Special Session on Vascular Imaging

Maite Termenon[1]

[1]Computational Intelligence Group

2012 March 2

Article

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

ELSEVIER

Evaluation framework for carotid bifurcation lumen segmentation and stenosis grading

K. Hameeteman [a,*], M.A. Zuluaga [b,c], M. Freiman [d], L. Joskowicz [d], O. Cuisenaire [e], L. Flórez Valencia [f], M.A. Gülsün [g], K. Krissian [h], J. Mille [i], W.C.K. Wong [k], M. Orkisz [b], H. Tek [g], M. Hernández Hoyos [c], F. Benmansour [j], A.C.S. Chung [k], S. Rozie [l], M. van Gils [l], L. van den Borne [l], J. Sosna [m], P. Berman [m], N. Cohen [m], P.C. Douek [n,b], I. Sánchez [n], M. Aissat [n], M. Schaap [a], C.T. Metz [a], G.P. Krestin [l], A. van der Lugt [l], W.J. Niessen [a,o], T. van Walsum [a]

[a] *Biomedical Imaging Group Rotterdam, Departments of Radiology & Medical Informatics, Erasmus MC, P.O. Box 2040, 3000 CA Rotterdam, The Netherlands*
[b] *CREATIS, Université de Lyon, Université Lyon 1, INSA-Lyon, CNRS UMR5220, INSERM U630, F-69621 Villeurbanne, France*
[c] *Grupo IMAGINE, Grupo de Ingeniería Biomédica, Universidad de los Andes Bogotá, Colombia*
[d] *Computer-Aided Surgery and Medical Image Processing Laboratory, School of Engineering and Computer Science, The Hebrew University of Jerusalem, Israel*
[e] *Philips Healthcare, Medisys Research Laboratory, 33 rue de Verdun, B.P. 313, F-92156 Suresnes Cedex, France*
[f] *Pontificia Universidad Javeriana, Facultad de Ingeniería, Departamento de Ingeniería de Sistemas, Bogota, Colombia*
[g] *Imaging and Visualization Department, Siemens Corporate Research, Princeton, NJ 08540, United States*
[h] *GIMET, Dept. de Informática y Sistemas, Univ. Las Palmas de Gran Canaria, Spain*
[i] *Université de Lyon, CNRS, Université Lyon 1, LIRIS, UMR5205, F-69622, France*
[j] *CEREMADE, UMR CNRS 7534, Université Paris Dauphine, F-75775, France*
[k] *Lo Kwee-Seong Medical Image Analysis Laboratory, Department of Computer Science and Engineering, The Hong Kong University of Science and Technology, Hong Kong*
[l] *Department of Radiology, Erasmus MC, Rotterdam, The Netherlands*
[m] *Hadassah Hebrew University Medical Centre, Jerusalem, Israel*
[n] *Hôpital Louis Pradel, Bron, France*
[o] *Imaging Science and Technology, Faculty of Applied Sciences, Delft University of Technology, Delft, The Netherlands*

Outline

Outline

1. Introduction
2. Evaluation Framework
   - Lumen segmentation and stenosis grading
   - CTA data
   - Reference standard
   - Evaluation measures and ranking
3. MICCAI workshop

Description

- Evaluation framework that allows a standardized and objective quantitative comparison of carotid artery lumen segmentation and stenosis grading algorithms.
- This framework has been introduced at the ***MICCAI 2009 workshop 3D Segmentation in the Clinic: A Grand Challenge III***, and we compare the results of eight teams that participated.
- Results show that automated segmentation of the vessel lumen is possible with a precision that is comparable to manual annotation.

Description

- Computed Tomography Angiography (CTA).
- We focus on the carotid bifurcation, where the Common Carotid Artery (CCA) splits into the External Carotid Artery (ECA) and Internal Carotid Artery (ICA).

