Research Article | Vol. 11, No. 6 / 1 June 2020 / Biomedical Optics Express 3374

Biomedical Optics EXPRESS

Automatic stent reconstruction in optical coherence tomography based on a deep convolutional model

PENG WU,[1] JUAN LUIS GUTIÉRREZ-CHICO,[2] HÉLÈNE TAUZIN,[3] WEI YANG,[4] YINGGUANG LI,[5] WEI YU,[1] MIAO CHU,[1] BENOÎT GUILLON,[3] JINGFENG BAI,[1,8] NICOLAS MENEVEAU,[3] WILLIAM WIJNS,[6] AND SHENGXIAN TU[1,7]

[1] *Biomedical Instrument Institute, School of Biomedical Engineering, Shanghai Jiao Tong University, No. 1954 Hua Shan Road, 200030 Shanghai, China*
[2] *Department of Interventional Cardiology, Campo de Gibraltar Health Trust, 11207 - Algeciras, Spain*
[3] *Department of Cardiology, University Hospital Jean Minjoz, EA3920, Boulevard Fleming, 25000 Besançon, France*
[4] *School of Biomedical Engineering, Southern Medical University, 510515 Guangzhou, China*
[5] *Kunshan Industrial Technology Research Institute Co.,Ltd., 215347 Kunshan, China*
[6] *The Lambe Institute for Translational Medicine and Curam, National University of Ireland Galway, University Road, H91 TK3 Galway, Ireland*
[7] *sxtu@sjtu.edu.cn*
[8] *jfbai@sjtu.edu.cn*

Abstract: Intravascular optical coherence tomography (IVOCT) can accurately assess stent apposition and expansion, thus enabling the optimisation of a stenting procedure to minimize the risk of device failure. This paper presents a deep convolutional based model for automatic detection and segmentation of stent struts. The input of pseudo-3D images aggregated the information from adjacent frames to refine the probability of strut detection. In addition, multi-scale shortcut connections were implemented to minimize the loss of spatial resolution and refine the segmentation of strut contours. After training, the model was independently tested in 21,363 cross-sectional images from 170 IVOCT image pullbacks. The proposed model obtained excellent segmentation (0.907 Dice and 0.838 Jaccard) and detection metrics (0.943 precision, 0.940 recall and 0.936 F1-score), significantly better than conventional features-based algorithms. This performance was robust and homogenous among IVOCT pullbacks with different sources of acquisition (clinical centres, imaging operators, type of stent, time of acquisition and challenging scenarios). In addition, excellent agreement between the model and a commercialized software was observed in the quantification of clinically relevant parameters. In conclusion, the deep-convolutional model can accurately detect stent struts in IVOCT images, thus enabling the fully-automatic quantification of stent parameters in an extremely short time. It might facilitate the application of quantitative IVOCT analysis in real-world clinical scenarios.

© 2020 Optical Society of America under the terms of the OSA Open Access Publishing Agreement

1. Introduction

Ischemic heart disease (IHD) is still today the first cause of mortality in the world, especially in developed countries [1–3]. The vast majority of cases are due to atherosclerosis, a complex systemic degenerative process resulting in cholesterol accumulation in the extra-cellular space of the arterial intima, with inflammation, foam-cells formation, and necrosis [4–7]. The clinical manifestations of coronary atherosclerosis comprise from stable angina, due to flow-limiting stenosis of the artery, to acute myocardial infarction or sudden death, when the atheroma gets complicated by thrombotic phenomena and occludes suddenly the vessel lumen [8]. Percutaneous

#390113 https://doi.org/10.1364/BOE.390113
Journal © 2020 Received 10 Feb 2020; revised 17 Apr 2020; accepted 17 May 2020; published 29 May 2020

Biomedical Optics EXPRESS

coronary intervention (PCI) with implantation of stents has become the treatment of choice for most cases of IHD in any clinical presentation [9]. However, the stent itself constitutes an insult for the vascular tissue, eliciting a healing reaction that might result in stent failure. An excessive vascular reaction, with excessive neointimal proliferation is the most common substrate for stent restenosis, whilst an insufficient neointimalisation and reendothelialisation may trigger stent thrombosis [10]. The optimisation of stent apposition and expansion during the implantation plays a critical role to achieve a balanced neointimalisation [11] and thus to minimise the risk of stent failure. Malapposition is associated with delayed neointimalisation [10,12,13], which is one of the pathological substrates for stent thrombosis [14], whilst underexpansion is associated with both restenosis [15–19] and stent thrombosis [20].

Intravascular optical coherence tomography (IVOCT) is an optical-based imaging modality with high axial resolution (10-15μm) that enables in vivo assessment of both apposition and expansion during the stent implantation, as direct monitoring of the vascular response and neointimalisation at follow-up [11,21], thus opening for the first time interesting possibilities for personalised medicine. Nonetheless, this assessment requires the accurate detection of all stent struts in each cross-section and the outlining of lumen contour. Since current IVOCT systems render multiple cross-sections (up to every 0.2mm, depending on rotation and pullback speeds), with multiple struts each, this results in a prohibitive amount of measurements and information per stent. If manually performed, IVOCT quantification becomes an extremely cumbersome and time-consuming task, subject to considerable inter- and intra-observer variability [22], impossible to undertake in real-time during the intervention for clinical decision-making.

Algorithms for automatic strut detection in metallic stents might expedite the quantification process while improving its reproducibility, thus making possible its routine implementation to optimise PCI and tailor the adjuvant treatment according to the patient's needs and the principles of precision Medicine. These algorithms take into account the characteristic appearance of stent struts in IVOCT: since metal acts as a perfect reflector of near-infrared radiation, the intense backscatter produces a thin bright line, perpendicular to the near-infrared beam; casting a shadow in perfectly straight line from the emitting source, as a result of the complete attenuation of the near-infrared radiation [11,21].

Most algorithms for automatic strut detection have focused on the A-lines hitherto, using different approaches. Classification based on machine learning method was mostly used like: classification using 4 parameters along the A-line (peak-intensity, presence of shadow, shadow length and speed of intensity rise and fall) [23], bagged decision trees with 12 features of strut and shadow [24], using the wavelet response of each strut, with feature extraction and classification by means of probabilistic neural networks [25], detection of the brightest pixel along the A-line and Prewitt compass filters to detect the trailing shadow and cluster the candidate pixels into struts [26], Bayesian networks and graph search to compute the probability of strut, reinforced by en-face views [27], covered or uncovered struts classification based on support vector machine and uncovered strut clusters detection using mesh growing [28]. Deep learning algorithms was also introduced in strut detection, in [29] artificial neural networks with one hidden layer and ten nodes was used to classify the strut. All these A-line-based methods share some drawbacks: the appearance of the struts may vary between different IVOCT systems (the manufacturer of OCT machine, the dose of contrast agent, the type of vessels and the type of imaging catheters); in addition, some artefacts like incomplete flushing or vascular structures like clots or neointimal flaps may be mistaken as struts. Finally, focusing on the A-line limits the reception field of detection, disregarding global and semantic information, like the consistency of the stent structure in adjacent frames.

Profiting from the extraordinary ability of convolutional neural networks (CNN) for feature extraction [30], we adopted CNN as the element method to extract features of stent struts automatically from a large amount of training data. Considering the stent strut in IVOCT images

Biomedical Optics EXPRESS

as a region instead of a single point, we transformed the detection of struts into an issue of image semantic segmentation. Following the principles of fully convolutional networks (FCN) [31] and U-Net [32], with successful application in biomedical segmentation, we constructed a deep convolutional model with similar hierarchical structures composed by encoder and decoder parts. As differential features, we designed our own four basic modules with bigger capacity and deeper layers. Pseudo-3D image inputs were adopted to aggregate consecutive information from adjacent frames. Furthermore, to mitigate the impact of spatial information lost in the process of feature extraction on tiny strut regions, multi-scale short connections were introduced in our method, contrary to U-Net. We trained our models on over 10,000 cross-sections with proper training strategy and optimal model fine-tuning. Testing of the model performance was carried out on 153 post-PCI pullbacks from the DOCTORS clinical trial [33] and 17 follow-up pullbacks from an core lab. We obtained satisfactory results with fast inference speed and significantly better performance than conventional feature-based methods.

2. Materials

The image dataset to train our metallic stent strut segmentation models came from our academic core lab (CardHemo, Med-X Research Institute, Shanghai Jiao Tong University). A total of 60 IVOCT pullbacks with metallic stent implanted were collected, comprising the three main coronary arteries: LAD, LCX and RCA. IVOCT pullbacks were acquired by the Dragonfly™ catheter using C7-XR™ or ILUMIEN OPTIS™ FD-OCT systems (Abbott, St Paul, Minnesota, USA). The IVOCT pullbacks contained the most commonly used stent platforms worldwide: Xience (Abbott, St Paul, Minnesota, USA), Taxus Liberté (Boston Scientific), Resolute (Medtronic, Santa Rosa, CA, USA), Promus (Boston Scientific) or Orsiro (Biotronik, Bülach, CH). In the ILUMIEN OPTIS™ system the rotation speed was 180 Hz and the pullback speed could be 18mm/sec (maximal scanned length 54mm) or 36mm/sec (maximal scanned length 75mm). In the C7-XR™, the rotation speed was 100Hz and all the pullbacks were performed at 20mm/sec (maximal scanned length 54mm). Cross-sections corresponding to non-stented segments or unacceptable image quality due to incomplete flushing or non-uniform rotational distortion artefacts were excluded from the analysis, resulting in a total of 10,417 cross-sections and 93,059 struts finally analysed. All struts were manually labelled by three experienced IVOCT analysts in the CardHemo core lab, using the ITK-SNAP software (version 3.8) [34]. And a senior analyst performed the quality control before using these labelled data.

