Intra-Slice Motion Correction of Intravascular OCT Images Using Deep Features

Atefeh Abdolmanafi ⓘ, Luc Duong, Nagib Dahdah ⓘ, and Farida Cheriet

Abstract—**Intra-slice motion correction is an important step for analyzing volume variations and pathological formations from intravascular imaging. Optical coherence tomography (OCT) has been recently introduced for intravascular imaging and assessment of coronary artery disease. Two-dimensional (2-D) cross-sectional OCT images of coronary arteries play a crucial role to characterize the internal structure of the tissues. Adjacent images could be compounded; however, they might not fully match due to motion, which is a major hurdle for analyzing longitudinally each tissue in 3-D. The aim of this study is to develop a robust tissue-matching-based motion correction approach from a sequence of 2-D intracoronary OCT images. Our motion correction technique is based on the correlation between deep features obtained from a convolutional neural network (CNN) for each frame of a sequence. The optimal transformation of each frame is obtained by maximizing the similarity between the tissues of reference and moving frames. The results show a good alignment of the tissues after applying CNN features and determining the transformation parameters.**

Index Terms—**Motion correction, tissue characterization, deep features, coronary artery, optical coherence tomography (OCT).**

I. INTRODUCTION

CORONARY arteries, which are responsible to deliver oxygenated blood to the heart muscles can be affected by arterial stenosis and lead to myocardial infarction [1], [2]. Functionality of the cardiac tissues significantly depends on the coronary blood flow to the myocardium. Therefore, coronary artery disease (CAD) is the main leading cause of myocardial infarction and sudden death [1], [2]. Angiographic images allow to visualize the trajectory of the contrast agent but they cannot provide any information on the underlying coronary tissue layers. Considering the limitations of coronary angiography to evaluate coronary artery disease, catheter based Intravascular Ultrasound (IVUS) has been used for many years to evaluate coronary artery tissue layers and pathological formations on different coronary artery layers [3]. The IVUS imaging is restricted by limited spatial image resolution (100-150 μm) to detect the thickness of various pathological formations, and low pullback speed. In contrast, Optical Coherence Tomography (OCT) is another catheter-based imaging system, which plays a significant role in the development of medical imaging modalities with interesting advantages over IVUS imaging modality [3]. OCT is an interferometric imaging modality that maps the backscattered near-infrared (NIR) light to create cross-sectional images of the tissues under review in micrometer scale [4]. The image-wire is inserted into the coronary artery using an over-the-wire balloon catheter from patient's groin. A sequence of cross-sectional images of coronary artery segment is recorded using the backscattered light from the arterial wall through each pullback. Considering the fact that light can be attenuated by blood before reaching the vessel wall, blood clearance is required before starting the image acquisition [5], [6]. OCT has been developed for the diagnosis and treatment guidance of coronary artery disease. It has high resolution ranging from 10 to 20 μm to characterize the internal structure of tissues such as vessel wall layers and plaque accumulation [7]. In cross-sectional view, normal coronary artery has a three-layered structure. Intima is the first layer, which is composed of endothelial collagen and is connected to lumen by a single layer of endothelial cells. The second layer, media, consists of muscle cells and is determined by internal and external elastic lamina. Media is enclosed by the outermost layer, adventitia [8].

Intravascular assessment of coronary artery tissues is a challenging task considering the pathological formations due to various coronary artery complications, limitations and the possible artifacts of the imaging system. 2D cross-sectional OCT images of coronary arteries play a crucial role to estimate the thickness variations of arterial wall layers and evaluate the severity of the disease by detecting the various scarring remodelling features [9]. But, the accurate assessment of pathological formations is obtained by considering the information of the adjacent frames and analyzing the volume variation of each tissue using 3D reconstruction. This can be useful to evaluate the dynamics of coronary artery motion and distensibility variation as a result of calcium deposits in a specific coronary artery segment. Volumetric measurements of different tissues are significant for

studying the progression and regression of various pathological formations particularly to evaluate the aneurysmal regions and stenotic segments of the arterial wall [10]–[12].

3D assessment of the morphological tissues is problematic because they are highly affected by motion artifacts resulted by the rotation and translation of the imaging catheter along the artery during image acquisition. Generally, heart beating and respiration during OCT acquisition are the main sources of both axial and longitudinal motions [13]. Furthermore, when the vessel dimensions are changed during a cardiac cycle, it results in larger lumen area in cross-sectional images [13]. The frequency-domain OCT with the pullback speed of 20 mm/sec is 10 - 20 times faster than the previous generation of OCT (time-domain OCT). Therefore, accelerated image acquisition within fewer numbers of cardiac cycles reduces motion artifacts but they cannot be entirely eliminated [14].

A. Related Works

1) Motion Correction: Since the OCT probe moves freely in coronary artery pathway, cross sectional images might be misaligned. This is problematic to evaluate longitudinally each tissue and to design robust clinical measurements. Since OCT is recently used in cardiology, motion correction methods are mostly focused on intracoronary IVUS images and OCT retinal images.

Intravascular Ultrasound (IVUS) has been used for many years as an intracoronary imaging modality in cardiology. Therefore, many of the motion correction and 3D reconstruction methods are focused on IVUS images. Zheng *et al.* focused on 3D reconstruction method of intracoronary IVUS images by reconstructing the pullback path using snake algorithm. The precision of the 3D reconstruction method is increased by focusing on 3D axial position, spatial orientation, and surface fitting [15]. 3D artery is registered with intracoronary IVUS cross-sections using distance mapping algorithm in the work of Tu *et al.* [16]. The other study is focused on 3D reconstruction of IVUS images using the biplane angiography to detect the 3D centerline. The IVUS frames are distributed along the reconstructed centerline in equivalent time intervals [17]. Athanasiou *et al.* proposed a 3D reconstruction of intracoronary IVUS images by estimating the lumen and the outer vessel wall borders using the approximation of the arterial centerline to evaluate plaque formations [18]. Karlas *et al.* proposed a 3D reconstruction method based on the fusion of IVUS and preoperative CT data. The alignment of the detected lumen borders is performed using the ellipse fitting technique and CT data is used to detect the arterial shape [19]. Zhao *et al.* developed a 3D reconstruction method using EM sensor to detect the catheter pose. Lumen contours are detected by applying a radial scan method [20].

Some works are performed on OCT images of coronary arteries. 3D reconstruction of coronary artery images is performed by Ellwein *et al.* using graph theory applied on computed tomography (CT) and OCT data of a single patient after stent placement [21]. Athanasiou *et al.* proposed a semi-automated 3D reconstruction method using OCT images and biplane angiography [22].

Nevertheless, very few studies addressed the problem of non-rigid, intra-slice motion correction in intravascular OCT images using tissue information, rather than the lumen outline. Contour based motion corrections are not reliable enough to solve the problem of motion correction, particularly in pathological cases. When the arterial wall tissue is affected by disease, the lumen does not show a complete circular shape and in some cases it is not detectable properly. Intimal hyperplasia with media disappearance is the common coronary artery complication as a result of coronary artery disease. In most of the pathological cases, there is no media-adventitia border to be detected. However, tissue based motion correction can be generalized to all the normal and pathological cases. Also, accelerated image acquisition within fewer numbers of cardiac cycles reduces the motion artifacts. Therefore, we need to have more accurate motion correction technique to deal with the small variations in the rotation and translation of the catheter from one frame to another due to the accelerated image acquisition. We looked at the problem from the perspective of tissue matching. As much as we can align the tissues, we have more accurate assessment of deformation and pathological changes from one tissue to another. In some cases, the tissue borders are not clear and the shape of the arterial wall is not specific and visible due to pathological formations in coronary artery tissue layers. Therefore, the motion correction model should be based on tissue information rather than the lumen outline to be generalizable to all the cases.

