MBFF-Net: Multi-Branch Feature Fusion Network for Carotid Plaque Segmentation in Ultrasound

Shiyu Mi[1], Qiqi Bao[1], Zhanghong Wei[2], Fan Xu[3], and Wenming Yang[1,3(✉)]

[1] Department of Electronic Engineering, Shenzhen International Graduate School, Tsinghua University, Shenzhen, Guangdong, China
yang.wenming@sz.tsinghua.edu.cn
[2] Department of ultrasound, Shenzhen People's Hospital, Shenzhen, Guangdong, China
[3] Peng Cheng Laboratory, Shenzhen, Guangdong, China

Abstract. Stroke is one of the leading causes of death around the world. Segmenting atherosclerotic plaques in carotid arteries from ultrasound images is of great value for preventing and treating ischemic stroke, yet still challenging due to the ambiguous boundary of plaque and intense noise in ultrasound. In this paper, we introduce a new approach for carotid plaque segmentation, namely Multi-Branch Feature Fusion Network (MBFF-Net). Inspired by the prior knowledge that carotid plaques generally grow in carotid artery walls (CAWs), we design a Multi-Branch Feature Fusion (MBFF) module with three branches. Specifically, the first two branches are well-designed to extract plaque features of multiple scales and different contexts, and the other branch is to exploit the prior information of CAWs. In addition, a boundary preserving structure is applied to alleviate the ambiguity of plaque boundary. With the proposed MBFF and the novel structure, our model is capable of extracting discriminative features of plaques and integrating the location information of CAWs for better segmentation. Experiments on the clinical dataset demonstrate that our model outperforms state-of-the-art methods. Code is available at https://github.com/mishiyu/MBFF.

Keywords: Carotid plaque · Segmentation · Ultrasound image · Deep learning · Stroke

1 Introduction

Stroke is the second leading cause of death in adults worldwide [1,2]. And atherosclerotic plaque in the carotid artery is an important cause of ischemic stroke. Currently, B-mode ultrasound is the routine imaging modality to examine carotid atherosclerotic plaques. Evaluating the condition of plaques requires

S. Mi and Q. Bao—These authors contributed equally to this work.

M. de Bruijne et al. (Eds.): MICCAI 2021, LNCS 12905, pp. 313–322, 2021.
https://doi.org/10.1007/978-3-030-87240-3_30

manually delineating lesion regions by radiologists. However, such subjective examination is prone to a high diagnostic error rate. Thus, an automatic and reliable method needs to be designed to segment plaques, improving clinical diagnosis efficiency.

The problem of carotid plaque segmentation in carotid longitudinal section (CLS) ultrasound images and transverse section ultrasound images has been extensively explored in the literature. Since carotid plaques generally grow in CAWs, there are many studies on the detection of CAW. Methods proposed in [6–8] were dedicated to solving the problem of arterial wall segmentation in images of transverse sections. For CLS images, Sifakis et al. [4] exploited basic statistics along with anatomical knowledge to recognize the carotid artery. The predicted result can further facilitate arterial segmentation in CLS ultrasound images. Golemati et al. [5] investigated the possibility of applying the Hough transform to segment the CAW from ultrasound image sequences. Azzopardi et al. [20] explored the boundary segmentation of the carotid artery in ultrasound images with convolutional neural networks. However, these methods only detect CAW and do not further segment plaques.

Many methods explicitly detect plaques in ultrasound images. Some of them [11–13] tried to segment plaque in images of transverse sections. For CLS images, Destrempes et al. [9] utilized motion estimation and Bayesian model to segment the plaque in CLS images. However, since it needs manual annotation of the first frame, the operator's subjectivity may affect the segmentation results. Yoneyama et al. [10] compared a semi-automatic method called morphology-enhanced probabilistic with the human reading method in plaque segmentation, showing a good correlation between these two methods. Nevertheless, these traditional methods need to extract features manually, which is not flexible and universal. With the rapid development of deep learning, many methods [14–19] have achieved great success in the image segmentation task. Recently, deep learning methods have also been applied to the segmentation of carotid plaques. Meshram et al. [21] presented a dilated U-Net architecture to segment plaques, finding the performance of the semi-automatic method with bounding box are significantly better than those of the fully automatic method. In general, the current plaque segmentation methods don't make good use of the prior information of CAW. Therefore, how to effectively utilize attributes of CLS ultrasound images for plaque segmentation remains a challenge.

