ORIGINAL ARTICLE

Automated lumen segmentation using multi-frame convolutional neural networks in intravascular ultrasound datasets

Paulo G.P. Ziemer[1,2], Carlos A. Bulant[2,3], José I. Orlando[3], Gonzalo D. Maso Talou[4], Luis A. Mansilla Álvarez [ORCID] **[1,2], Cristiano Guedes Bezerra[5], Pedro A. Lemos[2,5,6], Héctor M. García-García[7,8]*, Pablo J. Blanco[1,2]***

[1]National Laboratory for Scientific Computing, Av. Getúlio Vargas 333, 25651-075 Petrópolis, Brazil; [2]National Institute of Science and Technology in Medicine Assisted by Scientific Computing, Petrópolis, Brazil; [3]National Scientific and Technical Research Council, CONICET and Pladema Institute, National University of the Center of the Buenos Aires Province, Tandil, Argentina; [4]Auckland Bioengineering Institute, The University of Auckland, Auckland, New Zealand; [5]Department of Interventional Cardiology, Heart Institute (InCor) and the University of São Paulo Medical School, São Paulo, Brazil; [6]Hospital Israelita Albert Einstein, São Paulo, Brazil; [7]Interventional Cardiology Department, MedStar Washington Hospital Center, Washington, DC, USA and ; and [8]Georgetown University School of Medicine, Washington, DC, USA

Aims	Assessment of minimum lumen areas in intravascular ultrasound (IVUS) pullbacks is time-consuming and demands adequately trained personnel. In this work, we introduce a novel and fully automated pipeline to segment the lumen boundary in IVUS datasets.
Methods and results	First, an automated gating is applied to select end-diastolic frames and bypass saw-tooth artefacts. Second, within a machine learning (ML) environment, we automatically segment the lumen boundary using a multi-frame (MF) convolutional neural network (MFCNN). Finally, we use the theory of Gaussian processes (GPs) to regress the final lumen boundary. The dataset consisted of 85 IVUS pullbacks (52 patients). The dataset was partitioned at the pullback-level using 73 pullbacks for training (20 586 frames), 6 pullbacks for validation (1692 frames), and 6 for testing (1692 frames). The degree of overlapping, between the ground truth and ML contours, median (interquartile range, IQR) systematically increased from 0.896 (0.874–0.933) for MF1 to 0.925 (0.911–0.948) for MF11. The median (IQR) of the distance error was also reduced from 3.83 (2.94–4.98)% for MF1 to 3.02 (2.25–3.95)% for MF11-GP. The corresponding median (IQR) in the lumen area error remained between 5.49 (2.50–10.50)% for MF1 and 5.12 (2.15–9.00)% for MF11-GP. The dispersion in the relative distance and area errors consistently decreased as we increased the number of frames, and also when the GP regressor was coupled to the MFCNN output.
Conclusion	These results demonstrate that the proposed ML approach is suitable to effectively segment the lumen boundary in IVUS scans, reducing the burden of costly and time-consuming manual delineation.
Keywords	IVUS • Segmentation • Gating • Neural networks • Gaussian process

Introduction

Intravascular ultrasound (IVUS) is the gold standard imaging modality for the assessment of coronary artery disease.[1] Intravascular ultrasound provides a highly detailed view of the inner coronary structure, such as lumen, external elastic membrane (EEM), and plaque.

One of the most arduous tasks when analysing IVUS datasets is the delineation (segmentation) of the lumen boundary and EEM, for

which an expert has to manually outline them. This process is performed either one frame at a time using transversal contouring or at the dataset level by tracing a small number of longitudinal cutting planes. Given the large intra- and inter-observer variability for drawing tasks and its time-consuming nature, the translation of this procedure to online diagnosis pipelines is hindered and unreliable. Thus, a systematic, reproducible and automatic computational technique would enable a reliable translation for clinical usage.

Throughout the longitudinal view, a typical saw-tooth artefact is usually observed. This structure hinders the longitudinal analysis of the IVUS images. Therefore, a gating process is mandatory to ensure a better depiction of the vessel's inner structures. However, electrocardiogram (ECG)-synchronized images are not always available. Thus, a consistent and accurate alternative is required to select the end-diastolic frames automatically.[2]

In this article, we propose an automated workflow to segment lumen boundaries in IVUS datasets using a machine learning (ML) approach. We focus on lumen segmentation due to its clinical relevance in the definition of minimum lumen area (MLA) and percentage area of stenosis. Both features are used for decision making, e.g. to decide if a given lesion must be treated. The proposed pipeline includes an ECG-free gating procedure, lumen boundary segmentation using a multi-frame convolutional neural network (MFCNN) and further refinement of the lumen contour through a Gaussian Process (GP) regression stage.

Methods

This section describes the proposed pipeline for the automated processing of IVUS datasets (see *Figure 1*). Stages of the methodology are detailed in the Supplementary material online, *Files*.

