ORIGINAL ARTICLE

Open Access



PlaqueNet: deep learning enabled coronary artery plaque segmentation from coronary computed tomography angiography

(2024) 7:6

Linyuan Wang^{1†}, Xiaofeng Zhang^{2†}, Congyu Tian³, Shu Chen⁴, Yongzhi Deng^{1*}, Xiangyun Liao^{3*}, Qiong Wang³ and Weixin Si³

Abstract

Cardiovascular disease, primarily caused by atherosclerotic plaque formation, is a significant health concern. The early detection of these plaques is crucial for targeted therapies and reducing the risk of cardiovascular diseases. This study presents PlaqueNet, a solution for segmenting coronary artery plaques from coronary computed tomography angiography (CCTA) images. For feature extraction, the advanced residual net module was utilized, which integrates a deepwise residual optimization module into network branches, enhances feature extraction capabilities, avoid-ing information loss, and addresses gradient issues during training. To improve segmentation accuracy, a depthwise atrous spatial pyramid pooling based on bicubic efficient channel attention (DASPP-BICECA) module is introduced. The BICECA component amplifies the local feature sensitivity, whereas the DASPP component expands the network's information-gathering scope, resulting in elevated segmentation accuracy. Additionally, BINet, a module for joint network loss evaluation, is proposed. It optimizes the segmentation model without affecting the segmentation results. When combined with the DASPP-BICECA module, BINet enhances overall efficiency. The CCTA segmentation algorithm proposed in this study outperformed the other three comparative algorithms, achieving an intersection over Union of 87.37%, Dice of 93.26%, accuracy of 93.12%, mean intersection over Union of 93.68%, mean Dice of 96.63%, and mean pixel accuracy value of 96.55%.

Keywords Medical image segmentation, Coronary artery plaques, Deep residual networks, Attention mechanism

[†]Linyuan Wang and Xiaofeng Zhang contributed equally to this work.

*Correspondence: Yongzhi Deng olympicschina@163.com Xiangyun Liao xiangyun-l@163.com

¹ Department of Cardiovascular Surgery, the Affiliated Hospital of Shanxi Medical University, Shanxi Cardiovascular Hospital (Institute), Shanxi Clinical Medical Research Center for Cardiovascular Disease, Taiyuan 030024, Shanxi, China

² Department of Mechanical Engineering, Nantong University, Nantong 226019, Jiangsu, China

³ Shenzhen Institute of Advanced Technology, Chinese Academy

of Sciences, Shenzhen 518055, Guangdong, China

⁴ Department of Cardiovascular Surgery, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430022, Hubei, China

Introduction

Coronary artery plaque, is a prevalent cardiovascular disease that cause various significant health issues. The development of these plaques leads to a narrowing of the carotid artery, which disrupts the blood supply to the brain. This can result in transient interruptions in blood flow, potentially causing unpredictable damage to the brain. Consequently, carotid plaque segmentation research is crucial for advancing carotid atherosclerosis diagnosis and treatment. Disease severity can be evaluated with greater precision by achieving precise segmentation and localization of carotid plaques, providing a foundation for more effective treatment strategies. In cardiovascular disease treatment, early identification and segmentation timely interventions, ultimately reduce



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

disease-related mortality [1]. Consequently, this study is significant because it may enhance patient outcomes and improve the overall management of carotid atherosclerosis, thereby contributing to a healthier and more resilient population.

Medical image segmentation, which combines the power of medical imaging with advanced deep learning techniques, provides a highly effective and intuitive means of precisely delineating areas of interest, particularly in the context of lesion detection [2]. Medical image segmentation can be broadly categorized into two distinct approaches: traditional algorithm-based and deep learning-based [3]. However, when dealing with the inherent complexity of medical images, traditional segmentation methods may fall short and often require supplementary algorithms [4]. However, this can compromise the accuracy of the segmentation results. To address these challenges, a robust deep learning segmentation model based on neural networks has emerged as a superior alternative [5, 6]. These models can learn and leverage critical feature information for segmentation tasks, thereby significantly improving the accuracy and efficiency of the segmentation process. Consequently, they consistently outperformed traditional algorithms. In recent years, several prominent deep-learning segmentation methods have risen to the forefront of medical image analysis. Notably, models such as FCN, Deeplabv3, and Deeplab3plus have gained recognition for their exceptional performance and have become the go-to solutions in this evolving field [7-9]. These innovative approaches are poised to revolutionize medical image segmentation, enabling more precise diagnostics and treatment planning for a wide range of medical conditions.

In medical image segmentation, the development of a segmentation network model is a complex and multifaceted task that requires a nuanced understanding of medical images and their associated parameters [10, 11]. This underscores the need for a segmentation network that can handle the intricacies of medical imagery, while producing precise mask outputs. PlaqueNet introduced an innovative approach that employs a multi-path parallel residual optimization network for medical image feature extraction. By leveraging the multi-path parallel residual structure of ResNet, the robust deep-level information extraction capabilities can be harnessed [12, 13].

The proposed approach involves enhancing the network by incorporating a pooling mapping function along with the original parallel residual mapping function. This deepens the network' architecture, resulting in more effective feature extraction. The pooling mapping function mitigates the feature information loss that often occurs in deeper networks, thereby preserving the overall feature information. To improve the accuracy of the segmented areas further, this study proposes implementing of a deep bicubic attention-space separable convolution module. The proposed method capitalizes on deep separable atrous convolution to capture a more comprehensive range of pixel information while avoiding information loss in the network output segment. Additionally, the bicubic attention mechanism augments the network's capacity to identify local information, thereby enhancing the overall accuracy of the segmented regions. This study also introduces a novel technique designed to enhance the accuracy of medical image segmentation. It introduces a bilinear reflection filling upsampling network that incorporates reflection filling into the bilinear upsampling network. This, in conjunction with the depth bicubic sampling attention space separable convolution module, jointly evaluates the loss function, leading to an overall improvement in the network's segmentation accuracy. However, the proposed algorithm has certain limitations. This research primarily focused on the segmentation of vascular plaques in two-dimensional images, neglecting the segmentation of three-dimensional images. Consequently, the process of segmentation visualization overlooks the corresponding three-dimensional information structure, and the resulting two-dimensional structure fails to capture the complete three-dimensional characteristics of the vascular plaques. Future study will focus on three-dimensional image vascular plaque segmentation.