In this paper, we propose the MBFF-Net for automatic carotid plaques segmentation in CLS ultrasound images. It is known that CAW is divided into intima, media, and adventitia. Since plaques usually appear between the intima and the media, detecting the carotid intima-media (IM) region can be of vital importance for detecting plaques [3]. Our model is well-designed based on the above-mentioned pathological knowledge of plaques and achieves discriminative feature extraction and segmentation performance. The main contributions of our work are outlined as follows.

Firstly, as plaques usually exist between the IM regions of CAWs, we refer our previous work of IM detection and use it as prior information to assist plaque segmentation.

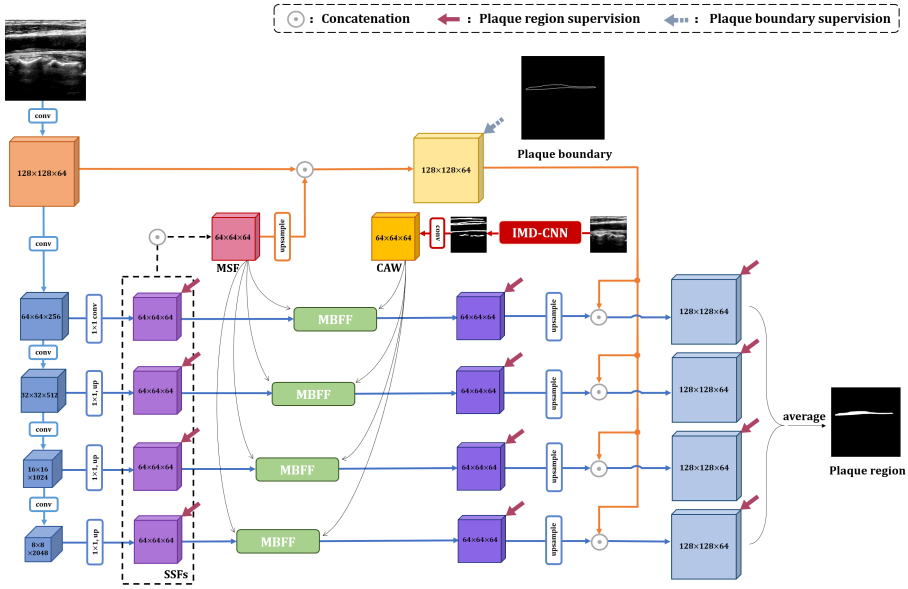

Fig. 1. The overview framework of the proposed MBFF-Net, in which the orange lines show the boundary preserving structure, the red lines show the extraction process of CAWs, and the blue lines show the extraction process of plaques. (Color figure online)

Secondly, we design a new MBFF module to use the prior information, fusing the features of multi-scales and multi-contexts.

Thirdly, we present a boundary preserving structure to enhance the plaque boundary, which diminishes the influence of blurred boundaries of plaque in ultrasound images.

Finally, through comparative experiments, we show that MBFF-Net is superior to existing general segmentation networks in plaque segmentation of CLS ultrasound images and achieves state-of-the-art.

2 Method

The overview of the proposed MBFF-Net is illustrated in Fig. 1. It takes carotid ultrasound images as inputs and outputs the segmentation results. Specifically, the original image is passed through the IMD-CNN and a CNN (ResNeXt101 [22] is used here) to detect the CAW and generate a set of feature maps with different resolutions, respectively. Then, the set of feature maps (except the feature maps at the shallowest layer) is used to generate Multi-Scale Feature (MSF). Afterwards, the MBFF module integrates MSF, CAW, and Single-Scale Feature (SSF) of each layer to produce the refined feature maps.

Furthermore, the feature maps at the shallowest layer are used for segmenting the plaque boundary after combining with the MSF. Subsequently, the feature

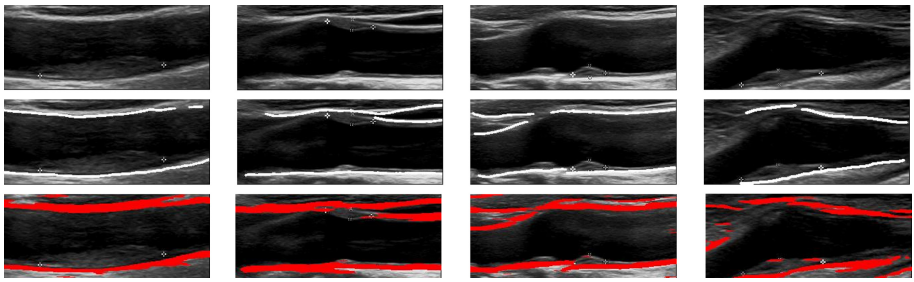

Fig. 2. Examples of the IM detection results. The first row shows the original images, the second row shows the manually labeled ground truth masks, and the third row shows the outputs of the IMD-CNN (the IM regions are shown in red color). (Color figure online)

of the boundary segmentation is also integrated into the refined features of other layers. At last, the final segmentation results of plaques are obtained by averaging the predicted maps of those layers. The following subsections present the details of each component in our MBFF-Net.