IVUS dataset and ground truth generation

All transversal images have the same resolution of 512×512 pixels (256×256 in the intrinsic polar space), and all pullbacks were performed with the same speed (0.5 mm/s). *Figure 1A* depicts a longitudinal slice of an IVUS pullback from our dataset. Ground truth (GT) lumen contour was manually delineated using four longitudinal slices (spaced by 45°) and smoothed polylines (cubic splines) to define the lumen contour in the transversal IVUS frames.

Automatic image processing and segmentation

The automatic ML segmentation consists of three stages, namely: gating, segmentation, and filtering. The automated ECG-free gating is performed as reported in ref.,[3] by locating the minimum of a motion signal constructed by a combination of inter-frame inverse correlation and intra-frame intensity gradients. *Figure 1B* illustrates the motion signal and *Figure 1C* depicts the outcome of an automatic end-diastolic gating. The segmentation is performed by a U-Net convolutional neural network[4] with image frames ordered in layers [stacked multi-frame (MF)], which allows us to feed the network not only with a single IVUS frame but also with its neighbouring frames. We term our network MFCNN (see *Figure 1D*). As usual, the stack of frames is given in a system of coordinates in which each point is determined by a distance from a centre point and an angle from a reference direction (polar coordinates), and the output (also in polar coordinates) is a binary segmentation, denoted by MFX, being X the number of frames employed in the stack. We tested different

MFCNN scenarios: a 1-frame approach (no neighbours) and MF strategies using 3- to 11-frames (all odd numbers in the range so neighbours are placed symmetrically to the frame under analysis). Finally, a GP regressor[5] is connected to the MFCNN output to convert the binary image into a lumen contour, filtering intrinsic noise from the high-dimensional image domain and adding periodic consistency. We denote the combined method simply as MFX-GP. The relationship between input and output can be observed in *Figure 1E*.

The outcome of this process is the segmentation of the lumen boundary, as observed in *Figure 1F and G*. An in-house software was developed to implement the complete workflow illustrated in *Figure 1*.

Comparison metrics

The predictive performance of the MFCNN is assessed with the intersection-over-union metric[6] of the lumen area (see *Figure 2A and B*). Here, we term this metric as degree of overlapping, since it measures the degree of coincident pixels present in the automatic ML segmentation and the GT segmentation mask (see *Figure 2A and B*).

With the GP regression algorithm, we move from image to contour domain (see *Figure 1E*). The performance in the contour domain is assessed through the mean and maximum distances between points belonging to the MFCNN-GP and GT contours (see *Figure 2A and B* for the definition of the metrics). In turn, the accuracy of the lumen area obtained by MFCNN-GP is characterized through relative area error and analyses of agreement with GT lumen area (see *Figure 3A and B*). Frames containing side branches or where the lumen is partially out of the field of view were excluded.

The reported statistical analysis includes mean, standard deviation, median and interquartile ranges (IQRs) for the different metrics. In all cases, the lumen area refers to the lumen area of individual frames, noting that the outcome of the ML algorithm is the segmentation of the central frame within the stack of IVUS frames given as input.

Results

The dataset consisted of 85 IVUS pullbacks (52 patients). The dataset was partitioned at the pullback-level using 73 pullbacks for training (20586 frames), 6 pullbacks for validation (1692 frames), and 6 for testing (1692 frames). The study population consisted of mostly male patients 44 (84.6%) and the average age was 63.7 ± 10.2. The majority of imaged vessels were left anterior descending 48 (56.5%), followed by left circumflex artery 17 (20%), right coronary artery 14 (16.5%), and other secondary branches such as diagonal, marginal or ramus intermedius 6 (7%).