2) Convolutional Neural Network (CNN): Convolutional Neural Networks (CNNs) have been demonstrated very effective in different study fields [23]–[26]. Recently, CNNs are widely used in medical image analysis to perform segmentation and classification tasks [27]–[29]. Considering the fact that it is rare to access sufficient data to train a network from scratch, specifically in the field of medical image analysis, it is wise to transfer the preserved knowledge of a pre-trained network to a new application. Transfer learning is one of the most efficient characteristics of CNNs, which is demonstrated by the work of Azizpour *et al.* [30] in the field of medical imaging. Pre-trained networks have been used as feature generators [31]–[34]. Also, fine-tuned pre-trained networks have been used recently for classification tasks for different applications of medical image analysis [35]. A framework of tissue classification is introduced in our previous work using deep feature learning. Fine-tuned network is used to generate features from each coronary artery layer to train random forest for classification of various arterial wall layers [34].

In this study, we demonstrate the application of deep features to solve the problem of motion correction for intracoronary OCT images considering the movement of the catheter during the image acquisition, and the challenges of the OCT imaging. Our motion correction algorithm is designed to correct the rotation and translation of the catheter as well as the arterial wall deformation through image acquisition by finding the best transformation parameters that will maximize the similarity between deep features extracted from the reference and moving frames at each transformation.
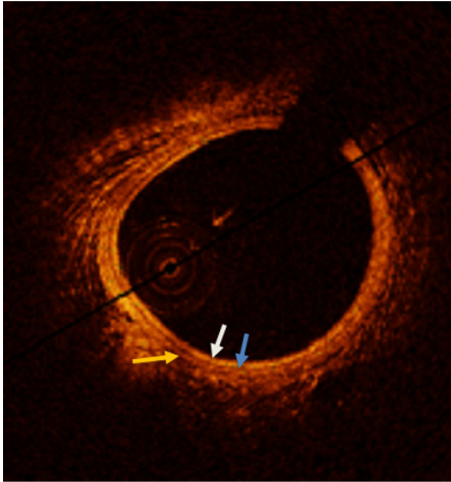
Fig. 1. The three layered structure of coronary artery: Intima is characterized as a signal rich pattern in OCT images, which is the first layer of coronary artery (white arrow). Media is the second layer, which is shown as signal poor layer in OCT imaging (blue arrow), and Adventitia is visualized as the signal rich layer in intracoronary OCT images (yellow arrow). All the three layers are covered by surrounding tissues.

The originality of this approach is as follows:

- Application of deep learning on intracoronary OCT images for motion correction for the first time.
- Intra-slice motion correction in intravascular OCT images using tissue information rather than the lumen outline.
- Using the similarity between deep features extracted by CNN to perform registration.
- This study is not limited to our application and it can be a key to evaluate different plaque formations, and tissue deformations.

This study is organized as follows: We start with a brief explanation of data and pre-processing steps in Section II-A. Then, we present the CNN feature learning process and introduce our motion correction approach in Section II-B. The results using patient dataset are shown and discussed in Section III. Finally, we conclude with possible extensions of this work in Section IV.

II. MATERIAL AND METHODS

A. Data Collection and Pre-Processing

Normal coronary artery has three-layered structure, intima is the first coronary artery layer, which is characterized as a signal rich well-delineated layer in OCT imaging. The second layer, media, appears as a homogeneous signal poor pattern in OCT images and the outermost layer is adventitia, which is visualized as a signal rich layer. Surrounding tissues are covered the three coronary artery layers as it is shown in Fig. 1 [1], [9], [11].

Kawasaki Disease (KD), mucocutaneous lymph node syndrome, is an acute vasculitis syndrome in infants and young children, which is characterized by fever, rash, conjunctivitis, and swollen erythematous hands and feet and leads to inflammation in the walls of medium-sized arteries throughout the body. Coronary arteries are affected by arterial inflammation. Intimal thickening is the most distinguished pathological feature of late

coronary artery lesions in Kawasaki disease. In severe cases, it can lead to localized stenosis, extensive intimal hyperplasia, and consequently disappearance of media [1], [2].

The experiments are performed on 26 retrospective cases comprising of pullbacks of intracoronary cross-sectional images obtained from different pediatric patients with KD using ILU-MIEN OCT system (St. Jude Medical Inc., St. Paul, Minnesota, USA). The axial and lateral resolutions of the OCT system are 12–15 μm and 20–40 μm respectively. FD-OCT is used for image acquisition with the pullback speed of 20 mm/sec and frame rate of 100 frames/sec. Each pullback consists of 270 frames of DICOM images per patient. Permission to conduct this study on retrospective OCT studies was granted by the institutional review board.

For the pre-processing, the approximate region of interest (ROI) including the lumen, arterial wall layers, and the catheter are recognized and extracted using active contour while the Gaussian filter is applied to smoothen the catheter. Active contour does not perform properly to remove the catheter and unwanted red blood cells, specifically when the catheter is located very close to the arterial wall. Therefore, for the last step of the pre-processing, we removed the catheter and unwanted red blood cells by applying the smallest connected components approach [36] (Fig. 2). The original OCT images (RGB images) are used for feature extraction using AlexNet (the OCT images after pre-processing with the size of $352 \times 352 \times 3$ are resized to $227 \times 227 \times 3$ to be used as the input of the network). We converted the images to gray scale for transformation and building the 3D model.

B. Motion Correction Model

1) Deep Feature Extraction: As it is mentioned in the related works, pre-trained CNNs work efficiently as fixed feature generators for OCT images by removing the uppermost fully connected layer (classification layer) from the network architecture and using the activations of the last fully connected layer, right before the classification layer, as deep features of the new images. The process of feature extraction using CNNs is briefly explained in the diagram of Fig. 3.

AlexNet is one of the pre-trained networks, which is used commonly in medical image analysis domain [37], [38]. It is categorized in the group of shallow and simple pre-trained networks. Since AlexNet is trained on natural images, to extract the features which can describe coronary artery tissues more accurately, we prepared the network for our application using transfer learning and fine-tuning. Using AlexNet, the fine-tuning and training process is fast since the network is not very deep. Therefore, we have fewer numbers of parameters for the fine-tuning process compared against deeper networks. In addition, extracted features from fine-tuned AlexNet are strong enough to describe coronary artery tissues as they are used in our previous work to classify different coronary artery layers with high precision [34]. In this study, fine-tuned AlexNet is applied to generate the features for our dataset for motion correction of intracoronary OCT images. The AlexNet model is built on 60 million parameters and 650000 neurons and is trained on 1.2 million

