



Interpretable Deep Biomarker for Serial Monitoring of Carotid Atherosclerosis Based on Three-Dimensional Ultrasound Imaging

Xueli Chen¹ , Xinqi Fan¹ , and Bernard Chiu^{1,2}

¹ Department of Electrical Engineering, City University of Hong Kong, Kowloon, Hong Kong, China

{xuelichen3-c,xinqi.fan}@my.cityu.edu.hk

² Department of Physics and Computer Science,
Wilfrid Laurier University, Waterloo, ON, Canada
bchiu@wlu.ca

Abstract. We developed an interpretable deep biomarker known as Siamese change biomarker generation network (SCBG-Net) to evaluate the effects of therapies on carotid atherosclerosis based on the vessel wall and plaque volume and texture features extracted from three-dimensional ultrasound (3DUS) images. To the best of our knowledge, SCBG-Net is the first deep network developed for serial monitoring of carotid plaque changes. SCBG-Net automatically integrates volume and textural features extracted from 3DUS to generate a change biomarker called *AutoVT* (standing for **A**utomatic integration of **V**olume and **T**extural features) that is sensitive to dietary treatments. The proposed *AutoVT* improves the cost-effectiveness of clinical trials required to establish the benefit of novel treatments, thereby decreasing the period that new anti-atherosclerotic treatments are withheld from patients needing them. To facilitate the interpretation of *AutoVT*, we developed an algorithm to generate change biomarker activation maps (CBAM) localizing regions having an important effect on *AutoVT*. The ability to visualize locations with prominent plaque progression/regression afforded by CBAM improves the interpretability of the proposed deep biomarker. Improvement in interpretability would allow the deep biomarker to gain sufficient trust from clinicians for them to incorporate the model into clinical workflow.

Keywords: 3D Ultrasound Imaging · Carotid Atherosclerosis · Deep Biomarker · Interpretable Machine Learning · Activation Map

1 Introduction

Although cardiovascular events are highly prevalent worldwide, it was estimated 75–80% of cardiovascular events in high-risk patients could be prevented through

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/978-3-031-43987-2_29.

lifestyle changes and medical/dietary interventions [22]. The opportunity to prevent cardiovascular events calls for the development of sensitive and cost-effective tools to identify high-risk patients and monitor serial changes in response to therapies. As carotid atherosclerosis is a major source of ischemic stroke and a major indicator of systematic atherosclerosis [9], the carotid artery has long served as a major imaging target for assessment of atherosclerotic diseases.

Carotid intima media thickness (IMT) is an early imaging biomarker measured from two-dimensional ultrasound (2DUS) images. The use of IMT in serial monitoring is limited by the small annual change (~ 0.015 mm) [4], which does not allow treatment effects to be measured in a clinically affordable timeframe. As plaques grow 2.4 times faster along the arteries than they thicken [3] and change circumferentially as well, volume measurements, such as total plaque volume (TPV) and vessel wall volume (VWV), afforded by 3DUS imaging techniques are more sensitive to treatment effects [1, 12]. Biomarkers derived from plaque textural features extracted from 3DUS were also shown to be sensitive to medical [2] and dietary treatments [5, 14]. However, few studies consider both volume and textural features, and the handcrafted textural features extracted in previous studies are independent of the subsequent biomarker generation. To address these issues, we propose an end-to-end Siamese change biomarker generation network (SCBG-Net) to extract features from the baseline and follow-up images for generating a biomarker, *AutoVT*, quantifying the degree of change in volume and texture automatically. Although deep networks have been proposed for carotid plaque composition characterization [13], plaque echogenicity classification [16, 21], and plaque recognition [15, 17], SCBG-Net is the first deep network developed for serial monitoring of carotid atherosclerosis.

A convolutional neural network (CNN) is typically represented as a black-box function that maps images to an output. However, a biomarker should be interpretable for it to be trusted by clinicians. One approach to promote the interpretability of the biomarker is to allow the visualization of regions that have a prominent effect on the biomarker. Class activation map (CAM) [27] and its variant [19, 24] highlight regions having a strong contribution to classification results. Interpretability is not only desired in classification networks but also in networks focusing on quantifying the similarity of images, such as person re-identification [20, 25]. The ranking activation map (RAM) [25] and its variant, such as gradient-weighted ranking activation map (CG-RAM) [20], highlight regions contributing to the similarity between a reference image and other images. For our application, there is a need to develop a technique to generate activation maps localizing regions with a prominent effect on the novel biomarker. Another contribution of this paper is the development of such an approach to generate change biomarker activation maps (CBAM).