2.1 CAW Detecting Module

To embed the prior information of CAWs in the plaque segmentation model, a CAW detecting module is included in the MBFF-Net. This module is from our previous published work denoted as IMD-CNN [23], which uses the patch-wise training method to segment the IM regions in the CAW. Since collecting training data of medical images is a time-consuming, cumbersome and costly task, IMD-CNN is trained in a patch-wise manner instead of an image-wise manner. More concretely, each pixel is taken as the center to crop the patch. Then, IMD-CNN takes these patches as inputs and outputs pixel-wise labels indicating whether a pixel is in the IM regions. The IM detection results are shown in Fig. 2.

In this study, we collect a new CLS ultrasound images dataset different from the previous dataset. The IMD-CNN trained in a patch-wise manner is not affected by the change of image size. Therefore, in this new dataset, CAWs are obtained by adopting the original network structure and parameters directly.

2.2 MBFF Module

The structure of the MBFF module is shown in Fig. 3. Given the SSF of the m-th layer, denoted as Φ_m, it is sent into three branches, i.e., the multi-scale branch, the multi-context branch and the CAW-guided branch, boosting the performance of plaque segmentation. Each branch is described in detail below.

The first branch aims to fuse multi-scale features, integrating feature maps at shallow and deep layers to contain detailed information as well as capturing highly semantic information. In this branch, the Φ_m and MSF are concatenated

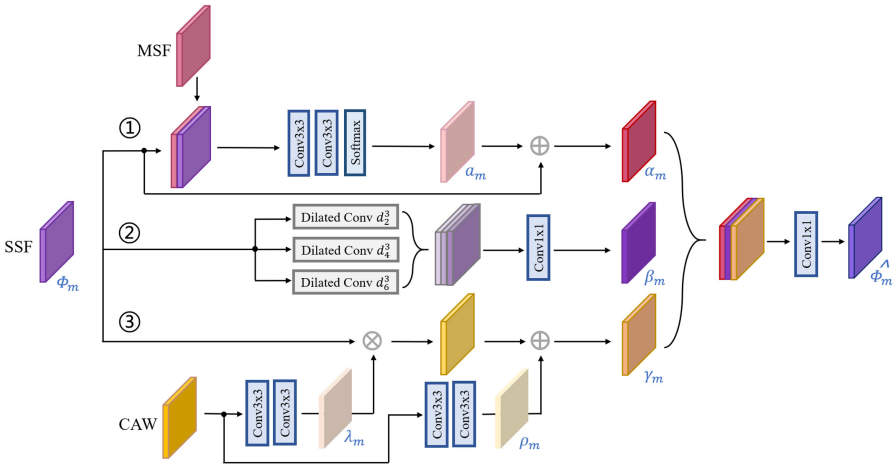

Fig. 3. The details of the MBFF Module at the m-th layer.

and sent into two convolutional layers followed by a Softmax operation to generate the intermediate feature a_m, and then the new feature α_m can be generated as

$$\alpha_m = \Phi_m + a_m \tag{1}$$

The second branch is designed to extract features of plaques considering multi-scale contextual information. By employing the dilated convolution, we extract features with a numerous range of receptive fields in this branch. Afterwards, these features with different receptive fields are concatenated across channel-wise. The generated feature β_m can be defined as

$$\beta_m = f_{conv}(d_2^3(\Phi_m) \odot d_4^3(\Phi_m) \odot d_6^3(\Phi_m)) \tag{2}$$

where $f_{conv}(\cdot)$ denotes the 1×1 convolution function, $d_q^p(\cdot)$ denotes dilated convolution function with a dilation rate q and filter size of $p \times p$, and $\odot$ denotes concatenation.

