IMR Press

Reviews in Cardiovascular Medicine

Original Research

Optical Coherence Tomography Vulnerable Plaque Segmentation Based on Deep Residual U-Net

Lincan Li[1,*] and Tong Jia[2]

[1] *College of Mechanical Engineering and Automation, Northeastern University, Shenyang, 110004, P. R. China*
[2] *College of Information Science and Engineering, Northeastern University, Shenyang, 110004, P. R. China*

*Correspondence: lilincan@stumail.neu.edu.cn (Lincan Li)

Automatic and accurate segmentation of intravascular optical coherence tomography imagery is of great importance in computer-aided diagnosis and in treatment of cardiovascular diseases. However, this task has not been well addressed for two reasons. First, because of the difficulty of acquisition, and the laborious labeling from personnel, optical coherence tomography image datasets are usually small. Second, optical coherence tomography images contain a variety of imaging artifacts, which hinder a clear observation of the vascular wall. In order to overcome these limitations, a new method of cardiovascular vulnerable plaque segmentation is proposed. This method constructs a novel Deep Residual U-Net to segment vulnerable plaque regions. Furthermore, in order to overcome the inaccuracy in object boundary segmentation which previous research has shown extensively, a loss function consisting of weighted cross-entropy loss and Dice coefficient is proposed to solve this problem. Thorough experiments and analysis have been carried out to verify the effectiveness and superior performance of the proposed method.

Keywords

Intravascular optical coherence tomography; image semantic segmentation; encoder-decoder architecture; residual block; boundary segmentation

1. Introduction

Advanced atherosclerosis in the coronary arteries is now one of the leading causes of death worldwide although it is preventable and treatable (Fleg et al., 2012). In order to accurately and effectively diagnose cardiovascular atherosclerosis (i.e. vulnerable plaque), Optical Coherence Tomography (OCT) imaging is often employed (Ambrose and Srikanth, 2010). However, the vast amount of OCT images acquired in a routine clinical exam makes it difficult and time-consuming for physicians to diagnose patients' images manually. Because of this, computer algorithms are needed to address this dilemma. Image segmentation is a crucial step in cardiovascular vulnerable plaque diagnose, and there are two issues which need to be solved.

First, OCT datasets are small compared to typical datasets in natural image domain; variations in data acquisition make OCT datasets even more heterogeneous. Deep learning networks can learn to recognize invariance towards image properties if the dataset is large enough. However, for a small dataset, deep learning network may soon over fit. To address this problem, a light convolutional neural network structure with fewer network parameters and high accuracy is needed. This will require a relatively small image dataset to train and accurately conduct network parameter optimization.

Second, Some standard clinical acquisition protocols in OCT still have limitations in visualizing the underlying anatomy due to imaging artifacts (e.g. guide-wire artifacts, blood artifacts) or operator-dependent errors (e.g. shadows, signal drop-outs). All of which increase the complexity of plaque region boundary segmentation. In these circumstances, new methods and functions should be introduced into the segmentation network, making the network better able to recognize the object boundary pixels.

1.1 Related Work

In the field of coronary artery plaque segmentation, Lu et al. (2014) proposed a method based on image feature extraction and Support Vector Machine (SVM), which realized semi-automatic segmentation of OCT images and achieved a 83% accuracy on a test dataset. Shalev et al. (2016) proposed a segmentation method based on hidden Markov random field (HMRF), which can detect plaques from OCT cardiovascular images. Wang et al. (2017) proposed a semi-automatic segmentation algorithm using K-means clustering to obtain points aggregation which is needed for random walk in the next stage, then used the obtained points aggregation as seed points, realizing plaque segmentation by random walk algorithm of weight function.

Deep learning techniques have made breakthroughs in medical imaging processing in recent years. Researchers have also applied deep learning models to the task of OCT image diagnosis. Gessert et al. (2018) proposed a novel adversarial training network for plaque classification with a small dataset. The presented classification network is able to learn invariant features from patient images, which achieved improvements in plaque classification accuracy. Abdolmanafi et al. (2017) developed an automatic algorithm using convolutional neural network as feature extractor, combined convolutional neural network, SVM and random forest method,

which is capable to classify the coronary artery (tunica adventitia, tunica media, tunica intima).

