2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
October 14-18, 2024. Abu Dhabi, UAE

3D Ultrasound Image Acquisition and Diagnostic Analysis of the Common Carotid Artery with a Portable Robotic Device

Longyue Tan, Zhaokun Deng, Mingrui Hao, Pengcheng Zhang, Xilong Hou, Chen Chen, Xiaolin Gu, Xiao-Hu Zhou, Zeng-Guang Hou, *Fellow, IEEE*, and Shuangyi Wang, *Member, IEEE*

***Abstract*— Ultrasound (US) imaging of the carotid artery (CA) is a non-invasive diagnostic tool widely used in the medical field to assess the condition of the carotid artery, thereby predicting the risk of cardiovascular and cerebrovascular diseases. However, implementing this method in primary healthcare can be challenging due to the requirement for professionally trained sonographers. With the adoption of US robotic devices, the probe pose can be acquired while scanning, offering the possibility for 3D reconstruction and providing analyses that are not dependent on operator experience. This article introduces a method to semi-automatically acquire serialized US images of the common carotid artery (CCA). The method involves a specially designed robotic device built with a 6-RSU parallel mechanism, which is controlled according to robot pose, force sensor data and synchronous US images. To validate the images acquired, a method is proposed to segment the intima-media of CCA and calculate the intima-media thickness (IMT), which is a key indicator for cerebrovascular events prediction. After that, we propose an algorithm to reconstruct CCA into 3D voxel data with patient movement and cardiac cycle compensated, and a longitudinal view US image of CCA can be resliced from the voxel. The methods are tested on human subjects and the results indicate that the system and workflow can provide both quantitative and qualitative information of CCA for further diagnosis.**

I. Introduction

The carotid arteries (CA), comprising the common, internal and external, are large blood vessels in the neck, supplying blood to brain, neck and face. The health of these arteries is critically important, as they are a common site for atherosclerotic plaque buildup, which can lead to significant cardiovascular diseases, including stroke, the leading cause of serious long-term disability [1]. Medical ultrasound (US), also known as sonography, has become an indispensable tool in modern healthcare, offering a non-invasive, cost-effective, and versatile means of visualizing internal body structures [2]. Medical ultrasound has emerged as a frontline diagnostic tool for the evaluation of the carotid arteries, allowing for the detailed assessment of the structure and function of these arteries. It helps in detecting plaques, measuring the degree of stenosis, and evaluating the flow of blood through the arteries [3]. Among all indicators, carotid intima-media thickness (IMT) can well reflect early signs of atherosclerosis and cardiovascular disease risk, which is a necessary mark in US examination of CA [4].

Freehand US has long been the common technique in US imaging, valued for its immediacy, flexibility, and non-invasive nature. However, the operator-dependence of freehand US can lead to variability in image quality and accuracy of the diagnostics [5]. Additionally, in complex examinations where precise probe positioning is critical, freehand US can be challenging, limiting the ability to consistently obtain high-quality images. In response to these challenges, the development of robot-assisted ultrasound has introduced a new dimension to US imaging.

Existing ultrasound robotic systems for CA scanning frequently utilize commercial robotic arms [6], [7], which tend to be unwieldy and not portable. The substantial size and weight of these arms may be risky to patients. With robotic US devices, various control methods are proposed to apply a CA scanning, like visual servo algorithm [8] and force control [9]. Meanwhile, reinforcement learning method is also applied for robot-assisted probe navigation [10], [11].

To evaluate atherosclerosis of CCA and predict the risk for stroke, IMT is a widely used predictive indicator. Currently, IMT is usually measured in longitudinal view US images, with various artificial intelligence based methods [12], [13], [14]. However, acquisition of longitudinal view US images is highly dependent on physician's skill, making it very complicated to robotize [15].

To support diagnosis and treatment, the 3D model of CA can be reconstructed using 2D US images, for which various methods have been proposed. Reconstruction algorithms of 3D US can be classified into three types: voxel-based methods (VBMs), pixel-based methods (PBMs), and function-based methods (FBMs) [16]. Specifically in the field of CA

*Research supported in part by the National Natural Science Foundation of China under grant 62373352, in part by the InnoHK program, and in part by the Xiaomi Young Talents Program. (*Corresponding authors: Zeng-Guang Hou and Shuangyi Wang*).

Longyue Tan, Zhaokun Deng, Mingrui Hao and Pengcheng Zhang are with the State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China, and also with the School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 101408, China (e-mail: tanlongyue2022@ia.ac.cn; dengzhaokun2021@ia.ac.cn; haomingrui2021@ia.ac.cn; zhangpengcheng2022@ia.ac.cn).

Xilong Hou is with the Center for Artificial Intelligence and Robotics, Hong Kong Institute of Science and Innovation, Chinese Academy of Sciences, Hong Kong (e-mail: xilong.hou@cair-cas.org.hk).

Xiaolin Gu is with the Lingshu Medical Company, Suzhou 215123, China (e-mail: jason.gu@lingshumed.com).

Zeng-Guang Hou, Xiao-Hu Zhou and Chen Chen are with the State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China (e-mail: zengguang.hou@ia.ac.cn; chen.chen@ia.ac.cn).