In summary, this study made four main contributions.

- The introduction of PlaqueNet for coronary artery plaque segmentation not only enhances the network's feature extraction capabilities but also utilizes a joint evaluation loss function to significantly improve both efficiency and accuracy.
- (2) The advanced residual net (AResNet) module, which excels at extracting input feature information, ensuring the retention of global information during the depth-based feature extraction process.
- (3) The depthwise atrous spatial pyramid pooling based on bicubic efficient channel attention (DASPP-BICECA) module expands the perceptual field range of feature information, reinforces local information connections, and increases the sensitivity of the network to feature information.
- (4) The BINet module was designed to simultaneously assess network output loss and elevate the overall network' segmentation performance.

With the increasing global prevalence of cardiovascular diseases, it is crucial to recognize that the primary catalyst underlying these conditions is the rupture of cardiovascular atherosclerotic plaques. Such ruptures can trigger a series of sudden and often devastating brain diseases with high mortality and disability rates, including strokes, cerebral infarctions, and cerebral hemorrhages [14]. The early detection of atherosclerotic plaques has the potential to significantly reduce the incidence of these brain-related diseases; and timely intervention following detection can help avert these sudden health crises [15]. Currently, deep learning-based medical and image processing techniques are key to precisely segmenting plaques. Extensive research has been conducted in the field of medical image segmentation. Xu et al. [16] proposed an automatic segmentation method for arterial vessel walls and plaques that would be beneficial for quantifying arterial morphology in magnetic resonance imaging, using the convolutional neural network VWISegNet model to extract features from MRVWI images and compute the class of each pixel to facilitate the segmentation of the vessel walls. Xu and Zhu [17] developed a semantic segmentation algorithm based on convolutional neural networks focusing on edge segmentation to segment arterial vessel walls and plaques to facilitate the quantitative assessment of plaques in patients with is chemic stroke.

A novel MSFA-U-Net segmentation method was introduced in the context of local radiotherapy for the thyroid segmentation of CT images. This method enhances the traditional U-Net model by incorporating multiple parallel channels, thereby enabling the fusion of feature information across different image resolutions. This strategic feature fusion approach prevents the generation of single-resolution information in U-Net during the downsampling process, thereby enhancing the accuracy and effectiveness of thyroid segmentation [18]. For the diagnosis and treatment of cancer, one of the key challenges lies in accurately delineating prostate sites from histopathological images obtained through cell puncturing. To address this issue, a BSP U-Net model is proposed to achieve precise prostate contour extraction. BSP U-Net builds upon the traditional U-Net network structure by incorporating prior knowledge of the prostate shape, resulting in more accurate and reliable prostate site localization [19, 20]. A pivotal step in automatic lung disease analysis is the accurate identification and segmentation of lung regions. To address this challenge, the VI-FCN algorithm was introduced to identify and segment the lung regions in frontal and lateral chest radiographs. This innovation is a critical contribution to the field of lung-disease analysis, aiding in the early diagnosis and treatment of such conditions [21]. Moreover, in the face detection domain, existing detectors often struggle to extract sufficient features, particularly from small-scale faces, which may result in missing detection data. To mitigate this issue, the R-FCN algorithm was proposed for small-scale face detection, offering a more robust solution for capturing facial features, even in challenging scenarios [22]. In diabetic retinopathy detection, which can be accurately identified through retinal fundus images, an enhanced object-detection algorithm basted on the R-FCN is introduced. This innovative approach incorporates a feature pyramid network and improved region structure, thereby bolstering the ability to recognize small-area objects with greater precision [23]. Position emission tomography imaging is one of the most effective methods for diagnosing malignant tumors. To alleviate the substantial workload on radiologists, a novel approach leveraging a multi-scale Mask R-CNN has been proposed, which significantly streamlines the diagnostic process [24]. In the domain of recognizing protein macromolecule crystallization, there has been a concerted effort to enhance the accuracy of classification algorithms. To achieve this, a groundbreaking strategy is presented: the application of the Mask R-CNN model to the detection of protein macromolecule crystallization. This innovative methodology also incorporates adaptive histogram techniques into Mask R-CNNs to mitigate issues, such as backlighting and precipitation effects, further refining the recognition process [25].

Many existing recognition algorithms overlook variations in spatial information within different perception fields. Some networks do not consider the relationships between the edge pixels in the target area, leading to misclassification and recognition errors. To mitigate this issue, the MR R-CNN addresses the problem by adjusting the step size of the region of interest alignment [26]. Deeplabv3plus is highly regarded as an exceptional segmentation algorithm in image segmentation, owing to its remarkable ability to effectively extract multi-scale information [27]. To pursue cerebrovascular and cranial nerve segmentation in medical images, an extended version of the Deeplabv3 algorithm was introduced. This extension incorporates a feature extraction module within the encoder structure and a shrinking pyramid pooling module into the decoder structure [28]. For the segmentation of glioblastoma tumor subregions, normal tissues, and the background, a novel algorithm named DeepNet was proposed. It leverages the structure of Deeplabv3plus and utilizes a predictively trained Resnet18 for weight initialization, resulting in more accurate and reliable segmentation results [29].