1.2 Contributions

In view of those studies, they still have not addressed our previously mentioned two concerns. First, this paper constructs a Deep Residual U-Net for segmentation task, using pre-trained ResNet101 as a backbone network of encoder and designed residual blocks as a decoder. The proposed Deep Residual U-Net achieves a fast convergence rate while the number of network layers is large. Since the network is very deep, it's able to learn more abundant image semantic features, thus providing more accurate segmentation results.

Secondly, a loss function composed of weighted cross-entropy loss and Dice coefficient is proposed to improve the network segmentations performance on the object boundary. During the algorithm training stage, the proposed loss function gives a greater penalty on boundary pixels which are inaccurately predicted by the network than a false predicted pixel within the object, so as to improve the accuracy of boundary segmentation.

In our experiment, the proposed method is applied to a segmented OCT cardiovascular vulnerable plaque dataset (Guo et al., 2018), which is provided by the Chinese Academy of Sciences, the First Affiliated Hospital of China Medical University, and the Beijing Health Promotion Association. Segmentation results are qualitatively and quantitatively evaluated, which shows the superiority and effectiveness of our method.

2. OCT Cardiovascular Image Preprocessing

2.1 Description of the OCT Cardiovascular Dataset

The dataset used in our research is collected, dealt, and labeled by the Chinese Academy of Sciences, the First Affiliated Hospital of China Medical University, and Beijing Health Promotion Association. All cardiovascular images in the dataset are manually labeled by several specialists. Fig. 1 shows four OCT cardiovascular image samples. Among them, imaging artifacts (e.g. guide-wire artifacts, blood artifacts and artifacts caused by operational errors) are marked with white lines and text, while vulnerable plaque regions are marked with red lines and red text.

The dataset comprises of 2000 images in polar coordinates, 1000 of them are positive samples (i.e. images which include vulnerable plaques), the remaining 1000 images are negative samples (i.e. images without vulnerable plaques). The size of each image is 720 * 352 pixels. For the convenience of algorithm design, the OCT images used in this paper are transformed from polar coordinate to Cartesian coordinates (Athanasiou et al, 2014), the size of converted image is 703 * 703 pixels.

2.2 Data Augmentation of OCT Cardiovascular Images

In OCT cardiovascular images, vulnerable plaque regions only account for a small part of the image, that is to say, the positive and negative class pixels are extremely unbalanced. The imbalance between positive and negative pixels makes the classifier inclined to classify an image pixel into a negative class (Lin et al., 2017), which makes segmentation boundaries inaccurate even identifying regions containing vulnerable plaque as a region without vulnerable plaque. Therefore, we discard negative samples directly and randomly select 800 positive samples as training set. The remaining 200 images were used as test data.
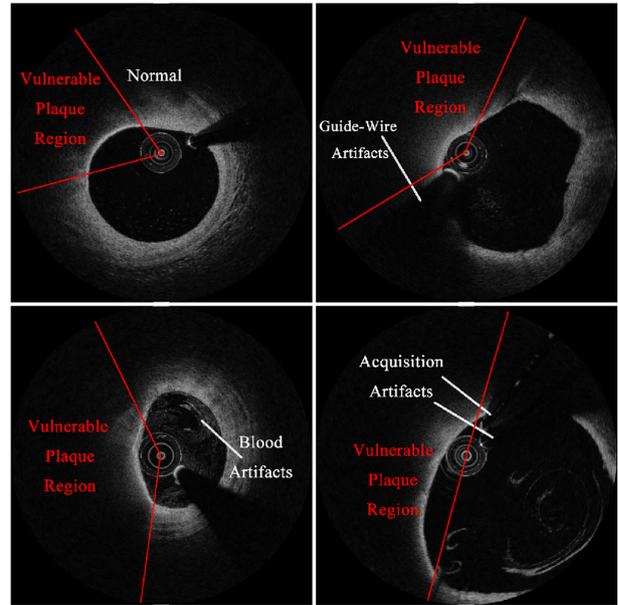

Figure 1. Samples of the OCT cardiovascular dataset (in Cartesian coordinate) and description of various imaging artifacts.
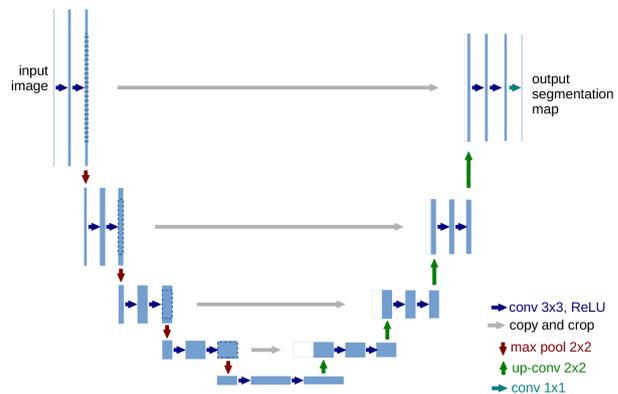

Figure 2. The structure of U-Net segmentation network.