3. Image analysis and algorithms

Instead of defining features of strut and trailing shadow for classification, like in previous approaches [23,24,26,27,29], we utilized deep convolutional network to extract automatically the key features for stent struts. We took the U-shape network [32] as the basis for our segmentation model and modified it to improve metallic stent segmentation in IVOCT. The designed model was trained and inferenced both in an end-to-end manner. Polar images of IVOCT were fed to the trained model and the output stent strut masks predicted by the trained model were reconstructed back to Cartesian images.

3.1. Image preprocessing

3.1.1. Paired polar images and mask generation

The two discriminative features of metallic strut, bright spot and trailing shadow, appear differently in Cartesian or polar views. In Cartesian view, struts are distributed around the lumen contour and shadows follow a radial direction in straight line from the optical catheter. Conversely, in polar view struts are distributed along the lumen contour and shadows follow the A-line as slim bars, perfectly perpendicular to the optical catheter side. Compared with Cartesian images, the

Biomedical Optics EXPRESS

strut features in polar coordinates are invariant to the irregular shape of the lumen (e.g. due to dissection or thrombosis) and less affected by artefacts and non-uniform rotational distortion. This invariance is preferable for deep convolutional method, thus polar images were selected to be the input of our models. Polar images were resized to 512×512 pixels to reduce the computation consumption while keeping the features of struts. The mask paired with each image was present in polar version and resized to 512×512 pixels as well, as shown in Fig. 1(A).

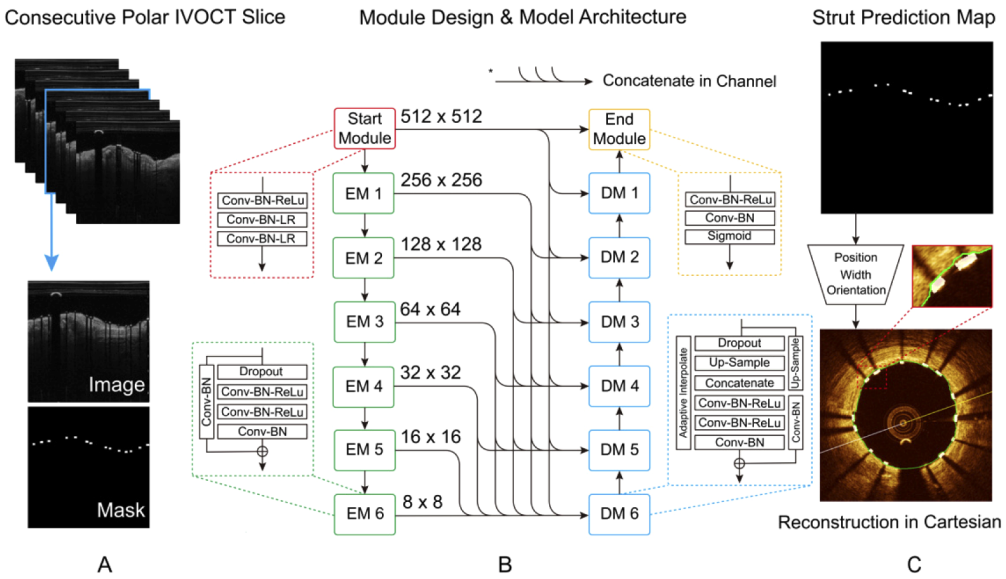

Fig. 1. (A) Pseudo 3D consecutive polar slices to feed the stent strut segmentation model as input. Paired polar image and manually labelled mask were generated. (B) The entire model architecture designed for stent strut segmentation. This model was composed of start module, six encode modules, six decode modules and end module, connected with multi-scale shortcut connection. (C) The strut map predicted from the model. Position, width and orientation of the struts were extracted in post-process for image reconstruction in Cartesian view.

3.1.2. Image normalization

Since IVOCT images were acquired from different OCT systems and different operators and exported with different settings, image quality and image intensity varied substantially among different IVOCT pullbacks. Mean intensity value and standard deviation were calculated for each polar cross-section. Then the polar image intensity was normalised by subtracting the mean value and dividing it by the standard deviation, to homogenise the intensity distribution of all images in the same range and reduce the variability intra-training of the dataset.

3.2. Metallic strut segmentation model

3.2.1. Module design and backbone architecture of model

Four basic modules were designed to compose the backbone architecture of our model: start module, encode module (EM), decode module (DM) and end module. Analogous to most deep learning methods tackling visual tasks, convolution operation (Conv) and batch normalization (BN) were exploited as element layers in each module for features extraction and to lessen the gradient vanishing in training, respectively [35,36].

Biomedical Optics EXPRESS

In the start module, one Conv-BN-ReLU element, and two Conv-BN [30,36], activated by LeakyReLU [37], composed the sequential structure of the module, as structure shown in Fig. 1(B) with red dashed box. Encode module was responsible for extracting features in different scale-levels, from shallow spatial information to high semantic features. Dropout layer was introduced in this module to avoid overfitting [38], then two Conv-BN-ReLU and one Conv-BN were included (with 2 strides in the first convolution to up-sample the features and enlarge the reception field). To accelerate the converging process of our model, we used residual connection introduced by He et al., which demonstrated to be effective in training deep networks [39] (green dashed box in Fig. 1(B)). Decode module was analogous to encode module, containing residual connection, dropout and Conv-BN-ReLU elements. In addition, necessary adaptive interpolation (up-sample or down-sample) and channel concatenating operation were needed to recover the image resolution from the features in different scale-levels extracted via encode modules, represented by the block enclosed by blue dashed box in Fig. 1(B). The last part was the end module, this module outputs the prediction probability map activated by sigmoid layer after Conv-BN-LeakyReLU and Conv-BN element (orange dashed box in Fig. 1(B)). High response probability is expected be found in strut region (low probability in background) for a well-trained model.

For the overall backbone architecture of our model, one start module, six encode modules, corresponding six decode modules and one end module were sequentially contained. Original polar images input first passed through the start module and generate the same resolution (512×512) feature maps, then underwent six encode modules successively, outputting down-sampling features in each encode module. In the last encode module (EM6), 8×8 highly semantic features were obtained. Six one-to-one scale-level matching decode modules composed the decoder of the model, corresponding to each encode module. These decode modules accepted multi-scale inputs acquired from encode modules above (described in detail in section 3.2.2). Decode modules recovered the original image resolution step by step, from 8×8 features to 512×512 maps. In the last step, the end module generated maps with probability of prediction of struts. The whole architecture was U-shape-like connected (as shown in Fig. 1(B) structure), similar to U-Net by Ronneberger et al. [32].

3.2.2. Multi-scale shortcut connection

In processes of features extraction, together with resolution down-sampling, image spatial information gets gradually lost, thus the contour or shape of the target segmentation will not be as good as expected, though with high scoring in some metrics. Shortcut connections adopted in U-Net and RefineNet are aimed at addressing this flaw by bridging the low spatial features from encode modules with high semantic features from decode modules and combine them in concatenating or summing manner [32,40]. As a result, the lost spatial information can be retrieved again along with recovering the resolution and the shape of the prediction map can be refined. However, unlike medical image segmentation with relatively larger targets (e.g. tumor or ventricle segmentation), stent struts are an extremely tiny target with regular shape. The loss of spatial information may affect more the segmentation of so small targets, as some details may vanish in the pooling process. Therefore, the requirement for fine-grained boundary in small targets is stricter.

As a consequence, shortcut connections from one single scale features in one encode module might not suffice to maintain enough spatial information. To overcome this problem, Yu et al. proposed the method of deep layer aggregation to combine iteratively and hierarchically more scale-level features [41]. Our method of aggregation, conversely, connected multi-scale features directly from all encode modules with lower feature level than the current decode level, on a concatenating way, shown as the connection between left and right sides of our model structure in Fig. 1(B).

Biomedical Optics EXPRESS

3.2.3. Input of pseudo-3D images to aggregate the context of adjacent cross-sections

Metallic stents in a coronary artery are 3D targets with specific structural design, (Fig. 2, showing a Xience stent, Abbott). For each single cross-section in an IVOCT sequence, the spatial and semantic information are both correlated with adjacent cross-sections for consistency of stent structures. Our method incorporates the consistency of adjacent frames to refine and restrain the segmentation results, especially in case of uncertain or ambiguous struts. It is common practice among expert cardiologists to check the preceding and subsequent frames to aid the interpretation of challenging images, including equivocal struts. Nevertheless, combining the whole stent segments as a 3D volume can challenge the computation capacity. Most metallic stent platforms, however, are designed as repetition of short modular elements along their longitudinal axis. These short modular elements correspond to a stack of consecutive cross-sections in the IVOCT pullback, dubbed pseudo-3D image. Input from pseudo-3D images, combining n adjacent slices, is enough to aggregate adjacent context. For computation, a multi-channels input feeds the model as shown in Fig. 1(A), and the slice for segmentation is the medium one.

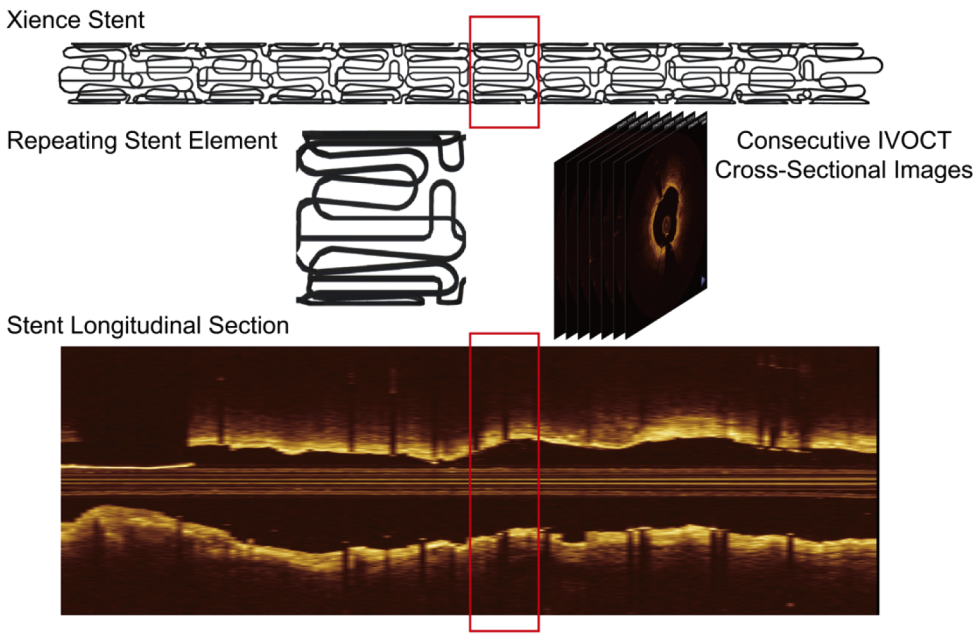

Fig. 2. The metallic stents are often composed of several repeating modular elements (as shown like the real stent structure of Xience, Abbott Inc. and red box outlines its repeating stent element), corresponding to the consecutive IVOCT slices in IVOCT pullbacks (pseudo-3D images).