Shuangyi Wang is with the State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China, also with the School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 101408, China, and also with the Center for Artificial Intelligence and Robotics, Hong Kong Institute of Science and Innovation, Chinese Academy of Sciences, Hong Kong (e-mail: shuangyi.wang@ia.ac.cn).

reconstruction, Lv et al. introduces a real-time interactive 3D ultrasound imaging system combining Bezier interpolation and PNN reconstruction algorithms, capable of displaying 3D human CA models [17]. Djukic et al. utilize a generalized human CA model and capture patient-specific image data to adjust the model, making it possible to assess the state of the arteries in clinical diagnostics [18]. In addition, implicit representation using neural networks has been introduced to reconstruct human CA [19]. However, most of current works consider human CA as a static and rigid object, ignoring the patient's movement and cardiac cycle during the image acquisition, which might cause pose estimation error of US probe and non-rigid deformation in US image.

In contrast to the previous work, we utilize a specially designed parallel robot structure to implement robot-assisted semi-automatic US scanning of human CCA. A control algorithm combined of force control and visual-servoing is designed to control the robot, aiming at acquiring continuous and clear US images of the cross-sectional view of CCA. After that, we propose a 3D reconstruction algorithm, adjusting CCA segmentation masks based on continuity prior assumption of the masks to compensate cardiac cycle and patient movement while scanning. Building on the work of our predecessors, we focus more deeply on the intima-media, which is an important indicator for cardiac diseases. Semantic information is extracted with a wavelet based segmentation algorithm, and then used to offer diagnostic information for physicians, while generating a US images of the longitudinal view, which are requirement of the standard diagnosis protocol.

The innovation and contribution of this paper are: a complete semi-automatic workflow is proposed to assess the state of health about human CCA with US, combining a portable robotic device, control method and both qualitative and quantitative analysis of acquired US images; a cardiac cycle compensation method is proposed to assist with 3D reconstruction of CCA from US images, providing a qualitative representation of CCA for physicians with temporal perturbation corrected; longitudinal view images of CCA are synthesized as diagnostic assistance with intima-media segmented and IMT calculated.

II. Methods

A. Hardware Setup

As illustrated in Fig. 1a, the proposed 6-Revolute-Spherical-Universal (as 6-RSU) parallel-mechanism-based ultrasound robot consists of a handheld pocket ultrasound probe, a moving platform, a static platform with six motors, and six kinematic chains with transmission mechanism. The static platform does not move while scanning, and the moving platform can translate and rotate driven by motors. The motor angles and data from the inertial measurement unit (IMU) within the moving platform are collected to calculate the pose of the moving platform. A build-in force sensor placed in the center of the moving platform can record the one-axis contact force between the probe and skin. Further details about the hardware designs are presented in our previous works [20], [21], [22].

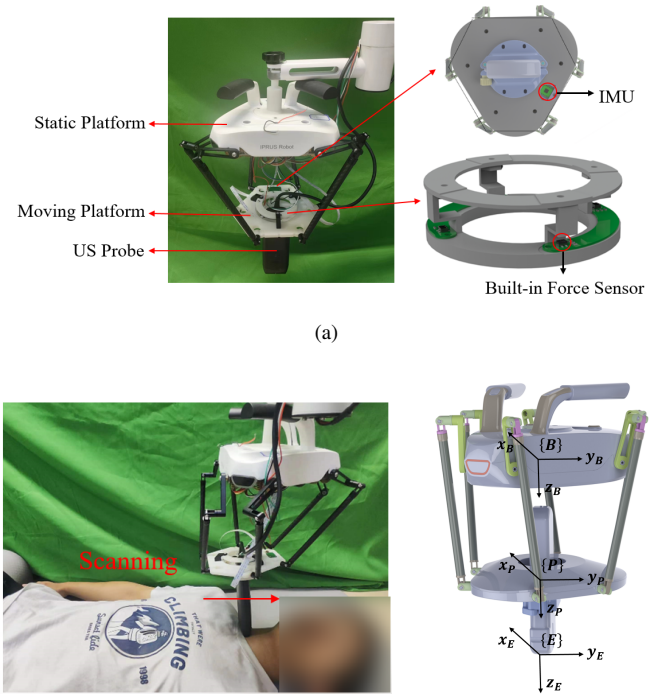

Fig. 1: (a) The structure of the parallel-mechanism-based US robot with built-in IMU and force sensor. (b) Example of a CCA scanning case. (c) Definition of coordinates.

B. Frames Definition

We define the base frame B fixed with the static platform and the moving platform frame P fixed with the moving platform. The transformation matrix between B and P at time t is noted as ${}^{B}_{P}T_t$, which is calculated from motor angles and IMU data. The probe end frame E is fixed with the end of probe, The transformation matrix between P and E is noted as ${}^{P}_{E}T$, which is a constant matrix only related to the distance between center of moving platform and probe end. The transform matrix between B and E can be calculated as:

$$ {}^{B}_{E}T_t = {}^{B}_{P}T_t {}^{P}_{E}T. \tag{1} $$

The acquired US image is within the x-z plane of the frame E. The definition of frames is shown in Fig. 1c.

