Investigating Temporal Features of Carotid Intima-Media Thickness from Ultrasound Imaging with Recurrent Neural Networks

Published in:
2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)

Publication Status:
Published online: 11/12/2023

DOI:
10.1109/EMBC40787.2023.10340661

Document Version
Author Accepted version

Investigating Temporal Features of Carotid Intima-Media Thickness from Ultrasound Imaging with Recurrent Neural Networks

Min Jing*[1], Kathryn Owen[2], Brian Mac Namee[3], Iab B. A. Menown[2] and James McLaughlin[1]

Abstract— **Measuring carotid intima-media thickness (cIMT) of the Common Carotid Artery (CCA) via B-mode ultrasound imaging is a non-invasive yet effective way to monitor and assess cardiovascular risk. Recent studies using Convolutional Neural Networks (CNNs) to automate the process have mainly focused on the detection of regions of interest (ROI) in single frame images collected at fixed time points and have not exploited the temporal information captured in ultrasound imaging. This paper presents a novel framework to investigate the temporal features of cIMT, in which Recurrent Neural Networks (RNN) were deployed for ROI detection using consecutive frames from ultrasound imaging. The cIMT time series can be formed from estimates of cIMT in each frame of an ultrasound scan, from which additional information (such as min, max, mean, and frequency) on cIMT time series can be extracted. Results from evaluation show the best performance for ROI detection improved 4.75% by RNN compared to CNN-based methods. Furthermore, the heart rate estimated from the cIMT time series for seven patients was highly correlated with the patient's clinical records, which suggests the potential application of the cIMT time series and related features for clinical studies in the future.**

Clinical relevance— **The temporal features extracted from cIMT time series provide additional information that can be potentially beneficial for clinical studies.**

I. INTRODUCTION

Studies have shown that measuring the carotid intima-media thickness (cIMT) of the Common Carotid Artery (CCA) by B-mode ultrasound can help to visualise the arterial walls, monitor the early stages of the atherosclerotic process [1], and serve as an important risk marker for ischemic stroke and myocardial infarction [2], [3]. Our recent work has demonstrated that cIMT is a valuable predictor of coronary artery disease (CAD) and an indication of its severity [4], in which the regions of interest (ROI) (the area of intima-media) was manually selected and cIMT was measured by using the Phillips ultrasound software QLAB.

Traditional approaches for cIMT measurement have been reviewed in [5] and include edge-detection, active contours, dynamic programming, local statistics, Hough transform and statistical modelling. Recently, some studies [6], [7] have proposed deep learning approaches that use Convolutional Neural Networks (CNN) for ROI detection in ultrasound image patches. For example, a study [6] applied a CNN to detect the ROI after cropping each ultrasound image into a small set of patches of size 75×75 pixels. The region of carotid far wall, however, was manually selected before applying CNN, so their approach was not fully automated. Another study [7] proposed a U-Net model combined with local energy optimisation for segmentation of intima-media and thickness measurement. Three types of patches were used to train the U-Net including lumen-intima, media-adventitia and background patches. These CNN-based approaches, however, treated each frame of an ultrasound independently and did not take advantage of patterns present in consecutive frames. One early study [8] evaluated the IMT over the cardiac cycle from ultrasound video but it was based on the conventional model-based segmentation without deep learning involved. Apart from these studies focused on the images, a study of note [9] proposed a novel method for cIMT estimation based on the decomposition of a radio frequency (RF) signal. In their work, a matching pursuit signal decomposition of the ultrasound RF signal was used to isolate the desired tissue interface echoes, which provided a new perspective on the problem and the potential for multi-modal data integration.

To date, most CNN-based work were applied only to single frames, which did not explore the temporal features of cIMT within ultrasound video sequences. The contribution of this study includes: (1) a novel framework using Recurrent Neural Networks (RNNs) that takes consecutive frames into account for ROI detection; (2) investigation of the temporal features of cIMT time series constructed from estimation in each frame, which may provide additional information valuable for clinical studies.