Data augmentation is necessary because of the relatively small training set. As shown in Fig. 1, the foreground of OCT image in Cartesian coordinate is in a circular shape. Considering the geometric properties of circle, it's very suitable to make a rotational transformation. To be specific, we can rotate the image foreground clockwise by 30 degrees, 60 degrees, 90 degrees, 120 degrees, 150 degrees, 180 degrees, 210 degrees, 240 degrees, 270 degrees, and 300 degrees. As a result, the amount of image data increases to 8000, 10 times that of the original number.

3. Methodology

3.1 The Design of Encoder

In the field of image semantic segmentation, there is a basic convolutional neural network structure named U-Net, which has been widely used in medical image segmentation, satellite image segmentation, and road scene segmentation. It's advantages and excellent performance have been analyzed and discussed by researchers.

U-Net structure consists of a convolutional encoding path and a symmetrical decoding path, also called "the Encoder-Decoder structure" by the academic world. An image uploaded to the encoding path will be processed by several repeated 3 * 3 convolution layers and 2 * 2 pooling layers. While the feature map is gradually down sampled, the number of channels increases by multiples. In the expansion path, however, the up sampling operation is conducted in each step, which increases the resolution of the feature map and reduces the number of channels by half. The feature map from the encoding path and decoding path are then concatenated together on a channel dimension through the "skip connection" structure, thus accurate positioning can be achieved. The last layer of U-Net is a 1 * 1 kernel convolutional layer, which maps the number of channels to the number of pixel categories. The number of pixel points in each channel suggests the probability that the pixel belongs to a certain category.
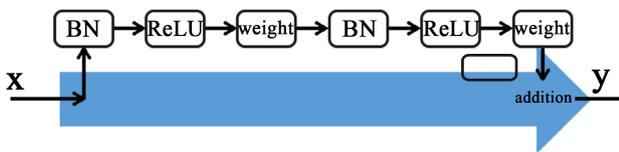

Figure 3. The structure of residual unit utilized in this paper.
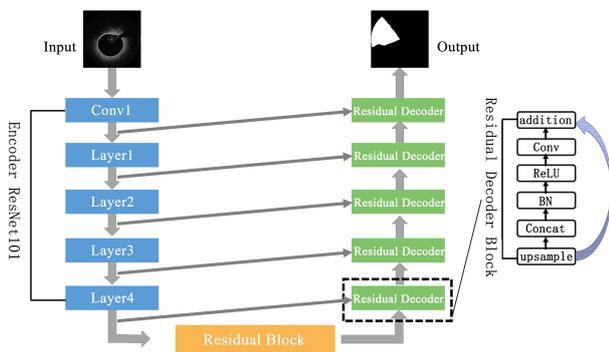

Figure 4. Deep Residual U-Net network structure graph.

U-Net structure utilizes several merits for image segmentation. First, the structure of U-Net is not complicated and easily understandable. It's easy to adjust the structure according to the requirements of specific problems. Second, U-Net can utilize both high-level semantic information and has global positioning ability, which is very effective for a segmentation task with limited imaging data.

The function of the encoder is to gradually decrease the resolution of the feature map and to learn upper level features from image semantic information (Long et al., 2014). Usually, pre-trained classification networks are adopted as encoders.

ResNet (He et al., 2016) was proposed in 2015 and won 1st place in an ImageNet classification competition. Because of its simplicity and practicability, many methods in detection, segmentation,

and recognition fields have been built on the basis of ResNet50 or ResNet101.

With the increase of network layers, traditional deep learning networks witness a reduction in training set accuracy while ResNet solves this problem very well (Russakovsky et al., 2015). The performance improvements of ResNet come from the introduction of a residual unit with an identity shortcut connection, which is able to cope with the gradient disappearance problem. In this paper, ResNet101 is selected as an encoder of segmentation model, which is capable of extracting complex and high level features and then input them into a decoder.