The third branch combines the location priors of CAWs, which are acquired from the feature around plaques. In this branch, spatial feature modulation is constructed, where affine transformation parameters are formulated with spatial feature transformation [24]. With the priors of the CAW, this branch learns a mapping function T to produce a modulation parameter pair (λ_m, ρ_m), which is written as:

$$(\lambda_m, \rho_m) = T(\Phi_m) \tag{3}$$

Afterwards, intermediate feature maps are adjusted by scaling and shifting according to the modulation parameters. The modified feature γ_m is defined as:

$$\gamma_m = \Phi_m \otimes \lambda_m + \rho_m \tag{4}$$

where $\otimes$ indicates element-wise multiplication.

In the end, the output feature $\hat{\Phi_m}$ of the MBFF module can be written as:

$$\hat{\Phi_m} = f_{conv}(\alpha_m \odot \beta_m \odot \gamma_m) \tag{5}$$

2.3 Boundary Preserving Structure

Due to the high noise and low quality of ultrasound images, the boundary region of the lesion is usually blurred, leading to difficulty in accurate segmentation. Therefore, to preserve the accurate information of the plaques' boundary in CLS ultrasound images, we embed a boundary preserving structure into the MBFF-Net. Specifically, we use the feature maps at the shallowest layer which retain the most detailed information to extract the plaque boundary. In detail, the MSF is concatenated with the feature of the shallowest layer to diminish the impact of missing semantic information. Then, a plaque boundary loss is used as a supervision to extract the boundary feature. In the end, the feature extracted from the boundary segmentation is integrated into the feature of each refined SSF to enhance the structure boundary of plaques.

3 Experiments

Dataset. Our experiments are conducted on the CLS ultrasound images of the carotid artery collected by Philips IU22 with the L9-3 probe and GE logiq E9 with the 9L probe. The two devices are of the same center frequency with 9MHz. A total of 430 images of different patients are included in this dataset with the size of 768×576. Since the ultrasound acquisition device generates ultrasound images surrounded by irrelevant parameter information, we crop the ultrasound images from the original images and then rescale them into the size of 256×256 to adjust the network's input size. In addition, randomly selected 330 images are used as the training dataset, and the remaining 100 images are used as the testing dataset.

Loss Function. Binary cross-entropy with logits loss is used as the loss function to optimize the whole network. The total loss (denoted as L_t) is defined as:

$$L_t = \sum_{i=1}^{n} \omega_i L_{1i} + \sum_{j=1}^{n} \omega_j L_{2j} + \sum_{k=1}^{n} \omega_k L_{3k} + \omega_b L_b \tag{6}$$

where L_{1i} denotes the loss of the feature before the MBFF module at the i-th layer, L_{2j} denotes the loss of the refined feature extracted by the MBFF module at the j-th layer, L_{3k} denotes the loss of the refined feature integrated with the boundary feature at the k-th layer, and L_b denotes the loss on plaque boundary. These supervised features are specifically identified by arrows in Fig. 1. Meanwhile, n denotes the number of layers, and ω denotes weight. All weights here are set as 1 empirically.

Evaluation Metrics. Widely-used evaluation indicators are employed to measure the performance of the proposed model, including Dice score, Intersection over Union (IoU), Precision, and Recall.

Implementation. Our model is implemented by PyTorch on the NVIDIA 2080Ti GPU. The SGD optimizer with the learning rate being 0.005 and the momentum being 0.9 is adopted to train the MBFF-Net. The maximum number of epochs is set to 100. Each epoch takes about forty seconds.

Ablation Experiment. To evaluate the efficiency of proposed components, we conduct ablation experiments and show the quantitative results in Table 1. It can be clearly observed that each component has a certain positive effect. We remove the MBFF module and the boundary preserving structure from MBFF-Net as the baseline. Results of the baseline are represented in the first row. Benefiting from the dilated convolution branch (denoted as DC), the second row indicates that the model is able to utilize contextual information for better segmentation. In the third row, it shows the results with the CAW-guided branch (denoted as CAW) added to the baseline. And the improved visualization results illustrate the effectiveness of the integrated prior knowledge. In the fourth row, the multi-scale feature fusion branch (denoted as MSF) is evaluated. Since the model is equipped with the ability of multi-scale feature extraction, it can be seen that the segmentation effect is improved. Then we combine these three branches together and obtain a further improvement, demonstrating the power of the proposed MBFF module. Finally, the boundary preserving structure (denoted as BPS) is included, and our proposed MBFF-Net achieves the best performance.

Table 1. Ablation results of different components.