Table 1 reports the statistical analysis of the degree of overlapping, relative distance error, maximum relative error, and relative area error for different ML models (MFX, X = 1, 3, 5, 7, 9, 11, 11-GP). The degree of overlapping median (IQR) systematically increased from 0.896 (0.874–0.933) for MF1 to 0.925 (0.911–0.948) for MF11. The median (IQR) of the relative distance error was also reduced from 3.83 (2.94–4.98)% for MF1, to 3.02 (2.25–3.95)% for MF11-GP. The median (IQR) of the maximum distance error was also reduced from 20.61 (10.53–22.75)% for MF1 to 16.77 (7.25–16.03)% for MF11-GP The corresponding median (IQR) in the lumen area error remained between 5.49 (2.50–10.50)% for MF1 and 5.12 (2.15–9.00)% for MF11-GP. The dispersion in the relative distance and area errors consistently decreased as we increased the number of frames, and also when the GP regressor was coupled to the MFCNN output.
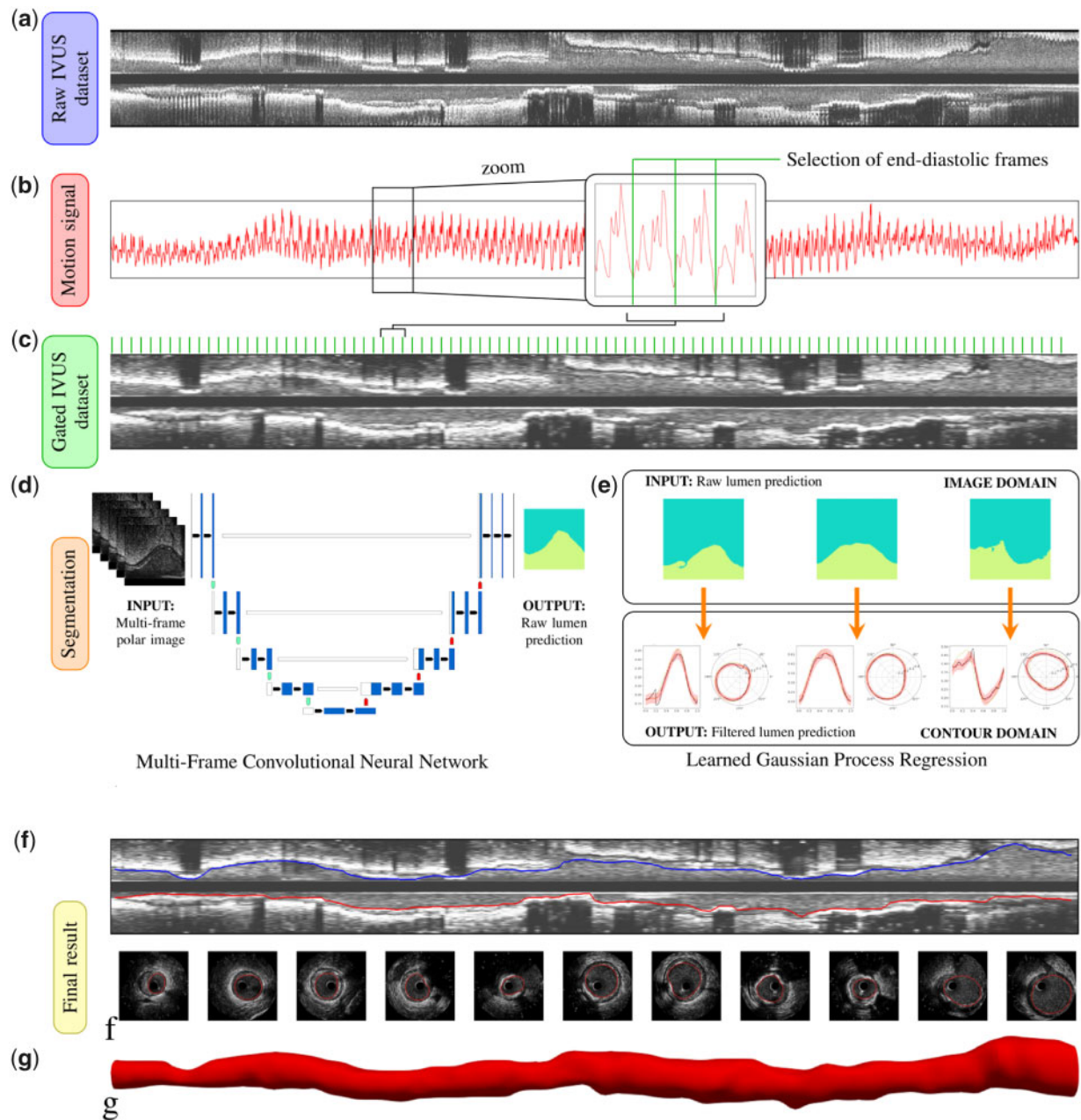
Figure 1 Workflow for the automatic processing of the lumen boundary in intravascular ultrasound (IVUS) datasets. (A) Raw IVUS dataset (longitudinal view); (B) combined signal to describe the motion contained in the IVUS images and select the end-diastolic frames; (C) gated volume constructed from the selected end-diastolic frames; (D) multi-frame convolutional neural network (MFCNN) architecture; (E) mapping from image to contour domain through Gaussian Process (GP) regression; (F) transversal and longitudinal lumen outline on top of the gated dataset; (G) final lumen boundary represented in three-dimensional (3D) space as a surface.

Figure 2 shows the violin plots with the distribution of the degree of overlapping and the relative distance error metrics (see *Figure 2A and B*), for the whole sample (*Figure 2D–F*), and at per-pullback level (*Figure 2G*). These results are shown for the different ML models (see *Figure 2C*). The IQRs, median, mean and standard deviation values are reported in *Table 1*. It is worth noting that increasing the number of frames and the further addition of the GP stage widens the lower

part of the error distribution, while shrinks the upper part, improving the results. By increasing the number of frames, the difference is statistically significant ($P < 0.05$), as shown by the horizontal bars.