C. Semi-automatic Scanning

The semi-automatic scanning requires the operator to teleoperate the probe pose until a clear cross-sectional US image of CCA is acquired. A real-time AI-assisted system to indicate the presence of target structures is utilized to assist the search, making it possible for less experienced operators to complete the scan. Corresponding contact force along the z_e axis are recorded, noted as F_d.

After that, as shown in Fig. 1c, the robot automatically scans along the neck. The movement of the probe can be decoupled as two sets: in-plane movement, which includes

translation along the x_e and z_e axis, and out-of-plane movement, which is translation along the y_e axis.

Translation along the z_e axis is controlled using admittance control according to the contact force data. For each control loop, desired acceleration $\ddot{z}_d$ is calculated:

$$\ddot{z}_d = \frac{1}{M}(F_e - F_d - B\dot{z} - Kz), \tag{2}$$

where symbols F_e, F_d represent the target force and detected force respectively. Meanwhile, $\dot{z}, z$ represent the velocity and displacement in the last loop. Position command to the robot z_d can be integrated from $\ddot{z}_d$. Target force is set as the recorded force in the probe initialization stage.

Translation along the x_e axis is based on the US image information, the purpose of which is to maintain the CCA in a fixed position of image. A YOLOv5 network [23] is trained to detect the CCA in US image. The bounding box center is noted as (x_b, y_b). Such movement is controlled as:

$$\dot{x}_e = K_d(x_b - x_{desire}), \tag{3}$$

where x_{desire} is the target horizontal location of bounding box in US image, which is set as the initial horizontal position of bounding boxes in the images.

Translation along the y_e axis is set as a linear translation, leading the probe to scan along the neck.

D. 3D Reconstruction of CCA

1) Overview: In the previous scanning process, a set of carotid US images for each subject are collected. Initially, a neural network based on the UNet architecture is trained to segment the carotid arteries from US images, acquiring a carotid mask for each frame. Subsequently, we introduce an intima-media segmentation algorithm for the carotid artery, based on wavelet transformation. The thickness of the intima-media as an indicator for diagnosis is calculated with the masks of intima-media.

Following this, to address the continuity of morphological characteristics of the carotid mask, a motion recognition and compensation algorithm is proposed. This algorithm compensates errors on the carotid mask due to robot pose errors during scanning, the patient's cardiac cycle and other movements. Using the original carotid masks and the corrected masks, a dense optical flow field between two masks is calculated. Subsequently, the optical flow field is used to perform non-rigid transformations on the original US images and the segmented intima-media masks.

Finally, to assist physicians in diagnosis, the acquired ultrasound voxel data was resampled into a 2D US image to present longitudinal view. The intima-media segmentation results were then adjusted to better align with the actual images, enhancing their accuracy and reliability for medical examination. The flow chart to illustrate our algorithm is presented as Fig. 2.

2) CCA and intima-media segmentation: To segment CCA in US images, we elaborate on the utilization of a U-Net architecture neural network [24] for segmentation tasks, employing a symmetric expanding path to enable precise localization, combined with a contracting path for contextual information capture.

After acquiring the mask of CCA, the mask of intima-media is segmented. The initial step involves determining the Region of Interest (ROI) for intima-media segmentation based on the CCA segmentation mask. Considering that the intima-media layer of the CCA can generally be seen around the bottom contour of CCA in a US image, the ROI is set as a rectangle area. The upper bound of the area is 2 mm above the bottom contour of the carotid mask and the lower bound is 0.5 mm below. The left and right bound is aligned on the corresponding contour of the mask. The new image, which is cut from the acquired US image within the obtained ROI, is referred to as I_0.

Inspired by [25], a set of convolution kernels is defined as:

$$k_1 = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}, k_i = \begin{bmatrix} 1 & \cdots & 2 & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 2 & \cdots & 4 & \cdots & 2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & \cdots & 2 & \cdots & 1 \end{bmatrix}, \tag{4}$$

where the size of k_i is $[2i+1, 2i+1], i = 1, 2, 3, ..., N$. In kernel k_i, nine positions are denoted with non-zero values, and the rest elements are set to be 0. This series of kernels are then applied to the image to capture multi-scale information:

$$I_i = I_{i-1} \otimes k_i, \tag{5}$$

and the differences between these multi-scale images then represent features in a specific scale:

$$d_i = I_i - I_{i-1}. \tag{6}$$

After that, a threshold segmentation is applied to segment intima-media, in which the threshold is selected as:

$$t_i = \mu_i - \alpha \cdot \sigma_i, \tag{7}$$

where μ_i and σ_i denote the average and standard deviation of d_i. The threshold of segmentation is denoted as t_i. We set $i = 3$ and $\alpha = 0.2$ as the hyper parameters.

Following the acquisition of the mask through a threshold segmentation algorithm, to mitigate the effects of ultrasound artifacts and noise on the segmentation outcome, the connected domains of the obtained mask are extracted. Domains located excessively far from the carotid artery wall are removed. Subsequently, the largest connected domain is selected to serve as the mask for the intima-media. After obtaining the intima-media mask, the thickness of the intima-media in a single frame of the US image can be calculated based on the masks of the CCA and the intima-media.