3.2.4. Training strategy

To enlarge the training dataset and avoid overfitting in the training process, some data argumentations were adopted. We randomly flipped part of images in horizontal direction with a probability of 0.5 and translated the image in a random extent limited in the range of -10% to 10% of image size. Besides, random grey-scale transform was adopted as well. In the process of model training, we randomly divided the 10,417 cross-sectional images into training set and validation set, in a proportion of 8 to 2, and trained the model for 300 epochs and optimised parameters with Adam optimiser.

Biomedical Optics EXPRESS

We designed a joint loss composed by binary cross entropy (BCE) loss and Tversky loss. The BCE loss is frequently used in classification tasks, calculated as Eq. (1):

$$loss_{BCE}=\frac{1}{N}\sum_{i=1}^{N}-g_i log(p_i)-(1-g_i)log(1-p_i) \tag{1}$$

To optimise the performance of our model in segmentation of struts in so highly unbalanced data compared to background, Tversky loss was adopted to achieve improved trade-off between precision and recall (false positive and false negative), which was shown to have good performance in segmenting extremely tiny lesions in brain MRI images [42]. Tversky loss is calculated as Eq. (2) and the trade-off is controlled by α and β:

$$loss_{Tversky}(\alpha,\beta) = 1 - \frac{\sum_{i=1}^{N} p_i g_i}{\sum_{i=1}^{N} p_i g_i + \alpha \sum_{i=1}^{N} p_i(1-g_i) + \beta \sum_{i}^{N} (1-p_i) g_i} \tag{2}$$

Thus, the joint loss is the weighted sum of BCE loss and Tversky loss as Eq. (3):

$$loss_{joint} = \lambda \cdot loss_{BCE} + (1-\lambda) \cdot loss_{Tversky} \tag{3}$$

The λ is the proportional weight of two objective functions.

3.3. Postprocessing of the segmentation results

To visualize the stent structure clearly in the original IVOCT images and further reconstruct the stent in 3D vessels, we projected the polar segmentation result to the original Cartesian view, restoring the real anatomical structures of stents in the vessels. Instead of converting the polar outputs to Cartesian images in a pixel-by-pixel manner, we only extracted the position, orientation and width of each strut segmented in the polar image. Since the bright spot in IVOCT only represents the leading edge of the adluminal side of the strut, we reconstructed the strut in the original Cartesian view for a given strut thickness and overlaid the strut mask on the corresponding strut area in IVOCT images, as shown in Fig. 1(C). Thus, subsequent quantitative assessment of the results of stent implantation can be performed based on this reconstruction.

4. Experimental methods

To evaluate the performance of our proposed strut segmentation model, the model was implemented in a commercialized software (OctPlus, version 2.0, Pulse Medical Imaging Technology, Shanghai, China) for testing of its effectiveness. A serial of testing experiments was designed in our study. Additionally, correlation and analysis of agreement was performed on some quantitative parameters related to the stent (like minimal stent area), comparing our model vs. semi-automatic method using in a commercialized software (QIvus, version 3.1, Medis Medical Imaging System BV, Leiden, The Netherlands), to assess the feasibility of our model for analysis of routinely acquired IVOCT images.

4.1. Independent testing data and experimental environment

The testing process was performed on an independent IVOCT dataset from the DOCTORS clinical trial [33]. These data were collected from multinational centres, totally independent from the training set. The DOCTORS trial included 120 patients with IVOCT performed immediately after stent implantation [33]. After excluding pullbacks of unacceptable image quality due to NURD or insufficient blood flushing, a total of 153 IVOCT pullbacks, with 187,068 struts from 19,494 cross-sectional images were analysed in the current study. The struts were manually labelled by two trained analysts using the ITK-SNAP software. In case of disagreement, an experienced cardiologist acted as referee to make a final decision. Finally, seventeen additional follow-up

Biomedical Optics EXPRESS

IVOCT pullbacks (1,869 cross-sectional images, 18,445 struts) from the CardHemo core lab were included to validate the performance of our model on detecting struts after neointimalisation (the struts were labelled manually). The experimental environment was under the platform of Intel i5-7500 with 32 GB memory and NVIDIA Geforce RTX 2080 graphic card.

4.2. Evaluation of segmentation results by ground truth

The performance of our model was evaluated by means of 5 quantitative metrics comparing the segmentation results from the model vs. the strut mask ground truth. Dice and Jaccard coefficients, reflecting the extent of overlap between segmentation result maps and the ground truth masks pixel-wise, were computed as Eq. (4) and (5):

$$Dice = \frac{2 \cdot |pred \cap mask|}{|pred| + |mask|} \tag{4}$$

$$Jaccard = \frac{|\ pred\ \cap\ mask\ |}{|\ pred\ \cup\ \text{mask}\ |} \tag{5}$$

* *pred: Prediction strut map from proposed model; mask: Ground truth strut map*

Similar to testing methods adopted in previous works [23,24,26,27], the strut detection was assessed strut-wise by the following parameters: precision, recall and F1-score, computed as Eq. (6–8). Precision represents the percentage of correctly detected struts among all predicted struts, whereas recall reflects the percentage of correctly detected struts among all

$$Precision(PC) = \frac{TP}{TP + FP} \tag{6}$$

$$Recall(RC) = \frac{TP}{TP + FN} \tag{7}$$

$$F1 - Score = \frac{2 \cdot PC \cdot RC}{PC + RC} \tag{8}$$

manually-labelled struts. F1-score is a synthetic index, combining precision and recall in a mutually restrictive manner.

The intersection over union (IOU) between the region of predicted strut and the region of manually-labelled strut was calculated, defining a correctly detected strut or true positive (TPs) as an IOU >50%. Thus, prediction, recall and F1-score can be derived. Since the pullbacks in our testing set came from eleven different sources, the results were evaluated per pullback instead of evaluating all unclustered struts as a whole. Subgroup analysis per centre was performed to validate the generalisability and robustness of our model in images generated from different OCT imaging systems and operators. An additional subgroup analysis per stent type was performed in the cases from the clinical site that contributed the greatest number of patients to the testing dataset. Furthermore, the reconstructed 3D platforms were compared with the real stent structure. Finally, a final subgroup analysis was performed on challenging cases for automatic detection: jailing struts over side branches, stent thrombosis, severe stent malapposition, overlapping stents and stent with residual blood.

4.3. Ablation experiments

The performance of our model for strut detection was compared vs. the conventional features-based algorithm used in the QIvus software that has been extensively used worldwide in quantitative evaluation of stent struts by different core labs [26]. Since QIvus only detects single strut points, Dice and Jaccard coefficients cannot be computed: only precision, recall and F1-score were calculated.

Biomedical Optics EXPRESS

Finally, two additional deep-convolution-based models were designed to elucidate the efficiency of the key differential components in our model: pseudo-3D image input and multi-scale shortcut connection. The performance of the whole model was compared vs. U-Net, as basic U-Net, and vs. ResU-Net with pseudo-3D image input but without multiscale shortcut connection, as pseudo-3D ResU-Net, using the same 5 metrics described in section 4.2. All models were designed with the same depth (same number of encoders and decoders), the same basic channels and were trained with the same dataset and the same training strategy. The conventional features-based algorithm and the three deep-convolution-based models were tested on the same testing set described in section 4.1, with 170 pullbacks in total. A prespecified subgroup analysis comparing the performance immediately post-PCI vs. that at follow-up was also performed.

4.4. Correlation and agreement between the model and QIvus in quantitative parameters

The quantitative assessment of an implanted stent comprises several parameters in the analysis per cross-section, like minimum stent area (MSA) and average stent area (ASA), and several parameters in the analysis per strut, like malapposition distance and coverage thickness. Underexpansion and malapposition can be defined after these quantitative parameters, so the cardiologist can estimate the risk for restenosis or thrombosis and adjust the therapy accordingly [43]. Semi-automatic measurements obtained with QIvus software were used as reference standard for the analysis of agreement in quantitative parameters. MSA and ASA were compared in 169 pullbacks, after excluding one case due to abnormal lumen shape; Malapposition distance was assessed in 28 pullbacks with 1,841 malapposed struts and coverage thickness in 17 follow-up pullbacks with 1,602 covered struts with a neointima thickness >0.05mm by visual estimation. The quantitative parameters were calculated by our segmentation model as follows:

(1) malapposition distance and coverage thickness:

The lumen contour was automatically delineated [44]. Malapposition distance was then calculated as the distance between the strut centre and the lumen contour, following a straight line connecting the strut centre with the lumen centre [10,12]. The detachment distance can be calculated by subtracting half of the strut thickness from the malapposition distance [13] (Fig. 3(A)). In the case of covered struts, malapposition distance renders a negative value. For these struts, coverage thickness is calculated as the absolute value of malapposition distance. (Figure 3(B)).