Figure 3 illustrates the relation between ML lumen area and GT lumen area for the specific models MF1, MF11, and MF11-GP. The metric employed in the assessment is described in *Figure 3A and B*. *Figure 3C* displays the relative lumen area error, which is consistently
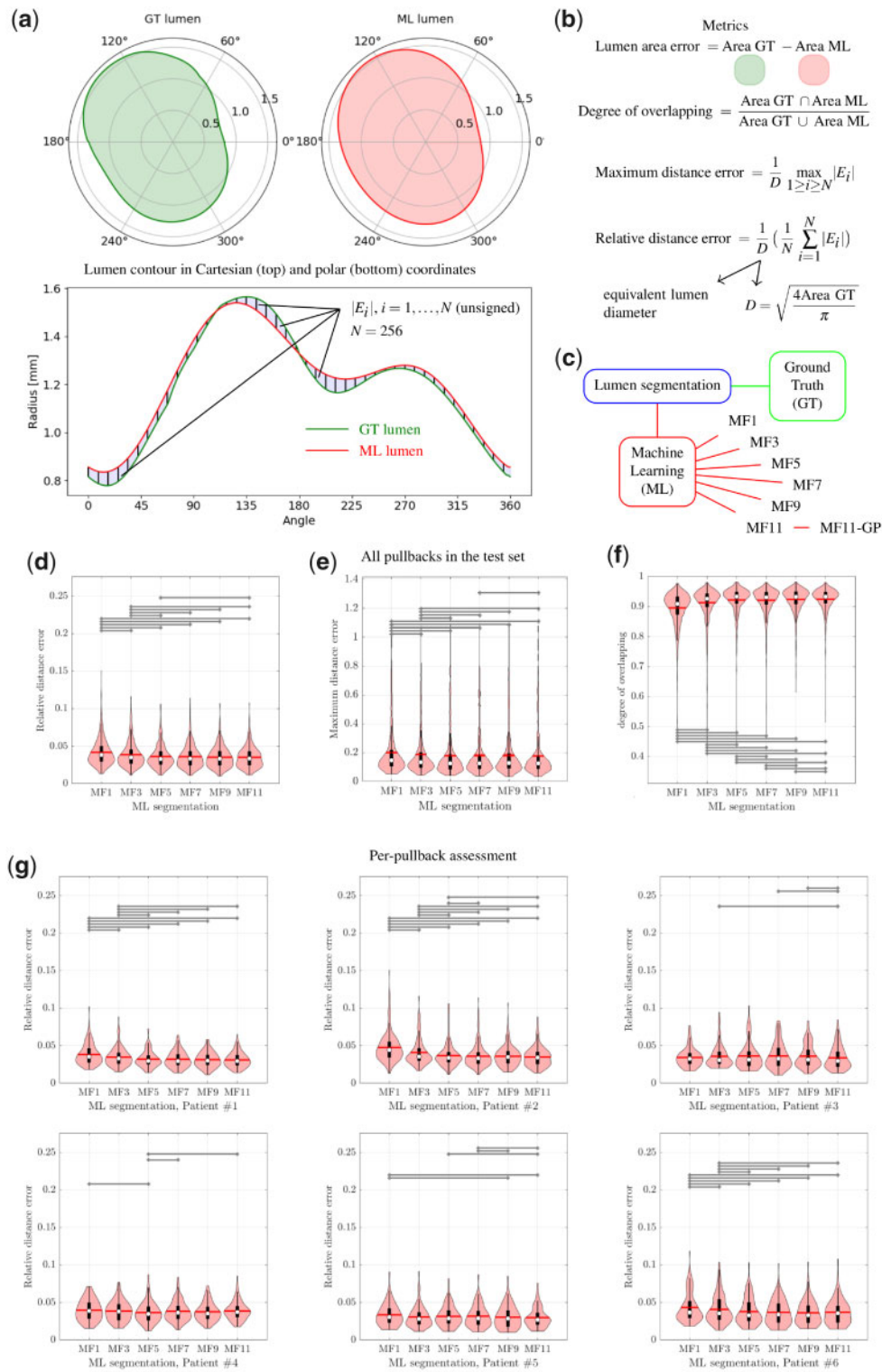
Figure 2 Segmentation error for the different machine learning (ML) models. (A) Lumen contour in Cartesian (top) and polar (bottom) coordinates illustrating the mismatch between ground truth (GT) and ML segmentations; (B) definition of relevant metrics (relative distance error, maximum distance error, and degree of overlapping) to assess the performance of the ML segmentation; (C) notation for the GT and ML models; (D) violin plot of the relative distance error for the entire test set; (E) violin plot of the maximum distance error for the entire test