It's time consuming to train a completely new neural network, and a common approach is to use a network trained by the ImageNet dataset as the source of weight initialization. This method is also called transfer learning (Pan and Yang, 2010). Based on this strategy, pre-trained ResNet101 is used as the encoder of the proposed segmentation network.

3.2 Deep Residual U-Net Segmentation Network

U-Net with deep layer can provide better segmentation results, but the increased network layer tends to decrease training set accuracy. Thus, we utilized the merits of residual network and added a residual unit to U-Net (Xie et al., 2016). In addition, the core idea of U-Net is to stitch low level features into corresponding high level features, adding low level texture features to high level semantic features (Ronneberger et al., 2015). This idea is similar to a residual network. A residual network is composed of many residual units. The general equation of residual unit can be expressed as

$$Y = f(h(x) + F(x, w)) \qquad (1)$$

Where x denotes the input of residual unit, h denotes identity mapping (i.e. $h(x) = x$), and F denotes residual function. The structure of a residual unit used in our research is displayed in Fig. 2.

How to add residual units into Deep Residual U-Net is described as follows: we utilize ResNet101 in network encoder part, this is to say that the encoder has already become a residual structure. In decoder part, we construct the residual decoder block, adding residual connections in the original decoder. Upsampling operation between two adjacent residual decoders is realized by bilinear interpolation. The last layer, before output, is a 1 * 1 kernel convolutional layer, which adjusts the number of channels and outputs a pixel-level probability map. The structure of a deep residual U-Net is demonstrated in Fig. 3. As Fig. 3 shows, the encoder is ResNet101, and we modify each decoder layer to residual decoder block. The structure of residual decoder block is also shown in Fig. 3.

3.3 Improved Loss Function for Accurate Boundary Segmentation

The original U-Net achieves superior segmentation results by calculating cross-entropy loss between feature map and the actual label in each pixel. However, in the OCT cardiovascular dataset, vulnerable plaque regions only account for a small part of whole image. Unbalanced foreground and background makes it easy for the network to predict a pixel as background, which leads to incomplete detection of vulnerable plaque regions. In order to im-
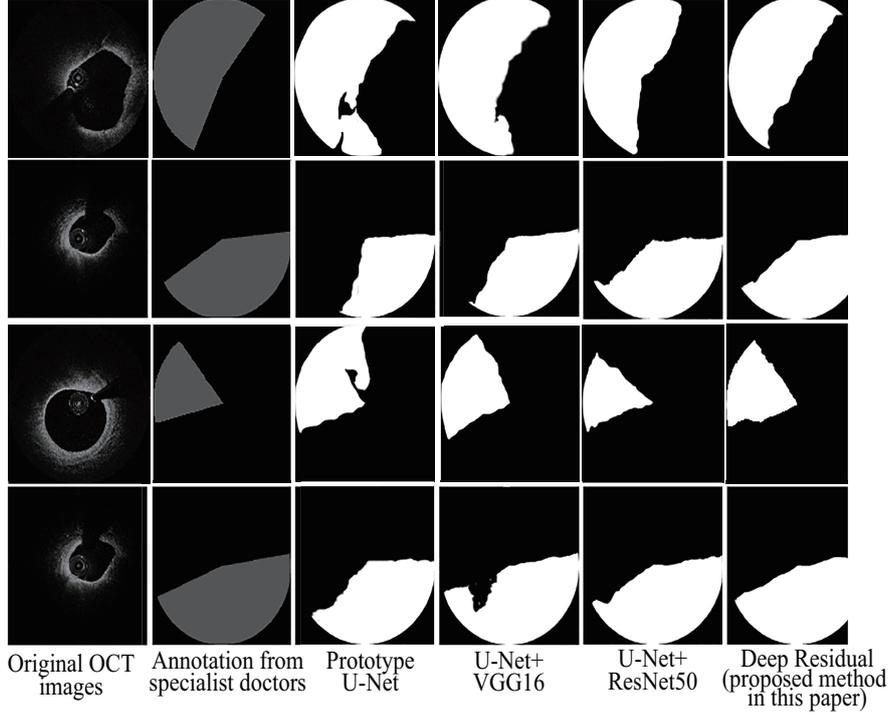
Figure 5. Qualitative evaluation of different segmentation results. Each column utilizes a different method, the quality of segmentation results can be observed directly.

prove segmentation results, a loss function comprised of weighted cross-entropy and Dice coefficient is adopted (Papandreou et al., 2015; Ronneberger et al., 2015).

