2018;20(11):1-11.
3. Ben-Dor I, Mahmoudi M, Deksissa T, Bui AB, Gaglia MA, Gonzalez MA, et al. Correlation between fractional flow reserve and intravascular ultrasound lumen area in intermediate coronary artery stenosis. J Am Coll Cardiol. 2011;57(14):E1855.
4. Unal G, Bucher S, Carlier S, Slabaugh G, Fang T, Tanaka K. Shape-driven segmentation of the arterial wall in intravascular ultrasound images. IEEE Trans Inf Technol Biomed. 2008;12(3):335-47.
5. Zhu X, Zhang P, Shao J, Cheng Y, Zhang Y, Bai J. A snake-based method for segmentation of intravascular ultrasound images and its in vivo validation. Ultrasonics. 2011;51(2):181-9.
6. Mendizabal-Ruiz G, Kakadiaris IA. A physics-based intravascular ultrasound image reconstruction method for lumen segmentation. Comput Biol Med. 2016;75:19-29.
7. Su S, Hu Z, Lin Q, Hau WK, Gao Z, Zhang H. An artificial neural network method for lumen and media-adventitia border detection in IVUS. Comput Med Imaging Graph. 2017;57:29-39.
8. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention, Pt lI. 2015;9351. pp. 234-41.
9. Yang J, Tong L, Faraji M, Basu A. IVUS-Net: an intravascular ultrasound segmentation network. arXiv:1806.03583 [Preprint], 2018. https://arxiv.org/abs/1806.03583 (accessed February 11, 2021).
10. Yang J, Faraji M, Basu A. Robust segmentation of arterial walls in intravascular ultrasound images using Dual Path U-Net. Ultrasonics. 2019;96:24-33.
11. Zhou Z, Siddiquee MMR, Tajbakhsh N, Liang J. UNet+++: a nested U-Net architecture for medical image segmentation. In: Deep learning in medical image analysis and multimodal learning for clinical decision support: 4th International Workshop, DLIMA 2018, and 8th International Workshop, ML-CDS 2018, held in conjunction with MICCAI 2018, Granada, Spain, 2018;11045, pp. 3-11.
12. Zhao C, Bober R, Tang H, Tang J, Dong M, Zhang C, et al. Semantic segmentation to extract coronary arteries in invasive coronary angiograms. J Adv Appl Comput Math. 2022;9:76-85.
13. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the Inception Architecture for Computer Vision. Paper Presented at: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, June 27-30, 2016.
14. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Commun ACM. 2017;60(6):84-90.
15. Sergey I, Christian S. Batch normalization: accelerating deep network training by reducing internal covariate shift. Paper Presented at: ICMI2015.
16. Otsu N. A threshold selection method from gray-level histograms. IEEE Trans Syst Man Cybern Syst. 1979;9(1):62-6.
17. Asundi A, Wensen Z. Fast phase-unwrapping algorithm based on a gray-scale mask and flood fill. Appl Opt. 1998;37(23):5416-20.
18. Kurbiel T, Khaleghian S. Training of deep neural networks based on distance measures using RMSProp. arXiv:1708.01911 [Preprint]. 2018. https://arxiv.org/abs/1708.01911 (accessed February 23, 2021).
19. Deng J, Dong W, Socher R, Lijia L, Feifei L. ImageNet: a large-scale hierarchical image database. Paper Presented at: IEEE Conference on Computer Vision & Pattern Recognition, 2009.
20. Mintz G. American College of Cardiology clinical expert consensus document on standards for acquisition, measurement and reporting of intravascular ultrasound studies (IVUS). A report of the American College of Cardiology Task Force on clinical expert consensus documents developed in collaboration with the European Society of Cardiology endorsed by the Society of Cardiac Angiography and Interventions. Eur J Echocardiogr. 2001;2(4):299-313.
21. Saba T, Khan MA, Rehman A, Marie-Sainte SL. Region extraction and classification of skin cancer: a heterogeneous framework of deep CNN features fusion and reduction. J Med Syst. 2019;43(9):289.
22. Faraji M, Cheng I, Naudin I, Basu A. Segmentation of arterial walls in intravascular ultrasound cross-sectional images using extremal region selection. Ultrasonics. 2018;84:356-65.
23. Downe R, Wahle A, Kovarnik T, Skalicka H, Lopez J, Horak J, et al. Segmentation of intravascular ultrasound images using graph search and a novel cost function. Paper Presented at: Proc. 2nd MICCAI Workshop on Computer Vision for Intravascular and Intracardiac Imaging, 2008, pp. 71-9.
24. Balocco S, Gatta C, Ciompi F, Wahle A, Radeva P, Carlier S, et al. Standardized evaluation methodology and reference database for evaluating IVUS image segmentation. Comput Med Imaging Graph. 2014;38(2):70-90.
25. Huang Y, Yan W, Xia M, Guo Y, Zhou G, Wang Y. Vessel membrane segmentation and calcification location in intravascular ultrasound images using a region detector and an effective selection strategy. Comput Methods Programs Biomed. 2020;189:105339.
26. Lee JG, Ko J, Hae H, Kang SJ, Kang DY, Lee PH, et al. Intravascular ultrasound-based machine learning for predicting fractional flow reserve in intermediate coronary artery lesions. Atherosclerosis. 2020;292:171-7.