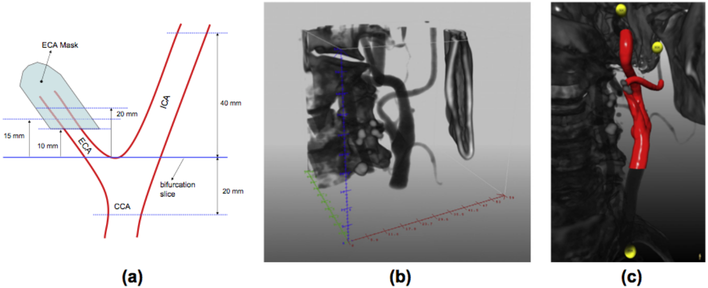

Fig. 1. Region around carotid bifurcation: (a) schematic depiction of the region of interest; (b) a rendering of this region for one of the datasets; (c) visualization of a dataset with the three initialization points and the reference segmentation.

Outline

Evaluation Framework

- Publicly accessible data repository.
- Set of standardized evaluation measures.
- Online evaluation system.
- Web: `http://cls2009.bigr.nl`.

Outline

Lumen segmentation

- The exact region where lumen must be segmented is defined around the bifurcation slice, as the first (caudal to cranial) slice where the lumen of the carotid artery appears as two separate lumens.
- Bifurcation slice is not revealed to the participants, who must ensure that their segmentation at least includes this region.
- The lumen segmentation must be represented as a partial volume segmentation (image where each voxel value represents the occupancy of the voxel by the vessel lumen)
- A value of 0 means no lumen present, and a value of 1 means fully occupied with lumen.

Stenosis grading

- Two stenosis grades must be determined for each ICA: an area-based and a diameter-based stenosis grade.
- We define the diameter of a cross-section as the shortest straight line that divides the contour in two equally-sized areas.

We use the following NASCET-like (NASCET Collaborators, 1991) definition for the area-based stenosis grade S_a:

$$S_a = 100\% \times \left(1 - \frac{a_m}{a_r}\right) \quad (1)$$

where a_m is the minimal cross-sectional area along the CCA and ICA, and a_r the average cross-sectional area over a distal reference part of the ICA. For the diameter-based stenosis grade S_d we similarly use:

$$S_d = 100\% \times \left(1 - \frac{d_m}{d_r}\right) \quad (2)$$

where d_m and d_r are the minimal and average reference cross-sectional diameter, respectively. We define the diameter of a

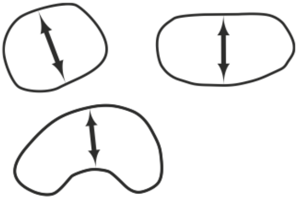

Fig. 2. Examples of our definition of the diameter of a cross-sectional contour.

Two Categories

- We distinguish two categories for the lumen segmentation and the stenosis grading task:
  - fully automatic methods.
  - semi-automatic methods where 3 initialization points may be used. To eliminate dependency on these initialization points, they are incorporated in the available data repository. The three points are located within the carotid artery proximal and distal to the defined ROI.

Outline

CTA data

- 56 CTA datasets from 3 different medical centers: Erasmus MC (Rotterdam, The Netherlands), Hôpital Louis Pradel (Bron, France) and Hadassah Hebrew University Medical Centre (Jerusalem, Israel).

Table 1
Distribution of the datasets over the five stenosis categories based on the reference standard.

Stenosis category	Stenosis degree (%)	Number of datsets
0	0	12
1	0–30	10
2	30–50	6
3	50–70	10
4	70–99	18

Datasets for training and testing

- 15 of the 56 datasets are made available for training.
- The reference standard for these training datasets is also available for download.
- The training datasets were chosen such that they were equally distributed over the 5 stenosis categories.

Table 2
Overview of scanning parameters of CTA datasets, EMC = Erasmus MC, Hd = Hadassah, LP = Louis Pradel.

	Scanner	In plane voxels	Slices	Pixel size (mm)	Z-spacing (mm)	Slice thick. (mm)	Kernel
EMC	Sensation 16	512 × 512	395–579	0.23–0.26	0.6	1	B30f
Hd	Brilliance 64	512 × 512	750	0.55	0.5	1	B
LP	Brilliance 64	512 × 512	636–827	0.414–0.547	0.45	0.9	B

Table 3
Datasets per center, and distribution over training and testing sets.

Center	Training	Testing	Total
Erasmus MC	9	27	36
Hadassah	3	7	10
Louis Pradel	3	7	10
Total (#)	15	41	56

Outline

Reference standard

- Reference standard was created by averaging 3 manual segmentations of the lumen in the CTA datasets.
- This section describes:
  - manual annotation process,
  - creation of a partial volume representation from each observers' annotations,
  - averaging of the observers' segmentations to obtain the reference standard.