(2) Stent contour fitting and stent area calculation:

After delineation of each individual strut in the cross-section, an ellipse-fitting algorithm was used to fit the stent contour. The elliptical stent contours are shown in Fig. 3(A) and Fig. 3(B) (white curves). Then, stent area per cross-section can be calculated, as MSA and ASA per pullback can be subsequently derived.

The statistical analysis of quantitative parameters started by a normality test, followed by a regression analysis, if appropriate, using manual measurements as reference standard. Correlation was measured by means of Pearson's correlation coefficient and agreement by means of Bland-Altman method.

4.5. 3D reconstruction of the stent structure and time required for analysis

The proposed model was integrated into the OctPlus software (Pulse Medical Imaging Technology, Shanghai, China) [44], allowing automated segmentation and reconstruction of the stents in 3D from IVOCT image pullbacks. The time required for analysis of the stent struts per IVOCT pullback was measured to assess the feasibility of our model in realistic clinical conditions.

Biomedical Optics EXPRESS

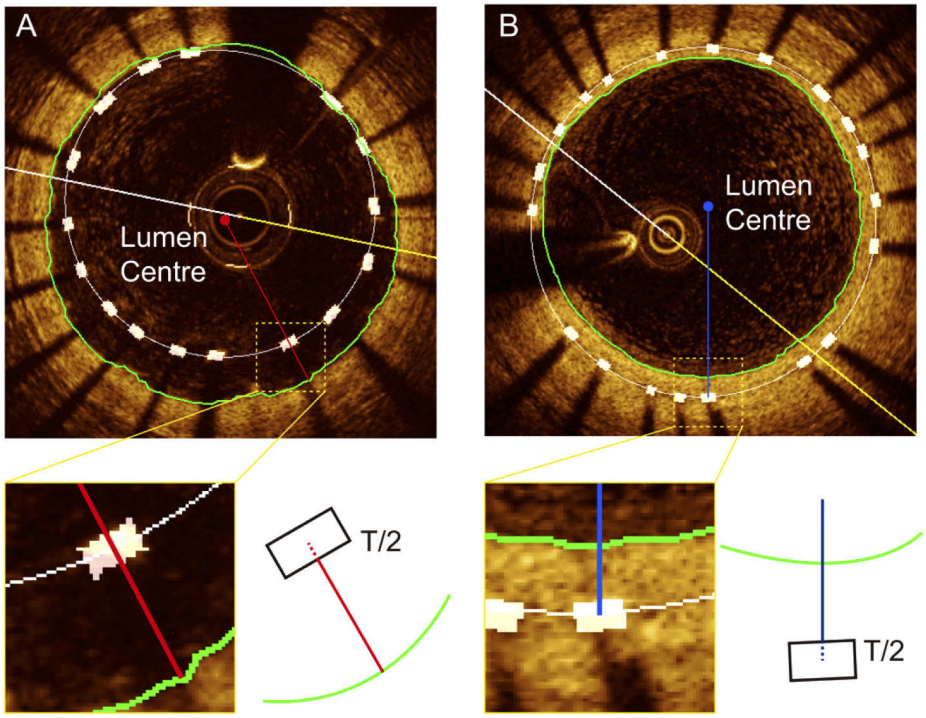

Fig. 3. (A) Malapposition distance is the distance from the midpoint of the strut to lumen contour, following a straight line connecting the midpoint of the strut with the centre of gravity of the vessel. Detachment distance is obtained by subtracting half of the strut thickness from malapposition distance. (B) In covered struts, the assessment of malapposition distance renders a negative value. Coverage thickness is defined then as the absolute value of malapposition distance in covered struts. The stent contour is delineated fitting an ellipse (white curve in A and B), thus enabling an easy calculation of stent area.

5. Results

5.1. Evaluation of segmentation results vs. ground truth

For segmentation metrics, mean Dice coefficient was 0.907 (SD 0.038) and mean Jaccard coefficient was 0.838 (SD 0.057). For detection metrics, mean precision was 0.943 (SD 0.036), mean recall 0.940 (SD 0.039) and mean F1-score 0.936 (SD 0.038). Subgroup analysis per clinical centre is summarised in Table 1, showing homogeneously excellent performance among all centres, thus validating the generalisability and robustness of the model throughout different IVOCT imaging systems and different operators. Considering that the 17 pullbacks from the core-lab corresponded to stents at follow-up, while the other centres provided stents immediately post-implantation, the homogeneous performance of both subgroups confirms the validity of the method for both scenarios.

Thirteen different types of stent were used in clinical centre 1. Results of the subgroup analysis per stent type are presented in Table 2. Segmentation metrics were homogeneous between the different stent types, except for the Omega (Boston) and Multilink (Abbott) stents, which displayed slightly lower segmentation metrics (Dice coefficient 0.876 and 0.867, respectively; Jaccard coefficient 0.792 and 0.781, respectively). For detection metrics, all stent types showed homogeneously high precision (>0.910), recall (>0.900) and F1-score (>0.900).

Biomedical Optics EXPRESS

Table 1. Subgroup analysis per clinical centre (CC)

Clinical Centre	Pullback Numbers	Segmentation Metrics		Detection Metrics (IOU 0.5)		
		Dice	Jaccard	Precision	Recall	F1-Score
Site 1	92	0.911(0.036)	0.845(0.055)	0.946(0.033)	0.948(0.032)	0.942(0.033)
Site 2	32	0.901(0.047)	0.828(0.068)	0.942(0.036)	0.935(0.056)	0.933(0.048)
Site 3	10	0.897(0.039)	0.824(0.053)	0.927(0.054)	0.929(0.031)	0.920(0.045)
Site 4-11[a]	19	0.900(0.034)	0.827(0.051)	0.936(0.045)	0.934(0.046)	0.931(0.046)
Core Lab	17	0.907(0.026)	0.841(0.043)	0.946(0.015)	0.919(0.018)	0.926(0.014)
Total	170	**0.907**(0.038)	**0.838**(0.057)	**0.943**(0.036)	**0.940**(0.039)	**0.936**(0.038)

[a]Due to small number of pullbacks from centres 4 to 11, pooled metrics for these centres were calculated

Table 2. Subgroup analysis per stent type in clinical centre 1

Stent type	Pullback Numbers	Segmentation Metrics		Detection Metrics (IOU 0.5)		
		Dice	Jaccard	Precision	Recall	F1-Score
Taxus Element	10	0.936	0.883	0.964	0.970	0.963
Omega	5	0.876	0.792	0.936	0.928	0.926
Resolute	10	0.912	0.846	0.950	0.956	0.948
Promus Premier	11	0.914	0.849	0.947	0.942	0.940
Pro Kinetic Energy	10	0.893	0.813	0.921	0.928	0.918
Xpedition	15	0.915	0.851	0.947	0.952	0.945
Nobori	8	0.921	0.858	0.953	0.969	0.958
Biomatrix	1	0.907	0.839	0.917	0.932	0.914
Xience prime	10	0.921	0.860	0.960	0.954	0.952
Ultimaster	3	0.936	0.882	0.974	0.966	0.968
Multilink	6	0.867	0.781	0.913	0.909	0.904
REBEL	3	0.925	0.864	0.948	0.945	0.941
Promus Element	1	0.933	0.879	0.939	0.969	0.949

Table 3 presents the subgroup analysis in challenging scenarios, showing similar performance to the general dataset. The most favourable scenarios were the jailing struts over side-branch and the overlapping stents, with Dice coefficients 0.917 and 0.922, Jaccard coefficients 0.853 and 0.863, precision 0.949 and 0.956, recall 0.953 and 0.964 and F1-score 0.947 and 0.955, respectively. Slightly poorer performance was observed in stent thrombosis and severe strut malapposition, whilst the worst performance was observed in cases with residual blood in the lumen (Dice 0.895, Jaccard 0.821, precision 0.931, recall 0.937 and F1-score 0.927). Several examples of strut segmentation in these challenging scenarios are displayed in Fig. 4(C) to Fig. 4(G), showing pretty accurate segmentation notwithstanding the handicap.

5.2. Ablation experiments

Table 4 shows the performance of the different segmentation methods applied in the testing set. All deep-convolution-based models outperformed the feature-based algorithm in detection metrics [26], showing smaller standard deviations. This was also observed in the 153 post-PCI pullbacks, in which the feature-based algorithm performed sensibly worse than any deep-convolutional model (precision 0.878, recall 0.841, F1-score 0.848). However, in the 17 pullbacks at follow-up,

Biomedical Optics EXPRESS

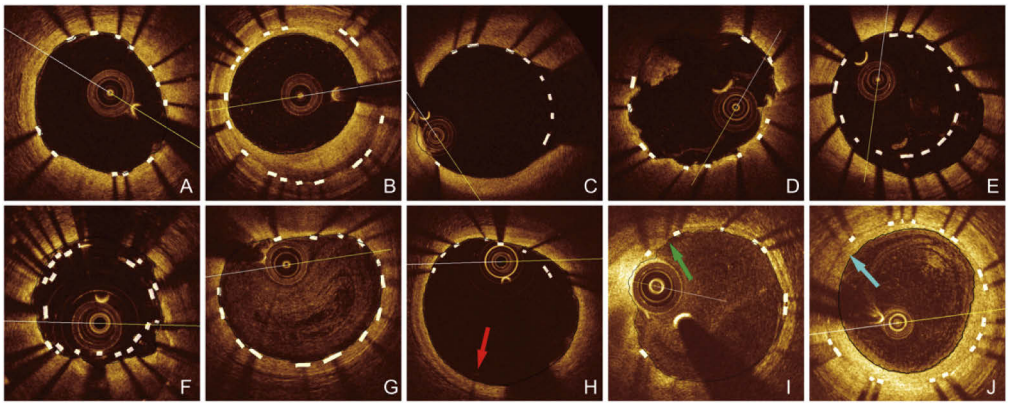

Fig. 4. Strut segmentation in a regular case post-PCI (A) and at follow-up (B). (C-G) Typical images of strut segmentation in five challenging scenarios: stent over side-branch (C), stent thrombosis (D), severe malapposition (E), overlapping stents (F) and residual blood due to incomplete flushing (G). (H-J) Strut detection failures in three cases: coloured arrows pointing to the missing struts.