2 Materials and Methods

2.1 3DUS Imaging and Preprocessing

In this work, we assessed the sensitivity of the proposed biomarker in evaluating the effect of pomegranate juice and tablets. Pomegranate is anti-oxidative, and

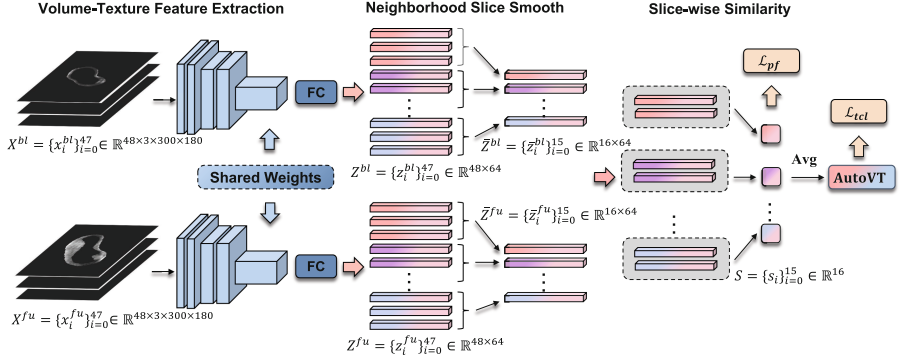


Fig. 1. Schematic of the Siamese change biomarker generation network (SCBG-Net).

previous studies have established that plaque texture features [5], the weighted average change of vessel-wall-plus-plaque thickness [26] and local vessel wall and plaque volume change [6] are able to detect the effect of pomegranate in the same cohort investigated in this study. Subjects were recruited by Stroke Prevention & Atherosclerosis Research Centre at Robarts Research Institute (London, Ontario, Canada) for a clinical trial (ISRCTN30768139). A total of 120 subjects involved in this study were randomized into two groups. 66 subjects received pomegranate extract and 54 subjects were given a placebo once daily for a year. There is no significant difference in the baseline characteristics between the two groups [26]. 3DUS images were acquired for participants at the baseline and a follow-up session, ranging from 283 to 428 days after the baseline scan. The reason for quantifying changes based on only two time points (i.e., baseline and follow-up) is that the rate of change of carotid plaque has been established as linear between the age of 50 to 75 in two studies involving over 6000 patients [11, 23]. The 3DUS images were obtained by translating an ultrasound transducer (L12-5, Philips, Bothell, WA, USA) mounted on a mechanical assembly at a uniform speed of 3 mm/s along the neck for about 4 cm. Ultrasound frames acquired using an ultrasound machine (ATL HDI 5000, Philips) were digitized at the rate of 30 Hz and reconstructed into a 3D image. The input image to SCBG-Net was obtained by masking the ultrasound image with the manually segmented boundaries (Supplementary Fig. 1). Each 3DUS image was resliced into a stack of 2D axial images with a 1 mm interslice distance (ISD) as described in [7].

2.2 Siamese Change Biomarker Generation Network (SCBG-Net)

Network Architecture. Figure 1 shows a schematic of the SCBG-Net. The baseline and follow-up image stacks consist of 48 axial images each, 24 from each of the left and right arteries. We denote the baseline (bl) and follow-up (fu) image stacks by $X^{bl} = \{x_i^{bl}\}_{i=0}^{47}$ and $X^{fu} = \{x_i^{fu}\}_{i=0}^{47}$, respectively, where x_i^{bl} and x_i^{fu} are axial slices with size 300×180 . The baseline and follow-up images are processed by a Siamese architecture. The volume-texture feature

extraction module utilizes an identical ResNet50 to extract features, resulting in 2048-dimensional vectors for each axial image. The vectors are then reduced to 64 dimensions using a fully connected (FC) layer with a rectified linear unit (ReLU) activation function, and these outputs are denoted as $z_i^{bl} = f(x_i^{bl})$ and $z_i^{fu} = f(x_i^{fu})$ for the two respective streams.