(a) Original image   (b) Gaussian filter to smoothen the catheter   (c) Active contour   (d) Smallest connected components
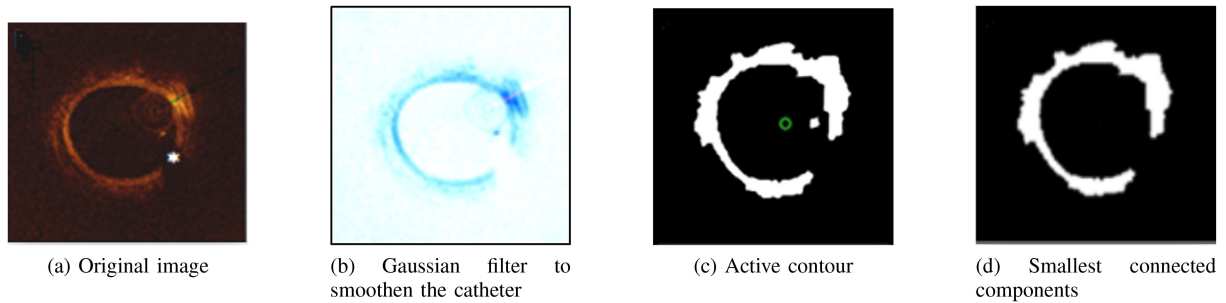
Fig. 2. Pre-processing steps: Original image is shown in (a). The star shows the shadow of the catheter in the image. (b) Gaussian filter is applied to smoothen the catheter to apply active contour and detect the approximate ROI including lumen, catheter, intima, media, and surrounding tissues as it is shown in (c). Then, catheter and unwanted blood cells are removed using connected components approach as shown in (d).

Fig. 3. Feature Extraction Using Convolutional Neural Networks (CNNs): Convolutional operation is performed by moving the filters with fixed stride through the input image and computing the convolution between each filter matrix and input image matrix. Non-linearity is introduced after each convolutional operation by applying a Rectified Linear Unit (ReLu) since convolution is a linear operator. Features are extracted from the fully connected layer right before the classification layer.
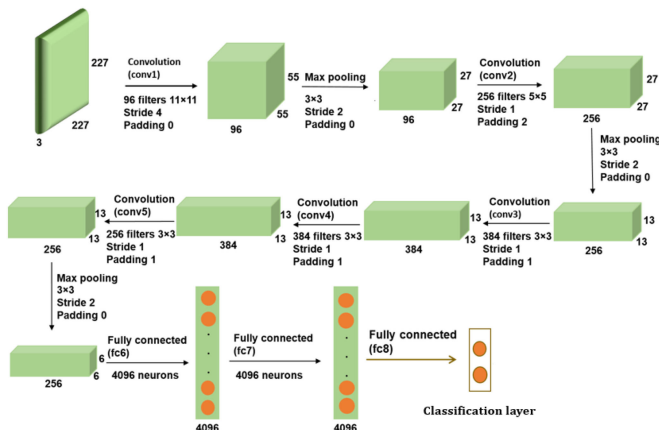

Fig. 4. Modified architecture of AlexNet with five convolutional layers and three fully connected layers, which is used in our experiments.

images from ImageNet dataset. It consists of five convolutional layers (conv1 to conv5) , three max pooling, and three fully connected layers (fc6, fc7, and fc8) [39]. Network architecture is shown in Fig. 4. Each convolutional filter or kernel has the role of feature detector of the images to create the feature maps by sliding through the whole image with defined stride and computing the convolution of the filter matrices and the input image matrix. The depth of the network is determined by the number of kernels. It is also important to mention that each layer extracts the features from the output of the previous layer [40]. The process of updating the weights at each layer and for each iteration, i, is as follows:

$$V_{i+1} = \mu V_i - \gamma_i \alpha \partial L / \partial W \quad (1)$$

$$W_{i+1} = W_i + V_{i+1} \quad (2)$$

Where μ is the momentum, α is the learning rate, γ is the scheduling rate which reduces the learning rate at the end of the iterations [34], [39]. L is the cost function which is aimed to be minimized with respect to the weights, w, at each layer and during the training process. To start inductive transfer, we used the same architecture as the pre-trained network and the weights of each layer are initialized by transferring the weights from the pre-trained network. To start fine-tuning, two facts are considered: 1. There is a risk of overfitting by deep fine-tuning the network, 2. Deep features extracted from the upper layers are more dataset specific, which characterize distinctive features of the images. To fine-tune the network, we evaluated the performance of the network using each set of parameters by calculating the accuracy of classification of coronary artery layers, intima and media, on validation set. For this purpose, different coronary artery layers are annotated manually for all the frames of all 26 pullbacks. Then, we divided the data into the three equal sets of training, validation, and test sets. The images are resized to $227 \times 227 \times 3$ to be used as an input of the network. For the first step, we removed the last three layers (fc8, prob, and classification layer), which are designed for classification task and replaced them by the new classification layers, which are designed for two class classification to classify coronary artery layers, intima and media. Therefore, fine-tuning is started from the last fully connected layer (fc7) using grid searching for an extensive interval of learning rates and we evaluated the performance of the network by measuring the accuracy of classification on validation set using each set of learning parameters. The optimal learning parameters are chosen when the classification accuracy stopped improving. After finding the optimal learning rates, network is retrained by applying the learning rate of 0.1 for fc7 and the learning rate values are decreased to 0.01 from the layer fc6. Also, fine-tuning is stopped at the third convolutional layer since continuing the fine-tuning does not improve the network performance. The learning rates of the first and the second convolutional layers are fixed at zero to keep the weights constant. Also, μ and γ are fixed at 0.9 and 0.95 respectively for the whole process of fine-tuning. Table I shows the learning rates of fine-tuned network. The features used for motion correction are extracted from the last fully connected layer (fc7).

2) Motion Correction: Translation and rotation should be considered in the problem of motion correction, since the

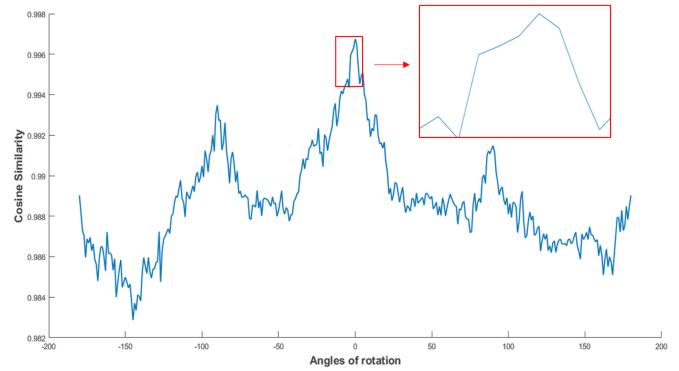| Layers	Learning Rates
fc7	0.10
fc6	0.01
Conv5	0.01
Conv4	0.01
Conv3	0.01
Conv2	0.00
Conv1	0.00

Fig. 5. Cosine similarity between the feature vectors of fixed and moving frames: The calculations are performed in a small interval of angles considering that the probe can not have a huge movement during image acquisition using OCT system.
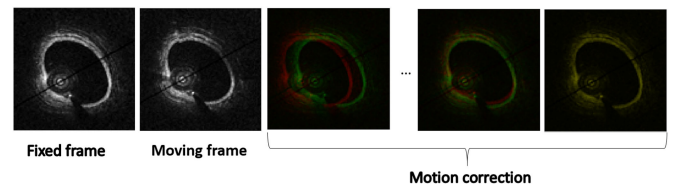

Fig. 6. Motion correction process for one frame of a sequence: The first frame is considered as the reference frame, the second one is moving based on ρ and θ. At each transformation the cosine similarity between the deep feature vectors of both fixed and moving frames are computed. The process is terminated when the maximum similarity between the feature vectors is reached.