Manual Annotations

- 3 different observers annotated the carotid lumen boundary and graded the stenosis in the ICA for each dataset.
- Manual annotations for the lumen segmentation and stenosis grading were performed with a custom made tool, based on MeVisLab (MeVis Research, Bremen, Germany).
- Procedure:
  1. Bifurcation point was identified and manually selected.
  2. Positions along the centerlines for both the ICA and ECA were clicked, starting in the CCA.
  3. Resampled centerlines were used to generate Curved Multi Planar Reformatted images (CMPRs).
  4. Cross-sectional contours orthogonal to the centerline were created at 1 mm intervals along the centerline.

Manual Annotations

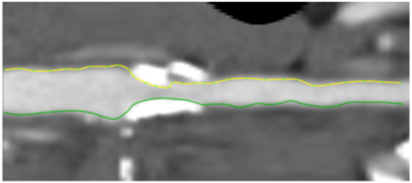

Fig. 4. Example of two longitudinal contours drawn on a CMPR image.

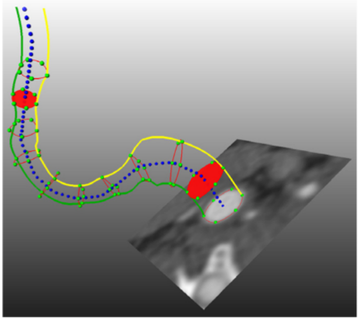

Fig. 5. Example of two longitudinal contours (yellow and green, which correspond to the ones shown in Fig. 4) with cross-sectional contours that are spline interpolations of the six intersection points (green dots) of longitudinal contours with the plane (shown image plane) that is perpendicular to the centerline (dark blue dots). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Observers' contours processing

- Contours of each observer were separately processed to obtain the partial volume (pv) segmentations.
    1. Contours (for ICA and ECA) were converted to pv segmentations.
    2. Signed distance maps (for ICA and ECA) were generated from the pv segmentations.
    3. The ICA and ECA signed distance maps were combined, resulting in a signed distance map for the complete bifurcation.
    4. The pv segmentations of the ICA and ECA were also combined by taking the voxel-wise maximum to obtain the pv segmentation of the bifurcation. (This pv segmentation is used to rate the observer in the same way as the contestants' segmentations.)

Observers' contours processing

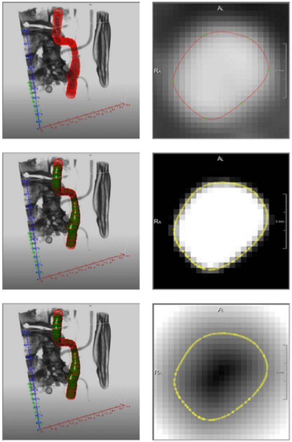

Fig. 6. Processing of observer annotations, left column shows a 3D visualization and right column a 2D visualization. From top to bottom: initial contours, partial volume from contours (left: isosurface at 0.5) and signed distance map from partial volume (left: isosurface at 0.0).

Combining observers' lumen segmentations

- The 3 segmentations of the observers were used to generate the data of the reference standard:
  - bifurcation slice number: computed by averaging the locations. This slice, and the bounding boxes of the contours, were used to determine the ROI for the evaluation.
  - evaluation ROI: bounding box of the contours, ex- tended with 15 mm both in x- and in y-direction.
  - lumen segmentation: contains three representations: a signed distance map, a surface representation and a pv segmentation. It also contains the mask for the distal part of the ECA.

Stenosis values

- The ICA and CCA contours were used for stenosis grading.
- Graphs of the contour area and diameter along the centerline were created, based on the corrected cross-sectional contours.
- The stenosis grade was determined using the values from these graphs.
- The 3 observer values for the stenosis were averaged to obtain the reference standard stenosis values.

Initialisation points

- 3 initialization points for the semi-automatic methods were annotated by one of the observers.