set; (F) violin plot of the degree of overlapping for the entire test set; (G) violin plots showing the relative distance error metric at a per-pullback level. In all plots, horizontal bars indicate scenarios with statistically different results.
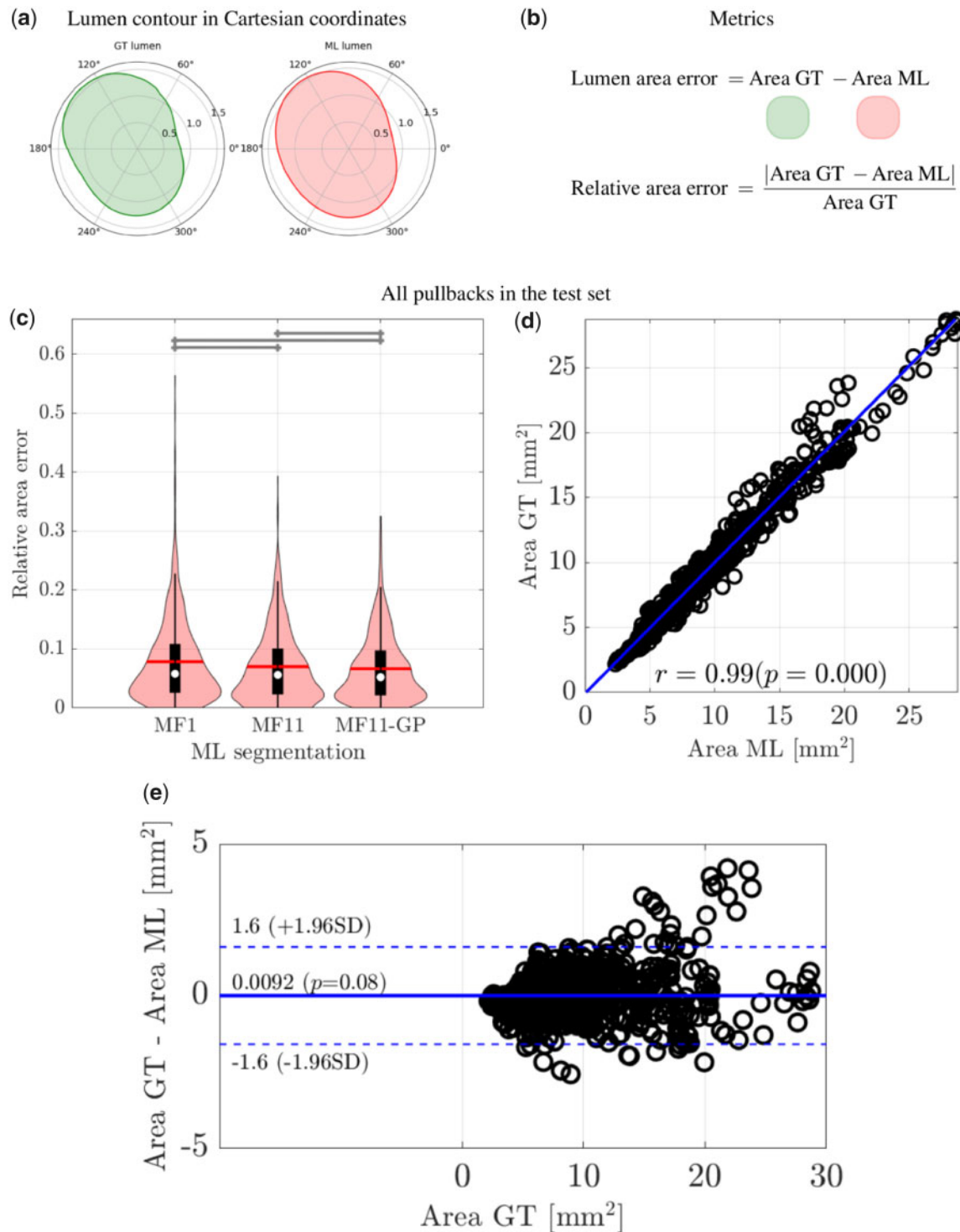
Figure 3 Lumen area error for the different machine learning (ML) models. (A) Lumen contour in Cartesian coordinates illustrating the definition of ground truth (GT) lumen area and ML lumen area; (B) definition of the metrics to assess the error in the lumen area; (C) violin plot of the relative lumen area error for different ML models (multi-frame—MF1, MF11, and MF11-Gaussian Process—GP) and the entire test dataset; (D) correlation plot between the ML lumen area (MF11-GP) and the GT lumen area; (E) Bland–Altman plot for the signed lumen area error as a function of the GT lumen area.