position of the catheter is generally out of the center due to the blood flow and high heart beat particularly in infants and children. Also, movements of the interventionist hand can cause a very small rotation of the catheter. Arterial wall as a soft tissue has some deformations due to the cardiac motion, which should be considered in the problem of motion correction. The motion correction is formulated as a non-rigid registration problem when the first frame is considered as the reference frame. The deep feature vector for the reference frame is extracted using the fine-tuned CNN as it is described in the previous section. In this step of the work, the original images before applying pre-processing are used to ensure that all the tissue information is considered during feature extraction. To perform the motion correction, other frames are considered as the moving frames. For the first step, by translating the images at each value of ρ, $\rho = \sqrt{x^2 + y^2}$, deep feature vectors are extracted using CNN model and the cosine similarity is computed as the dot product of the feature vectors of the fixed and moving frames at each translation. To capture the rotation, at each rotational angle, the CNN features are extracted and the cosine similarity between the extracted feature vectors from reference and moving frames is computed. The same process is performed for scaling and shearing in both x and y directions to find the optimal parameters to correct vessel wall deformation. When we evaluated the performance of our proposed method for each translation, rotation, shearing and scaling separately, the final optimal transformation matrix is obtained by performing all the steps simultaneously. The final transformation matrix to map a pair of input coordinates P = (x,y) to a pair of output coordinates P$\prime$ = (x$\prime$, y$\prime$) can be summarized in the following equation,

$$P\prime = \mathbf{M}P \tag{3}$$

Where $\mathbf{M}$ is the final transformation matrix, which is defined as follows,

$$\mathbf{M} = M_{\text{shear}} * M_{\text{scale}} * M_{\text{rigid}}$$

$$= \begin{bmatrix} 1 & \lambda_y & 0 \\ \lambda_x & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_x & 0 & 0 \\ 0 & v_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta & t_x \\ \sin\theta & \cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix} \tag{4}$$

Where $M_{\text{rigid}}, M_{\text{scale}},$ and M_{shear} are rigid transformation, scaling, and shearing matrices respectively. t_x and t_y are translation

parameters along x and y axes respectively, θ is the angle of rotation in 2D, v_x and v_y are scaling parameters, λ_x and λ_y are shearing parameters along x and y directions.

The objective function is defined based on cosine similarity between the feature vectors of reference and moving frames as follows,

$$\text{argmin}_\psi f(\psi) = \text{argmin}_\psi (1 - \cos(\psi)) \tag{5}$$

$$\cos(\psi) = (\mathbf{a}.\mathbf{b})/ \mid \mathbf{a} \parallel \mathbf{b} \mid \tag{6}$$

Where ψ is the angle between the extracted feature vectors from reference and moving frames (**a** and **b** respectively) at each transformation. Unlike other similarity measures, cosine similarity is a measure of the direction-length correspondence between vectors, which is mostly used in high dimensional space and large-scale studies.

Since we aimed to investigate the robustness of CNN features to be applied for motion correction, the optimal transformation parameters are solved using exhaustive search for all the possible transformations. At each translation, rotation, and deformation, we look for the angle which can maximize the cosine similarity of two feature vectors. Since using the OCT system, it is unlikely to have a large rotation of the probe during image acquisition. Therefore, we bounded our exhaustive search in the interval of $[-10, 10]$ although we know that the rotation of the probe is even less than that as it is demonstrated by the experiments (see Figs. 5 and 6). After obtaining the optimal transformation parameters, all the pre-processing steps are performed on the
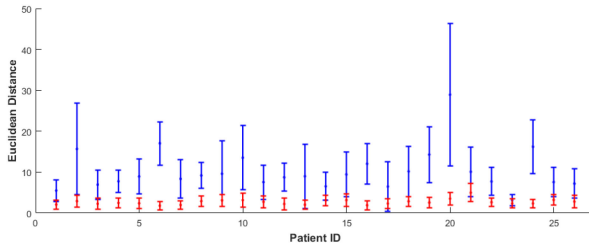
Fig. 7. Comparison of Euclidean distances between the centerline of the 3D model and the reference centerline before and after motion correction, which are shown in blue and red respectively. Error bars show the standard deviation. Euclidean distance is reported as mean± std for each patient.
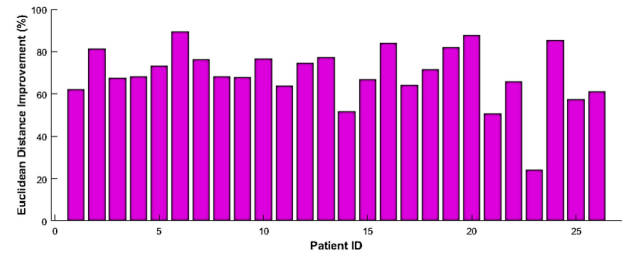

Fig. 8. Percentage of improvement in the calculated mean Euclidean distance between the centerline of the 3D models before and after motion correction for each OCT pullback.
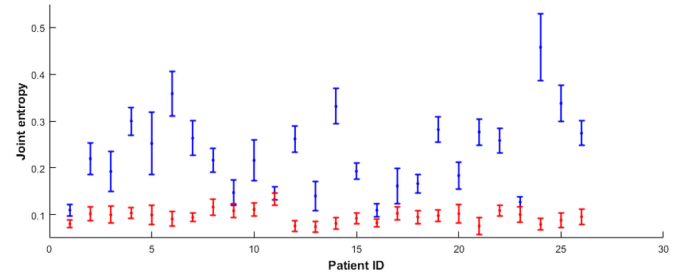

Fig. 9. Intra-slice joint entropy to evaluate tissue matching before and after motion correction, which are shown in blue and red respectively for all 26 pullbacks. Error bars show the standard deviation. Joint entropy is reported as mean± std for each patient.

moving frame. The 3D matrix is built based on the information of the reference frame and the moving frames after applying the optimal transformations. The number of frames in each pullback is considered as the length of the z-axis.

3) Validation: Since OCT is a new imaging system which is recently used in cardiology, to the best of our knowledge, there is no ground-truth available for OCT images. The motion correction of intracoronary OCT images is visually validated by the expert cardiologist. Also, the quantitative validations are performed in two steps:

• *The centerlines of the 3D models are compared against each other before and after motion correction:* For each sequence of images, the approximate center of the lumen for each frame is calculated before and after motion correction. Since a reference centerline is required to evaluate the centerlines before and after motion correction, the Euclidean distance between the lumen center and the corresponding point on the fixed straight line centered on the lumen of the reference frame (the first frame of each pullback) is calculated using equation (7). The results are reported in Fig. 6.

$$D(x,y) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \qquad (7)$$

Where D is the Euclidean distance between the points (x_1, y_1) and (x_2, y_2), which are respectively the lumen center of each frame and the corresponding point on the fixed straight line centered on the lumen of the reference frame.

• *Intra-slice tissue alignment for each 3D model before and after motion correction:* To validate the alignment of tissues, we evaluated the alignment between the tissues of each frame and the previous one by estimating the joint entropy of two consecutive frames before and after motion correction (equation (8)). The results are shown in Fig. 7 for all 26 pullbacks.

$$H(f_1, f_2) = -\Sigma_j \Sigma_k P(j,k) log_2[P(j,k)] \qquad (8)$$

Where f_1 and f_2 are two consecutive frames with pixel values of j and k. P(j,k) is the joint probability of appearing the pixel values j and k at corresponding pixels in the two consecutive frames. Joint entropy is calculated for every two consecutive frames. Then, the result is reported as

mean± std of all the calculations for each volume before and after motion correction.