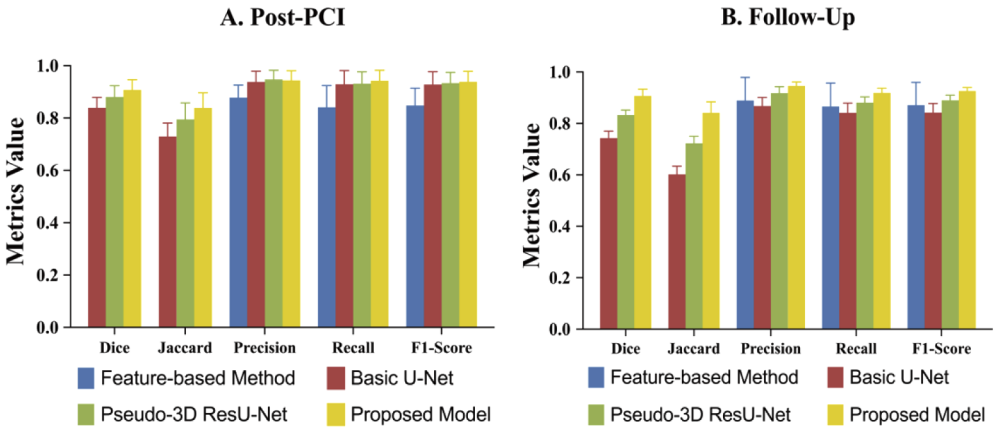

Fig. 5. (A) Performance of the different methods in post-PCI cases. All deep-convolutional-based models outperformed the feature-based algorithm in detecting struts [26], with the best results obtained by the hereby proposed model, particularly in segmentation. (B) Performance of the different methods in follow-up cases. The hereby proposed model outperformed all others methods, especially in segmentation again.

Table 3. Subgroup analysis in challenging scenarios

Challenging Scenarios	Pullback Numbers	Segmentation Metrics		Detection Metrics (IOU 0.5)		
		Dice	Jaccard	Precision	Recall	F1-Score
stent over side-branch	74	0.917(0.029)	0.853(0.044)	0.950(0.029)	0.953(0.023)	0.947(0.027)
stent thrombosis	88	0.905(0.040)	0.835(0.059)	0.943(0.032)	0.941(0.038)	0.938(0.036)
severe stent malapposition	28	0.903(0.046)	0.833(0.065)	0.944(0.035)	0.937(0.047)	0.936(0.043)
overlapping stents	12	0.922(0.015)	0.863(0.021)	0.956(0.018)	0.964(0.015)	0.955(0.020)
residual blood	51	0.895(0.036)	0.821(0.053)	0.931(0.037)	0.937(0.038)	0.927(0.038)

Biomedical Optics EXPRESS

the basic U-Net showed substantially poor performance (Fig. 5(A)), even lower than the feature-based algorithm. The hereby proposed model showed the best performance of all methods for follow-up pullbacks (precision 0.946, recall 0.919 and F1-score 0.926) (Fig. 5(B)).

Table 4. Performance of the different segmentation methods used in the testing set

Stent State	Methods	Segmentation Metrics		Detection Metrics (IOU 0.5)		
		Dice	Jaccard	Precision	Recall	F1-Score
Post-PCI	Feature-based	N/A	N/A	0.878(0.048)	0.841(0.083)	0.848(0.066)
	Basic U-Net	0.839(0.040)	0.729(0.052)	0.938(0.041)	0.929(0.052)	0.928(0.049)
	Pseudo-3D ResU-Net	0.880(0.043)	0.794(0.063)	**0.947**(0.035)	0.931(0.045)	0.933(0.041)
	Proposed	**0.907**(0.039)	**0.838**(0.058)	0.943(0.037)	**0.942**(0.040)	**0.938**(0.040)
Follow-Up	Feature-based	N/A	N/A	0.889(0.090)	0.866(0.091)	0.871(0.089)
	Basic U-Net	0.743(0.027)	0.602(0.032)	0.868(0.033)	0.841(0.038)	0.842(0.036)
	Pseudo-3D ResU-Net	0.833(0.019)	0.723(0.027)	0.918(0.025)	0.881(0.022)	0.890(0.020)
	Proposed	**0.907**(0.026)	**0.841**(0.043)	**0.946**(0.015)	**0.919**(0.018)	**0.926**(0.014)
Total	Feature-based	N/A	N/A	0.879(0.054)	0.843(0.084)	0.850(0.069)
	Based U-Net	0.829(0.048)	0.717(0.063)	0.931(0.046)	0.920(0.057)	0.920(0.054)
	Pseudo-3D ResU-Net	0.875(0.044)	0.787(0.064)	**0.944**(0.035)	0.926(0.046)	0.929(0.041)
	Proposed	**0.907**(0.038)	**0.838**(0.057)	0.943(0.036)	**0.940**(0.039)	**0.936**(0.038)

Regarding the efficiency of the different components of the model, the pseudo-3D ResU-Net and the proposed model were consistently superior to the basic U-Net in the total dataset and in all subgroups, thus confirming the efficiency of the basic modules, the input of pseudo-3D images and the multiscale shortcut connection in the hereby proposed method. Nonetheless, for analysis immediately post-PCI there was no significant difference in detection metrics between the pseudo-3D ResU-Net and the proposed model. Still, segmentation metrics were clearly better in the proposed model than in the pseudo-3D ResU-Net in all subgroups. These results are reassuring of the efficiency of multi-scale shortcut connections in preserving better spatial information for final prediction maps. Finally, the hereby proposed model resulted in smaller standard deviations for strut detection.

5.3. Correlation and agreement between the model and QIvus in quantitative parameters

Our proposed model showed good correlation with QIvus semi-automatic measurements for MSA ($r = 0.95$; $p < 0.001$, Fig. 6(A)) and Bland-Altman analysis showed excellent agreement (95% limits of agreement -0.07 ± 1.15 mm^2 (Fig. 6(B)). Also for ASA our model showed excellent correlation ($r = 0.99$, $p < 0.001$, Fig. 6(C)) and agreement (95% limits of agreement -0.36 ± 0.42 mm^2) with semi-automatic measurements (Fig. 6(D)).

For parameters at strut-level, excellent correlation and agreement was observed for both malapposition distance ($r = 0.98$, $p < 0.001$, Fig. 6(E); 95% limits of agreement 0.02 ± 0.05mm, Fig. 6(F)) and coverage thickness ($r = 0.99$, $p < 0.001$, Fig. 6(G); 95% limits of agreement 0.00 ± 0.05 mm, Fig. 6(H)).

The prespecified subgroup analysis, pooled by clinical centres, showed no heterogeneity in the agreement, with I^2 statistic = 0.00 for all quantitative parameters analysed, suggesting negligible variability between centres/operators.

Biomedical Optics EXPRESS

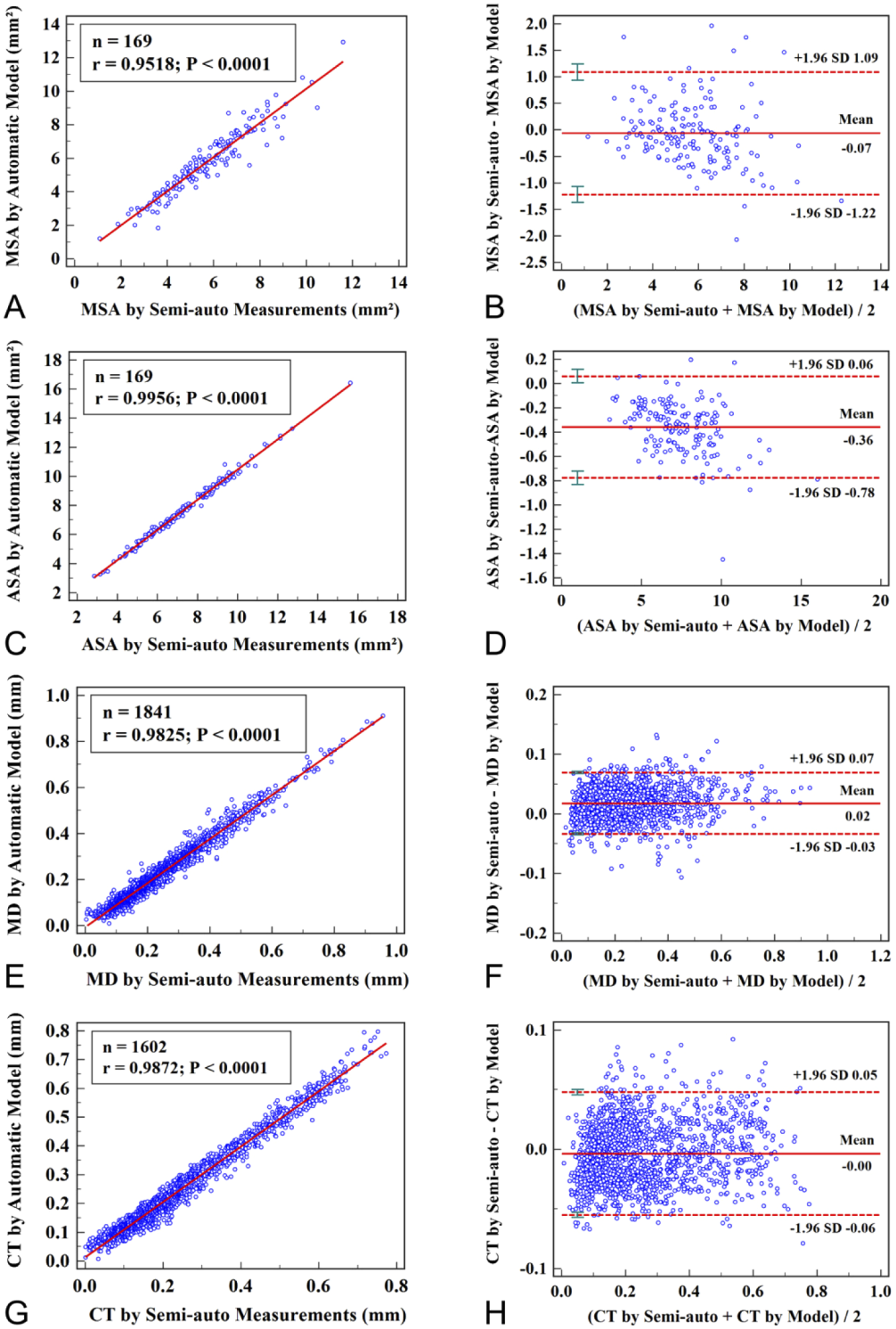

Fig. 6. Correlation and agreement between the model and semi-automatic measurements in minimal stent area (A-B),average stent area (C-D),malapposition distance (E-F), neointimal coverage thickness (G-H).

