



Simultaneous segmentation and anatomical labeling of the cerebral vasculature



David Robben^{a,c,*}, Engin Türetken^{b,d}, Stefan Sunaert^{a,e,g}, Vincent Thijs^{a,f,h,i}, Guy Wilms^e, Pascal Fua^b, Frederik Maes^{a,c}, Paul Suetens^{a,c}

^a Medical Imaging Research Center (MIRC), KU Leuven, Leuven, Belgium

^b Computer Vision Laboratory, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

^c Medical Image Computing (MIC), ESAT-PSI, Department of Electrical Engineering, KU Leuven, Leuven, Belgium

^d Swiss Center for Electronic and Microtechnology (CSEM), Switzerland

^e Department of Radiology, UZ Leuven, Leuven, Belgium

^f Department of Neurology, University of Leuven, Leuven, Belgium

^g Translational MRI, Department of Imaging & Pathology, KU Leuven, Leuven, Belgium

^h Leuven Research Institute for Neuroscience & Disease (LIND), KU Leuven, Leuven, Belgium

ⁱ Laboratory of Neurobiology, Vesalius Research Center, Leuven, Belgium

ARTICLE INFO

Article history:

Received 20 August 2015

Revised 20 January 2016

Accepted 16 March 2016

Available online 1 April 2016

Keywords:

Cerebral vasculature

Segmentation

Centerline extraction

Anatomical labeling

Circle of Willis

Integer programming

ABSTRACT

We present a novel algorithm for the simultaneous segmentation and anatomical labeling of the cerebral vasculature. Unlike existing approaches that first attempt to obtain a good segmentation and then perform labeling, we optimize for both by simultaneously taking into account the image evidence and the prior knowledge about the geometry and connectivity of the vasculature. This is achieved by first constructing an overcomplete graph capturing the vasculature, and then selecting and labeling the subset of edges that most likely represents the true vasculature. We formulate the latter problem as an Integer Program (IP), which can be solved efficiently to provable optimality. We evaluate our approach on a publicly available dataset of 50 cerebral MRA images, and demonstrate that it compares favorably against state-of-the-art methods.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Automated segmentation and anatomical labeling of blood vessels is an important problem with many practical applications. In clinical settings, it can give an interventional radiologist extra guidance when navigating through the vasculature of a patient or enable automatic quantification of specific vessel segments. In a research context, it can be used to detect patterns that may be correlated to the incidence of vascular pathologies.

In this work, we focus on the cerebral vasculature and more specifically on the Circle of Willis (CoW) and its adjacent vessels. The CoW is a circle of arteries at the base of the skull that connects the left and right side of the anterior cerebral circulation with the posterior cerebral circulation (Fig. 1). It is supplied with blood by three large arteries, namely the left and right internal carotid arteries (ICA) and the vertebrobasilar artery (VBA).

The CoW plays a crucial role in several vascular pathologies, notably hemorrhagic and ischemic stroke. Cerebral aneurysms are balloon-like bulges on the wall of cerebral vessels and their rupturing is the main cause of subarachnoid hemorrhagic stroke (van Gijn and Rinkel, 2001). About 90 % of all cerebral aneurysms are found on the CoW (Brisman et al., 2006). The specific topology of the CoW determines the redundancy in blood supply to the brain and is associated with the prevalence of border zone infarcts in patients with unilateral ICA stenosis (Hendrikse et al., 2001). Although the CoW has a very characteristic morphology, it is highly variable: less than half of the population has a complete circle because one or more arteries are usually missing (Krabbe-Hartkamp et al., 1998). This variability makes both segmentation and labeling challenging.

In the following section, we provide an overview of the state-of-the-art techniques for vascular segmentation and anatomical labeling. Subsequently, we propose a novel method that tackles both problems simultaneously. Finally, we evaluate our approach and demonstrate its advantage over current techniques.

* Corresponding author.

E-mail address: david.robben@esat.kuleuven.be (D. Robben).