The remainder of the paper is structured as follows. Data acquisition is introduced in Section II. The proposed framework and cIMT estimation are explained in Section III. Section IV provides the experimental results before conclusions in Section V.

II. DATA ACQUISITION

This study was approved by the Health Research Authority and the Southern Health and Social Care Trust (REC Reference:18/LO/2086). The subjects were recruited in adherence with the Declaration of Helsinki [14] in the Craigavon Area Hospital, Northern Ireland. Patients with a history of ischemic-type chest pain or angina equivalent undergoing Invasive Coronary Angiography or Computed Tomography Coronary Angiography underwent high-resolution B-mode ultrasound to measure cIMT and plaque. The images were acquired using the Philips Affiniti 70G US, L12-3 Philips transducers (with a frequency range of 3-12 MHz) and Philips QLAB software. The data used in this study includes 47 ultrasound scans acquired from Left Distal CCA. The

*This work was supported by the European Union's INTERREG VA Programme managed by the Special EU Programmes Body (SEUPB).
[1]M. Jing and J. McLaughlin are with Nanotechnology and Integrated BioEngineering Centre (NIBEC), School of Engineering, Ulster University, UK, *m.jing@ulster.ac.uk; [2]K. Owen and I. Menown are with Southern Health & Social Care Trust, Craigavon Area Hospital, Craigavon, UK; [3]B. Mac Namee is with School of Computer Science, University College Dublin, Republic of Ireland.

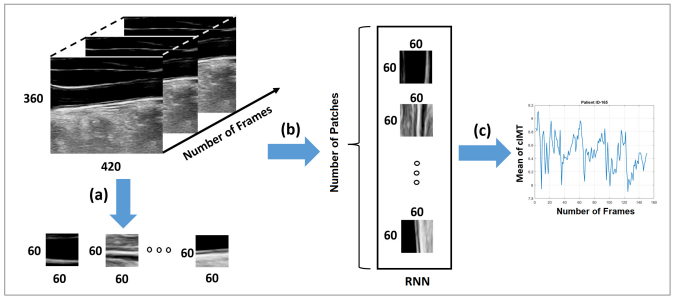

Fig. 1. Overview of the proposed framework.

frame rate was 30 frame per second (fps) and 151 B-mode frames were acquired from each ultrasound scan. More details can be found in [11].

III. Proposed Methods

A. ROI Detection via RNN

Fig. 1 illustrates the overview of the proposed framework: (a) The original ultrasound scan was cropped to exclude the text annotation and the final image size obtained was 360×420 pixels. Each frame was divided into 6×7 patches and each patch size was 60×60 pixels. Each patch was named with the index indicating its location at the original image; (b) To take advantage of dynamic patterns presented in ultrasound data, an RNN was used for ROI detection using patch images from consecutive frames. This is framed as a binary classification problem with any patches covering a region of carotid intima-media considered a positive ROI detection (examples of both positive and negative patches are shown in Fig. 2); (c) To study the temporal features of cIMT, after ROI detection, cIMT was estimated at each frame to form a cIMT time series, from which further information can be extracted for investigation.

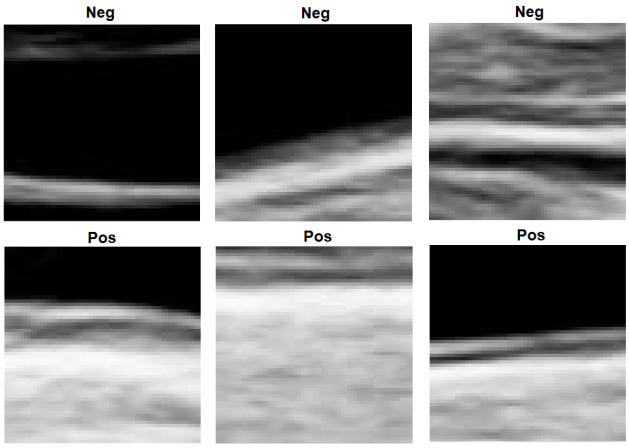

Fig. 2. Examples of patch images for positive (Pos) and negative (Neg) classes.