3) Motion Compensation and 3D Reconstruction: In Section C, the probe's pose is estimated while moving along the neck. The acquired US images and corresponding masks are firstly translated according to the change of probe end position. However, the patient movement during the scanning process cannot be measured, thus necessitates motion compensation. In addition, the carotid artery expands and

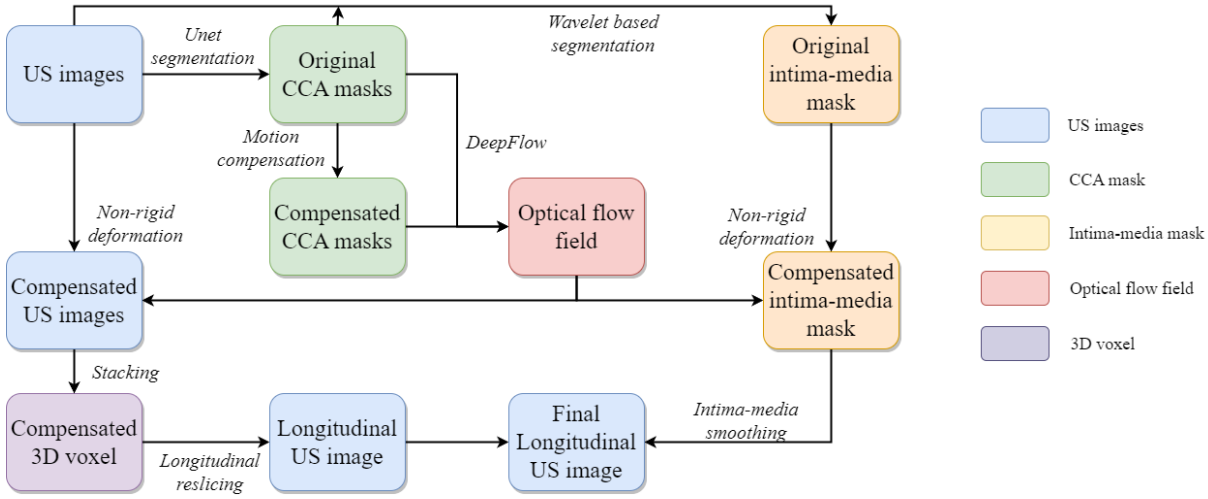

Fig. 2: The flow chart of our motion compensation and longitudinal view reslicing algorithm.

contracts with the heartbeat, leading to the temporal perturbation of the acquired US image series, which requires cyclical correction.

Based on the CCA segmentation results, this study employs the centroid, area, and boundary continuity of the carotid mask for mask correction. For an US image frame U_i, its carotid artery segmentation result is a binary mask M_i, with its foreground centroid calculated as $c_i = (x_i, y_i)$, and the area as A_i. Based on the continuity of the mask centroid, a low-pass filter is performed on sequences x_i and y_i to obtain a new centroid $\hat{c}_i = (\hat{x}_i, \hat{y}_i)$. Similarly, based on the continuity of the mask area, low-pass filter is performed on the sequence A_i to obtain a new area $\hat{A}_i$. For the mask M_i, an initial translation transformation is applied, with a translation amount of $(\hat{x}_i - x_i, \hat{y}_i - y_i)$. Then, a scaling transformation centered on the centroid $\hat{c}_i$ with a scale of $\hat{A}_i/A_i$ is performed, resulting in a new mask denoted as $\tilde{M}_i$. Subsequently, the newly obtained mask undergoes Gaussian filtering:

$$\hat{M}_i = g(\tilde{M}_{i-w}, \cdots, \tilde{M}_i, \cdots, \tilde{M}_{i+w}). \tag{8}$$

This implies that the value of a given pixel is derived from Gaussian smoothing of the values at corresponding positions in several adjacent masks. It is important to note that the mask at this stage is no longer binary. The elements in the mask are updated with re-thresholding, where the thresholds are selected by

$$\sigma = \arg\min\left(\left|\sum_{i=1}^{n} Area(\widehat{M}_i) - \sum_{i=1}^{n} \hat{A}_i\right|\right), \tag{9}$$

indicating that we select the threshold to minimize the volume error of CCA before and after the re-thresholding.

In the next stage, we need to deform the original US image based on the calculated masks, so an algorithm based on optical flow is proposed. To ensure the continuity of the optical flow, both masks need to be expanded. For pixel $M(x, y)$ in mask M, its distance from the foreground is calculated and noted as d. The value of the pixel is reset as:

$$M(x,y) \leftarrow \begin{cases} 1, d = 0 \\ 1 - d/d_{max}, 0 < d < d_{max} \\ 0, d \geq d_{max} \end{cases}, \tag{10}$$

where d_{max} is a hyper-parameter of the algorithm, which is set as 100.

Based on the expansion, a dense optical flow field is calculated between the expanded original CCA mask and the compensated and expanded mask using the DeepFlow method [26]. By applying the dense optical flow field to the original US images, we can obtain the desired US images with motion compensation.

Considering the movement of the probe between each frame of the US images have been considered, we do not employ a traditional 3D ultrasound reconstruction algorithm. Instead, the 2D images, after motion compensation, are directly stacked to form 3D voxel data.