To circumvent the risk of early glaucoma-related visual impairment, the Deeplabv3plus architecture was harnessed for optic disc segmentation in the initial screenings with the specific aim of achieving accurate detection. This involves substituting multiple encoder modules in Deeplabv3plus with convolutional layers to enhance segmentation performance [30]. In thyroid segmentation in ultrasound images, a novel approach capitalizes on spatiotemporal recurrent deep learning networks that incorporate time series information. Specifically, it leverages an LSTM model based on Deeplabv3plus to conduct semantic segmentation, thereby facilitating the automatic identification of thyroid components [31]. For the real-time segmentation of bladder lesions in cystography, a range of neural network models were employed during the training phase. The results demonstrated that the PAN model outperformed the other models, thereby demonstrating its superior performance in this context [32]. To enhance the precision of pressure sore diagnosis and overcome the limitations of manually marking feature points in traditional machine learning, a novel superpixel-assisted classification image-labeling method rooted in a regional organization was introduced [33]. To address the need for more accurate eye detection and segmentation, an enhanced Deeplabv3plus network architecture was proposed [34]. In prostate cancer screening, where efficiency and precision are paramount, a deep learning-based approach for swift and accurate detection of abnormal cells was proposed [35]. For early pneumonia diagnosis using lung X-rays, a segmentation model based ResNet was developed to reduce the error rates associated with traditional methods [36]. In the domain of brain tumor detection, a modified ResNet architecture was presented to augment the watershed model, distinguishing it from conventional machine learning techniques [37]. Furthermore, in the pursuit of improved diagnostic tools for pneumonia detection, an automated pneumonia detection and diagnostic tool based on a pre-trained deep learning CNN architecture was introduced [38].

Methods

The PlaqueNet architecture introduced in this study features a multi-path parallel residual network structure, complemented by a deep-pooling mapping function to enhance feature extraction. This deep-pooling mapping function was seamlessly integrated into each residual structure within the multi-path parallel residual network, thereby maintaining the integrity of the feature information during the transmission of pixel data. This approach enables the model to gain more valuable insights from the input data. In the final stages of the segmentation mask output, the innovation includes the introduction of a deep bicubic attention space separable convolution module. This module leverages deep separable dilated convolution to expand the scope of feature information capture and effectively minimize information loss. Simultaneously, the bicubic attention mechanism augments the relevance of the local information, resulting in the generation of more contiguous segmented regions. To further augment the network segmentation performance, an auxiliary prediction network loss module was introduced. This module combines the bilinear reflection-filling up sampling network with the deep bicubic attention space-separable convolution module, and collaboratively address the network's segmentation loss function. This comprehensive approach significantly enhances the accuracy and effectiveness of the segmentation process.

Deepwise parallel residual optimization module

Enhancing the computational prowess of a neural network model for processing input data typically involves augmenting or modifying the depth and width of the network. However, such operations place significant demands on the network's design and computing capacity. The Resnet network comprises a series of identical residual mapping functions structured in parallel, with all the residual blocks sharing the same topological configuration. In total, there are 32 identical residual structures. This study introduces the AResNet network, which builds on the Resnet architecture by incorporating a deep residual optimization structure within each residual mapping component (Fig. 1). This structural feature is referred to as the deep parallel residual mapping optimization function, denoted by Y(x). Y(x)comprises two key components: the initial feature point extraction result $G_i(x)$, which is derived from the feature extraction module, and the optimization information $H_i(x)$ obtained through the deep residual optimization structure. $G_i(x)$ represents the initial outcome of the feature point extraction, whereas $H_i(x)$ maps, filters, and extracts feature information from the initial residual results, thereby effectively eliminating redundant information during dimension reduction. This process ensures that edge information is accurately extracted for the target region.

The structure denoted by $H_i(x)$ comprises several key components: an average pooling layer, a convolutional layer, a batch normalization (BN) layer, and an activation function. This structural composition further extracts and optimizes feature information for each residual structure. By considering feature points from adjacent areas and computing their averages, the average pooling layer contributes to the preservation of the background in medical images. This preserved background served as a valuable reference point for comparing segmentation results and facilitating disease diagnosis. The combination of the convolutional and BN layers mitigates the training challenges associated with the depth of the residual structure. It effectively addresses issues such as gradient disappearance and explosion, which can hinder the model's performance during the training process.



Fig. 1 AResNet network module

Additionally, the activation layer enhances the adaptability of the feature information, making it easier for the segmentation result to fit the data. Moreover, it helps reduce the number of model parameters, thereby optimizing the feature extraction module.

 $H_i(x)$ is the result of Deepwise residual optimization structure.

$$H_i(x) = Relu\left(\sum_{i=1}^{M} \left(k_{ci}x_i + k_{APi}x_i\right) + b_i\right)$$
(1)

Where x_i represents the input feature information, k_{ci} represents the convolution kernel of the convolution layer, k_{APi} represents the convolution kernel of the average pooling layer, b_i represents the bias function, *Relu* represents the activation function, and M represents the number of convolutional layers.

 Y_x is the result of Deepwise parallel residual optimization functions.

$$Y_x = \sum_{i=1}^{C} (G_i(x) + H_i(x)) + x$$
(2)

where $G_i(x)$ represents the convolutional result in the parallel residual network structure, $H_i(x)$ represents the result of the depthwise pooling feature extraction function, and *C* represents the cardinality in the parallel mapping residual network.

DASPP-BICECA module

The DASPP-BICECA module plays a pivotal role in predicting the segmentation output stage. DASPP, which employs deep atrous convolution operations with varying atrous convolution rates, extends the scope of regional information perception during convolution. Deepwise separable convolution effectively segregates the regional input information from the channel convolution points, thereby reducing the number of parameters during model transmission. The integration of BICECA further enhanced the preservation of feature information throughout the sampling process, resulting in outstanding segmentation outcomes.

The network architecture for forecasting the output of the carotid plaque segmentation region is illustrated in Fig. 2. In this structure, the input image undergoes an initial convolution at various dilation rates through the atrous convolution layer. Subsequently, depthwise convolution and pointwise convolution employing depthwise separable convolution techniques were applied to finetune the model parameters. The target region is obtained through bicubic interpolation, yielding a high-resolution segmentation region.