Biomedical Optics EXPRESS

5.4. 3D reconstruction of stent structure and time required for analysis

Figure 7(A) to Fig. 7(D) show reconstructions of four different types of stent platforms (A. Resolute, B. Taxus Element, C. Xpedition, D. Pro Kinetic Energy). The characteristic pattern of each type of stent can be easily recognised. This accurate 3D-rendering can be particularly useful in some scenarios, like bifurcations (Fig. 7(E), red arrow pointing out side branch) or severely malapposed struts (Fig. 7(F), green arrow pointing out malapposed struts and the gap between lumen and stent). The time required for stent detection and segmentation was 0.02 seconds per cross-sectional images. The average computational time required for 3D reconstruction of an IVOCT image pullback was 9.22 ± 2.82 seconds.

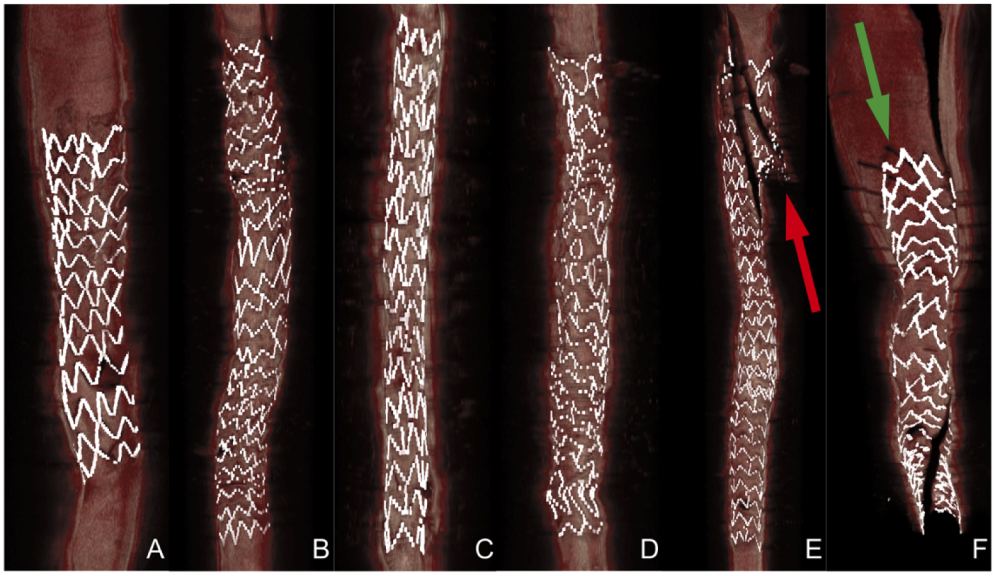

Fig. 7. (A-D) Four different types of stent platforms (A. Resolute, B. Taxus Element, C. Xpedition, D. Pro Kinetic Energy) rendered by our proposed model in 3D visualisation. 3D visualization of bifurcation stent (E, red arrow) and severely malapposed stent (F, green arrow).

6. Discussion

In this study we proposed a novel deep convolution model for fully automatic segmentation of stent struts from IVOCT images. The model was based on the FCN and U-shape architecture [31,32], but was designed to resolve specific challenges in stent strut segmentation. In particular we designed four structured modules with residual connection to automatically extract optimal features of stent strut for segmentation. Furthermore, we adopted pseudo-3D input and multi-scale shortcut connection for more accurate detection and finer strut contours. A large-scale training dataset were used to train and finetune our proposed model. The results of testing on the independent, extensive testing set showed both satisfactory segmentation and detection performance of the hereby proposed model, and the model outperformed other methods irrespective of the acquisition timing after stent implantation, especially exceeded the feature-based method by a big margin, suggesting all proposed steps were effective in improving the accuracy of the strut detection. In addition, excellent correlation and agreement between the model and semi-automatic measurements in quantitative parameters including MSA, ASA, malapposition distance, and coverage thickness were observed.

Biomedical Optics EXPRESS

Compared with previous studies on automatic detection of stent struts [23,24,26,27,29] that validated different algorithms on limited number of OCT image pullbacks with a few thousands struts, the current study collected a sizable sample of data, merging 153 post-PCI IVOCT pullbacks from the DOCTORS clinical trial [33] and 17 follow-up pullbacks from the core lab, resulting in 21,363 cross-sectional images with 205,513 struts for independent testing. This extensive and comprehensive dataset comprised most real clinical and anatomical scenarios of IVOCT imaging, including both high and low image quality, with various artefacts increasing the difficulty for a correct detection of struts. Very high segmentation accuracy was obtained with the hereby proposed model, with a Dice coefficient of 0.907 and a Jaccard coefficient of 0.839. The Dice coefficient value over 0.9 indicates a high overlapping extent of the segmented strut and the corresponding labels under a strict pixel-wise standard. Similar to previous studies in the field, precision, recall and F1-score were calculated strut-wise as detection metrics. Although the differences in sample size preclude a direct comparison with previous studies, a precision of 0.943, a recall of 0.940 and an F1-score of 0.936 compare quite well with the results from prior approaches and validate our deep convolution model. Moreover, the model is robust and reliable throughout different IVOCT systems and different operators, as suggested by the homogeneity in the subgroup analysis per clinical centre.

Likewise, the subgroup analysis per type of stent also suggested that the model was scarcely influenced by the type of stent, with consistent and homogeneous results in most stent types. Nonetheless, the performance in the Omega and Multilink stents was slightly but intriguingly lower than in other stent platforms. Oddly curious was the differential metrics between the Multilink and the Xience stents, because both stent types share exactly the same stent platform, only differing in the polymer and drug coating (present in Xience and absent in Multilink). The pullbacks corresponding to these two kinds of stents were purposely reanalysed in search for potential causes of poorer performance, founding some cases affected by artefacts of incomplete flushing or thrombotic clots. In the case of the Multilink stent, we observed a distinctly lower brightness of severely malapposed struts or struts jailing side-branches, precluding an accurate recognition of the strut. In these cases, our model might fail to detect the struts or they might be deleted in the post-processing, due to the low probability in the strut maps (Fig. 4(H) and Fig. 4(J)). Notwithstanding this consideration and in absence of dedicated studies to elucidate this finding, the subgroup analysis of challenging scenarios for accurate strut detection resulted in overall excellent performance of the model under the most adverse conditions, including severe malapposition, jailing struts over side-branches, presence of intraluminal thrombi or overlapping stents. The challenge in jailing, malapposed or overlapping struts stems from the dissimilar appearance of the strut or of its anatomical context as compared with regular struts. This problem can be easily overcome in a deep convolution model by providing enough typical images for model training. Nonetheless, in cases with residual blood due to incomplete flushing, the likelihood of detection failure is higher than in other scenarios and lower evaluation metrics were observed (Table 3). This finding is consistent with previous studies on the topic and can be easily explained because the residual blood affects the overall quality of the image and it is more difficult to overcome by providing enough training to the model. It simply depends on the extent at which quality is deteriorated by the artefact. Indeed, in cases with residual blood, weakening the signal but still allowing the recognition of structures, the model performed reasonably well, as in the example shown in Fig. 4(G). Conversely, when the presence of residual blood eclipsed the image formation, preventing adequate recognition of anatomical structures, the model failed (Fig. 4(I)), with no room for improvement and no chance for training, because the model cannot extract information which was not acquired and is hence missing. Our results suggest that image quality itself influences the performance of the model more than irregular features with acceptable image quality.

Biomedical Optics EXPRESS

All deep convolutional models outperformed the conventional feature-based algorithm that was implemented in the commercialized QIvus software in all segmentation and detection metrics. These results indicate that the features automatically extracted by deep convolutional models are more discriminating and comprehensive than conventional fixed features, therefore the detection results based on the former are more accurate and robust than those based on the latter. Still some nuances must be explained depending on whether IVOCT images were acquired immediately post-PCI or at follow-up. While in post-PCI images all deep-convolutional models performed better than the conventional feature-based algorithm, in follow-up cases the conventional feature-based algorithm was even better than the basic U-Net model (precision 0.889 vs. 0.868, recall 0.866 vs. 0.841 and F1-score 0.871 vs. 0.842, respectively). This might be explained because the vast majority of images in the training dataset were post-PCI and only a few were follow-up images. Thus, the basic model, requiring more data for adequate training, would have accused the effect of insufficient training data more than the Pseudo-3D ResU-Net model and our proposed model, both designed in four modules and with pseudo-3D input, conferring them greater net capacity and more information for strut prediction, thus being less dependent on training data feeding. Interestingly, the hereby proposed model showed no significant superiority to the Pseudo-3D ResU-Net model for detection performance in post-PCI cases, but it demonstrated a clear advantage in follow-up cases, which are more challenging and demanding than the analysis immediately post-PCI. Furthermore, our proposed model obtained the highest metrics in segmentation performance irrespective of the moment when the pullback was acquired, suggesting the efficiency of the multi-scale shortcut connection hereby adopted, enabling the generation of finer strut contours. In some cases where two struts were close to each other, the intermediate model tended to segment them as a long single strut, while our proposed model discriminated them as two different struts. This may be attributed to the multi-scale shortcut connection, rendering more accurate struts recognition and stent contour fitting.