4) Longitudinal US Image Reslicing: To better assist the physicians in diagnosis, we extract longitudinal 2D images of the carotid artery through resampling within the 3D voxels. The resampling method is as follows: we first calculate the centroid of the carotid mask for each frame, denoted as $c_i = (x_i, y_i)$. and then we take the ith column of the resampled image corresponding to the ith column of the $[y_i]$th frame.

Since in the previous motion compensation step we have not considered the alignment of intima-media, the resulting images could not continuously and clearly display the carotid intima-media. The dense optical flow field obtained previously can be applied to the intima-media segmentation mask, and the same stacking and resampling processes are then applied to the obtained mask, allowing the intima-media mask to be visualized on the resampled images. Contour detection methods are used to identify several bounding boxes containing the intima-media. Boxes are merged according to their distance to form the ROI of the intima-media in the resampled image. Horizontal Gaussian smoothing within the ROI results forms a smoother appearance of the intima-media

in the resampled images.

III. Experimental Results

A. Data Acquisition

The system to perform semi-automatic CCA scanning consisted of three parts: a specially designed parallel robot which is described in [20], a portable hand-held US scanner (Q7, VINNO Technology (Suzhou)Co., Ltd.) and a PC laptop (Intel(R) Core(M) i5-8300H, 16GB RAM and Nvidia GeForce GTX 1050Ti). B-mode images were collected by the US probe and transferred to PC by an image acquisition card. Five volunteers were invited to attend the experiment. Manual scan of CCA was firstly done by an expert to acquire images for detection and segmentation models training. These images were labeled by the expert. Semi-automatic CCA scanning were sequentially implemented according to our method, with the acquired images labelled by the expert, including IMT. Ethics approval was granted from the Institutional Review Board of Institute of Automation, Chinese Academy of Sciences (IA11-2401-020301).

B. Robot Control

After the manual start point localisation, the initial contact force and horizontal position of CCA in the US images were recorded, and then force control and visual-servoing method were both applied to control the robot to automatically scan along the neck. In our experiment, the initial contact force from five subjects was between 4 N to 9 N, and the force error during the scanning was below 1 N. Meanwhile, the horizontal position error of the bounding boxes in the US images was less than 50 pixels, which is about 2 mm.

C. CCA Detection and Segmentation

1961 US images of CCA were collected on the five volunteers as the dataset for the detection and segmentation task. They were randomly seperated with the test set including 10% of the images. An example result of detection and segmentation is shown in Fig. 3.

For detection task, we have achieved 99.5% for AP_{50} metric (Average Precision at 50% Intersection over Union threshold) and 93.2% for $mAP_{50:95}$ metric (mean Average Precision at different Intersection over Union thresholds ranging from 50% to 95%) in the test set. Meanwhile, the inference efficiency has achieved 21 frames per second in the laptop, which sufficiently satisfies the requirements of our control method.

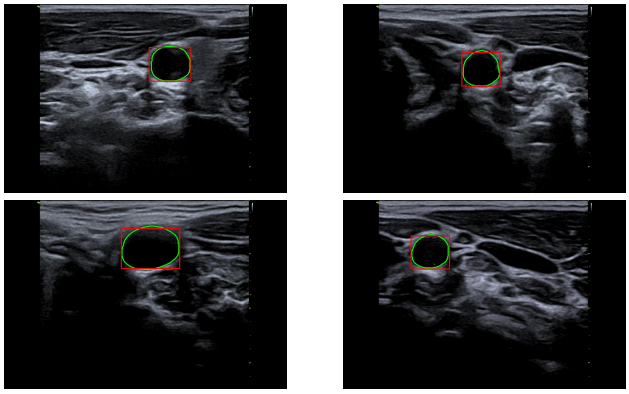

Fig. 3: Example for CCA detection and segmentation. The red squares represent the bounding boxes of CCA detection, and the green lines represent the contour of CCA segmentation masks.

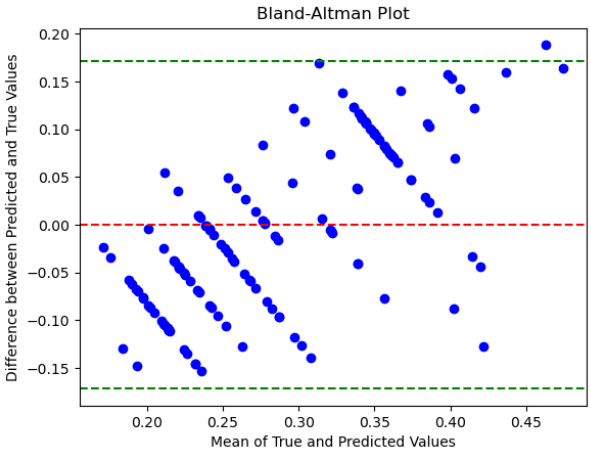

Fig. 4: Measurement result of IMT: Bland-Altman plot of measurement values and ground truth.