III. RESULTS AND DISCUSSION

Motion correction is performed on 26 OCT pullbacks obtained from 26 different patients with KD. The results are shown in Fig. 7 demonstrate the measured Euclidean distance between all the points along the centerline of the 3D model and the corresponding points on the fixed straight line centered on the lumen of the first frame before and after motion correction. For each patient, the mean value of all the measured Euclidean distances and standard deviations are calculated. As it is illustrated in Fig. 7, there is a considerable improvement of measured values of Euclidean distance after motion correction. The percentage of improvement in the Euclidean distance between the centerlines of the 3D models before and after motion correction is shown in Fig. 8. To ensure that all the aspects of the quantitative validation of motion correction are considered, the alignment of the corresponding tissues are evaluated by estimating the joint entropy between the region of interests of every two consecutive frames for each volume before and after motion correction. As it is illustrated in Fig. 9, the lower joint entropy after motion correction shows the higher dependency and good alignment of the tissues between the frames for each patient. Fig. 10 demonstrates the percentage of improvement in joint entropy after motion correction.

The results of motion correction of intracoronary OCT images are visually illustrated in Figs. 11 and 12 for two different patients. The centerlines of the 3D models show a good
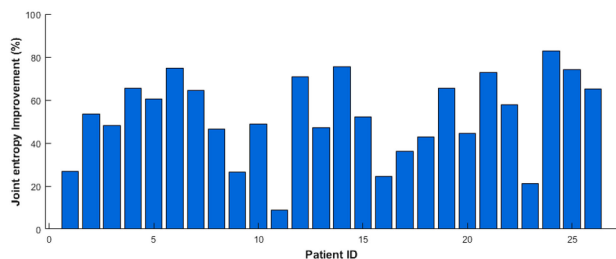
Fig. 10. Percentage of improvement in the calculated mean joint entropy for each OCT pullback before and after motion correction.

alignment of the frames, since the center of the frames are well aligned. Joint entropy is estimated for each two consecutive frames, which demonstrates the robustness of the proposed motion correction approach to align the tissues. The longitudinal cross-section of the 3D model after motion correction shows a good alignment of different coronary artery layers. The parameters of transformation are changed slightly from one frame to the other as it is expected based on very small movements of the catheter and probe along the artery.

The whole process of finding the optimal transformation parameters takes $\sim$ 119 minutes for each OCT pullback. Matlab 2017b is used for all the experiments in this study. The computer configuration is as follows: Intel core i7-6700k, 16 GB of RAM. The experiments are performed on GPU (GeForce Titan X, RAM: 12GB), Windows 10 (64 bit).

The proposed motion correction method is based on deep features extracted from fine-tuned CNN and is aimed to find the best possible correspondence between tissues to facilitate the 3D assessment of tissues and pathological formations. The results show the susceptibility of deep features to describe and characterize different coronary artery tissues while the images obtained from young children with very small size of blood vessels, high heart rate, and lack of collaboration of the patient during the imaging process. Considering different artifacts caused by the imaging system, which definitely leave some deteriorating effects on the image quality, intravascular assessment of coronary artery tissues is a challenging task. In some cases, the cross-section is positioned at a side branch that can cause fold over artifact and can complicate the tissue alignment process.

Since, this is the first time that tissue information is used to solve the problem of motion correction, our method is not comparable with any of the existing methods. Therefore, we discuss some limitations of the existing models, which motivated us to propose tissue based motion correction model. First of all, tissue information is more reliable than lumen and media-adventitia borders to perform the registration, since it can give us detailed information of the tissues. In this study, the experiments are performed on 26 sequences of intracoronary OCT images with three-layered structure. In all the 26 cases, which are used for the experiments, the lumen and media-adventitia borders can be detected, although they are affected by Kawasaki disease. Because as a result of disease, the arterial wall layers can be thickened while the three-layered structure is maintained. We proposed a motion correction approach, which does not contradict with

other existing methods since in normal arterial structure, those methods can also work while their performance is not as robust as tissue based motion correction as it is shown in Fig. 13. Although the contour based motion correction can be applied for the frame registration of all 26 OCT pullbacks, but the results show lower joint entropy using our proposed method. This is expectable since more accurate information regarding the tissues are used to match them properly frame by frame.

However, the feature based motion correction, which is proposed in this study can be generalized to all the cases when the lumen does not show a complete circular shape, media destruction happens due to the disease, and when we have various imaging artifacts. Some examples of these cases are shown in Fig. 14. As an example, intimal hyperplasia with media disappearance is one of the significant coronary artery complications caused by Coronary artery disease (Fig. 15). Therefore, we do not have any information regarding the media-adventitia border. As it is shown in Fig. 16(a), motion correction is performed using our proposed method by extracting deep features from each frame of the pullback of Fig. 15. Also, we used the information of lumen, and media-adventitia borders to solve the problem of motion correction (Fig. 16(b)). As it is shown in the Fig. 16(a), tissue based motion correction approach has a good performance for all the frames of the OCT pullback. But, contour based motion correction is failed in performance after forty three frames since intimal hyperplasia and disappearance of media border caused by KD on coronary artery layers started from the frame forty four and it continued to the end of the sequence. Therefore, there is no information regarding the media-adventitia border to be used for registration (Fig. 16(b)).

To demonstrate the robustness of the tissue based motion correction, we selected 11 KD patients with intracoronary pathological formations while we do not have the three-layered structure of the arterial wall in the whole pullback or part of the pullback. The results of measured intra-slice joint entropy for motion correction using contour based and tissue based methods are shown in Fig. 17. The calculated joint entropy is considerably lower using tissue based motion correction compared against contour based registration. The reason is that the contour based motion correction fails in performance when the lumen or media-adventitia borders are not detectable due to coronary artery disease. Fig. 18 shows the percentage of improvement in joint entropy using tissue based motion correction than contour based motion correction.

In addition, the shadow of the catheter on arterial wall causes loss of information in OCT imaging. As a result, even in the least challenging cases, the lumen border does not appear in a complete circular shape. When the registration is based on the geometry of the lumen border, the algorithm tries to find transformation parameters that can match the lost part of the lumen border along each other. This sometimes causes large rotational angles, which is not possible in the reality of OCT imaging, since the probe has very small rotations during image acquisition. Also, Gaussian filter that is used in some techniques to detect the lumen border is not reliable enough, since it is not edge preserving. In some studies, EM sensors are attached to the