The neighborhood slice smoothing module was designed to reduce the effect of potential image misalignment. Although the baseline and follow-up 3DUS images were aligned according to the bifurcation position and the manually identified common carotid artery (CCA) axis as previously described [7], the internal carotid artery (ICA) of the two images may still be misaligned due to different bifurcation angles and non-linear transformations resulting from different head orientations [18]. To reduce the effect of potential image misalignment, the features of three neighboring slices in the baseline and follow-up images are averaged and denoted by $\bar{z}_i^{bl} = \frac{1}{3} \sum_{j=3i}^{3i+2} z_j^{bl}$ and $\bar{z}_i^{fu} = \frac{1}{3} \sum_{j=3i}^{3i+2} z_j^{fu}$, for $i = 0, 1, \dots, 15$. The slice-wise cosine “dissimilarity” for a baseline-follow-up slice pair was defined by $d_c(\cdot, \cdot) = 1 - \cos(\cdot, \cdot)$. To represent disease progression and regression, the sign of the slice-wise vessel wall volume change ΔVol_i from baseline to follow-up was used to determine the sign of the slice-wise score. ΔVol_i of each smoothed slice pair was computed by averaging vessel wall volume change (i.e., area change \times 1mm ISD) for a group of three neighbouring slices involved in the smoothing operation. The slice-wise score was obtained by:

$$s_i = s(\bar{z}_i^{bl}, \bar{z}_i^{fu}) = \text{sgn}(\Delta Vol_i) d_c(\bar{z}_i^{bl}, \bar{z}_i^{fu}), \quad (1)$$

where sgn represents the signed function. The use of ReLU in FC layers results in non-negative z_i , thereby limiting $d_c(\bar{z}_i^{bl}, \bar{z}_i^{fu})$ and $s(\bar{z}_i^{bl}, \bar{z}_i^{fu})$ to the range of $[0, 1]$ and $[-1, 1]$, respectively. Defined as such, s_i integrates vessel-wall-plus-plaque volume change with textural features extracted by the network. Finally, the *AutoVT* biomarker was obtained by averaging 16 slice-wise scores (i.e., $\text{AutoVT} = \frac{1}{16} \sum_{i=0}^{15} s_i$).

Loss Functions. We developed a treatment label contrastive loss (TCL) to promote discrimination between the pomegranate and placebo groups and a plaque-focus (PF) constraint that considers slice-based volume change.

(i) *Treatment Label Contrastive Loss.* The contrastive loss [8] maps similar pairs to nearby points and dissimilar pairs to distant points. In our biomarker learning problem, instead of separating similar and dissimilar pairs, we aim to discriminate baseline-follow-up image pairs of the pomegranate and placebo subjects. As changes occur in all patients, the baseline-follow-up image pairs are in general dissimilar for both groups. However, pomegranate subjects tend to experience a smaller plaque progression or even regression, whereas the placebo subjects have a larger progression [26]. As such, our focus is more on differentiating the two groups based on the signed difference between the baseline and follow-up images. We designed a treatment label contrastive loss (TCL) specifically for our biomarker learning problem:

$$\mathcal{L}_{tcl} = y \max(\text{AutoVT}, 0)^2 + (1 - y) \max(m - \text{AutoVT}, 0)^2, \quad (2)$$

where y is the group label of the input subject (pomegranate = 1, placebo = 0). For pomegranate subjects, instead of assigning a penalty based on the squared distance as in [8] (i.e., $AutoVT^2$ in our context), in which the penalty applies to both positive or negative $AutoVT$, we penalize only positive $AutoVT$ since it is expected that some pomegranate subjects would have a larger regression, which would be represented by a negative $AutoVT$. For placebo subjects, the penalty is applied only if $AutoVT$ is smaller than m .