Outline

Lumen segmentation

3.4.1. Lumen segmentation

The partial volume lumen segmentations as supplied by a participant are evaluated using the following three performance measures[1]:

- The Dice similarity index D_{si}:

$$D_{si} = \frac{2 \times |pv_r \cap pv_p|}{|pv_r| + |pv_p|} \quad (3)$$

where pv_r and pv_p are the reference and a participant's partial volumes, respectively, the intersection operation is the voxel-wise minimum operation, and $|\cdot|$ is the volume, *i.e.* the integration of the voxel values over the complete image.

- The mean surface distance D_{msd}:

$$D_{msd} = \frac{1}{2} \times \left(\frac{\int_{S_r} |sdm_p| ds}{A_r} + \frac{\int_{S_p} |sdm_r| ds}{A_p} \right) \quad (4)$$

where $|sdm_p|$ and $|sdm_r|$ are the absolute signed distance maps of the reference and a participant's segmentation, respectively, S_r and S_p are the lumen boundary surfaces (isosurfaces of the signed distance map at the value 0) and A_i is the surface area of surface S_i, *i.e.* $A_i = \int_{S_i} ds$.

- The Hausdorff[2] distance D_{hd}:

$$D_{hd} = \max \left(\max_{x \in S_r} |sdm_p(x)|, \max_{x \in S_p} |sdm_r(x)| \right) \quad (5)$$

Both distance measures are symmetric, and all measures are only evaluated in the region of interest that is specified in Section 3.1.

Stenosis grading

- Stenosis grading error: absolute difference between the reference standard value and the value determined by a participant.
- Stenosis errors are not communicated per dataset, but only per ensemble (training or testing).
- Final ranking is determined by averaging the (hidden) errors per dataset and stenosis grade (diameter and area).
- We developed a simple standard stenosis grading algorithm and applied it to all segmentation results.
- The algorithm takes the partial volume lumen segmentation of the participants as input and determines the desired stenosis measure.

Outline

Description

- For the lumen segmentation, 8 teams participated in the category for semi-automatic methods and one team submitted results of a fully automatic method.
- Only 3 teams submitted results for the stenosis grading, none of these methods were fully automatic.

Evaluated algorithms (I)

- Cuisenari:
  - fully automated method that segments the carotid bifurcation and the vessels from the aortic arch to the circle of Willis (COW), and the external branch and the vertebral arteries.
  - Centerline of each vessel is extracted using a local adaptive fast marching algorithm that is both seeded and constrained by an anatomical model.
  - These constraints are adapted to the individual patient using both registration of the brain and segmentation of the brain and spine.
  - Seeds are automatically placed in the COW and the lower part of the neck.
  - Vessel lumen is segmented using 3D active objects initialized as a tube around the centerline.

Evaluated algorithms (II)

- Gülsün and Tek:
  - It uses graph-cuts optimization technique together with centerline models for segmenting the carotid arteries.
  - It first detects the centerline representations between user placed seed points (based on a minimal path detection method which operates on a medialness map).
  - Lumen of carotid arteries is then extracted by a graph-cut optimization technique using the detected centerlines as input.

Workshop results

Table 4
Lumen segmentation performance of all submissions. The average value and rank is shown for each of the three similarity measures as defined in Section 3.4.1. The algorithm by Cuisenaire (in **bold**) is the only automatic method.

Method name	Total success	Dice		Msd		Hausdorff		Processing time	Total rank
		%	**rank**	**mm**	**rank**	**mm**	**rank**		
Observer A	41	95.1	2.61	0.10	2.68	0.65	2.73	–	2.67
Observer B	41	94.6	3.34	0.11	3.37	0.83	3.07	–	3.26
Observer C	41	94.4	3.46	0.12	3.32	0.97	3.85	–	3.54
M.A. Gülsün and H. Tek	41	91.8	5.95	0.18	5.68	1.5	5.27	50 s.	5.63
K. Krissian et al.	41	87.3	6.05	0.54	7.02	4.4	6.83	15 min	6.63
O. Cuisenaire	33	89.6	8.05	0.17	6.90	1.7	7.00	1–2 min	7.32
J. Mille et al.	41	83.5	8.66	0.74	8.68	10	8.90	90 s	8.75
M. Freiman et al.	41	82.9	8.85	0.75	9.05	9.2	8.90	2 min	8.94
W.C.K. Wong et al.	41	77.5	9.34	1.1	9.73	11	9.07	90 s	9.38
M.A. Zuluaga et al.	41	80.9	9.63	0.82	9.68	10	9.07	4 min	9.46
L. Florez Valencia et al.	37	53.6	11.1	3.4	10.9	12	10.2	2 min	10.7