(a)   (b)   (c)
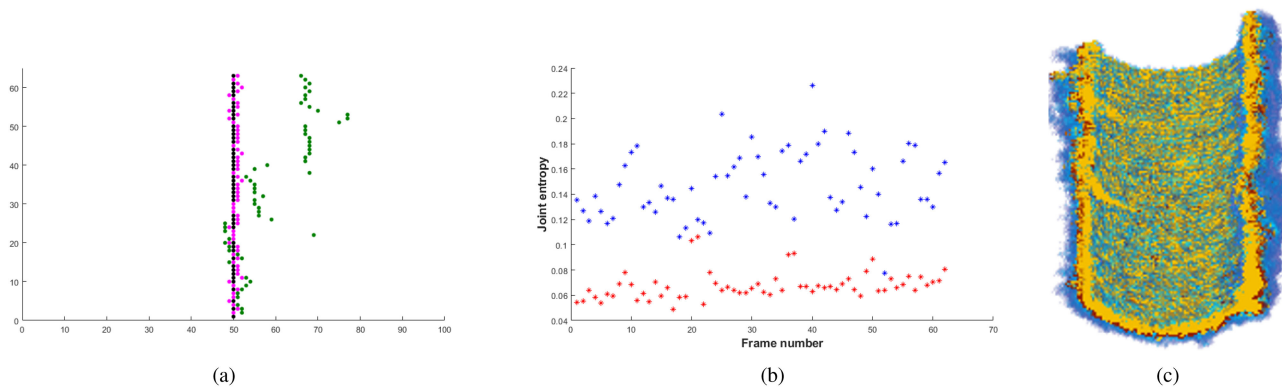
Fig. 11.   Patient 1: Motion correction of intracoronary OCT images: (a) shows the x-z projection of centerlines in green and purple before and after motion correction respectively compared against the fixed straight line centered on the lumen of the first frame (black). Joint entropy between each frame and the previous one is calculated in (b) for the whole sequence of frames before and after motion correction (blue and red respectively). The Longitudinal Cross-section is shown in (c) to illustrate tissue matching after motion correction.
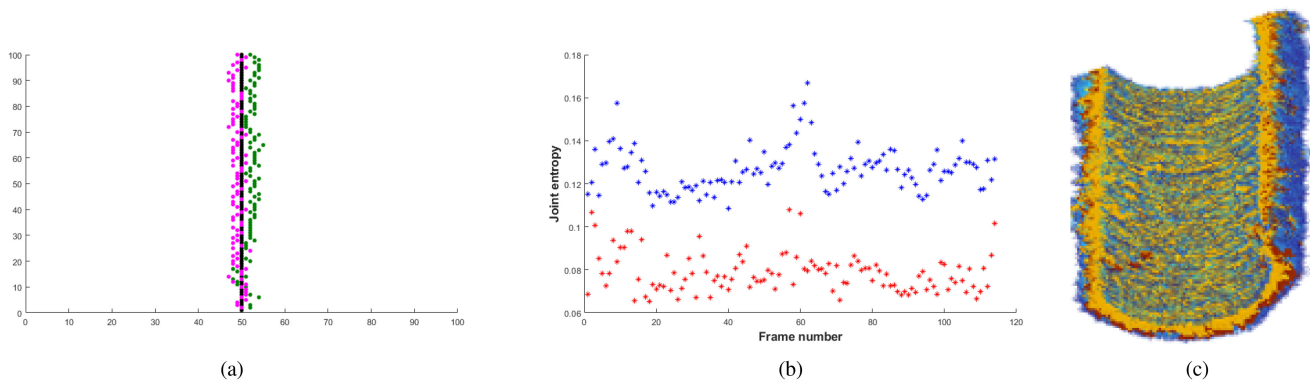

(a)   (b)   (c)

Fig. 12.   Patient 2: Motion correction of intracoronary OCT images: (a) shows the x-z projection of centerlines in green and purple before and after motion correction respectively compared against the fixed straight line centered on the lumen of the first frame (black). Joint entropy between each frame and the previous one is calculated in (b) for the whole sequence of frames before and after motion correction (blue and red respectively). The Longitudinal Cross-section is shown in (c) to illustrate tissue matching after motion correction.
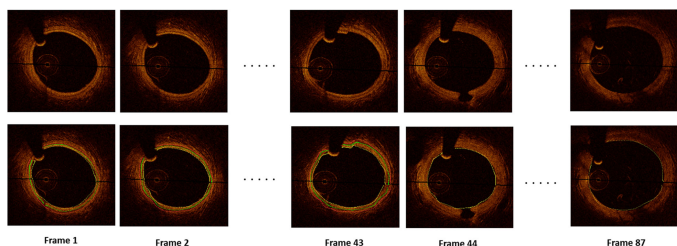

Fig. 13.   Intra-slice joint entropy to evaluate tissue matching before and after motion correction, which are shown in blue and red respectively for all 26 pullbacks. The joint entropy using contour based motion correction is shown in green. Error bars show the standard deviation. Joint entropy is reported as mean± std for each patient.

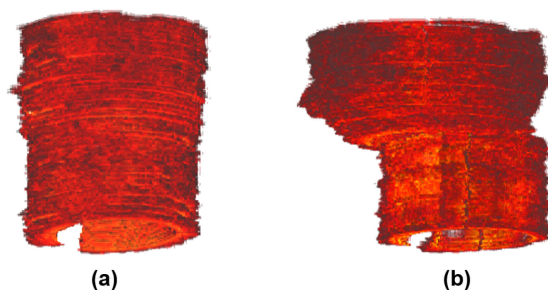(a)   (b)

Fig. 14.   Three examples of lumen deformation caused by the disease and the artifacts of the imaging system.

tip of the catheter to facilitate the accessibility of spacial position of the catheter. But, using EM sensors can reduce the manipulability of the catheter. Using mutual information and joint entropy as objective functions for registration has some limitations. Both measures have low performance with increasing noise. In monomodal image registration, both mutual information and joint entropy can be insensitive to intensity variations between the images. Therefore, they are used for multimodal image registration. From another perspective, deep features give us detailed information of the tissues, which are more reliable than distribution of pixel values to be considered for motion correction and registration problem. The limitations of mentioned studies demonstrate the advantages of using deep features and tissue information for the problem of motion correction in OCT imaging. As the limitation of the proposed motion correction
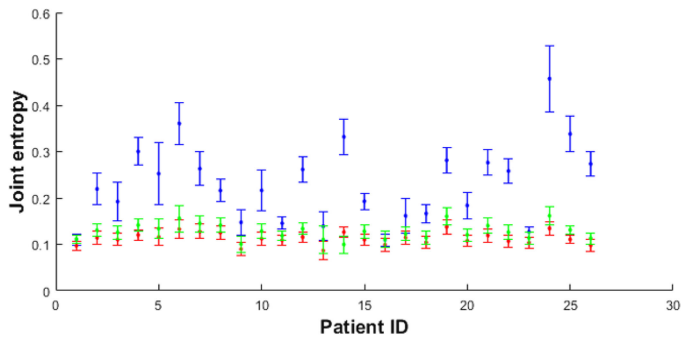
Fig. 15. Intracoronary OCT pullback with three layered arterial wall structure for the first 43 frames and intimal hyperplasia with media disappearance as a result of KD from frame 44 to the end of the sequence. The original frames and corresponding manual segmentation of different arterial wall layer borders are shown for some frames of the sequence.
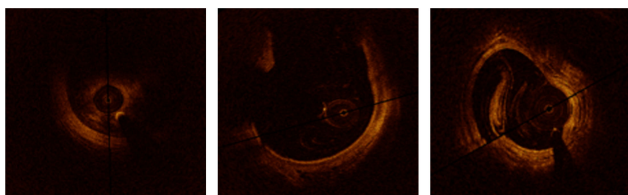

Fig. 16. Intra-slice motion correction is performed (a) using the similarity between deep tissue features, and (b) using lumen, and media-adventitia contours. Using contour information, the registration fails in performance when media-adventitia border is disappeared as a result of the coronary artery disease, particularly KD in this case.
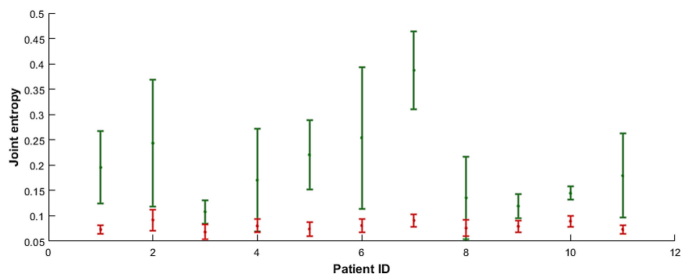

Fig. 17. Intra-slice joint entropy to evaluate the motion correction using tissue-based and contour-based models (red and green respectively). Error bars show the standard deviation. Joint entropy is reported as mean± std for each patient.

method in this study, the proposed non-rigid motion correction approach may lead to a deformation of the artery that should be considered in future works.