Weighted cross-entropy loss provides more attention to object boundary pixels, making segmentation boundaries more accurate. While Dice coefficients provide high accuracy of pixel classification, ensuring the general segmentation is in good quality. Details of the proposed loss function is described as follows. We use weighted cross-entropy loss to counteract the unbalanced foreground and background. The expression of weighted cross-entropy loss is displayed in Eqn. 2:

$$E = \sum_{x \in \Omega} w(x) \log \left(p_{l(x)}(x) \right) \tag{2}$$

Where the weight $w(x)$ is an approach to counteract unbalanced foreground and background, hoping the algorithm will learn more information about vulnerable plaque boundary. The weight $w(x)$ is calculated using Eqn. 3:

$$w(x) = w_c(x) + w_0 \exp \left(-\frac{(d_1(x) + d_2(x))^2}{2\delta^2} \right) \tag{3}$$

Where denotes the frequency of classes, $d_1(x)$ denotes the distance between a pixel and the nearest vulnerable plaque boundary, and $d_2(x)$ denotes the distance between a pixel and the second nearest vulnerable plaque boundary. The value of constant parameterand are determined following literature (Han and Ye, 2018;

Man et al., 2019; Ronneberger et al., 2015). We set $w_0 = 10$ and $\sigma = 5$ pixel.

Dice coefficient is derived from dichotomy and is essentially a measure of the overlapping parts of two samples. The index ranges from 0 to 1, in which 1 represents completely overlap. Dice loss is also appropriate for unbalanced foreground and background. Eqn. 4 shows the expression of Dice coefficient loss:

$$D = 1 - \frac{2 \sum\limits_{pixels} y_{true} y_{pred}}{\sum\limits_{pixels} y_{true}^2 + \sum\limits_{pixels} y_{pred}^2} \tag{4}$$

Where y_{true} denotes the real value of a pixel and $y_{predict}$ denotes the predicted value of a pixel. Combining the two kinds of losses together, the total loss is obtained as shown in Eqn. 5:

$$L_{\text{total}} = E + D \tag{5}$$

4. Experiments and Analysis

Experiments were performed using the OCT cardiovascular dataset as previously described. 800 samples were randomly selected from 1000 positive images, and the training set was augmented to 8000 images using the data augmentation method stated above. The remaining 20% of original dataset is used as a test set.

4.1 Evaluation Index of Experiments

The evaluation indexes in our experiment are shown in Table 1.

1) Pixel Accuracy (PA): This is a basic and commonly used segmentation performance evaluation index, which calculates the

Table 1. Segmentation evaluation indexes and explanation.

Evaluation indexes	Explanation
Pixel Accuracy (PA)	The proportion of correctly classified pixels to total pixels.
Mean Pixel Accuracy (MPA)	An improved index based on PA. It calculates the correct proportion of each kind of pixel classification by category, then takes the average value of all categories.
Mean Intersection over Union (MIoU)	The proportion of the union of predicted region and ground truth region to predicted region plus ground truth region.
Frequency Weight IoU (FWIoU)	An improved index based on MOU. FWIoU balance the weight of each category according to the frequency of each category.
Precision Rate (P)	Proportion of actual positive samples in the positive samples predicted by the algorithm.
Recall Rate (R)	Proportion of positive samples predicted by the algorithm in total actual positive samples.

proportion of correctly classified pixels to total pixels. The calculate equation is the following:

$$PA = \frac{\sum_{i=1}^{k} n_{ii}}{\sum_{i=1}^{k} t_i} \qquad (6)$$

2) Mean Pixel Accuracy (MPA): This is an improved segmentation evaluation index from PA. MPA calculates the proportion of correct classified pixels to total number of pixels in each pixel category, then calculates the average value from all categories. The equation is described as Eqn. 7:

$$MPA = \frac{1}{k} \sum_{i=0}^{k} \frac{n_{ii}}{t_i} \qquad (7)$$

3) Mean Intersection over Union (MIoU): This index is a measurement which calculates the overlap degree of the algorithm segmented area and the reality area. The equation is listed as equation Eqn. 8:

$$MIoU = \frac{1}{k} \sum_{i=0}^{k} \frac{n_{ii}}{t_i - n_{ii} + \sum_{j=1}^{k} n_{ji}} \qquad (8)$$

4) Frequency weighted IoU (FWIoU): This is improved index from MIoU, it balances the weight of each category according to occurrence frequency of each category. The equation is given as below:

$$FWIoU = \left(\sum_{i=1}^{k} t_i\right)^{-1} \frac{1}{k} \sum_{i=1}^{k} \frac{n_{ii}}{t_i - n_{ii} + \sum_{j=1}^{k} n_{ji}} \qquad (9)$$

Where k denotes the total number of categories in the image dataset, t_i denotes the total number of pixel belongs to i-th category, and n_{ij} denotes pixels that belong to i-th category, but incorrectly predict as j-th category.