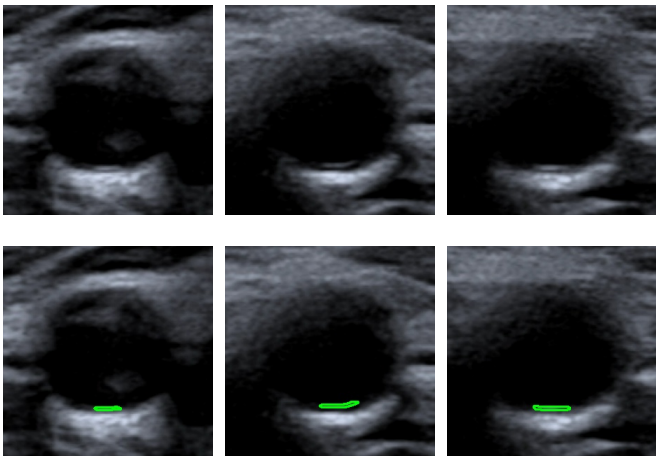

Fig. 5: Example for intima-media segmentation. The first row shows original images and the second shows the contours of intima-media using green lines.

For segmentation task, we achieved a Dice loss of 0.00535 and an intersection over union (IoU) score of 0.991 in the training set, 0.0076 and 0.986 respectively in the test set. This outcome demonstrates the algorithm's capability to accurately segment and identify CCA within the images. It has to be noticed that the segmentation task does not have to be operated in real time in the same laptop, so the inference time is not under consideration.

D. Intima-media Segmentation

It has to be noticed that intima-media can not be identified in some of cross-sectional US images. Among all the images we have acquired during semi-automatic scanning, 171

TABLE I: IMT MEASUREMENT RESULTS OF FIVE SUBJECTS.

Subject id.	Num. of images	Num. of positive images	true value (mm)	predicted value (mm)	Recall	MSE
1	239	40	0.25±0.03	0.23±0.04	0.85	0.005
2	200	23	0.27±0.02	0.28±0.08	0.91	0.007
3	318	33	0.35±0.05	0.36±0.08	0.79	0.009
4	245	27	0.28±0.04	0.23±0.03	0.67	0.005
5	242	48	0.29±0.04	0.32±0.09	0.93	0.008
total	1244	171	/	/	0.86	0.008

images were labelled by experts as positive, where intima-media can be identified. The aggregate rate of identification by our segmentation algorithm is quantified at 86.0%, indicating that 147 images were predicted as positive.

In these images, IMT was measured according to segmentation masks. Fig. 4 illustrates the Bland-Altman plot of measurement error between ground truth values and calculated ones by our algorithm, and measurement results of each subject are listed in Table I. Specifically, 66.0% of the solutions exhibited an error less than 0.1 mm; furthermore, all of the solutions demonstrated an error margin not exceeding 0.2 mm. Additionally, Fig. 5 shows various examples to facilitate a qualitative evaluation for intima-media segmentation.

E. 3D Reconstruction Results

The purpose of our 3D reconstruction algorithm is to reconstruct a smooth 3D voxel mask of human CCA, with compensation of positioning error and cardiac cycle. Five volunteers were scanned according to the scanning protocol described in Section II.C, and the scanning data were then processed with the algorithm described in section II.D. After that, the 3D voxel mask was utilized to deform the original US images, aiming at synthesising a longitudinal view US image of CCA. The reconstruction and reslicing results are showed in Fig. 6. As can be seen, our algorithm successfully compensated patient movement and cardiac cycle for CCA reconstruction and longitudinal view synthesis.

To further evaluate the reconstruction results quantitatively, we calculated the CCA area of each cross-sectional view of the 3D voxel. For four of the subjects, the area of each frame is plotted in Fig. 7. As can be seen, before compensated, the area of CCA shows a periodicity due to the cardiac cycle, which causes discontinuity in reconstructed 3D voxels. Our algorithm successfully compensates this impact and attains a smooth voxel of CCA. Meanwhile, with intima-media area adjusted, we can obtain longitutional view images, which are often seen as better view for diagnosis.

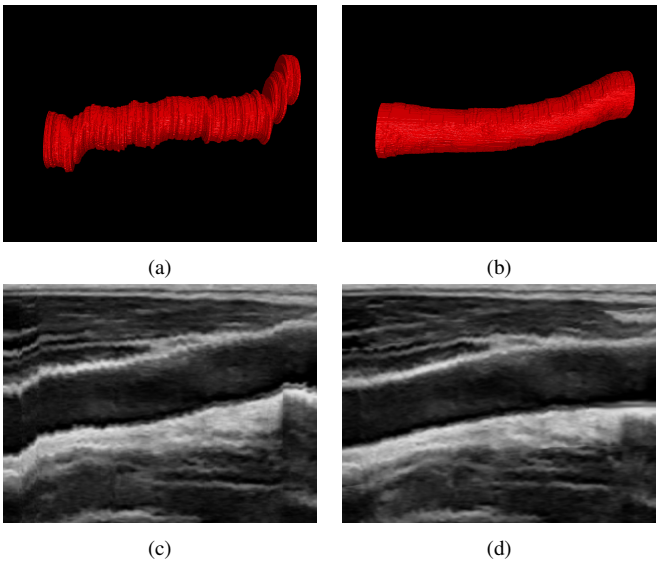

Fig. 6: 3D reconstruction results of human CCA: (a) 3D voxel without compensation. (b) 3D voxel with compensation. (c)Longitudinal view without our algorithms. (d) Longitudinal view with our algorithms.