The DASPP module reduces the number of network parameters during model training. ASPP through hole convolution effectively extends the perceptual field range by processing the input image with various atrous convolution rates. This process divided the features sampled at each unique hole convolution rate into separate branches for subsequent processing. By employing different atrous rates, the model can capture a broader context of image information while avoiding adverse effects on image resolution that may result from large step sizes during convolution. Equation 3 defines the ASPP hole pooling process, where Q(c, d) represents the output outcome of the hole convolution in coordinates (c, d). Given that depthwise separable convolution



Fig. 2 Network output prediction using the DASPP-BICECA module

varies its convolution kernels based on the different channels of the input network, the convolution process is bifurcated into two segments. The depthwise convolution operation is expressed in Eq. 4, followed by the pointwise convolution operation in Eq. 5. Equation 6 encapsulates the ultimate output result of depthwise separable convolution. Here, $DConv(\beta_d, Q)_{(i,j)}$ denotes the output of depthwise convolution, $PConv(\beta_p, Q)_{(i,j)}$ stands for the output of pointwise convolution, and $DSConv(\beta_d, \beta_p, Q)_{(i,j)}$ represents the output of the depthwise separable convolution.

$$Q(\mathbf{c},\mathbf{d}) = \sum_{i,j} q(\mathbf{c} + \mathbf{e} \cdot \mathbf{i},\mathbf{d} + \mathbf{e} \cdot \mathbf{j}) \cdot k(\mathbf{i},\mathbf{j})$$
(3)

where q is the input information, e is the rate of atrous convolution, k is the filter, and (i, j) is the location of the atrous convolution layer where the convolution is performed.

$$DConv(\beta_d, Q)_{(i,j)} = \sum_{m,n}^{M,N} \beta_d(m,n) Q(i+m,j+n)$$
(4)

$$PConv(\beta_p, Q)_{(i,j)} = \sum_{\nu}^{\nu} \beta_p(\nu) Q(i,j)$$
(5)

$$DSConv(\beta_d, \beta_p, Q)_{(i,j)} = PConv(\beta_p, Q)_{(i,j)}(\beta_p, DConv(\beta_d, Q)_{(i,j)})$$
(6)

where Q represents the input information; β_d represents the convolutional layer weight of the channel convolution; β_p represents the convolutional layer weight of the point convolution; M and N represent the dimensions of the convolutional layer, respectively; and V represents the point convolution of a channel. To increase the model's sensitivity to the feature points within the target region, BICECA was introduced into the prediction mask output segment. This attention mechanism safeguards critical information during the convolution process by allocating distinct weight coefficients to the input feature regions, and subsequently selecting the region information to be segmented. ECA employs dynamic convolution kernels, treating a one-dimensional convolution as a nonfully connected layer, with each convolution operation affecting only a fraction of the convolution layers.

The input data pass through a global average pooling layer with activation, converting two-dimensional convolution into a one-dimensional counterpart, as expressed in Eq. 7. The local cross-channel interaction operation of the one-dimensional convolution, detailed in Eqs. 8 and 9, combines the input information to derive the attention factor. This attention factor is then integrated with the two-dimensional input information through the activation function to yield the attention channel output, denoted by $Q_{(i,j)}^{(H*W*C)}$, as shown in Eq. 10. Bicubic interpolation leverages the grayscale values surrounding the sampled pixel points for interpolation. It fits the grayscale influence of 16 neighboring pixel points in an adjacent area and determines the pixel value of the target pixel through a weighted summation of the surrounding pixel values. Equation 11 outlines the dual-cubic interpolation process, where G(x, y) is the output of the dual-cubic linear function.

$$q_{avg(ij)}^{(1*1*C)} = \sum_{i,j} F\left(Relu\left(GAP_{avg}\left(DSConv(Q)^{(H*W*C)}\right)\right)\right)_{(ij)}$$
(7)

$$\phi_{c(i,j)}^{(1*1*C)} = \sum_{i,j} \alpha \left(LCCI\left(q_{avg}^{(1*1*C)}\right) \right)_{(i,j)}$$
(8)

$$\eta_{A(i,j)}^{(H*W*C)} = \sum_{i,j} \left(\phi_{c(i,j)}^{1*1*C} \otimes DSConv(Q)^{(H*W*C)} \right)_{(i,j)}$$
(9)

$$Q_{(ij)}^{(H*W*C)} = \sum_{ij} Relu \left(\eta_A^{(H*W*C)} \oplus DSConv(Q)^{(H*W*C)} \right)_{(ij)}$$
(10)

where $q_{avg(i,j)}^{(1*1*C)}$ is the one-dimensional output after a global average pooling function, $\phi_{c(i,j)}^{(1*1*C)}$ is the output after a local cross-channel interaction operation, α is the sigmoid function, $\eta_{A(i,j)}^{(H*W*C)}$ expresses the attention factor, and *Relu* is the activation function.

$$G(x,y) = \sum_{i}^{3} \sum_{j}^{3} Q_{(i,j)} \delta(y - y_j) \delta(x - x_i)$$
(11)

segmentation mask area without affecting the final model output. The difference between the predicted and actual mask areas served as the basis for calculating the loss.

This loss value was then combined with the loss derived from the DASPP-BICECA network to form a comprehensive loss function. This refined function provides a more accurate representation of the gap between the predicted and actual values, resulting in an enhanced segmentation model, as illustrated in Fig. 3.