To discuss the usefulness of the CNN features in our motion correction model, we investigated the features that the network learns. We extracted features using CNN from one of the OCT images, which is shown in (Fig. 19). Then, we projected activations at each layer to the input pixel space to compare the activations with the original image. The channels in each layer learn different features, which starts from the abstract level features in the earlier layers and ends up to the complex features in the deeper layers. Since there are many channels in each layer and there are many images to investigate while we project the activations of each layer, we focused on projecting the 16 strongest
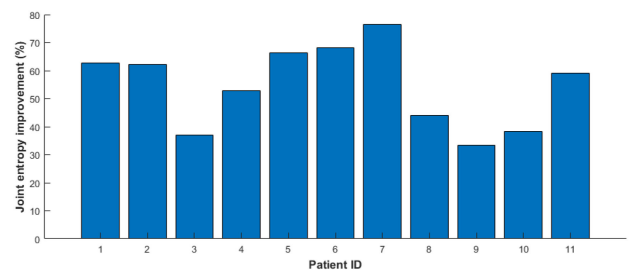

Fig. 18. Percentage of improvement in the calculated mean joint entropy for each OCT pullback using tissue-based motion correction compared against contour based motion correction.
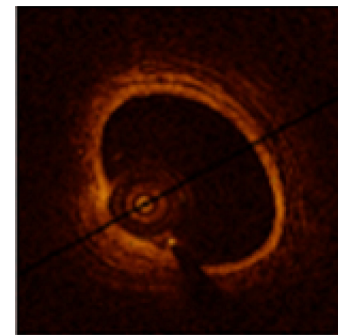

Fig. 19. The original OCT image, which is used for visualization of the CNN features.
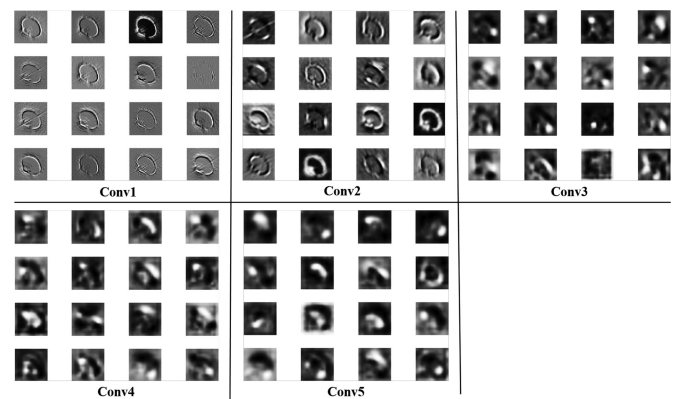

Fig. 20. Projection of the 16 strongest activations of the first to fifth convolutional layers using AlexNet for feature extraction of intracoronary OCT images.

activations at each convolutional layer. Considering the fact that only the positive activations are used to build the final feature map because of applying the Rectified linear unit (ReLu) after each convolutional layer, we visualized the activations of both convolutional layers and ReLu layers to project the positive activations as well. As it is shown in the Figs. 20 and 21, layers one and two excite the information regarding the shape, corners, and edges of the original image and deeper layers (layers four and five); demonstrate the significant and more complex invariances by capturing the similar textures. Therefore, all the details of the images are considered from the edges and borders of the images to the complex texture information of the tissues. Therefore,
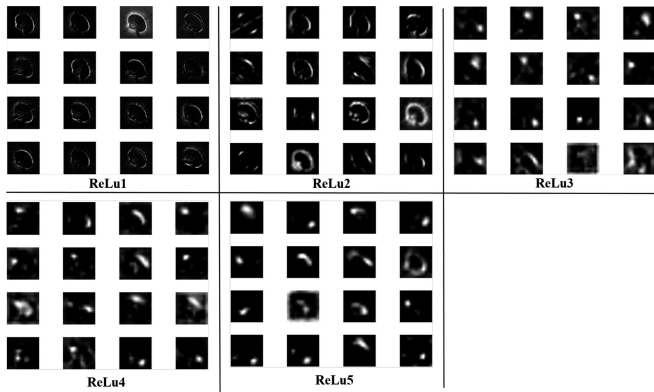
Fig. 21. Projection of the 16 strongest activations of the first to fifth ReLu layers using AlexNet for feature extraction of intracoronary OCT images.

when one information is missing, there are other details that can be considered and the problem of motion correction is not restricted to using a specific feature in the images. The information of contours and borders can be used in normal cases, but in the pathological cases if there is no information, regarding the contours, combination of texture information can be used to find the maximum similarity between the tissues. Therefore, the motion correction model can be generalized to all the cases, normal and disease.

The analysis of coronary artery tissues is a broad study field, which consists of four main steps: 1. Classification of coronary artery layers to recognize characteristic attributes of each layer. 2. Characterizing the abnormalities caused by disease on coronary artery tissues. 3. Measuring the dimensions of the abnormal segments of the artery, which requires both longitudinal and transversal assessment of different pathological formations. 4. Evaluating the functionality of coronary arteries by estimating the stiffness of coronary artery tissues, which affects vascular elasticity and can reduce the vascular distensibility. This work is another complimentary study of our framework for tissue analysis of coronary artery in pediatric cardiology, which is open for any improvement.

IV. CONCLUSION

In this study, a framework for motion correction of the OCT images is proposed. The main contribution of this work is the application of deep features in solving the problem of motion correction of intracoronary OCT images. This will contribute to evaluate the functionality of coronary arteries by analyzing the volume variation and considering the motion of the vessel. Also, it is a robust method to assess the pathological formations by finding the correlation between the tissues of adjacent frames which are recognized using deep features. Since there is no technical method proposed in the literature for longitudinal assessment of the coronary artery abnormalities caused by KD in pediatric patients, this study paves the way for identifying and precisely evaluating vascular wall abnormalities to prevent future complications in young adults who suffered from KD.

REFERENCES

[1] J. W. Newburger *et al.*, "Diagnosis, treatment, and long-term management of Kawasaki disease," *Circulation*, vol. 110, no. 17, pp. 2747–2771, 2004.

[2] M. Hauser *et al.*, "Myocardial blood flow and coronary flow reserve in children with normal epicardial coronary arteries after the onset of kawasaki disease assessed by positron emission tomography," *Pediatric Cardiol.*, vol. 25, no. 2, pp. 108–112, 2004.

[3] K. S. Rathod, S. M. Hamshere, D. A. Jones, and A. Mathur, "Intravascular ultrasound versus optical coherence tomography for coronary artery imaging–apples and oranges," *Interventional Cardiol. Rev.*, vol. 10, no. 1, pp. 8–15, 2015.