(ii) *Plaque-focus Constraint*. We observe that slice pairs with high volume change are typically associated with a large plaque change (Supplementary Fig. 2). To incorporate such volume change information into $AutoVT$, we assigned a pseudo label ζ_i to each of the 16 smoothed slice pairs indexed by i based on the volume change ΔVol_i . $|\Delta Vol_i|$ was then ranked with the K_l slices with largest $|\Delta Vol_i|$ assigned $\zeta_i = 1$ and the K_s slices with smallest $|\Delta Vol_i|$ assigned $\zeta_i = 0$. The PF constraint was defined to promote the magnitude of s_i associated with slices with large $|\Delta Vol_i|$ and suppress that associated with slices with small $|\Delta Vol_i|$:

$$\mathcal{L}_{pf} = \frac{1}{K_s} \sum_{i \in \{i|\zeta_i=0\}} |s_i|^2 + \frac{1}{K_l} \sum_{j \in \{j|\zeta_j=1\}} (1.0 - |s_j|)^2. \quad (3)$$

The overall loss \mathcal{L} is a weighted combination of \mathcal{L}_{tcl} and \mathcal{L}_{pf} (i.e., $\mathcal{L} = \mathcal{L}_{tcl} + w\mathcal{L}_{pf}$, where w is the weight).

2.3 Change Biomarker Activation Map (CBAM)

Figure 2 shows a schematic of the proposed CBAM developed for visualizing important regions contributing to the $AutoVT$ scoring. Like previous CAM methods, CBAM generates activation maps by linearly weighting feature maps at different levels of a network. However, the weights associated with the attention maps in CBAM are novel and tailored for the proposed SCBG-Net.

The importance of a feature map is determined by how much it affects the absolute value of slice-wise score s (Eq. 1). The reason for focusing on $|s|$ is that we would highlight regions that contribute to both positive and negative changes. We denote $A_{L,i}^{p,k}$ as the k th channel of feature maps from the inner convolution layer L of an image slice x_i^p , where $p \in \{bl, fu\}$. The importance of $A_{L,i}^{p,k}$ towards the slice-wise score is defined in a channel-wise pair-associated manner by:

$$R(A_{L,i}^{p,k}) = |s(f(x_i^p \circ M_{L,i}^{p,k}), f(x_i^q))| = d_c(f(x_i^p \circ M_{L,i}^{p,k}), f(x_i^q)), \quad (4)$$

with $M_{L,i}^{p,k} = \text{Norm}(\text{Up}(A_{L,i}^{p,k}))$, \circ representing the Hadamard product and $(p, q) \in \{(bl, fu), (fu, bl)\}$. $\text{Up}(\cdot)$ upsamples $A_{L,i}^{p,k}$ into the size of x_i^p , and $\text{Norm}(\cdot)$ is a min-max normalization function mapping each element in the matrix into $[0, 1]$.

$A_{L,i}^{p,k}$ is first upsampled and normalized to $M_{L,i}^{p,k}$, which serves as an activation map to highlight regions in the input image. The importance of $A_{L,i}^{p,k}$ to the slice-wise score s is quantified by the cosine dissimilarity between the feature vectors

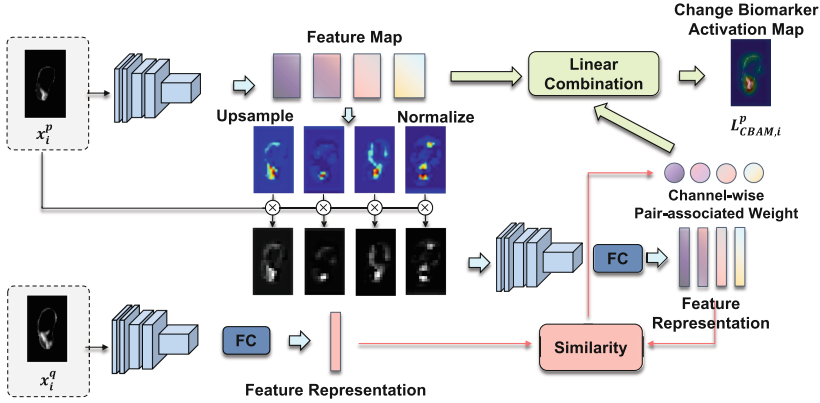


Fig. 2. Schematic of change biomarker activation map (CBAM).

generated by SCBG-Net for the highlighted input image and the corresponding image slice in the baseline-follow-up image pair. If the input image is a baseline image, the corresponding slice would be from the follow-up image, and vice versa.