Equation 12 illustrates the output following the reflection-filled convolution operation, whereas Eq. 13 shows the results after the BINet upsampling of the output. Here, r^N denotes the output of the reflection-filled convolution layer, and R^n represents the output of the upsampling module.

$$r^{N} = \sum_{1}^{N} \left(\sum_{i}^{K-1} \sum_{j}^{K-1} r(x+i, y+i) k_{(i,j)} \right)$$
(12)

$$R^{n} = F_{Avg} \left((1 - \alpha_{x}) \left(1 - \alpha_{y} \right) r_{(m,n)}^{N-1} + \alpha_{x} \left(1 - \alpha_{y} \right) r_{(m+1,n)}^{N-1} + (1 - \alpha_{x}) \alpha_{y} r_{(m,n+1)}^{N-1} + \alpha_{x} \alpha_{y} r_{(m+1,n+1)}^{N-1} \right)$$
(13)

where $Q_{(i,j)}$ is the input information and $\delta(x_i)$ and $\delta(y_j)$ are the interpolation weighting factors in the horizontal and vertical directions, respectively.

Joint assessment of network loss module

BINet is introduced as a collaborative evaluation-loss module for PlaqueNet. It uses the feature information extracted by AResNet and employs it to predict the where *N* is the number of convolutional layers, r(x + i, y + i) is the pixel value at (x + i, y + i), and *avg* is a pooling function.

BINet incorporates a reflective filling convolution operation combined with bilinear upsampling. This method ensures the consistency of the nearest-neighbor interpolation throughout the upsampling procedure, effectively preventing any disruptions in the prediction mask within the segmentation region.



Fig. 3 Loss of BINet network joint assessment model

Algorithm 1 is used as an example to offer an intuitive depiction of the process of jointly evaluating PlaqueNet's loss function.

Algorithm 1 Joint evaluation of cross-entropy loss function

Input: Actual split output network DASPP-ECABIC, evaluation segmented output network BINET. Real mask Y. D and r are the feature information entering the two networks, respectively, Output: The joint output of the cross-entropy loss $CELoss(\omega)$ 1: While each $n_1 \in to [1, N_1]$ do 2: $Q(D) \leftarrow \sum_{i,j} q(c + eD_i, d + eD_j)k$ $DSConv(D) \leftarrow \begin{cases} (DConv(D) - \sum_{(m,n)} \beta_d Q(D)) \\ (PConv(D) - \sum_{(m,n)} \beta_p Q(D)) \end{cases}$ 3: $q(1,1,C) \leftarrow F(Relu) \leftarrow GAP(Q(D_{(H,W,C)}))$ 4: 5: $\eta_A \leftarrow LCCI(q_{(1,1,C)}) \otimes Q(D_{(H,W,C)})$ 6: $Q_{(H,W,C)} \leftarrow Relu(\eta_A) \oplus Q(D_{(H,W,C)})$ 7: $G(D) \leftarrow \sum_{i}^{3} \sum_{j}^{3} Q \delta_{xi} \delta_{yj}$ 8: end while 9: While each $n_2 \in to [1, N_2]$ do 10: $r \leftarrow \sum_i \sum_j (r_{xi}) (r_{yi})k$ 11: $R \leftarrow F(AVG) \leftarrow g(\alpha_x, \alpha_y)r$ 12: end while $13:CELoss(G) \leftarrow \frac{-1}{N} (\sum_{i} \sum_{k} Y(i, k) lnG(i, k))$ $14: CELoss(D) \leftarrow \frac{-1}{N} (\sum_{i} \sum_{k} Y(i, k) ln D(i, k))$ $15:CELoss(\omega) \leftarrow F(CELoss(G), CELoss(D))$

Results

The goal was to extract and convert these slices into planar images for segmentation and recognition. The research dataset included 742 images, divided into a training set of 519 images and a test set of 223 images. The PlaqueNet segmentation model we trained using these two-dimensional slice images of the vascular plaque model. In addition, three established segmentation algorithms (FCN, Deeplabv3, and Deeplabv3plus) were evaluated by considering their parameter configurations and segmentation results to assess the performance of PlaqueNet.

Four control experiments were conducted using the dataset presented in this study, focusing on FCN, Deeplabv3, Deeplabv3plus, and PlaqueNet. The results are shown in Fig. 4, with rows one to four representing the employed segmentation algorithms and columns one to nine displaying the outcomes produced by these algorithms. FCN's segmentation results revealed substantial discontinuous segmentation regions. Deeplabv3 exhibits both discontinuous and over-segmented results. Deeplabv3plus's results suffer from excessive segmentation. In contrast, PlaqueNet's segmentation results surpassed those of the previous three segmentation algorithms. There were no discontinuous or over



Fig. 4 Comparison of four segmentation algorithms, each represented by a row, and the columns display the segmentation result plots. The first row corresponds to the FCN algorithm, which exhibited noticeable segmentation gaps. The second row represents the Deeplabv3 algorithm, which shown a significant over-segmentation in its results. The third row showcases the Deeplabv3plus algorithm, which demonstrates a lower degree of over-segmentation. The fourth row presents the PlaqueNet algorithm proposed in this study

segmented areas, and the entire segmented region effectively covered the target area.

In this study, the segmentation results generated by the four different algorithms were evaluated using six key evaluation metrics: intersection over Union (IoU), Dice, accuracy, mean IoU (mIoU), mean Dice (mDice)

Algorithm	IoU (%)	Dice (%)	Accuracy (%)	mloU (%)	mDice (%)
FCN	61.72±5.24	76.33±4.22	67.86±7.32	80.88±2.63	88.15±2.11
Deeplabv3	69.04 ± 5.87	81.66±6.11	77.32±6.57	84.49±2.94	90.83±2.23
Deeplabv3plus	72.87±4.49	84.31±3.07	78.52 ± 4.33	86.42±2.25	92.14±1.54
PlaqueNet	87.37±10.38	93.26±8.36	93.12±10.66	93.68 ± 5.20	96.63±4.19