Let TP be the positive class, which the algorithm also predicts as positive class, FP be the negative class, which the algorithm

also predicts as positive class. FN is the positive class which the algorithm predicts as negative class. Then the equations of precision rate and recall rate can be obtained as following:

$$precision rate : P = \frac{TP}{(TP+FP)} \qquad (10)$$

$$recall rate : R = \frac{TP}{(TP+FN)} \qquad (11)$$

4.2 Experiment Detail of Deep Residual U-Net

Input images are normalized after subtracting the mean value. We utilize the SGD optimization algorithm, where the parameter momentum of SGD is set to 0.0005, and the parameter weight_decay is set to 0.0002. Learning rate of the network is set to 0.001 and batch-size is set to 1 during training.

The segmentation results of deep residual U-Net are evaluated qualitatively and quantitatively. Qualitative evaluation is shown in Fig. 4. and quantitative evaluation is shown in Table 2.

4.3 Experiment and Analysis of Boundary Segmentation

To verify the effectiveness of the proposed loss function on object boundary, several experiments are conducted as described below:

1) Prototype U-Net and prototype U-Net with proposed loss function.

2) Deep Residual U-Net + ResNet101 and Deep Residual U-Net + ResNet101 with proposed loss function.

The experiment results are shown as follows:

From Fig. 6, it can be seen that (d) has improved boundary segmentation results than (c). Comparing (e) with (f), both of them utilized Deep Residual U-Net proposed in this paper, the difference is that (e) only used common loss function while (f) used the proposed loss function. Results demonstrate that (f) has a smoother boundary shape and is closer to manual labeling .

After qualitative evaluation, we investigate further to see how much the proposed function improves boundary segmentation accuracy. Quantitative evaluation is conducted as following. We contour object boundary pixels from manual image labeling as shown in Fig. 7, and pay close attention to the boundary pixels only.

We recorded IoU rate of boundary pixels during network training iterations and plotted the IoU value-Iteration curve, which is

Table 2. Quantitative evaluation of different U-Net based segmentation methods.

|	Pixel Accuracy (PA)	Mean Pixel Accuracy (MPA)	Mean Intersection over Union (MIoU)	Frequency Weight IoU (FWIoU)	Precision Rate (P)	Recall Rate (R)
Prototype U-Net	0.6253	0.6514	0.4207	0.5023	0.8245	0.7438
U-Net + VGG16	0.7829	0.8042	0.6511	0.7356	0.8633	0.8267
U-Net + ResNet50	0.8931	0.824	0.7748	0.8627	0.8721	0.8384
U-Net + ResNet101	0.9195	0.9005	0.777	0.8587	0.8965	0.8691
Deep Residual U-Net (The proposed method in this paper)	**0.9331**	**0.9011**	**0.8548**	**0.864**	**0.9433**	**0.9135**
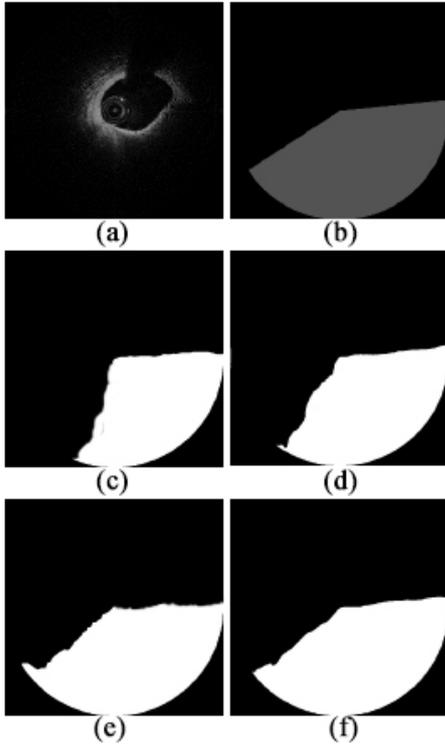
Figure 6. Original image in OCT cardiovascular image dataset is shown in (a), the manual label segmentation boundary from specialists is shown in (b), prototype U-Net segmentation result are shown in (c), prototype U-Net with proposed loss function segmentation result are shown in (d), Deep Residual U-Net + ResNet101 segmentation result are shown in (e) and Deep Residual U-Net + ResNet101 with proposed loss function segmentation result are shown in (f).