IV. DISCUSSIONS AND CONCLUSION

The main contribution of this paper is that we introduce a method to semi-automatically acquire 3D US images with movement of patients and cardiac cycle compenseted, while offering diagnostic information of CCA, such as longitudinal view images and IMT. To be more specific, this work utilizes a portable robotic US device and a semi-automatic scanning protocol to acquire serialized US images of human CCA, enabling both quantitative diagnostic indication analysis and qualitative 3D modeling and visualization of human CCA. The device and protocol are tested on human subjects, the results of which prove that clear US images of continuous position can be acquired. After that, an algorithm is proposed to compensate deformation of US images due to unexpected movement and cardiac cycle of patients, allowing accurate 3D reconstruction of CCA. To provide better assistance for physicians, we propose an algorithm to precisely calculate the IMT of CCA, which is an significant indicator for cerebrovascular events prediction. The quantitative results of IMT calculation shows that error is acceptable. Meanwhile, we reslice the 3D voxel and modify intima-media area to get longitudinal view image of CCA, where more diagnostic information can be acquired.

In practical scenarios, with the framework we proposed, it is possible for patients to get US examination in community medical institution, supervised by less experienced personnel. Subsequently, the 3D model and corresponding diagnostic information like IMT can be sent to physicians for further diagnosis and treatment.

To further expand our framework, in the future, our work will focus on multi-modality fusion of US, like elastography and Doppler, to construct person-specific 3D model of CA based on multi-modality information. In addition, we will analyze more clinical features like plaque and stenosis can be analyzed based on US images acquired.

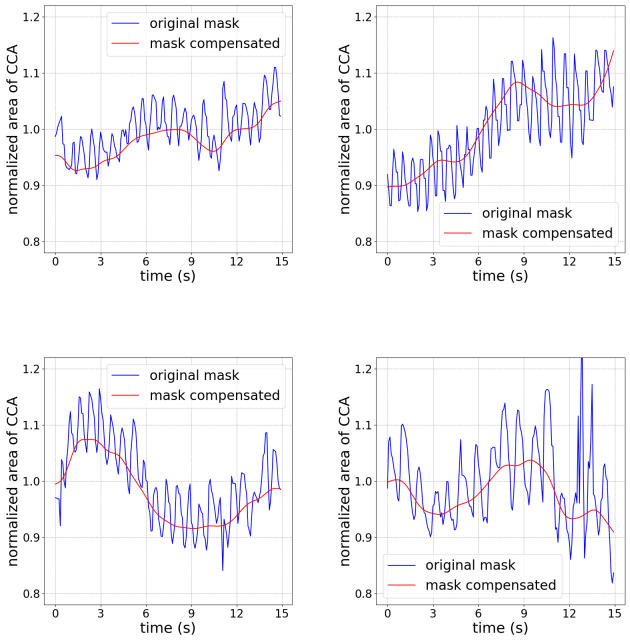

Fig. 7: Comparison of normalized area of CCA cross section between original CCA masks and compensated ones.

REFERENCES

[1] E. V. Ratchford and N. S. Evans, "Carotid artery disease," *Vascular Medicine*, vol. 19, no. 6, pp. 512–515, 2014.

[2] K. Li, Y. Xu, and M. Q.-H. Meng, "An overview of systems and techniques for autonomous robotic ultrasound acquisitions," *IEEE Trans. on Med. Robot. and Bionics*, vol. 3, no. 2, pp. 510–524, 2021.

[3] C. S. G. Murray, T. Nahar, H. Kalashyan, H. Becher, and N. C. Nanda, "Ultrasound assessment of carotid arteries: Current concepts, methodologies, diagnostic criteria, and technological advancements," *Echocardiography*, vol. 35, pp. 2079–2091, Dec. 2018.

[4] P.-J. Touboul, D. E. Grobbee, and H. den Ruijter, "Assessment of subclinical atherosclerosis by carotid intima media thickness: technical issues," *Eur. J. OF Preventive Cardiol.*, vol. 19, pp. 18–24, 8 2012.

[5] X. Guan, H. Wu, X. Hou, Q. Teng, S. Wei, T. Jiang, J. Zhang, B. Wang, J. Yang, and L. Xiong, "Study of a 6dof robot assisted ultrasound scanning system and its simulated control handle," in *2017 IEEE International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM)*, pp. 469–474, 2017.

[6] D. Huang, Y. Bi, N. Navab, and Z. Jiang, "Motion magnification in robotic sonography: Enabling pulsation-aware artery segmentation," in *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 6565–6570, 2023.

[7] J. Xia, F. Deng, G. Yun, Q. Zhang, and Z. Liu, "Automatic ultrasound scanning for the Carotid Artery by robotic system," in *2022 International Conference on Service Robotics (ICoSR)*, (Chengdu, China), pp. 170–174, June 2022.

[8] Z. Jiang, H. Wang, Z. Li, M. Grimm, M. Zhou, U. Eck, S. V. Brecht, T. C. Lueth, T. Wendler, and N. Navab, "Motion-aware robotic 3d ultrasound," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 12494–12500, 2021.