For each slice x_i^p , the activation map from the convolutional layer L was generated as $L_{CBAM,i}^p = \text{ReLU}(\sum_k \alpha_{L,i}^{p,k} A_{L,i}^{p,k})$, where the weight $\alpha_{L,i}^{p,k}$ is $R(A_{L,i}^{p,k})$ normalized by the softmax function: $\alpha_{L,i}^{p,k} = \frac{\exp(R(A_{L,i}^{p,k}))}{\sum_h \exp(R(A_{L,i}^{h,k}))}$.

3 Experiments and Results

Statistical Evaluation. The discriminative power of biomarkers was evaluated by p-values from two-sample t-tests for normally distributed measurements or Mann-Whitney U tests for non-normally distributed measurements. P-values quantify the ability of each biomarker to discriminate the change exhibited in the pomegranate and placebo groups.

Experimental Settings. Our model was developed using Keras on a computer with an Intel Core i7-6850K CPU and an NVIDIA RTX 1080Ti GPU. The ResNet50 was initialized by the ImageNet pretrained weights. The SGD optimizer was applied with an initial learning rate of 3×10^{-3} . An exponential decay learning rate scheduler was utilized to reduce the learning rate by 0.8 every 10 epochs. We set the number of slices with top/last $|\Delta Vol_i|$ in the definition of the PF constraint as $K_l = K_s = 3$. All models were evaluated by three-fold cross-validation with 80 labeled subjects and 40 test subjects. Labeled subjects are further partitioned into training and validation sets with 60 and 20 subjects, respectively. For the proposed SCBG-Net, the margin m and loss function weight w were tuned using the validation set. In all three trials, the optimized m and w were 0.8 and 0.15, respectively.

Comparison with Traditional Biomarkers. Table 1 shows p-values for *AutoVT* and traditional biomarkers. The proposed biomarker based on the

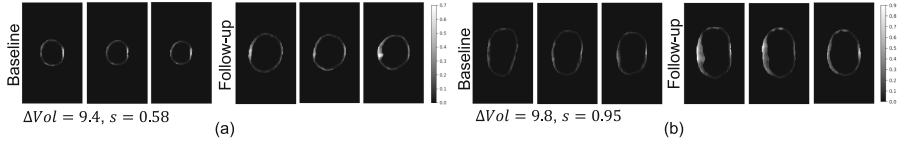


Fig. 3. Examples showing the sensitivity of *AutoVT* in highlighting texture-based plaque change.

overall loss function \mathcal{L} , $\text{AutoVT}(\mathcal{L})$, was the most sensitive to the effect of pomegranate with the lowest p-value. This biomarker has learned the volumetric information from the input images, as demonstrated by the correlation coefficient of 0.84 between $\text{AutoVT}(\mathcal{L})$ and ΔVWV . $\text{AutoVT}(\mathcal{L})$ has also learned the texture information, as demonstrated in Fig. 3. While the slice pairs in Fig. 3(a) and (b) had similar ΔVol , the changes in slice pairs shown in (a) were more related to expansive remodeling of the vessel wall, whereas the changes in slice pairs shown in (b) were caused by plaque growth. This difference is characterized by the slice-wise score, which was higher for the example in Fig. 3(b) than in (a).

Comparison with Different Deep Models. We compared treatment effect sensitivity attained by SCBG-Net to those attained by two alternative deep models: the Aver-Net and Atten-Net. Given the smooth features $\{\bar{z}_i^{bl}\}_{i=0}^{15}$ and $\{\bar{z}_i^{fu}\}_{i=0}^{15}$, Aver-Net first computes a uniform average vector for each of the baseline and follow-up images by $z_{aver}^p = \frac{1}{16} \sum_{i=0}^{15} \bar{z}_i^p$ where $p \in \{bl, fu\}$ and then generates a biomarker by $bio_{aver} = \text{sgn}(\Delta VWV) d_c(z_{aver}^{bl}, z_{aver}^{fu})$, where VWV in the baseline/follow-up session was computed by summing the vessel wall areas at each axial image of the left and right arteries and multiplied by the 1mm ISD. ΔVWV is the difference between VWV obtained from the baseline and follow-up images. In contrast, Atten-Net computes a weighted

Table 1. Comparison of *AutoVT* with other biomarkers.