Table 1 Statistics for the different metrics considered for the different machine learning (ML) models (multi-frame—MF1, MF3, MF5, MF7, MF9, MF11, MF11-Gaussian Process—GP)

Variable/scenario	MF1	MF3	MF5	MF7	MF9	MF11	MF11-GP
Degree of overlapping (*n* = 1692)							
Mean	0.896	0.913	0.922	0.921	0.924	0.925	NA
Standard deviation	0.547	0.523	0.452	0.423	0.389	0.383	NA
Min. value	0.446	0.414	0.439	0.477	0.613	0.514	NA
Max. value	0.977	0.980	0.982	0.980	0.982	0.980	NA
Median	0.909	0.926	0.933	0.931	0.932	0.933	NA
Start of Q2	0.874	0.899	0.910	0.907	0.909	0.911	NA
Start of Q4	0.933	0.943	0.949	0.948	0.950	0.948	NA
Relative distance error (*n* = 736)							
Mean	0.0413	0.0381	0.0356	0.0355	0.0350	0.0347	0.0321
Standard deviation	0.0172	0.0161	0.0151	0.0147	0.0137	0.0134	0.0132
Min. value	0.0131	0.0114	0.0114	0.0107	0.0101	0.0117	0.00834
Max. value	0.151	0.116	0.106	0.114	0.107	0.108	0.102
Median	0.0383	0.0347	0.0326	0.0330	0.0330	0.0329	0.0302
Start of Q2	0.0294	0.0268	0.0256	0.0246	0.0247	0.0247	0.0225
Start of Q4	0.0498	0.0449	0.0418	0.0424	0.0426	0.0420	0.0395
Maximum distance error (*n* = 736)							
Mean	0.206136	0.191395	0.183159	0.186508	0.186625	0.18236	0.167697
Standard deviation	0.167561	0.167939	0.169514	0.172888	0.168031	0.173319	0.179095
Min. value	0.050249	0.042731	0.03965	0.037524	0.032608	0.033117	0.020035
Max. value	1.080349	1.07283	1.081207	1.096364	1.096364	1.080349	1.067945
Median	0.148047	0.132288	0.126469	0.126746	0.131104	0.125005	0.103314
Start of Q2	0.105294	0.096762	0.085048	0.088773	0.090754	0.08798	0.072494
Start of Q4	0.227504	0.205034	0.190995	0.189249	0.198662	0.179076	0.160346
Relative area error (*n* = 736)							
Mean	0.0765	0.0711	0.0719	0.0734	0.0695	0.0681	0.0643
Standard deviation	0.0721	0.0680	0.0646	0.0661	0.0606	0.0598	0.0559
Min. value	0.000113	0.000127	0.000321	0.000239	0.000157	0.00006	0.000011
Max. value	0.565	0.515	0.403	0.442	0.429	0.394	0.325
Median	0.0549	0.0533	0.0527	0.0534	0.0536	0.0534	0.0512
Start of Q2	0.0250	0.0233	0.0238	0.0251	0.0220	0.0224	0.0215
Start of Q4	0.105	0.0978	0.101	0.104	0.0988	0.0955	0.0900

Degree of overlapping: degree of coincident pixels between ML segmentation and ground truth (GT) images. Relative distance error: mean distance from ML lumen to GT lumen. Maximum distance error: maximum distance from ML lumen to GT lumen. Relative area error: difference in the lumen area between ML lumen and GT lumen. Start of Q2 and of Q4 define the interquartile range (IQR). NA: not applicable.

reduced as we move towards a more complex model for the lumen boundary segmentation. Particularly, *Figure 3D and E* present the correlation and Bland–Altman plots between GT lumen area and ML lumen area given by MF11-GP model.

Representative frames featuring the raw image, GT lumen boundary (green contour) and ML lumen boundary (red contour) are displayed in *Figure 4*. These frames were randomly selected to illustrate the segmentation capabilities for different morphological features of the lumen contour in different pullbacks and for a wide range of relative area error magnitude. The relative area error is also reported next to each row, ranging between 0.05 and 0.10 for these frames.

Discussion

Automatic gating and segmentation are mandatory tools towards improving the reliability in the assessment of IVUS datasets, and reducing the intra- and inter-observer variability in the catheterization lab.