[4] A. M. Zysk, F. T. Nguyen, A. L. Oldenburg, D. L. Marks, and S. A. Boppart, "Optical coherence tomography: A review of clinical development from bench to bedside," *J. Biomed. Opt.*, vol. 12, no. 5, pp. 051 403–051 403, 2007.

[5] H. G. Bezerra, M. A. Costa, G. Guagliumi, A. M. Rollins, and D. I. Simon, "Intracoronary optical coherence tomography: A comprehensive review: Clinical and research applications," *JACC: Cardiovascular Interventions*, vol. 2, no. 11, pp. 1035–1046, 2009.

[6] W. Drexler and J. G. Fujimoto, *Optical Coherence Tomography: Technology and Applications*. New York, NY, USA: Springer, 2015.

[7] G. Ferrante, P. Presbitero, R. Whitbourn, and P. Barlis, "Current applications of optical coherence tomography for coronary intervention," *Int. J. Cardiol.*, vol. 165, no. 1, pp. 7–16, 2013.

[8] E. Regar, J. Ligthart, N. Bruining, and G. van Soest, "The diagnostic value of intracoronary optical coherence tomography," *Herz*, vol. 36, no. 5, p. 417, 2011.

[9] A. Dionne *et al.*, "Coronary wall structural changes in patients with Kawasaki disease: New insights from optical coherence tomography (OCT)," *J. Amer. Heart Assoc.*, vol. 4, no. 5, 2015, Art. no. e001939.

[10] K. C. Harris *et al.*, "Feasibility of optical coherence tomography in children with Kawasaki disease and pediatric heart transplant recipients clinical perspective," *Circulation: Cardiovascular Imag.*, vol. 7, no. 4, pp. 671–678, 2014.

[11] J. M. Orenstein *et al.*, "Three linked vasculopathic processes characterize kawasaki disease: A light and transmission electron microscopic study," *PloS One*, vol. 7, no. 6, 2012, Art. no. e38998.

[12] G. J. Tearney *et al.*, "Three-dimensional coronary artery microscopy by intracoronary optical frequency domain imaging," *JACC: Cardiovascular Imag.*, vol. 1, no. 6, pp. 752–761, 2008.

[13] S. Balocco, M. Zuluaga, G. Zahnd, S. Demirci, and S. Lee, *Computing and Visualization for Intravascular Imaging and Computer-Assisted Stenting*. New York, NY, USA: Elsevier, 2016. [Online]. Available: https://books.google.ca/books?id=AAE0vgAACAAJ

[14] S. Takarada *et al.*, "Advantage of next-generation frequency-domain optical coherence tomography compared with conventional time-domain system in the assessment of coronary lesion," *Catheterization Cardiovascular Interventions*, vol. 75, no. 2, pp. 202–206, 2010.

[15] S. Zheng, "Three-dimensional reconstruction of vessels from intravascular ultrasound sequence and x-ray angiograms," in *Proc. 2nd Int. Conf. Biomed. Eng. Inform.*, 2009, pp. 1–5.

[16] S. Tu, N. R. Holm, G. Koning, Z. Huang, and J. H. Reiber, "Fusion of 3D QCA and IVUS/OCT," *Int. J. Cardiovascular Imag.*, vol. 27, no. 2, pp. 197–207, 2011.

[17] H. T. Ma, H. Wang, C. Wang, and W. K. Hau, "3D reconstruction of coronary arteries using intravascular ultrasound (IVUS) and angiography," in *Proc. IEEE Region 10 Conf.*, 2013, pp. 1–4.

[18] L. Athanasiou *et al.*, "Three-dimensional reconstruction of coronary arteries and plaque morphology using CT angiography—Comparison and registration with IVUS," *BMC Med. Imag.*, vol. 16, no. 1, p. 9, 2016.

[19] A. Karlas and S.-L. Lee, "Towards an IVUS-driven system for endovascular navigation," in *Proc. IEEE 12th Int. Symp. Biomed. Imag.*, 2015, pp. 1324–1327.

[20] L. Zhao, S. Giannarou, S.-L. Lee, and G.-Z. Yang, "SCEM+: real-time robust simultaneous catheter and environment modeling for endovascular navigation," *IEEE Robot. Automat. Lett.*, vol. 1, no. 2, pp. 961–968, Jul. 2016.

[21] L. M. Ellwein *et al.*, "Optical coherence tomography for patient-specific 3D artery reconstruction and evaluation of wall shear stress in a left circumflex coronary artery," *Cardiovascular Eng. Technol.*, vol. 2, no. 3, pp. 212–227, 2011.

[22] L. Athanasiou *et al.*, "3D reconstruction of coronary arteries using frequency domain optical coherence tomography images and biplane angiography," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2012, pp. 2647–2650.

[23] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 1–9.

[24] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.

[25] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 818–833.

[26] D. Eigen, J. Rolfe, R. Fergus, and Y. LeCun, "Understanding deep architectures using a recursive convolutional network," 2013, *arXiv:1312.1847*.

[27] H. R. Roth, A. Farag, L. Lu, E. B. Turkbey, and R. M. Summers, "Deep convolutional networks for pancreas segmentation in CT imaging," *Proc. SPIE*, vol. 9413, 2015, Art. no. 94131G.

[28] F. Ciompi *et al.*, "Automatic classification of pulmonary peri-fissural nodules in computed tomography using an ensemble of 2D views and a convolutional neural network out-of-the-box," *Med. Image Anal.*, vol. 26, no. 1, pp. 195–202, 2015.

[29] M. Havaei *et al.*, "Brain tumor segmentation with deep neural networks," *Med. Image Anal.*, vol. 35, pp. 18–31, 2017.

[30] H. Azizpour, A. Sharif Razavian, J. Sullivan, A. Maki, and S. Carlsson, "From generic to specific deep representations for visual recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, 2015, pp. 36–45.

[31] B. van Ginneken, A. A. Setio, C. Jacobs, and F. Ciompi, "Off-the-shelf convolutional neural network features for pulmonary nodule detection in computed tomography scans," in *Proc. IEEE 12th Int. Symp. Biomed. Imag.*, 2015, pp. 286–289.

[32] Y. Bar, I. Diamant, L. Wolf, and H. Greenspan, "Deep learning with non-medical training used for chest pathology identification," *Proc. SPIE*, vol. 9414, 2015, Art. no. 94140V.

[33] J. Arevalo, F. A. González, R. Ramos-Pollán, J. L. Oliveira, and M. A. G. Lopez, "Convolutional neural networks for mammography mass lesion classification," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2015, pp. 797–800.

[34] A. Abdolmanafi, L. Duong, N. Dahdah, and F. Cheriet, "Deep feature learning for automatic tissue classification of coronary artery using optical coherence tomography," *Biomed. Opt. Express*, vol. 8, no. 2, pp. 1203–1220, 2017.

[35] H. Chen *et al.*, "Standard plane localization in fetal ultrasound via domain transferred deep neural networks," *IEEE J. Biomed. Health Inform.*, vol. 19, no. 5, pp. 1627–1636, Sep. 2015.

[36] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, vol. 455, 2nd ed. Beijing, China: Publishing House Electronics Industry, 2002.

[37] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, 2017.

[38] N. Tajbakhsh *et al.*, "Convolutional neural networks for medical image analysis: Full training or fine tuning?" *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1299–1312, May 2016.

[39] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.

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