	Measurements	Placebo	Pomegranate	P-value
		(n = 54)	(n = 66)	
Traditional	$\Delta TPV(mm^3)$	25.0 ± 76.5	16.1 ± 83.9	0.24
	$\Delta VWV(mm^3)$	112.4 ± 118.5	58.9 ± 104.9	0.01
Deep	bio_{aver}	0.151 ± 0.253	0.070 ± 0.138	0.028
	bio_{atten}	0.148 ± 0.227	0.065 ± 0.155	0.021
	$\text{AutoVT}(\mathcal{L}_{tcl})$	0.234 ± 0.262	0.119 ± 0.206	0.0087
Loss	$\text{AutoVT}(\mathcal{L}_{ce})$	0.216 ± 0.255	0.107 ± 0.243	0.018
	$\text{AutoVT}(\mathcal{L}_{bd})$	0.135 ± 0.165	0.060 ± 0.160	0.013
	$\text{AutoVT}(\mathcal{L}_{tcl})$	0.234 ± 0.262	0.119 ± 0.206	0.0087
	$\text{AutoVT}(\mathcal{L})$	0.241 ± 0.255	0.111 ± 0.209	0.0029

average vector based on the attention weight generated by a multiple instance learning-based attention module [10]. The baseline/follow-up attention-based weighted average vectors are computed by $z_{atten}^p = \frac{1}{16} \sum_{i=0}^{15} \gamma_i^p \bar{z}_i^p$, where γ_i^p is the attention weight for Slice i . The biomarker generated by Atten-Net is $bio_{atten} = \text{sgn}(\Delta V W V) d_c(z_{atten}^{bl}, z_{atten}^{fu})$. Aver-Net and Atten-Net do not involve slice-by-slice comparison, whereas slice-by-slice comparison was involved in two components of SCBG-Net: (i) the slice-wise score s_i (Eq. 1) and (ii) the PF constraint (Eq. 3). In this section, we focus on investigating the effect of Component (i) and that of Component (ii) will be studied in the next section focusing on loss functions. For this reason, the three models compared in this section were driven only by \mathcal{L}_{tcl} (Eq. 2) for a fair comparison. Table 1 shows that SCBG-Net is the most sensitive to treatment effects among the three models.

Comparison with Different Losses. We compared our proposed loss with another two losses, including cross-entropy loss and bi-direction contrastive loss. Cross-entropy loss is expressed as $\mathcal{L}_{ce} = -y \log(\sigma(1 - \text{AutoVT})) - (1 - y) \log(\sigma(\text{AutoVT}))$, where $\sigma(\cdot)$ is a sigmoid function. The bi-direction contrastive loss is a symmetric version of \mathcal{L}_{tcl} , expressed as $\mathcal{L}_{bd} = y \max(m + \text{AutoVT}, 0)^2 + (1 - y) \max(m - \text{AutoVT}, 0)^2$. The margin m in \mathcal{L}_{bd} was tuned in the same way as the proposed \mathcal{L} , with $m = 0.4$ being the optimized parameter in all three cross-validation trials. Table 1 shows p-values for different losses. Our proposed loss \mathcal{L}_{tcl} has a higher sensitivity than \mathcal{L}_{ce} and \mathcal{L}_{bd} , with further improvement attained by the incorporation of \mathcal{L}_{pf} . Pomegranate, as a dietary supplement, confers a weaker beneficial effect than intensive medical treatment [5, 26]. \mathcal{L}_{tcl} was designed to better model the weak benefit by not forcing the *AutoVT*s of pomegranate patients to get too negative; the *AutoVT*s of pomegranate patients would not be penalized as long as it is smaller than 0. In contrast, \mathcal{L}_{ce} and \mathcal{L}_{bd} promote more negative *AutoVT*s for pomegranate patients. \mathcal{L}_{tcl} was designed to account for the weak beneficial effect of pomegranate, which may not lead to significant plaque regression in pomegranate patients compared to high-dose atorvastatin. Moreover, \mathcal{L}_{pf} improves the discriminative power of *AutoVT* by using the ranking of $|\Delta V V o l_i|$ among different axial images of the same patient.