In this work, several innovative ML segmentation strategies were assessed using multiple frames. The proposed MFCNN approach combined with GP regression resulted in an average distance error with a median (IQR) of 3.02 (2.25–3.95)%. The median error (IQR) in the lumen area was 5.12 (2.15–9.00)%. Adding information about neighbouring frames surrounding the frame of interest consistently
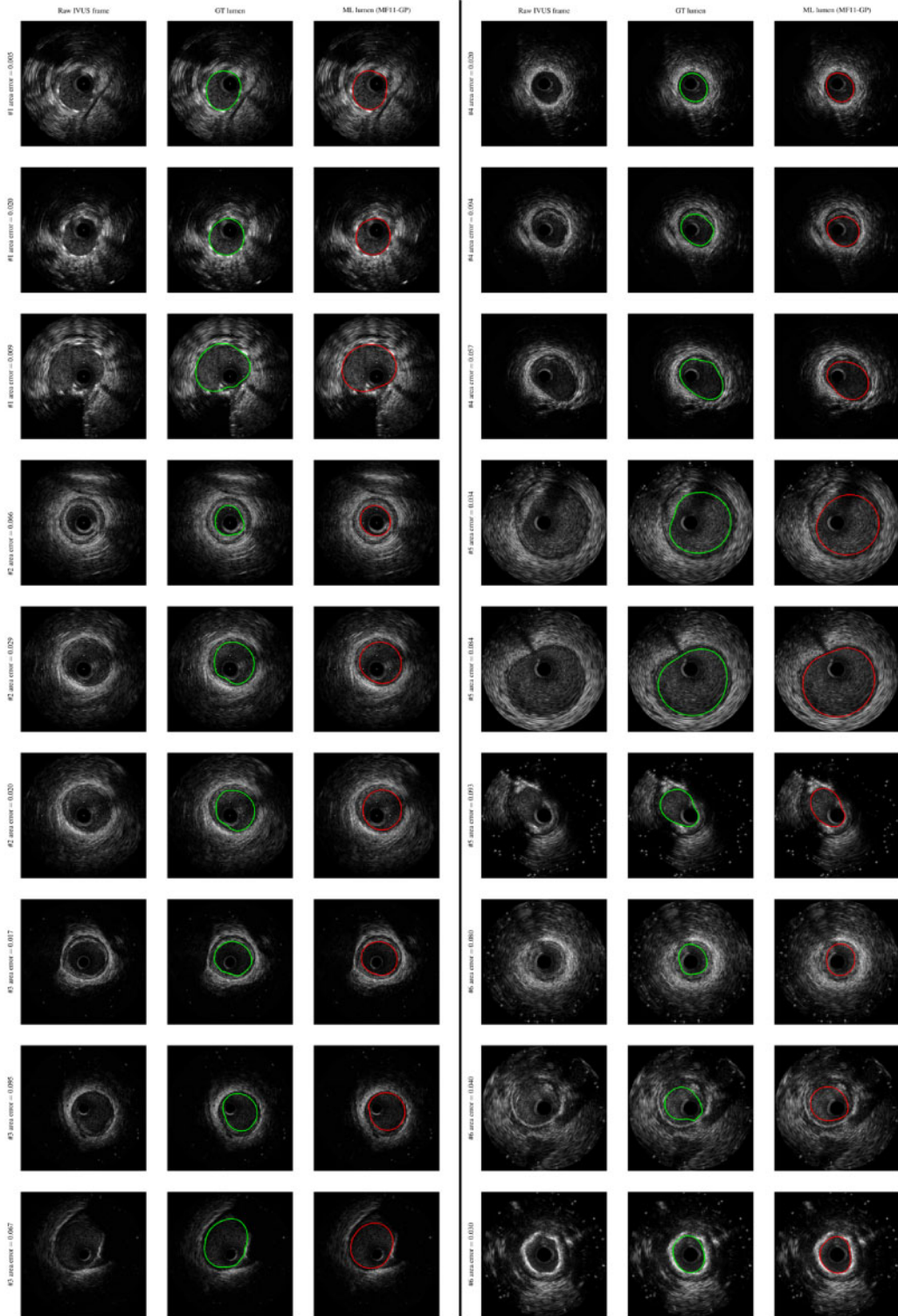
Figure 4 Randomly selected frames showing a raw image, ground truth (GT) lumen boundary (green contour), and machine learning (ML) lumen boundary (multi-frame—MF11-Gaussian process—GP, red contour).

improved the segmentation performance. Moreover, the use of the GP regressor improved the resulting segmentation by dealing with high-frequency noise and enforcing contour continuity (periodicity) of the lumen boundary, yielding anatomically coherent lumen delineations.

Literature concerning automatic segmentation for IVUS datasets using ML approaches is surprisingly scarce. Only a handful of methods relying on standard ML approaches have been proposed,[7–9] and there was not consensus about the best approach to the problem. In one, 109 IVUS frames were used to train a standard convolutional neural network and, in another one, 77 sequences of five adjacent IVUS frames, from a publicly available dataset, were used to train a kernel learning algorithm. The predictive capabilities reported in this work are in the range reported in these contributions. However, direct comparison with these techniques is not possible because the datasets are not the same, and because of the lack of clarity in the definition of the metrics employed in each case. Certainly, our method naturally accounts for the longitudinal coherence through the MF strategy, and deals with the noisy lumen definition in the image domain through the GP regressor. Hence, we strongly believe this approach of combining automatic gating, MF convolutional neural network segmentation and GP regression provides a consistent and reliable framework to account for the longitudinal and transversal coherence encountered in IVUS datasets.