Figure 7. (a) is an original image, (b) is the manual label of (a), (c) displays contoured object boundary.

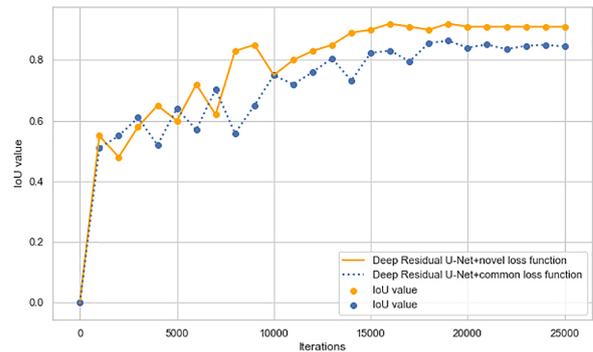Figure 8. IoU value-Iteration curve of Deep Residual U-Net with the proposed loss function and Deep Residual U-Net with common loss function.

also known as a ``learning curve''. As discussed before, IoU rate is a common object segmentation indicator, giving quantitative evaluation of segmentation accuracy. Fig. 8 shows the IoU value-Iteration curve of Deep Residual U-Net with the proposed loss function and Deep Residual U-Net with common loss function. It's not difficult to see that Deep Residual U-Net with proposed loss function gets a higher IoU value than with common loss function. Also, Deep Residual U-Net with the proposed loss function converges faster and gets less fluctuation in the curve.

Based on the experiment above, we can draw a conclusion that a novel loss function combining weighed cross entropy loss and Dice coefficient can improve the classification accuracy of boundary

pixels, reduce error classification, thus providing more accurate boundary segmentation.

5. Conclusion

In this paper, a Deep Residual U-Net segmentation network is proposed for OCT cardiovascular image segmentation. We focused on solving two problems in the application of deep learning to vulnerable plaque diagnose in OCT images. (1) Compared with typical datasets in natural image domain, the OCT cardiovascular dataset is smaller and variables in the image acquisition process make the OCT cardiovascular dataset more complex. In this case, deep learning network will overfit easily. (2) Owing to imaging artifacts (e.g. guide-wire artifacts and blood artifacts) or operational error, which cause image information loss and partial shadow, vulnerable plaque segmentation becomes more complex (Prakash et

al., 2013). To solve these problems, a Deep Residual U-Net segmentation network is proposed where the backbone network of encoder is replaced by pre-trained ResNet101 and the decoder is comprised of designed residual blocks. The deepening of the network layer and the introduction of residual block provide superior segmentation results. Furthermore, a novel loss function consists of weighted cross entropy and Dice coefficient is proposed to improve the segmentation accuracy of object boundary.

We conducted qualitative and quantitative evaluation to the proposed Deep Residual U-Net. Qualitative evaluation by randomly selecting four images from the test set and compared segmentation results using different methods. Quantitative evaluation uses Pixel Accuracy (PA), Mean Pixel Accuracy (MPA), Mean Intersection over Union (MIoU), Frequency Weight IoU (FWIoU), Precision Rate (P) and Recall Rate (R) as indicators and compared the value of different methods.

The proposed Deep Residual U-Net and loss function which consisted of weighted cross entropy and Dice coefficient received the highest score in all of the indicators we choose for quantitative evaluation. Besides, in qualitative evaluation, the proposed approach demonstrated the most accurate segmentation results. Both quantitative and qualitative evaluation prove the feasibility and advantage of the proposed method. Finally, the conclusion can be made that the method proposed in this paper is valuable and useful for automatic OCT cardiovascular image segmentation.

Acknowledgment

Conflict of Interest

The authors declare that they have no financial and personal relationships with other people or organizations that can inappropriately influence their work, there is no professional or other personal interest of any nature or kind in any product, service and/or company that could be construed as influencing the position presented in, or the review of the manuscript entitled.