Comparison with Other Activation Maps. Figure 4 compares the activation maps generated by CBAM and CG-RAM from features maps in the second convolutional layer. CBAM localizes regions with plaque changes accurately,

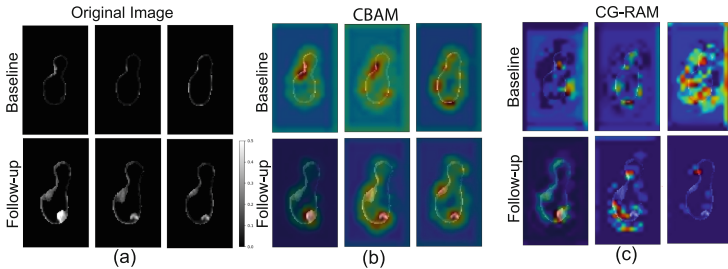


Fig. 4. Comparison between CBAM and CG-RAM.

whereas CG-RAM is less successful in highlighting regions with plaque changes. A possible reason for this observation is that CG-RAM is driven by gradient and may be adversely affected by gradient saturation issues, whereas CBAM is gradient-free and not affected by artifacts associated with gradient saturation.

4 Conclusion

We, for the first time, developed a deep biomarker to quantify the serial change of carotid atherosclerosis by integrating the vessel wall and plaque volume change and the change of textural features extracted by a CNN. We showed that the proposed biomarker, *AutoVT*, is more sensitive to treatment effect than vessel wall and plaque volume measurements. SCBG-Net involves slice-based comparison of textural features and vessel wall volume (Eq. 1) and we showed that this architecture results in a biomarker that is more sensitive than Aver-Net and Atten-Net that quantify global change for the left and right arteries. This result is expected as atherosclerosis is a focal disease with plaques predominantly occurring at the bifurcation. For the same reason, PF constraint that involves local slice-based assessment further improves the sensitivity of *AutoVT* in detecting treatment effects. We developed a technique to generate activation maps highlighting regions with a strong influence on *AutoVT*. The improvement in the interpretability of *AutoVT* afforded by the activation maps will help promote clinical acceptance of *AutoVT*.

Acknowledgement. Dr. Chiu is grateful for the funding support from the Research Grant Council of HKSAR, China (Project nos. CityU 11203218, CityU 11205822). The authors thank Dr. J. David Spence for providing the 3D ultrasound images investigated in this study.

References

1. Ainsworth, C.D., Blake, C.C., Tamayo, A., Beletsky, V., Fenster, A., Spence, J.D.: 3D ultrasound measurement of change in carotid plaque volume: a tool for rapid evaluation of new therapies. *Stroke* **36**(9), 1904–1909 (2005)
2. Awad, J., Krasinski, A., Parraga, G., Fenster, A.: Texture analysis of carotid artery atherosclerosis from three-dimensional ultrasound images. *Med. Phys.* **37**(4), 1382–1391 (2010)
3. Barnett, P.A., Spence, J.D., Manuck, S.B., Jennings, J.R.: Psychological stress and the progression of carotid artery disease. *J. Hypertens.* **15**(1), 49–55 (1997)
4. Bots, M.L., Evans, G.W., Riley, W.A., Grobbee, D.E.: Carotid intima-media thickness measurements in intervention studies: design options, progression rates, and sample size considerations: a point of view. *Stroke* **34**(12), 2985–2994 (2003)
5. Chen, X., et al.: Three-dimensional ultrasound evaluation of the effects of pomegranate therapy on carotid plaque texture using locality preserving projection. *Comput. Methods Programs Biomed.* **184**, 105276 (2020)
6. Chen, X., Zhao, Y., Spence, J.D., Chiu, B.: Quantification of local vessel wall and plaque volume change for assessment of effects of therapies on carotid atherosclerosis based on 3-D ultrasound imaging. *Ultrasound Med. Biol.* **49**(3), 773–786 (2023)