Unlike previous studies calculating correlation and agreement on stent area at the cross-section level [24], our study analysed minimum stent area and average stent area per pullback. This approach is more sensitive to potential inaccuracies and it is therefore more demanding than the analysis at cross-section level to obtain high correlation and agreement. A single mistake in only one frame due to incorrect strut detection may result in a wrong MSA of the pullback, having bigger impact on correlation and agreement than in the analysis per cross-section, where the error gets diluted in a myriad of frame measurements. The hereby proposed model showed excellent correlation and agreement with manual measurements on MSA, with no significant bias in Bland-Altman plots. The correlation on MSA ($r = 0.95$) was slightly lower than on ASA ($r = 0.99$), probably reflecting the higher sensitivity of MSA to potential inaccuracies, because it is a parameter depending on a single summary cross-section, while ASA averages the stent areas throughout the pullback, thus buffering the impact of sporadic errors. The excellent agreement on stent area calculation is utmost relevant for clinical applications, because accurate calculation of stent areas is the basis for the assessment of stent expansion, defined as the relation between stent area and reference vessel area [11]. Underexpansion has been associated with stent restenosis [15–19] and with stent thrombosis [20], therefore its reliable assessment will be instrumental for OCT-guided PCI and for the appraisal of stent failure cases.

The Tversky loss we used to train our model can be regarded as a kind of generalized version of Dice loss that is often used in segmentation tasks. For segmentation of stent strut, the target is tiny compared to targets like lumen or plaque, and at the same time, there can be many targets in single slice. Thus, the balance between precision and recall can be hard to reach. The hyper-parameter Alpha the Tversky loss brings can control the precision and recall. At the angle of hyper-parameters fine-tune, Tversky loss promotes to reach to a precision-recall balance and optimized model parameters.

Biomedical Optics EXPRESS

Although the overall performance in the 17 follow-up cases was also excellent, follow-up images represented only a small proportion of the testing dataset. This uneven distribution might have affected the assessment of coverage thickness and the binary detection of covered/uncovered struts, because coverage is exclusively assessed at follow-up. As a consequence, the hereby reported results might not reflect the true performance of our model in detecting covered struts. Our results should be confirmed in future studies, specifically focused on follow-up datasets. It cannot be excluded that the performance would increase when increasing the number of follow-up data in the training or using more dedicated networks.

The poor performance of the model due to suboptimal quality of the images is often caused by unclear features of the strut or of the dark trailing shadows. In the cases of detection failure in the testing dataset, the model was more sensitive to problems with the strut than to problems with the trailing shadow. If the signal of the strut is simply too weak, even if the trailing shadow is totally missing, the model learns to detect correctly the struts. Nevertheless, in absence of the bright signal of the strut, the likelihood of detection failure of the strut increases substantially, irrespective of the presence of a typically casted shadow. This caveat of our model raises some concerns regarding the assessment of covered struts, in which the bright of the strut fades substantially as compared with the images post-PCI, until practically disappearing in some cases. The dark shadow is often the only landmark pointing out the presence of a covered strut and such cases can be disregarded by the current model (Fig. 4(J)). Following a deep convolutional architecture, our model extracts features automatically according to the label given, with no further human intervention or feedback. An additional supervision level on feature extraction, weighting the casted shadows properly, might improve the accuracy of the current model and should be considered in future upgrades and studies.

The current training process combined 7 adjacent slices in pseudo-3D input, as the context along the z-axis. The number of adjacent slices was arbitrarily chosen and might not be necessarily the optimal number or might be inappropriate for some types of stent. Actually, the adaptive selection of the optimal number of adjacent slices, depending on different scenarios, deserves to be considered in future algorithms, as it might improve the efficiency of the model. The guidewire casts often a shadow over a random sector of the cross-section, hiding the struts behind and rendering a dark gap in the 3D reconstruction of the stent structure. The virtual reconstruction of this invisible sector of the stent by extrapolation methods might be subject of further research in the future.

This analysis time is short enough to encourage the application of our model routinely in the cathlab, making IVOCT guidance of PCI feasible for real-time decision-making in realistic conditions. The combination of accurate morphoanatomical details and coronary physiology [44,45] offers the most comprehensive and precise collection of information currently available to guide the optimisation of coronary interventions. Accurate and prompt detection of stent struts enables the assessment of expansion and apposition, thus providing the cardiologist with unique information to minimise the risks of stent failure, while OCT-based FFR (OFR) estimates the functional impact of the intervention on coronary physiology [44,45]. The synthesis of all this information will be also instrumental for precision PCI and for the personalised treatment of cases of stent failure.

7. Conclusion

A fully-automatic, deep convolutional segmentation model for detection of struts in metallic stents was hereby designed. Contrary to conventional features-based methods, our model extracted optimal features automatically from a huge number of training data. The model was subsequently validated on a large-scale testing dataset stemming from multiple clinical centres, with excellent results. All proposed steps in the model proved efficiency to improve the accuracy of the strut detection: 1) deep-convolutional models outperformed conventional features-based

Biomedical Optics EXPRESS

algorithm, thus proving the superiority of comprehensive features automatically extracted on a deep-convolutional approach; 2) deep-convolutional models with pseudo-3D input outperformed the basic deep-convolutional model without it, thus confirming the efficiency of pseudo-3D input to aggregate consecutive information and improve the detection accuracy; 3) the hereby proposed model, with multi-scale shortcut connection, outperformed other deep-convolutional models without it, resulting in finer and more precise segmentation contours and confirming the advantage of multi-scale shortcut connection. The excellent agreement between the model and semi-automatic measurements in quantitative parameters, together with the short time required for analysis, suggests the feasibility of the method for routine stent assessment in the cathlab to guide clinical decision-making.

Funding

National Natural Science Foundation of China (Grant No. 81871460); Program of Shanghai Technology Research Leader; Shanghai Jiao Tong University (Grant No. ZH2018ZDB10).

Disclosures

ST: Pulse medical imaging technology. (I,F). The other authors declare no conflicts of interest related to this article.

References

1. C. J. Murray and A. D. Lopez, "Mortality by cause for eight regions of the world: Global Burden of Disease Study," Lancet **349**(9061), 1269–1276 (1997).
2. C. J. Murray and A. D. Lopez, "Alternative projections of mortality and disability by cause 1990–2020: Global Burden of Disease Study," Lancet **349**(9064), 1498–1504 (1997).
3. C. D. Mathers, T. Boerma, and D. Ma Fat, "Global and regional causes of death," Br. Med. Bull. **92**(1), 7–32 (2009).
4. H. C. Stary, D. H. Blankenhorn, A. B. Chandler, S. Glagov, W. Insull, M. Richardson, M. E. Rosenfeld, S. A. Schaffer, C. J. Schwartz, and W. D. Wagner, "A definition of the intima of human arteries and of its atherosclerosis-prone regions. A report from the Committee on Vascular Lesions of the Council on Arteriosclerosis, American Heart Association," Circulation **85**(1), 391–405 (1992).
5. H. C. Stary, A. B. Chandler, S. Glagov, J. R. Guyton, W. Insull, M. E. Rosenfeld, S. A. Schaffer, C. J. Schwartz, W. D. Wagner, and R. W. Wissler, "A definition of initial, fatty streak, and intermediate lesions of atherosclerosis. A report from the Committee on Vascular Lesions of the Council on Arteriosclerosis, American Heart Association," Circulation **89**(5), 2462–2478 (1994).
6. H. C. Stary, A. B. Chandler, R. E. Dinsmore, V. Fuster, S. Glagov, W. Insull, M. E. Rosenfeld, C. J. Schwartz, W. D. Wagner, and R. W. Wissler, "A Definition of Advanced Types of Atherosclerotic Lesions and a Histological Classification of Atherosclerosis A Report From the Committee on Vascular Lesions of the Council on Arteriosclerosis, American Heart Association," Circulation **92**(5), 1355–1374 (1995).
7. R. Virmani, F. D. Kolodgie, A. P. Burke, A. Farb, and S. M. Schwartz, "Lessons From Sudden Coronary Death: A Comprehensive Morphological Classification Scheme for Atherosclerotic Lesions," Arterioscler., Thromb., Vasc. Biol. **20**(5), 1262–1275 (2000).
8. J. Schaar, J. Muller, E. Falk, R. Virmani, V. Fuster, P. Serruys, A. Colombo, C. Stefanadis, S. Casscells, P. Moreno, A. Maseri, and A. van der Steen, "Terminology for high-risk and vulnerable coronary artery plaques Report of a Meeting on the Vulnerable Plaque, June 17 and 18, 2003, Santorini, Greece," Eur. Heart J. **25**(12), 1077–1082 (2004).
9. F. Neumann, M. Sousauva, A. Ahlsson, F. Alfonso, A. P. Banning, U. Benedetto, R. A. Byrne, J. Collet, V. Falk, and S. J. Head, "2018 ESC/EACTS Guidelines on myocardial revascularization," Eur. Heart J. **40**(2), 87–165 (2019).
10. J. L. Gutiérrez-Chico, J. J. Wykrzykowska, E. Nuesch, R. J. Van Geuns, K. T. Koch, J. J. Koolen, C. D. Mario, S. Windecker, G. Van Es, and P. Gobbens, "Vascular Tissue Reaction to Acute Malapposition in Human Coronary Arteries Sequential Assessment With Optical Coherence Tomography," Circ. Cardiovasc. Interv. **5**(1), 20–29 (2012).
11. J. L. Gutiérrez-Chico, E. Alegría-Barrero, R. Teijeiro-Mestre, P. Chan, H. Tsujioka, R. De Silva, N. Viceconte, A. C. Lindsay, T. Patterson, and N. Foin, "Optical coherence tomography: from research to practice," Europ. Heart J. - Cardiovasc. Imaging **13**(5), 370–384 (2012).
12. J. L. Gutiérrez-Chico, E. Regar, E. Nuesch, T. Okamura, J. J. Wykrzykowska, C. D. Mario, S. Windecker, G. Van Es, P. Gobbens, and P. Juni, "Delayed Coverage in Malapposed and Side-Branch Struts With Respect to Well-Apposed Struts in Drug-Eluting Stents In Vivo Assessment With Optical Coherence Tomography," Circulation **124**(5), 612–623 (2011).
13. N. Foin, J. L. Gutiérrez-Chico, S. Nakatani, R. Torii, C. V. Bourantas, S. Sen, S. Nijjer, R. Petraco, C. A. Kousera, and M. Ghione, "Incomplete Stent Apposition Causes High Shear Flow Disturbances and Delay in Neointimal