B. Estimation of cIMT

Fig. 3(a) shows a patch image containing the ROI from the carotid far wall with annotations indicating the structure of the carotid: lumen, lumen–intima (LI), and media–adventitia (MA). The IMT is defined as the distance between the LI and MA interfaces (as shown in red). Since the focus of this study was not to improve the accuracy of cIMT estimation

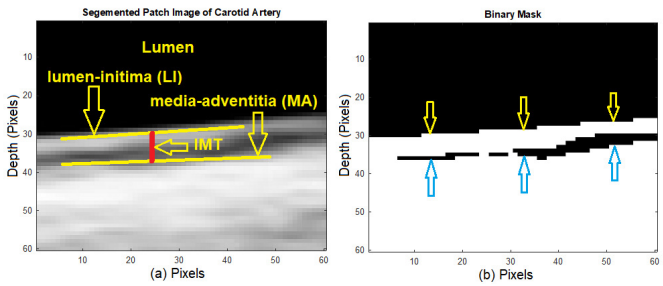

Fig. 3. (a) A positive patch showing the structure of carotid far wall with lumen, LI and MA; (b) a mask image created to estimate the cIMT at each column.

but rather to explore the temporal information in cIMT time series, we developed a simple approach for cIMT estimation.

From Fig. 2, it can be observed that positive patches have a similar structural pattern, such that the lumen is positioned above LI and MA. A threshold was applied to image frames to obtain a binary mask for LI and MA. An initial threshold was obtained using Ostu's method [12], then multiplied by a factor according to the image intensity (empirically in a range of $0.5 \sim 0.9$). For example in Fig. 3(b), the mask was based on a factor of 0.9.

In Fig. 3(b), non-zero values were detected first, and cIMT was then estimated at each column. From the top to bottom, after the first non-zero point, if any gap (zeros) was found, then cIMT was determined by the distance between the first non-zero point to the last zero point. If no gap was found after the first non-zero point, cIMT was set as zero. The final outcome of cIMT from one patch was the average of all non-zero cIMT values.

IV. Experiments & Results

A. CNN-based ROI Detection

For CNN-based experiments, the first frame of each scan was cropped to generate 1,974 patches. After selection, 154 positive and 198 negative patches were used in the experiments. Image augmentation was applied to increase the training data size by reflecting the image along the x-axis (horizontally). The final data size and data partition for CNN training and testing are presented in Table I. To take advantage of well-known pre-trained image classification networks, transfer learning was applied to fine tune six networks, which include googlenet, vgg16, vgg19, resnet18, resnet50, and resnet101. The performance of each resulting model, measured using classification accuracy, is given in Table II. The best performance was from vgg19 and resnet50 (94.85%) followed by vgg16 (93.38%).

B. RNN-based ROI Detection

RNN-based approaches, such as Long Short-Term Memory networks (LSTM) [10], have been used effectively for image classification by taking the temporal information within images into account [13]. Unlike CNNs which use only a single frame, a number of consecutive frames were used as input to the LSTM. As described in Section III.A, after cropping, one frame was divided into 6×7 patches, for N frames, the dimension of the image array is $6 \times 7 \times N$. There were 300 patches in total from the first frame (150 per

TABLE I

DATA PARTITION FOR CNN TRAINING & TESTING

Class	Size	Training (70%)	Training After Augmentation	Test (30%)
Pos	154	108	216	46
Neg	198	108	216	90

TABLE II

RESULTS OF CLASSIFICATION BY CNN

Classifiers	Number of Layers	Training Time (mins)	Accuracy (%)
googlenet	144	12	91.91
vgg16	41	61	93.38
vgg19	47	87	**94.85**
resnet18	71	15	88.24
resnet50	177	78	**94.85**
resnet101	347	111	92.65

class), which were also used in the experiments for CNN. The patches at the same location from subsequent frames were concatenated to form the sequences used in LSTM training and testing.