7. Egger, M., Spence, J.D., Fenster, A., Parraga, G.: Validation of 3D ultrasound vessel wall volume: an imaging phenotype of carotid atherosclerosis. *Ultrasound Med. Biol.* **33**(6), 905–914 (2007)
8. Hadsell, R., Chopra, S., LeCun, Y.: Dimensionality reduction by learning an invariant mapping. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 1735–1742. IEEE (2006)
9. Hennerici, M., Hübner, H.B., Hefer, H., Lammerts, D., Rautenberg, W.: Natural history of asymptomatic extracranial arterial disease: results of a long-term prospective study. *Brain* **110**(3), 777–791 (1987)
10. Ilse, M., Tomczak, J., Welling, M.: Attention-based deep multiple instance learning. In: *International Conference on Machine Learning*, pp. 2127–2136. PMLR (2018)
11. Johnsen, S.H., et al.: Carotid atherosclerosis is a stronger predictor of myocardial infarction in women than in men: a 6-year follow-up study of 6226 persons: the Tromsø study. *Stroke* **38**(11), 2873–2880 (2007)
12. Krasinski, A., Chiu, B., Spence, J.D., Fenster, A., Parraga, G.: Three-dimensional ultrasound quantification of intensive statin treatment of carotid atherosclerosis. *Ultrasound Med. Biol.* **35**(11), 1763–1772 (2009)
13. Lekadir, K., et al.: A convolutional neural network for automatic characterization of plaque composition in carotid ultrasound. *IEEE J. Biomed. Health Inform.* **21**(1), 48–55 (2016)
14. Lin, M., et al.: Longitudinal assessment of carotid plaque texture in three-dimensional ultrasound images based on semi-supervised graph-based dimensionality reduction and feature selection. *Comput. Biol. Med.* **116**, 103586 (2020)
15. Liu, J., et al.: Deep learning based on carotid transverse B-mode scan videos for the diagnosis of carotid plaque: a prospective multicenter study. *Eur. Radiol.* **33**, 1–10 (2022)
16. Ma, W., et al.: Multilevel strip pooling-based convolutional neural network for the classification of carotid plaque echogenicity. *Comput. Math. Methods Med.* **2021** (2021)
17. Ma, W., Zhou, R., Zhao, Y., Xia, Y., Fenster, A., Ding, M.: 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. IEEE (2019)
18. Nanayakkara, N.D., Chiu, B., Samani, A., Spence, J.D., Samarabandu, J., Fenster, A.: A “twisting and bending” model-based nonrigid image registration technique for 3-D ultrasound carotid images. *IEEE Trans. Med. Imaging* **27**(10), 1378–1388 (2008)
19. Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D.: Grad-cam: visual explanations from deep networks via gradient-based localization. In: *Proceedings of the IEEE International Conference on Computer Vision*, pp. 618–626 (2017)
20. Shen, D., Zhao, S., Hu, J., Feng, H., Cai, D., He, X.: ES-Net: erasing salient parts to learn more in re-identification. *IEEE Trans. Image Process.* **30**, 1676–1686 (2020)
21. Skandha, S.S., et al.: 3-D optimized classification and characterization artificial intelligence paradigm for cardiovascular/stroke risk stratification using carotid ultrasound-based delineated plaque: atheromaticTM 2.0. *Comput. Biol. Med.* **125**, 103958 (2020)
22. Spence, J.D.: Intensive management of risk factors for accelerated atherosclerosis: the role of multiple interventions. *Curr. Neurol. Neurosci. Rep.* **7**(1), 42–48 (2007)
23. Spence, J.D.: Determinants of carotid plaque burden. *Atherosclerosis* **255**, 122–123 (2016)

24. Wang, H., et al.: Score-CAM: score-weighted visual explanations for convolutional neural networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 24–25 (2020)
25. Yang, W., Huang, H., Zhang, Z., Chen, X., Huang, K., Zhang, S.: Towards rich feature discovery with class activation maps augmentation for person re-identification. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1389–1398 (2019)
26. Zhao, Y., Spence, J.D., Chiu, B.: Three-dimensional ultrasound assessment of effects of therapies on carotid atherosclerosis using vessel wall thickness maps. *Ultrasound Med. Biol.* **47**(9), 2502–2513 (2021)
27. Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., Torralba, A.: Learning deep features for discriminative localization. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2921–2929 (2016)