DC	CAW	MSF	BPS	Dice	IoU	Precision	Recall
				0.666	0.577	0.900	0.627
✓				0.683	0.597	0.852	0.685
	✓			0.688	0.611	0.840	0.707
		✓		0.710	0.625	0.910	0.671
✓	✓	✓		0.736	0.647	0.896	0.710
✓	✓	✓	✓	0.780	0.702	0.849	0.797

Comparative Experiment. We quantitatively compare the results of our MBFF-Net with some state-of-the-art methods, including PSPNet [14], ResUNet [15], U-Net [16], SegNet [17] and DenseUNet [18]. The results are reported in Table 2. For fair comparison, we retrain these models with common implementations and adjust training parameters to obtain the best segmentation results. It can be observed in Table 2 that our method is superior to other methods in

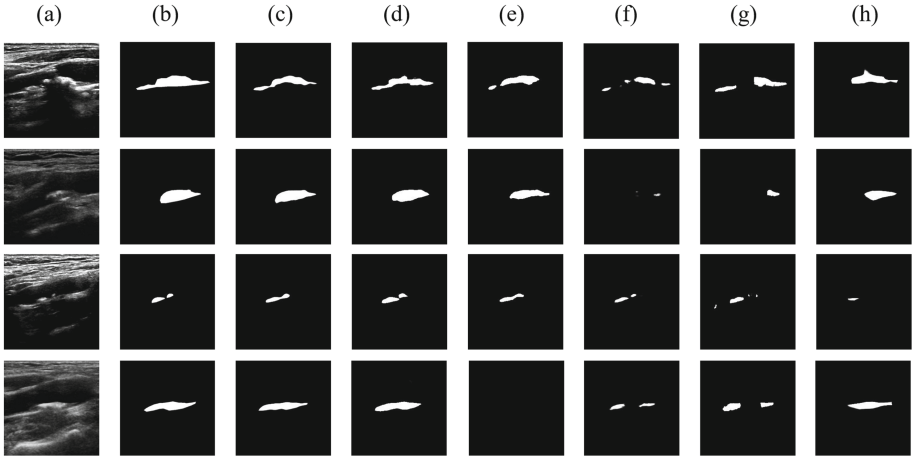

Fig. 4. Examples of plaque segmentation results produced by different methods. (a): CLS ultrasound images; (b): ground truth masks; (c)–(h): corresponding segmentations from our method, DenseUNet [18], SegNet [17], U-Net [16], ResUNet [15] and PSPNet [14], respectively.

almost all indicators. Figure 4 shows some of the segmentation results. Obviously, the plaque regions obtained by our method are more convincing.

Table 2. Metric results of diffrerent methods.

Method	Dice	IoU	Precision	Recall
PSPNet[14]	0.585	0.472	0.760	0.586
ResUNet[15]	0.621	0.521	0.755	0.635
U-Net[16]	0.630	0.552	0.817	0.658
SegNet[17]	0.687	0.593	0.856	0.676
DenseUNet[18]	0.778	0.690	**0.872**	0.763
Ours	**0.780**	**0.702**	0.849	**0.797**

4 Conclusion

In this paper, we propose a novel framework for carotid plaque segmentation in CLS ultrasound images. Complementary information is extracted from the features of multiple scales and different contexts. Besides, inspired by the physiological characteristics of plaque, we use the location priors of CAWs to guide plaque segmentation. Furthermore, the boundary preserving structure is designed to handle the fuzzy boundary of lesions and the noise in ultrasound. Experiments on the challenging CLS images demonstrate the superior performance of the MBFF-Net compared to the most advanced methods.

Acknowledgments. This work was supported by the Natural Science Foundation of Guangdong Province (No. 2020A1515010711) and the Special Foundation for the Development of Strategic Emerging Industries of Shenzhen (Nos. JCYJ20170818161845824, JCYJ20200109143010272 and JCYJ20200109143035495).