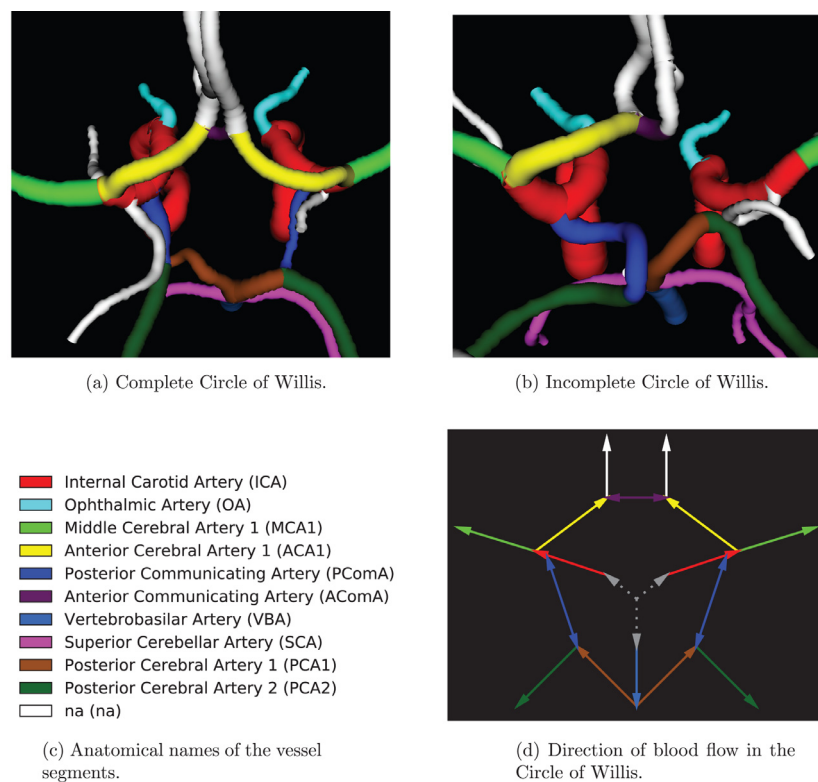


Fig. 1. (a,b) Anatomically labeled Circle of Willis in two healthy subjects. The colors refer to the anatomical names of the vessel segments as listed in (c). The CoW in (a) is complete, while the one in (b) misses several segments: right ACA1, right PCoMA and left PCA1. (d) Direction of blood flow in the Circle of Willis. The gray dotted arrows indicate the points where blood enters the CoW. Note that in the ACoMA and PCoMA the direction is unknown.

1.1. Related works on anatomical labeling of tubular structures

The problem of anatomical labeling of the vasculature can be posed more generally as labeling of a branched tubular structure. Most existing approaches formulate the problem in a graph-based setting, where the segmented object of interest is represented by a graph. In this graph, the vertices represent the branch and end-points and the edges represent the branches. In the following, we present an overview of such methods.

A conventional approach to labeling is to anatomically match an unlabeled graph with one or more labeled graphs. This approach was successfully followed by [Graham \(2006\)](#), [Buelow et al. \(2006\)](#) and [Feragen et al. \(2012\)](#) to label the bronchi. However, as remarked by [Bogunović et al. \(2013\)](#), the reported results for approaches that learn a statistical labeling model are generally better.

[Tschirren et al. \(2005\)](#) label the bronchi of a patient by matching the graph edges with a labeled, probabilistic model containing for each label the mean and standard deviation of geometric properties. They take into account the length and direction of the edges and the angle and distance between pairs of edges. A learnt, fixed topology between the different labels is enforced. [van Ginneken et al. \(2008\)](#) label the bronchi with a probabilistic model that contains for each label the mean and standard deviation of the orientation, radius and angle with the parent edge. For all edges, label probabilities are calculated and the labels are assigned in a top-down fashion to the bronchial segments. [Mori et al.](#) published several works about bronchial labeling. Their latest approach ([Mori et al., 2009](#)) labels the bronchial branches in an edge matching approach. A trained classifier assigns a probability to each possible pair of edge and label. The algorithm then searches for the globally optimal assignment of edge labels taking into account several topological constraints. In another line of work, they label the abdominal arteries ([Mori et al., 2010](#); [Matsuzaki et al., 2015](#)), which