References

Abdolmanafi, A., Duong, L. and Dahdah, N. (2017) Deep feature learning for automatic tissue classification of coronary artery using optical coherence tomography. *Biomedical Optics Express* **8**, 120-133.

Acharya, U. R., Fujita, H., Lih, O. S., Adam, M., Tan, J. H. and Chua, C. K. (2017) Automated detection of coronary artery disease using different durations of ecg segments with convolutional neural network. *Knowledge-Based Systems* **12**, 634-651.

Ambrose, J.A. and Srikanth, S. (2010) Vulnerable plaques and patients: improving prediction of future coronary events. *American Journal of Medicine* **123**, 0-16.

Athanasiou, L. S., Bourantas, C. V., George, R., Sakellarios, A. I., Exarchos, T. P. and Siogkas, P. K. (2014) Methodology for fully automated segmentation and plaque characterization in intracoronary optical coherence tomography images. *Journal of Biomedical Optics* **311**, 172-176.

Fleg, J. L., Stone, G. W., Fayad, Z. A., Granada, J. F., Hatsukami, T. S. and Kolodgie, F. D. (2012) Detection of high-risk atherosclerotic plaque: report of the nhlbi working group on current status and future directions. *JACC. Cardiovascular imaging* **5**, 56-85.

Gessert, N., Heyder, M. and Latus, S. (2018) Adversarial training for patient-independent feature learning with ivoct data for plaque classification. *2018 IEEE International Conference on Bioinformatics and Biomedicine*, 1-2.

Guo, X., Tang, D., Molony, D., Yang, C., Samady, H. and Zheng, J. (2018) A machine learning-based method for intracoronary oct segmentation and vulnerable coronary plaque cap thickness quantification. *International Journal of Computational Methods* **6**, 196-209.

Han, Y. and Ye, J. C. (2018) Framing u-net via deep convolutional framelets: application to sparse-view ct. *IEEE Transactions on Medical Imaging* **4**, 15-30.

He, K., Zhang, X., Ren, S. and Sun, J. (2016) Deep residual learning for image recognition. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770-778.

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F. and Ghafoorian, M. (2017) A survey on deep learning in medical image analysis. *Medical Image Analysis* **42**, 60-88.

Lin, T. Y., Goyal, P., Girshick, R., He, K. and Dollár, Piotr. (2017) Focal loss for dense object detection. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2999-3007.

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

Lu, H., Jakl, M., Wang, Z. and Tanaka, K. (2014) Tct-371 evaluation of highly automated software for analyzing intravascular optical coherence tomography pullbacks of stents. *Journal of the American College of Cardiology* **64**, B108.

Man, Y., Huang, Y., Li, J. F. X. and Wu, F. (2019) Deep Q learning driven ct pancreas segmentation with geometry-aware u-net. *IEEE Transactions on Medical Imaging* **6**, 83-100.

Pan, S. J. and Yang, Q. (2010) A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering* **22**, 1345-1359.

Papandreou, G., Chen, L. C., Murphy, K. P. and Yuille, A. L. (2015) Weakly-and Semi-Supervised Learning of a Deep Convolutional Network for Semantic Image Segmentation. *IEEE International Conference on Computer Vision (ICCV)*, 1742-1750.

Prakash, A., Hewko, M., Sowa, M. and Sherif, S. (2013) Texture based segmentation method to detect atherosclerotic plaque from optical tomography images. *European Conference on Biomedical Optics* **8802**, 1-12.

Ronneberger O., Fischer P. and Brox T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. (eds.) *Medical Image Computing and Computer-Assisted Intervention (MICCAI)* **9351**, 234-241 .

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S. and Ma, S. (2015) Imagenet large scale visual recognition challenge. *International Journal of Computer Vision* **115**, 211-252.

Shalev, R., Gargesha, M., Prabhu, D. and Tanaka, K. (2016) Processing to determine optical parameters of atherosclerotic disease from phantom and clinical intravascular optical coherence tomography three-dimensional pullbacks. Journal of Medical Imaging **3**, 62-78.

Wang, G., Wang, P. and Han, Y. (2017) plaque segmentation of intracoronary optical coherence tomography images based on k-means and improved random walk algorithm. *Journal of Biomedical Engineering* **34**, 869-875.

Xie, S., Girshick, R., Dollár, Piotr., Tu, Z. and He, K. (2016) Aggregated residual transformations for deep neural networks. *IEEE Transactions on Pattern Analysis & Machine Intelligence,* 5987-5995.