Table 1 Comparison of segmentation performance of different residual networks

 Table 2
 Comparison of evaluation metrics for the six segmentation algorithms

Algorithm	Precision (%)	Recall (%)	F1score (%)
ResNet	89.943±5.146	76.330±4.479	67.860±3.075
ResNest	89.992 ± 2.875	71.399 ± 10.525	79.305 ± 8.558
ResNetVC	89.081 ± 3.944	80.586 ± 4.429	84.797±3.512
ResNetVD	88.717 ± 3.944	80.185 ± 3.124	84.800 ± 3.844
ResNext	90.200 ± 3.818	70.123 ± 9.483	79.407 ± 8.896
AResNet	93.423 ± 1.936	90.623±1.918	92.354±1.452

and mean pixel accuracy (MPA). The Dice, which is a pixel-level similarity measure, is commonly employed to assess segmentation performance, with higher values indicating more accurate segmentation. The IoU indicates the degree of overlap between segmented and actual regions. The accuracy, a measure of model accuracy, quantifies the proportion of correct classifications in the entire dataset, offering insights into the model's quality. The mIoU, which is the average intersection ratio, provides an overview of the IoU across the entire dataset and represent the average IoU across all categories. Similarly, the mDice was computed by averaging the Dice coefficients across all dataset categories. MPA is an improvement in pixel accuracy. It calculates the proportion of correctly classified pixels in each class, and then calculates the average of all classes. Table 1 illustrates that PlaqueNet outperformed the other three segmentation algorithms across all evaluation metrics, underscoring the enhancement in segmentation accuracy achieved by PlaqueNet.

To further validate the segmentation performance of PlaqueNet as presented in this study, ten experiments were conducted to compare FCN, Deeplabv3, Deeplabv3plus, and PlaqueNet. The evaluation metrics included accuracy, Dice, IoU, and precision. Figure 5 clearly demonstrates that PlaqueNet's segmentation performance surpasses that of the other three algorithms. To confirm the superior performance of the proposed AResNet structure in image segmentation compared to other residual structures, five common residual structures were selected for comparison. Their performances were evaluated using metrics such as precision, recall, F1score, and loss. As indicated in Table 2, AResNet outperformed all of the five structures. To illustrate the influence of joint evaluation network loss on segmentation performance, experiments were conducted using PlaqueNet and compare its performance with and without the inclusion of joint evaluation loss. As shown in Fig. 6, the incorporation of the joint evaluation loss into the segmentation algorithm significantly improved its overall performance.

Discussion

This study focused on the two-dimensional image segmentation of coronary artery plaques. To solve this problem, PlaqueNet, which segments coronary artery plaques from CCTA images, was introduced. In the initial stage of feature information extraction, a multi-path, parallel residual structure was introduced. This innovative structure significantly bolsters the feature extraction capacity of segmentation networks. This is accomplished by leveraging both the pooling mapping function and the original residual mapping function. Notably, this design mitigates the issue of gradient disappearance that often occurs in deep networks.

To further enhance the ability of segmentation network to capture feature information and minimize data loss, DASPP-BICECA module was presented. The BICECA component amplifies local feature sensitivity by addressing potential shortcomings of the network output segmentation, whereas the DASPP component expands the network's information-gathering scope. Additionally, BINet for joint network loss evaluation was introduced. It optimizes the segmentation model and enhances its overall efficiency when used in conjunction with the DASPP-BICECA module.

Vascular plaques have diverse shapes and sizes across various scales. To effectively capture the feature information at different scales, a multi-scale module was employed in the algorithm used in this study. This module enables the model to accurately locate detailed information about plaques while maintaining robust segmentation performance under different scales and transformation conditions. The multi-scale module is designed to extract and incorporate features from multiple scales during segmentation. By considering information from different scales, the model effectively captures the intricate details of vascular plaques, regardless of their varying sizes and shapes. This capability enhances the ability



Fig. 5 Analysis of segmentation evaluation metrics for algorithms



on segmentation performance

of the model to accurately segment plaques across different scales and under different transformation conditions. The integration of the multi-scale module into the algorithm ensures that the model can effectively adapt to the variability in vascular plaque characteristics. This adaptability is essential for achieving reliable and consistent segmentation results, because plaque morphology can vary significantly across different patients, imaging modalities, and acquisition techniques. By leveraging the multi-scale module, the algorithm demonstrated improved performance in accurately locating and segmenting vascular plaques across a range of scales. These advancements have contributed to the development of more precise and clinically relevant techniques for diagnosing and treating of vascular diseases. The advantages of employing PlaqueNet for detecting coronary artery plaques are substantial. This enables the early detection of these plaques through image segmentation, thus facilitating proactive treatment and reducing the risk of cardiovascular diseases. The proposed segmentation algorithm offers precise segmentation of coronary artery plaques. This information assists healthcare professionals in evaluating disease progression and conducting personalized plaque analyses. This sets the stage for the development of more tailored clinical treatment plans.

This study has several limitations: Owing to its focus on two-dimensional image segmentation of coronary artery plaques, this approach omits critical three-dimensional structural information, which can provide a more comprehensive view of these plaques. The selected twodimensional slices may not fully represent the entire spectrum of characteristics and intricate details of the plaques. Future studies will explore the real-time threedimensional segmentation of medical data format images, with a specific focus on coronary artery plaques.

Conclusions

This study introduced PlaqueNet, a novel approach to carotid plaque segmentation. PlaqueNet's feature extraction component employs a deep parallel residual optimization mapping network that integrates a deep residual optimization structure into each residual structure in ResNet. This optimization helps maintain global information in the input feature point field, addressing issues such as gradient disappearance and explosion caused by the network depth. The DASPP-BICECA module was used in the prediction mask output component of PlaqueNet. This module employs depth separable spatial convolution pyramid operations to expand the receptive field range of the target area during information upsampling. By initially conducting channel convolution followed by point convolution, the model training process reduces the parameter count. The BICECA module enhanced the network's sensitivity to feature points in the target area and mitigated losses during training. Bicubic interpolation helps prevent discontinuous segmentation of adjacent feature-point areas.