[9] M. Chen, Y. Huang, J. Chen, T. Zhou, J. Chen, and H. Liu, "Fully robotized 3d ultrasound image acquisition for artery," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 2690–2696, 2023.

[10] Y. Bi, Z. Jiang, Y. Gao, T. Wendler, A. Karlas, and N. Navab, "VesNet-RL: Simulation-Based Reinforcement Learning for Real-World US Probe Navigation," *IEEE Robot. and Automat. Lett.*, vol. 7, pp. 6638–6645, July 2022.

[11] Y. Huang, W. Xiao, C. Wang, H. Liu, R. Huang, and Z. Sun, "Towards Fully Autonomous Ultrasound Scanning Robot With Imitation Learning Based on Clinical Protocols," *IEEE Robot. and Automat. Lett.*, vol. 6, pp. 3671–3678, Apr. 2021.

[12] S. Yang, T. Li, Y. Lv, Y. Xia, and S. Li, "Carotid lumen diameter and intima-media thickness measurement via boundary-guided pseudo-labeling," *IEEE Signal Process. Lett.*, vol. 30, pp. 1027–1031, 2023.

[13] Y. Yuan, C. Li, L. Xu, S. Zhu, Y. Hua, and J. Zhang, "Joint segmentation of intima-media complex and lumen in carotid ultrasound images," in *2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 1796–1799, 2022.

[14] M. Jing, K. Owen, B. M. Namee, I. B. A. Menown, and J. McLaughlin, "Investigating temporal features of carotid intima-media thickness from ultrasound imaging with recurrent neural networks," in *2023 45th Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, pp. 1–4, 2023.

[15] Q. Huang, B. Gao, and M. Wang, "Robot-assisted autonomous ultrasound imaging for carotid artery," *IEEE Trans. on Instrum. and Measurement*, pp. 1–1, 2024.

[16] Q. Huang and Z. Zeng, "A Review on Real-Time 3D Ultrasound Imaging Technology," *BioMed Res. Int.*, vol. 2017, pp. 1–20, 2017.

[17] Y. Lv, Y. Ning, Y. Shen, J. Wang, and M. Zhang, "A Real-time Interactive 3D Ultrasound Imaging System," in *2022 WRC Symposium on Advanced Robotics and Automation (WRC SARA)*, (Beijing, China), pp. 113–119, IEEE, Aug. 2022.

[18] T. Djukic, B. Arsic, S. Djorovic, N. Filipovic, and I. Koncar, "Validation of the machine learning approach for 3D reconstruction of carotid artery from ultrasound imaging," in *2020 IEEE 20th International Conference on Bioinformatics and Bioengineering (BIBE)*, (Cincinnati, OH, USA), pp. 789–794, IEEE, Oct. 2020.

[19] S. Song, Y. Huang, J. Li, M. Chen, and R. Zheng, "Development of Implicit Representation Method for Freehand 3D Ultrasound Image Reconstruction of Carotid Vessel," in *2022 IEEE International Ultrasonics Symposium (IUS)*, (Venice, Italy), pp. 1–4, IEEE, Oct. 2022.

[20] Z. Deng, X. Hou, M. Hao, and S. Wang, "System design and workspace optimization of a parallel mechanism-based portable robot for remote ultrasound," in *2023 8th International Conference on Automation, Control and Robotics Engineering (CACRE)*, pp. 263–269, 2023.

[21] Z. Deng, X. Hou, P. Zhang, M. Hao, C. Chen, and S. Wang, "A novel imu-based 6-dof pose tracking method for parallel-mechanism-based ultrasound robot," in *2023 IEEE International Conference on Mechatronics and Automation (ICMA)*, pp. 852–857, 2023.

[22] X. Bao, S. Wang, L. Zheng, R. J. Housden, J. V. Hajnal, and K. Rhode, "A novel ultrasound robot with force/torque measurement and control for safe and efficient scanning," *IEEE Trans. on Instrum. and Meas.*, vol. 72, pp. 1–12, 2023.

[23] G. Jocher, A. Chaurasia, A. Stoken, J. Borovec, NanoCode012, Y. Kwon, K. Michael, TaoXie, J. Fang, imyhxy, Lorna, Z. Yifu), C. Wong, A. V, D. Montes, Z. Wang, C. Fati, J. Nadar, Laughing, UnglvKitDe, V. Sonck, tkianai, yxNONG, P. Skalski, A. Hogan, D. Nair, M. Strobel, and M. Jain, "ultralytics/yolov5: v7.0 - YOLOv5 SOTA Realtime Instance Segmentation," Nov. 2022.

[24] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pp. 234–241, 2015.

[25] G. Sciortino, D. Tegolo, and C. Valenti, "Automatic detection and measurement of nuchal translucency," *Comput. in Biol. and Med.*, vol. 82, pp. 12–20, 2017.

[26] P. Weinzaepfel, J. Revaud, Z. Harchaoui, and C. Schmid, "DeepFlow: Large Displacement Optical Flow with Deep Matching," in *2013 IEEE International Conference on Computer Vision*, pp. 1385–1392, 2013.