To study how the number of frames may affect the LSTM performance, evaluation was carried out varying the number of frames from 5 to 25 (since no significant improvement was observed for more than 25 frames). The number of sequences for LSTM training and testing under the different numbers of frames are shown in Table III. The performance, measured using classification accuracy, of LSTM models using different numbers of frames and hidden layers (HDL) are shown in Table IV. Overall LSTM models achieve better accuracy than the CNN-based models and the best performance improved 4.75% by LSTM. Furthermore, the average performance is improved with the increase of the number of frames from 5 to 15, then remains stable from 20 to 25, which suggests there might be the optimal number of frames to maximise the performance of LSTM (this can be further investigated in the future).

TABLE III

NUMBER OF SEQUENCES IN LSTM UNDER DIFFERENT NUMBERS OF FRAMES

Frames	5	10	15	20	25
Training	1050	2100	3150	4200	5250
Testing	450	900	1350	1800	2250

TABLE IV

ACCURACY (%) BY LSTM UNDER DIFFERENT NUMBER OF FRAMES AND HDL

HDL	Number of Frames				
	5	10	15	20	25
2	94.67	96.44	99.04	94.94	95.42
5	97.11	97.78	99.56	97.83	98.22
10	98.22	99.67	99.78	99.44	99.60
25	99.33	98.67	99.56	99.83	98.71
50	99.33	99.00	99.70	99.28	99.91
75	98.89	99.33	100.00	99.83	98.62
Mean	97.93	98.48	**99.60**	98.53	98.41

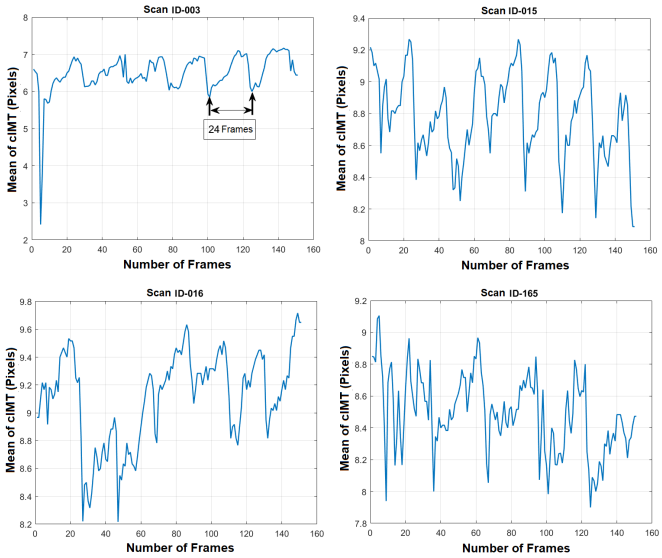

Fig. 4. Examples of cIMT times series obtained from four scans.

C. Study of Temporal Information of cIMT

1) cIMT Time Series: In our recent study [4], ROI was manually selected by a clinician, the cIMT was measured from the frames at three peaks of R wave and then averaged to provide one cIMT value for each scan. In this study, the ROI was automatically identified by LSTM at the first frame, then cIMT was estimated from the following 150 frames at the same location of ROI patch, from which a cIMT time series can be obtained and additional temporal features can be extracted.

Fig. 4 presents four examples of cIMT time series obtained from one identified ROI patch in each scan. The x-axis shows the frame number, and the y-axis shows the mean of cIMT estimated from the ROI patch per frame, which was the average of all non-zero cIMT values estimated per column in the patch (as described in Section III.B). Note the unit of cIMT is pixel not *mm* because the focus of this study was to extract the temporal features of cIMT, not to improve the accuracy of estimation (which can be considered in future work). It is observed that the time series are not very smooth, which can be due to the inconsistency of the ROI (such as change of image intensity or movement of ROI location) in adjacent frames of the ultrasound sequence. For improvement, adaptive threshold and ROI tracking will be considered in the future.