Furthermore, the BINet fitting evaluation network loss module collaborates with the DASPP-BICECA module to optimize the segmentation network model. The proposed segmentation algorithm was compared with three others. The experimental results demonstrate that the proposed method achieves impressive metrics: an IoU value of 87.37%, a Dice value of 93.26%, an accuracy value of 93.12%, an mIoU value of 93.68%, an mDice value of 96.63%, and an MPA value of 96.55%. The proposed algorithm outperforms the others in terms of segmentation accuracy, avoids discontinuous or over segmented areas, and demonstrates robust segmentation performance.

Abbreviations

CCTA	Coronary computed tomography angiography
AResNet	Advanced residual net
DASPP	Depthwise atrous spatial pyramid pooling
BICECA	Bicubic efficient channel attention
BN	Batch normalization
loU	Intersection over union
mloU	Mean IoU
mDice	Mean Dice
MPA	Mean pixel accuracy

Acknowledgements

We gratefully thank the doctors in the Affiliated Hospital of Shanxi Medical University for labeling the raw data.

Authors' contributions

LYW and XFZ proposed experimental methods and complete the whole experiment; CYT and SC complete manuscript writing and revise the manuscript; YZD, XYL and QW complete the organization of the graphs and tables in the manuscript; WXS and LZ make modifications to the overall arrangement of the manuscript.

Funding

This study was supported by the Major Science and Technology Scheme under Key Medical Research Project of Shanxi Province, No. 2021XM04; the National Natural Science Foundation of China, Nos. U22A2034 and 62072452; the Shenzhen Fundamental Research Program, Nos. JCYJ20200109110420626, JCYJ20200109110208764, and JCYJ20200109115201707.

Availability of data materials

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Declarations

Competing interests

The authors declare no conflict of interest.

Received: 25 October 2023 Accepted: 3 March 2024 Published online: 22 March 2024

References

- Wang H, Wang H, Huang ZL, Su HJ, Gao X, Huang FF (2021) Deep learning-based computed tomography images for quantitative measurement of the correlation between epicardial adipose tissue volume and coronary heart disease. SciProgram 2021:9866114. https://doi.org/10. 1155/2021/9866114
- Ramasamy A, Safi H, Moon JC, Andiapen M, Rathod KS, Maurovich-Horvat P et al (2020) Evaluation of the efficacy of computed tomographic coronary angiography in assessing coronary artery morphology and physiology: rationale and study design. Cardiology 145(5):285–293. https://doi. org/10.1159/000506537
- Han N, Ma Y, Li Y, Zheng Y, Wu C, Gan TJ et al (2023) Imaging and hemodynamic characteristics of vulnerable carotid plaques and artificial intelligence applications in plaque classification and segmentation. BrainSci 13(1):143. https://doi.org/10.3390/brainsci13010143
- Lee J, Kim JN, Gomez-Perez L, Gharaibeh Y, Motairek I, Pereira GTR et al (2022) Automated segmentation of microvessels in intravascular OCT images using deep learning. Bioengineering 9(11):648. https://doi.org/10. 3390/bioengineering9110648
- Li L, Jia T (2019) Optical coherence tomography vulnerable plaque segmentation based on deep residual U-Net. Reviews in cardiovascular medicine 20(3):171–177. https://doi.org/10.31083/j.rcm.2019.03.5201
- Yoon H, Park M, Yeom S, Kirkcaldie MTK, Summons P, Lee SH (2021) Automatic detection of amyloid beta plaques in somatosensory cortex of an Alzheimer's disease mouse using deep learning. IEEE Access 9:161926– 161936. https://doi.org/10.1109/ACCESS.2021.3132401
- Navon E, Miller O, Averbuch A (2005) Color image segmentation based on adaptive local thresholds. Image Vis comput 23(1):69–85. https://doi. org/10.1016/j.imavis.2004.05.011
- Bayá AE, Larese MG, Namías R (2017) Clustering stability for automated color image segmentation. Expert Syst Appl 86:258–273. https://doi.org/ 10.1016/j.eswa.2017.05.064
- Aslan MF (2022) A robust semantic lung segmentation study for CNNbased COVID-19 diagnosis. Chemom Intell Lab Syst 231:104695. https:// doi.org/10.1016/j.chemolab.2022.104695
- Wang EK, Chen CM, Hassan MM, Almogren A (2020) A deep learning based medical image segmentation technique in internet-of-medicalthings domain. Future Gener Comput Syst 108:135–144. https://doi.org/ 10.1016/j.future.2020.02.054
- Zhang ZY, Li Y, Shin BS (2022) Robust color medical image segmentation on unseen domain by randomized illumination enhancement. Comput Biol Med 145:105427. https://doi.org/10.1016/j.compbiomed.2022.105427
- Pant G, Yadav DP, Gaur A (2020) ResNeXt convolution neural network topology based deep learning model for identification and classification of pediastrum. Algal Res 48:101932. https://doi.org/10.1016/j.algal.2020. 101932
- 13. Khan MM, Uddin MS, Parvez MZ, Nahar L (2022) A squeeze and excitation ResNeXt-based deep learning model for Bangla handwritten compound character recognition. King Saud Univ Comput Inf Sci 34(6):3356–3364. https://doi.org/10.1016/j.jksuci.2021.01.021
- Zhan JY, Wang J, Ben ZF, Ruan HD, Chen SJ (2019) Recognition of angiographic atherosclerotic plaque development based on deep learning. IEEE Access 7:170807–170819. https://doi.org/10.1109/ACCESS.2019. 2954626
- Csippa B, Mihály Z, Czinege Z, Németh MB, Halász G, Paál G et al (2021) Comparison of manual versus semi-automatic segmentations of the stenotic carotid artery bifurcation. Appl Sci 11(17):8192. https://doi.org/ 10.3390/app11178192
- Xu WJ, Yang X, Li YK, Jiang GH, Jia S, Gong ZH et al (2022) Deep learningbased automated detection of arterial vessel wall and plaque on magnetic resonance vessel wall images. Front Neurosci 16:888814. https://doi. org/10.3389/fnins.2022.888814
- Xu W, Zhu Q (2022) A semantic segmentation method with emphasis on the edges for automatic vessel wall analysis. Appl Sci 12(14):7012. https:// doi.org/10.3390/app12147012