In the catheterization laboratory, MLAs are commonly used to inform the clinical decision whether the lesion requires revascularization particularly in the left main coronary artery.[10] Currently, this assessment is performed by visually inspecting the pullback and selecting what by eye seems to be the smallest lumen area which then by manual tracing *a number* is obtained representing the so-called MLA. This approach assumes first that by visual inspection of the pullback the right frame is selected (i.e. represents the smallest lumen in that lesion) and second that the manual tracing of the lumen is properly done. Unfortunately, there is a large variability in the selection of the MLA frame and in the tracings of the lumen area as well. More importantly, this is time-consuming and demands adequately trained personnel. Machine learning solutions overcome all these issues by providing a fast, accurate and precise assessment of the lumen areas. Thereby, ML methods would make this procedure more efficient and safer in a way since the error in the contour tracing would be minimized dramatically.

Limitations

Machine learning methods need to be customized for each IVUS catheter type and also need to account for manual vs. automatic pullback acquisition. The former would require new training of the algorithm which may be a simpler task compared to the training of an ML algorithm on a manually acquired pullback since the ML relies on a selection of end-diastolic frames to speed up segmentation.

Conclusions

These results demonstrate that the proposed machine learning approach is suitable to effectively segment the lumen boundary in IVUS scans, reducing the burden of costly and time-consuming manual delineation.

Supplementary material

Supplementary material is available at *European Heart Journal – Digital Health* online.

Data availability

Authors are not sharing any data.

Conflict of interest: H.M.G.G.: his institution has received research grants from Phillips which an IVUS company. His institution also receives research grants Biotronik, Medtronic, Neovasc, Boston Scientific, Abbott, Cook, Shockwave and Corflow. All other authors declare no conflict of interest.

References

1. Mintz GS, Nissen SE, Anderson WD, Bailey SR, Erbel R, Fitzgerald PJ, Pinto FJ, Rosenfield K, Siegel RJ, Tuzcu EM, Yock PG. American College of Cardiology Clinical Expert Consensus Document on Standards for Acquisition, Measurement and Reporting of Intravascular Ultrasound Studies (IVUS). A report of the American College of Cardiology Task Force on Clinical Expert Consensus Documents. *J Am Coll Cardiol* 2001;**37**:1478–1492.
2. De Winter SA, Hamers R, Degertekin M, Tanabe K, Lemos PA, Serruys PW, Roelandt JR, Bruining N. Retrospective image-based gating of intracoronary ultrasound images for improved quantitative analysis: the intelligate method. *Catheter Cardiovasc Interv* 2004;**61**:84–94.
3. Maso Talou GD, Larrabide I, Blanco PJ, Bezerra CG, Lemos PA, Feijoo RA. Improving cardiac phase extraction in IVUS studies by integration of gating methods. *IEEE Trans Biomed Eng* 2015;**62**:2867–2877.
4. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In N Navab, J Hornegger, WM Wells, AF Frangi, eds. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015 [Internet]*. Cham: Springer International Publishing; 2015 [cited 2020 Sep 1]. p. 234–241 (Lecture Notes in Computer Science; vol. 9351). Available from: http://link.springer.com/10.1007/978-3-319-24574-4_28.
5. Rasmussen CE, Williams CKI. *Gaussian Processes for Machine Learning*. Cambridge, MA: MIT Press; 2006. 248 p (Adaptive computation and machine learning).
6. Taha AA, Hanbury A. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. *BMC Med Imaging* 2015;**15**:29.
7. Yang J, Tong L, Faraji M, Basu A. IVUS-Net: an intravascular ultrasound segmentation network. arXiv:180603583 [cs, eess, stat] [Internet]. 2018 Jun 14 [cited 2020 Sep 1]; Available from http://arxiv.org/abs/1806.03583.
8. Balakrishna C, Dadashzadeh S, Soltaninejad S. Automatic detection of lumen and media in the IVUS images using U-Net with VGG16 Encoder. arXiv:180607554 [cs] [Internet]. 2018 Jun 20 [cited 2020 Sep 1]; Available from http://arxiv.org/abs/1806.07554.
9. Cui H, Xia Y, Zhang Y. Supervised machine learning for coronary artery lumen segmentation in intravascular ultrasound images. *Int J Numer Meth Biomed Engng* 2020;**36**:e3348. Available from https://onlinelibrary.wiley.com/doi/abs/10.1002/cnm.3348.
10. Neumann FJ, Sousa-Uva M, Ahlsson A, Alfonso F, Banning AP, Benedetto U, Byrne RA, Collet JP, Falk V, Head SJ, Jüni P, Kastrati A, Koller A, Kristensen SD, Niebauer J, Richter DJ, Seferovic PM, Sibbing D, Stefanini GG, Windecker S, Yadav R, Zembala MO; ESC Scientific Document Group. 2018 ESC/EACTS Guidelines on myocardial revascularization. *Eur Heart J* 2019;**40**:87–165.