2) Frequency of cIMT time series vs Heart Rate: From cIMT time series Fig. 4, min, max, and mean can be obtained. It is also noticed that the cIMT time series has a similar pattern resembles the pulse wave. For instance, time series from scan ID-003 has approximately a period of 24 frames (0.8sec based on frame rate 30fps), hence a frequency of 1.25, which is equal to 75 beats per minute (bpm). The clinical record shows that patient (for scan ID-003)'s heart rate is 76bpm. For further validation, Fig. 5 presents the seven patients' heart rates (x-axis) compared to those estimated from their cIMT time series (y-axis) with corresponding values. The correlation coefficients is 0.95, which indicates that the cIMT time series may have a

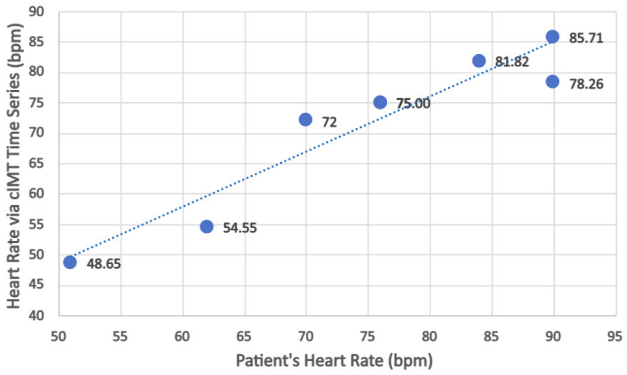

Fig. 5. Comparison of patients' heart rate with that estimated from the cIMT time series.

similar dynamic pattern as the heart beat. The finding from evaluation of cIMT over the cardiac cycle provide additional information in the assessment of arterial dynamics, which may potentially improve the risk assessment of CAD in future work.

3) Distribution of cIMT Time Series: A boxplot for cIMT time series from seven patients is given in Fig. 6 (scan ID-015 and ID-016 belong to one patient), in which the time points with zero cIMT values were excluded. It can be seen that the cIMT values from ID-003, ID-015, ID-016 and ID-165 are more tightly grouped (hence less variability) than the others, except ID-003 has two outlier points. The observation is accordant with their time series plotted in Fig. 4. The reason for the lower value for ID-003 was image intensity changes within the frames and a fixed threshold adjust factor (0.9) used in this experiment here was not suitable for the frame with relatively low intensity, which can result in zero or low cIMT values. Adaptive thresholds or more advanced approaches such as semantic segmentation by deep learning will be considered in future. The reason for higher values (in ID-009) was because the ROI detected at the first frame shifted to other neighbour patches in the following frames which caused poor cIMT estimation. Incorporating ROI tracking within the consecutive frames will also be considered in future work.

V. CONCLUSIONS

This study presented a novel framework to investigate the temporal features of cIMT time series constructed from B-mode ultrasound imaging. Unlike CNN-based approaches based on patch images from single frames, we have proposed using RNNs that can use consecutive frames to detect the ROI for cIMT estimation, in which the best performance was improved by 4.75% compared to CNN-based methods. From the cIMT time series, not only additional information such as min, max and mean of cIMT can be extracted, but also the frequency of cIMT time series was found to closely correlate with heart rate, which may further imply the potential values and applications of cIMT time series for clinical studies.

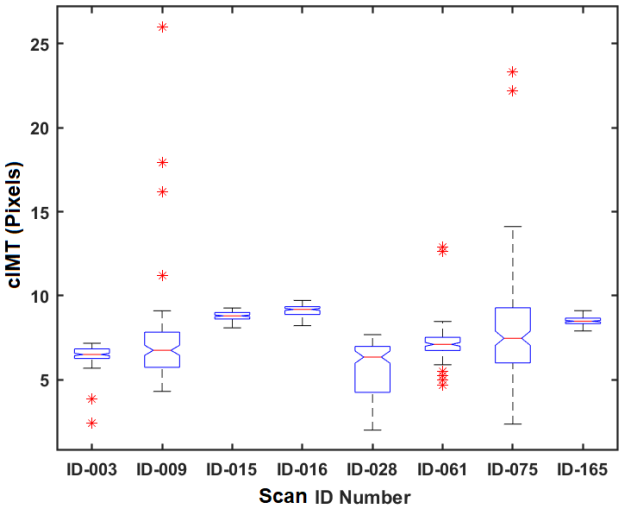

Fig. 6. Boxplot of distribution of the cIMT time series from seven patients.