References

1. Feigin, V.L., Nguyen, G., Cercy, K., et al.: Global, regional, and country-specific lifetime risks of stroke, 1990 and 2016. N. Engl. J. Med. **379**(25), 2429–2437 (2018)
2. Johnson, C.O., Nguyen, M., Roth, G.A.: Global, regional, and national burden of stroke, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol. **18**(5), 439–458 (2019)
3. Touboul, P.J., Hennerici, M.G., Meairs, S., et al.: Mannheim carotid intima-media thickness and plaque consensus (2004-2006-2011). Cerebrovasc. Dis. **34**(4), 290–296 (2012)
4. Sifakis, E.G., Golemati, S.: Robust carotid artery recognition in longitudinal b-mode ultrasound images. IEEE Trans. Image Process. **23**(9), 3762–3772 (2014)
5. Golemati, S., Stoitsis, J., Sifakis, E.G., et al.: Using the Hough transform to segment ultrasound images of longitudinal and transverse sections of the carotid artery. Ultrasound Med. Biol. **33**(12), 1918–1932 (2008)
6. Zhang, J., Teng, Z., Guan, Q., et al.: Automatic segmentation of MR depicted carotid arterial boundary based on local priors and constrained global optimisation. IET Image Proc. **13**(3), 506–514 (2019)
7. China, D., Nag, M.K., Mandana, K.M., et al.: Automated in vivo delineation of lumen wall using intravascular ultrasound imaging. In: 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 4125–4128 (2016)
8. Samber, D.D., Ramachandran, S., Sahota, A., et al.: Segmentation of carotid arterial walls using neural networks. World J. Radiol. **12**, 1–9 (2020)
9. Destrempes, F., Soulez G., Giroux, M.F., et al.: Segmentation of plaques in sequences of ultrasonic B-mode images of carotid arteries based on motion estimation and Nakagami distributions. In: IEEE International Ultrasonics Symposium, pp. 2480–2483 (2010)
10. Yoneyama, T., et al.: In vivo semi-automatic segmentation of multicontrast cardiovascular magnetic resonance for prospective cohort studies on plaque tissue composition: initial experience. Int. J. Cardiovasc. Imaging **32**(1), 73–81 (2015). https://doi.org/10.1007/s10554-015-0704-0
11. Wei, M., Ran, Z., Yuan, Z., et al.: Plaque recognition of carotid ultrasound images based on deep residual network. In: IEEE 8th Joint International Information Technology and Artificial Intelligence Conference, pp. 931–934 (2019)
12. Van't, K.R., Naggara, O., Marsico, R., et al.: Automated versus manual in vivo segmentation of carotid plaque MRI. Am. J. Neuroradiol. **33**(8), 1621–1627 (2012)
13. Bonanno, L., Sottile, F., Ciurleo, R., et al.: Automatic algorithm for segmentation of atherosclerotic carotid plaque. J. Stroke Cerebrovasc. Dis. **26**(2), 411–416 (2017)
14. Zhao, H., Shi, J., Qi, X., et al.: Pyramid scene parsing network. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 2881–2890 (2017)
15. Zhang, Z., Liu, Q., Wang, Y.: Road extraction by deep residual U-Net. IEEE Geosci. Remote Sens. Lett. **15**(5), 749–753 (2018)

16. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing & Computer Assisted Intervention, pp. 234–241 (2015)
17. Badrinarayanan, V., Kendall, A., Cipolla, R.: SegNet: a deep convolutional encoder-decoder architecture for image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. **39**(12), 2481–2495 (2019)
18. Li, X., Chen, H., Qi, X., et al.: H-DenseUNet: hybrid densely connected UNet for liver and tumor segmentation from CT volumes. IEEE Trans. Med. Imaging **37**(12), 2663–2674 (2018)
19. Wang, Y., Deng, Z., Hu, X., et al.: Deep attentional features for prostate segmentation in ultrasound. In: International Conference on Medical Image Computing & Computer Assisted Intervention (2018)
20. Azzopardi, C., Hicks, Y.A., Camilleri, K.P.: Automatic carotid ultrasound segmentation using deep convolutional Neural Networks and phase congruency maps. In: IEEE 14th International Symposium on Biomedical Imaging (2017)
21. Meshram, N.H., Mitchell, C.C., Wilbrand, S., et al.: Deep learning for carotid plaque segmentation using a dilated U-Net architecture. Ultrason. Imaging **42**(4–5), 221–230 (2020)
22. Xie, S., Girshick, R., Dollár, P., et al.: Aggregated residual transformations for deep neural networks. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 5987–5995 (2017)
23. Mi, S., Wei, Z., Xu, J., et al.: Detecting carotid intima-media from small-sample ultrasound images. In: Annual International Conference of the IEEE Engineering in Medicine & Biology Society, pp. 2129–2132 (2020)
24. Wang, X., Yu, K., Dong, C., et al.: Recovering realistic texture in image super-resolution by deep spatial feature transform. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 606–615 (2018)