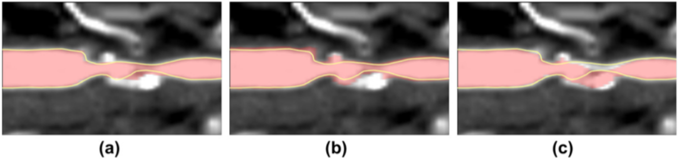

Fig. 7. Visual impression of reference standard (yellow line) and user segmentations (red) with different Dice measures: (a) 94.5, example from M.A. Gülsün and H. Tek, (b) 88.4, example from O. Cuisenaire and (c) 88.1, example from L. Florez Valencia et al. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Workshop results

Table 6
The average rank of all the methods specified for the different stenosis categories.

Method name	Stenosis category				
	None	0–30	30–50	50–70	70–99
Observer A	2.58	2.73	2.57	2.50	2.96
Observer B	3.58	3.30	3.43	3.12	2.88
Observer C	3.62	3.63	3.19	3.88	3.33
M.A. Gülsün and H. Tek	6.08	5.90	5.09	5.62	5.33
K. Krissian et al.	7.46	6.03	6.38	6.29	7.12
O. Cuisenaire	7.12	6.50	6.71	8.54	7.83
J. Mille et al.	8.62	8.47	9.81	8.42	8.62
M. Freiman et al.	8.46	8.87	8.95	9.00	9.42
W.C.K. Wong et al.	9.50	10.2	9.52	8.75	8.79
M.A. Zuluaga et al.	8.62	9.83	9.57	9.46	9.75
L. Florez Valencia et al.	10.8	10.7	11.3	10.6	10.3

Table 5
Lumen measure of the three observers with respect to each other.

Observer	AB	BC	CA	Average
Dice	0.92	0.91	0.91	0.92
Msd	0.18	0.20	0.19	0.19
Hausdorff	1.23	1.39	1.41	1.34

Table 8
Pearson correlation coefficient between the standard stenosis measures and the lumen measures.

Stenosis measure	Dice	Msd	Hausdorff	Total rank
Diam	−0.87	0.78	0.97	0.93
Area	−0.87	0.79	0.97	0.92

Table 7
Stenosis measures (as defined in Section 3.4.2) for all submissions. For both stenosis measures, the average difference with the reference stenosis and the rank is shown. The values in italics have been calculated with the standard stenosis grading algorithm and have not been supplied by the participants. The lumens column shows the number of lumen segmentations from which the standard values are calculated.

Method Name	Total Success	Diameter		Area		Standard algorithm			Total Rank
		Δ%	Rank	Δ%	Rank	Diameter*, Δ%	Area*, Δ%	Lumens*	
Observer A	41	3.40	2.15	2.90	2.02	*6.70*	*8.20*	*41*	2.09
Observer B	41	5.40	2.51	4.30	2.32	*9.00*	*8.70*	*41*	2.41
Observer C	41	5.70	2.66	5.00	2.68	*11.2*	*10.3*	*41*	2.67
M.A. Gülsün and H. Tek	41	–	–	–	–	*9.70*	*12.8*	*41*	–
K. Krissian et al.	41	–	–	–	–	*14.3*	*19.8*	*41*	–
O. Cuisenaire	33	–	–	–	–	*16.7*	*16.5*	*33*	–
J. Mille et al.	41	–	–	–	–	*31.2*	*35.7*	*41*	–
M. Freiman et al.	41	–	–	–	–	*24.5*	*29.4*	*41*	–
W.C.K. Wong et al.	41	31.4	4.20	25.2	4.17	*27.8*	*27.5*	*41*	4.18
M.A. Zuluaga et al.	41	17.0	3.56	16.9	3.71	*30.0*	*31.4*	*41*	3.63
L. Florez Valencia et al.	41	43.3	4.73	37.0	4.39	*36.2*	*39.7*	*37*	4.56