Biomedical Optics EXPRESS

Coverage as a Function of Strut to Wall Detachment Distance Implications for the Management of Incomplete Stent Apposition," Circ. Cardiovasc. Interv. **7**(2), 180–189 (2014).
14. A. V. Finn, M. Joner, G. Nakazawa, F. D. Kolodgie, J. B. Newell, M. C. John, H. K. Gold, and R. Virmani, "Pathological Correlates of Late Drug-Eluting Stent Thrombosis Strut Coverage as a Marker of Endothelialization," Circulation **115**(18), 2435–2441 (2007).
15. S. Sonoda, Y. Morino, J. Ako, M. Terashima, A. Hassan, H. N. Bonneau, M. B. Leon, J. W. Moses, P. G. Yock, and Y. Honda, "Impact of Final Stent Dimensions on Long-Term Results Following Sirolimus-Eluting Stent Implantation Serial Intravascular Ultrasound Analysis From the SIRIUS Trial," J. Am. Coll. Cardiol. **43**(11), 1959–1963 (2004).
16. S. Kang, G. S. Mintz, D. Park, S. Lee, Y. Kim, C. W. Lee, K. Han, J. Kim, S. Park, and S. Park, "Mechanisms of In-Stent Restenosis After Drug-Eluting Stent Implantation Intravascular Ultrasound Analysis," Circ. Cardiovasc. Interv. **4**(1), 9–14 (2011).
17. S. Kang, J. Ahn, H. Song, W. Kim, J. Lee, D. Park, S. Yun, S. Lee, Y. Kim, and C. W. Lee, "Comprehensive Intravascular Ultrasound Assessment of Stent Area and Its Impact on Restenosis and Adverse Cardiac Events in 403 Patients With Unprotected Left Main Disease," Circ. Cardiovasc. Interv. **4**(6), 562–569 (2011).
18. H. Song, S. Kang, J. Ahn, W. Kim, J. Lee, D. Park, S. Lee, Y. Kim, C. W. Lee, and S. Park, "Intravascular ultrasound assessment of optimal stent area to prevent in-stent restenosis after zotarolimus-, everolimus-, and sirolimus-eluting stent implantation," Cathet. Cardiovasc. Intervent. **83**(6), 873–878 (2014).
19. L. Song, G. S. Mintz, D. Yin, M. H. Yamamoto, C. Y. Chin, M. Matsumura, A. J. Kirtane, M. Parikh, J. Moses, and Z. A. Ali, "Characteristics of early versus late in-stent restenosis in second-generation drug-eluting stents: an optical coherence tomography study," Eurointervention **13**(3), 294–302 (2017).
20. K. Fujii, S. G. Carlier, G. S. Mintz, Y. Yang, I. Moussa, G. Weisz, G. Dangas, R. Mehran, A. J. Lansky, and E. M. Kreps, "Stent underexpansion and residual reference segment stenosis are related to stent thrombosis after sirolimus-eluting stent implantation: An intravascular ultrasound study," J. Am. Coll. Cardiol. **45**(7), 995–998 (2005).
21. 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: Cardiovasc. Interven. **2**(11), 1035–1046 (2009).
22. S. Brugaletta, H. M. García-García, J. Gómez-Lara, M. D. Radu, R. Pawar, J. Khachabi, N. Bruining, M. Sabate, and P. W. Serruys, "Reproducibility of qualitative assessment of stent struts coverage by optical coherence tomography," Int. J. Cardiovasc. Imaging **29**(1), 5–11 (2013).
23. G. J. Ughi, T. Adriaenssens, K. Onsea, P. Kayaert, C. Dubois, P. Sinnaeve, M. Coosemans, W. Desmet, and J. Dhooge, "Automatic segmentation of in-vivo intra-coronary optical coherence tomography images to assess stent strut apposition and coverage," Int. J. Cardiovasc. Imaging **28**(2), 229–241 (2012).
24. H. Lu, M. Gargesha, Z. Wang, D. Chamie, G. F. Attizzani, T. Kanaya, S. Ray, M. A. Costa, A. M. Rollins, and H. G. Bezerra, "Automatic stent detection in intravascular OCT images using bagged decision trees," Biomed. Opt. Express **3**(11), 2809–2824 (2012).
25. S. Tsantis, G. C. Kagadis, K. Katsanos, D. Karnabatidis, G. C. Bourantas, and G. Nikiforidis, "Automatic vessel lumen segmentation and stent strut detection in intravascular optical coherence tomography," Med. Phys. **39**(1), 503–513 (2011).
26. A. Wang, J. Eggermont, N. Dekker, H. M. Garciagarcia, R. Pawar, J. H. C. Reiber, and J. Dijkstra, "Automatic stent strut detection in intravascular optical coherence tomographic pullback runs," Int. J. Cardiovasc. Imaging **29**(1), 29–38 (2013).
27. Z. Wang, M. W. Jenkins, G. C. Linderman, H. G. Bezerra, Y. Fujino, M. A. Costa, D. L. Wilson, and A. M. Rollins, "3-D Stent Detection in Intravascular OCT Using a Bayesian Network and Graph Search," IEEE Trans. Med. Imaging **34**(7), 1549–1561 (2015).
28. H. Lu, J. Lee, S. Ray, K. Tanaka, H. G. Bezerra, A. M. Rollins, and D. L. Wilson, "Automated stent coverage analysis in intravascular OCT (IVOCT) image volumes using a support vector machine and mesh growing," Biomed. Opt. Express **10**(6), 2809–2828 (2019).
29. H. S. Nam, C. S. Kim, J. J. Lee, J. W. Song, J. W. Kim, and H. Yoo, "Automated detection of vessel lumen and stent struts in intravascular optical coherence tomography to evaluate stent apposition and neointimal coverage," Med. Phys. **43**(4), 1662–1675 (2016).
30. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *neural information processing systems*, 2012), 1097–1105.
31. J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *computer vision and pattern recognition*, 2015), 3431–3440.
32. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *medical image computing and computer assisted intervention*, 2015), 234–241.
33. N. Meneveau, G. Souteyrand, P. Motreff, C. Caussin, N. Amabile, P. Ohlmann, O. Morel, Y. Lefrançois, V. Descotes-Genon, J. Silvain, N. Braik, R. Chopard, M. Chatot, F. Ecarnot, H. Tauzin, E. V. Belle, L. Belle, and F. Schiele, "Optical Coherence Tomography to Optimize Results of Percutaneous Coronary Intervention in Patients with Non–ST-Elevation Acute Coronary Syndrome," Circulation **134**(13), 906–917 (2016).
34. P. A. Yushkevich, J. Piven, H. C. Hazlett, R. G. Smith, S. Ho, J. C. Gee, and G. Gerig, "User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability," NeuroImage **31**(3), 1116–1128 (2006).

Biomedical Optics EXPRESS

35. Y. Lecun, B. E. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," Neural Comput. **1**(4), 541–551 (1989).
36. S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," arXiv: Learning (2015).
37. A. L. Maas, A. Y. Hannun, and A. Y. Ng, "Rectifier nonlinearities improve neural network acoustic models," in *Proc. icml*, 2013), 3.
38. N. Srivastava, G. E. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," J. Mach. Learn. Res. **15**, 1929–1958 (2014).
39. K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *computer vision and pattern recognition*, 2016), 770–778.
40. G. Lin, A. Milan, C. Shen, and I. Reid, "Refinenet: Multi-path refinement networks for high-resolution semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017), 1925–1934.
41. F. Yu, D. Wang, and T. Darrell, "Deep Layer Aggregation," arXiv: Computer Vision and Pattern Recognition (2017).
42. S. S. M. Salehi, D. Erdogmus, and A. Gholipour, "Tversky loss function for image segmentation using 3D fully convolutional deep networks," arXiv: Computer Vision and Pattern Recognition (2017).
43. L. Raber, G. S. Mintz, K. C. Koskinas, T. W. Johnson, N. R. Holm, Y. Onuma, M. D. Radu, M. Joner, B. Yu, and H. Jia, "Clinical use of intracoronary imaging. Part 1: guidance and optimization of coronary interventions. An expert consensus document of the European Association of Percutaneous Cardiovascular Interventions," Eur. Heart J. **39**(35), 3281–3300 (2018).
44. W. Yu, J. Huang, D. Jia, S. Chen, O. Raffel, D. Ding, F. Tian, J. Kan, S. Zhang, and F. Yan, "Diagnostic accuracy of intracoronary optical coherence tomography-derived fractional flow reserve for assessment of coronary stenosis severity," Eurointervention **15**(2), 189–197 (2019).
45. F. Tian, W. Yu, J. Huang, T. Zhang, Y. Chen, and S. Tu, "First Presentation of Integration of Intravascular Optical Coherence Tomography and Computational Fractional Flow Reserve," Int. J. Cardiovasc. Imaging **35**(4), 601–602 (2019).