- Shin CI, Park SJ, Kim JH, Yoon YE, Park EA, Koo BK et al (2021) Coronary artery lumen segmentation using location-adaptive threshold in coronary computed tomographic angiography: A proof-of-concept. Korean J Radiol 22(5):688–696. https://doi.org/10.3348/kjr.2020.0296
- Wen XB, Zhao B, Yuan MF, Li JZ, Sun MZ, Ma LS et al (2022) Application of multi-scale fusion attention U-Net to segment the thyroid gland on localized computed tomography images for radiotherapy. Front Oncol 12:844052. https://doi.org/10.3389/fonc.2022.844052
- Bi H, Sun JW, Jiang YB, Ni XY, Shu HZ (2022) Structure boundary-preserving U-Net for prostate ultrasound images segmentation. Front Oncol 12:900340. https://doi.org/10.3389/fonc.2022.900340
- Xi YH, Zhong LM, Xie WJ, Qin GG, Liu YB, Feng QJ et al (2021) View identification assisted fully convolutional network for lung field segmentation of frontal and lateral chest radiographs. IEEE Access 9:59835–59847. https:// doi.org/10.1109/ACCESS.2021.3074026
- Tang CW, Chen SY, Zhou X, Ruan S, Wen HT (2020) Small-scale face detection based on improved R-FCN. Appl Sci 10(12):4177. https://doi.org/10. 3390/app10124177
- Wang JL, Luo JX, Liu B, Feng R, Lu LN, Zou HD (2020) Automated diabetic retinopathy grading and lesion detection based on the modified R-FCN object-detection algorithm. IET Comput Vis 14(1):1–8. https://doi.org/10. 1049/iet-cvi.2018.5508
- Zhang R, Cheng C, Zhao XH, Li XC (2019) Multiscale mask R-CNN-based lung tumor detection using pet imaging. Mol imaging 18. https://doi. org/10.1177/1536012119863531
- Qin JP, Zhang Y, Zhou H, Yu F, Sun B, Wang QS (2021) Protein crystal instance segmentation based on mask R-CNN. Crystals 11(2):157. https:// doi.org/10.3390/cryst11020157
- Zhang YQ, Chu J, Leng L, Miao J (2020) Mask-refined R-CNN: A network for refining object details in instance segmentation. Sensors 20(4):1010. https://doi.org/10.3390/s20041010
- Zhang XF, Bian HN, Cai YH, Zhang KY, Li H (2022) An improved tongue image segmentation algorithm based on deeplabv3 + framework. IET Image Processing 16(5):1473–1485. https://doi.org/10.1049/ipr2.12425
- Bai RF, Jiang S, Sun HJ, Yang YF, Li GJ (2021) Deep neural network-based semantic segmentation of microvascular decompression images. Sensors 21(4):1167. https://doi.org/10.3390/s21041167
- Khodadadi Shoushtari F, Sina S, Dehkordi ANV (2022) Automatic segmentation of glioblastoma multiform brain tumor in MRI images: Using deeplabv3 + with pre-trained resnet18 weights. Phys Med 100:51–63. https://doi.org/10.1016/j.ejmp.2022.06.007
- Sreng S, Maneerat N, Hamamoto K, Win KY (2020) Deep learning for optic disc segmentation and glaucoma diagnosis on retinal images. Appl Sci 10(14):4916. https://doi.org/10.3390/app10144916
- Webb JM, Meixner DD, Adusei SA, Polley EC, Fatemi M, Alizad A (2020) Automatic deep learning semantic segmentation of ultrasound thyroid cineclips using recurrent fully convolutional networks. IEEE Access 9:5119–5127. https://doi.org/10.1109/ACCESS.2020.3045906
- Varnyú D, Szirmay-Kalos L (2022) A comparative study of deep neural networks for real-time semantic segmentation during the transurethral resection of bladder tumors. Diagnostics 12(11):2849. https://doi.org/10. 3390/diagnostics12112849
- Chang CW, Christian M, Chang DH, Lai F, Liu TJ, Chen YS et al (2022) Deep learning approach based on superpixel segmentation assisted labeling for automatic pressure ulcer diagnosis. PLos one 17(2):e0264139. https:// doi.org/10.1371/journal.pone.0264139
- Hsu CY, Hu R, Xiang Y, Long X, Li ZY (2022) Improving the Deeplabv3+ model with attention mechanisms applied to eye detection and segmentation. Mathematics 10(15):2597. https://doi.org/10.3390/math1 0152597
- Huang HY, You ZY, Cai HY, Xu JF, Lin DX (2022) Fast detection method for prostate cancer cells based on an integrated ResNet50 and YoloV5 framework. Comput Methods Prog Biomed 226:107184. https://doi.org/ 10.1016/j.cmpb.2022.107184
- Çınar A, Yıldırım M, Eroğlu Y (2021) Classification of pneumonia cell images using improved ResNet50 model. Trait du Signal 38(1):165–173. https://doi.org/10.18280/ts.380117
- Sharma AK, Nandal A, Dhaka A, Koundal D, Bogatinoska DC, Alyami H (2022) Enhanced watershed segmentation algorithm-based modified ResNet50 model for brain tumor detection. BioMed Res Int 2022:7348344. https://doi.org/10.1155/2022/7348344.

 Elpeltagy M, Sallam H (2021) Automatic prediction of COVID-19 from chest images using modified ResNet50. Multimed tools Appl 80(17):26451–26463. https://doi.org/10.1007/s11042-021-10783-6

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.