REFERENCES

[1] J. M. George, R. Bhat, K. Mohan Pai, S. Arun, and J. Jeganathan. "The carotid intima media thickness: a predictor of the clincal coronary events." Journal of clinical and diagnostic research: JCDR 7, no. 6 (2013): 1082.

[2] M. L.Bots and D. E. Grobbee. "Intima media thickness as a surrogate marker for generalised atherosclerosis." Cardiovascular drugs and therapy 16, no. 4 (2002): 341-351.

[3] P-J. Touboul, M. G. Hennerici, S. Meairs, H. Adams, P. Amarenco, N. Bornstein, L. Csiba et al. "Mannheim carotid intima-media thickness consensus (2004–2006)." Cerebrovascular diseases 23, no. 1 (2007): 75-80.

[4] K.Owen, I. Menown and J. McLaughlin. "34 Use of ultrasound derived carotid intima-media thickness and plaque volume to predict single or multivessel coronary artery disease." Heart 106, no. Suppl 4 (2020): A23-A23.

[5] F. Molinari, G. Zeng and J. S. Suri. "A state of the art review on intima–media thickness (IMT) measurement and wall segmentation techniques for carotid ultrasound." Computer methods and programs in biomedicine 100, no. 3 (2010): 201-221.

[6] C. Rajasekaran, K. B. Jayanthi, S. Sudha and R. Kuchelar. "Automated diagnosis of cardiovascular disease through measurement of Intima media thickness using deep neural networks." The 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 6636-6639. IEEE, 2019.

[7] S. Devaraj,"Deep Unet network for the Intima Media Thickness Measurement in Ultrasound Images.". In 2020 IEEE 17th India Council International Conference (INDICON), pp. 1-4. IEEE, 2020.

[8] D. E. Ilea, C. Duffy, L. Kavanagh, A. Stanton and P. F. Whelan. "Fully automated segmentation and tracking of the intima media thickness in ultrasound video sequences of the common carotid artery." IEEE Trans. on Ultrasonics, Ferroelectrics, and Frequency Control, 60, no.1, pp.158-177, 2012.

[9] S. Steinberg, K. Tran, S. Rajan and Y. Ono. "Estimation of Intima-Media Thickness of Carotid Artery by Ultrasound Radiofrequency Signal Decomposition Using Matching Pursuit." In 2021 IEEE International Ultrasonics Symposium (IUS), pp. 1-4. IEEE, 2021.

[10] S. Hochreiter and J. Schmidhuber, "Long short-term memory", Neural Comput., vol. 9, no. 8, pp. 1735-1780, 1997.

[11] Kathryn Owen,"The Use of Ultrasound Derived Carotid Measurements and Plaque Volume to Predict Coronary Artery Disease", PhD Thesis, 2022.

[12] N. Otsu "A Threshold Selection Method from Gray-Level Histograms." IEEE Transactions on Systems, Man, and Cybernetics, vol.9, no.1, pp.62–66, 1979.

[13] M. Jing, D. Mclaughlin, S. E. Mcnamee, S. Raj, B. Mac Namee, D. Steele, D. Finlay and J. Mclaughlin, "A Novel Method for Quantitative Analysis of C-Reactive Protein Lateral Flow Immunoassays Images via CMOS Sensor and Recurrent Neural Networks", IEEE Journal of Translational Engineering in Health and Medicine, vol.9, pp. 1-15, 2021.

[14] M. D. Goodyear, K. Krleza-Jeric, and T. Lemmens, "The declaration of Helsinki." Bmj, 335, no.7621, pp. 624-625, 2007.