they consider more difficult than labeling bronchi due to the larger variability. They use an application-specific algorithm which they do not expect to work on vasculature of other organs. The method of [Bogunović et al. \(2011\)](#) aims at labeling five bifurcations in the anterior circulation of the CoW. This is done by explicitly enumerating all possible isomorphisms between a given graph and a pre-determined atlas, and then calculating a probabilistic cost function, which combines a data term based on bifurcation morphology with a prior term that imposes a certain ordering of the bifurcations. Although the performance of the method is very good, it is not computationally scalable: a preprocessing step is required to prune the graph to about 20 candidate vertices. [Robben et al. \(2013\)](#) label the full CoW by matching vertices to a probabilistic atlas. The approach relies on both unary potentials of the bifurcations and also pairwise potentials between them. [Bogunović et al. \(2013\)](#) also label the full CoW by matching bifurcations to an atlas. They use the bifurcation properties and have several reference graphs to model the topology of the bifurcations. The method is evaluated on ground truth segmentations as it requires topologically correct segmentations without loops, disconnected regions or spurious branches. [Bilgel et al. \(2013\)](#) label the anterior part of the cerebral vasculature using belief propagation on a Bayesian network. [Ghanavati et al. \(2014\)](#) label the complete vasculature of a mouse model. The labeling problem is formulated as a Markov Random Field and the optimal solution is sought through simulated annealing stochastic relaxation. It should be noted that these methods, except [Robben et al. \(2013\)](#), [Bogunović et al. \(2013\)](#) and [Ghanavati et al. \(2014\)](#), assume that the graph is a tree.

All these approaches rely on a pre-existing segmentation in the form of a graph of blood vessels or bronchi. When assigning the anatomical labels, they account for the fact that the vasculature and the airway system are not random sets of tubular structures but organs with specific connectivity patterns. However they do

not exploit this knowledge to improve the segmentations. An exception is the work of Lu et al. (2009) where three non-branching coronary arteries are segmented and labeled by generating many possible segmentations and selecting for each label the most likely one, based solely on geometry.

1.2. Related works on vascular segmentation

Vascular segmentation refers to estimating both the vessel geometry and topology. The former involves finding vessel centerline points and their associated radii, while the latter involves recovering the connectivity of these points. This kind of segmentation is also called vascular reconstruction and can be contrasted with voxelwise segmentation, where each voxel in a given volume is assigned either to the vasculature or to the background, and where there is no strict notion of connectivity or morphology.

One could attempt to recover the centerlines from a voxelwise segmentation using skeletonization techniques. However, when the distance between the boundaries of two vessels is smaller than the imaging resolution—a phenomenon called kissing vessels—the connectivity of the voxelwise segmentation does not match that of the underlying vasculature. As a result, skeletonization yields erroneous connections.

Other methods aim at extracting the centerlines directly. An overview is given in Lesage et al. (2009, Section 4.4), where two broad categories are distinguished. Tracking methods (Aylward and Bullitt, 2002; Wong and Chung, 2007; Gülsün and Tek, 2008; Friman et al., 2010) begin at a given starting point and follow the vasculature by iteratively searching an adjacent point on the vessel centerline using local information. Typically heuristic rules detect branch and endpoints. The second category contains minimal path techniques (Deschamps and Cohen, 2001; Hua and Yezzi, 2007). The resulting centerline connects two given points and minimizes the integral of an intensity function along the centerline. Both approaches typically use only local information and are as such prone to creating shortcuts in the presence of kissing vessels.

In a more recent line of research, the vascular centerline segmentation problem is approached in a more global fashion by finding the globally optimal subset of edges in a graph of potential centerlines (Türetken et al., 2010; González et al., 2010; Türetken et al., 2013). More specifically, the approach first involves building an overcomplete centerline segmentation by calculating the minimal paths that connect a large set of seed points, i.e. points that have a high probability of lying on the vasculature. Subsequently, the optimal subset of these centerlines is selected according to a cost function that incorporates image evidence and a prior on the presence of junctions. This is done subject to a set of constraints that impose connectedness and—if applicable—the tree-like topology of the resulting solution. Rempfler et al. (2015) extend this approach and make the probability of the presence of a junction depend on its geometric properties as well.

All these approaches use knowledge of the local appearance of blood vessels and possibly some knowledge about their global connectivity. We extend this and we leverage the available anatomical information inherent to many vascular systems to further improve the segmentation performance as we describe in the following.

1.3. Contributions

Our approach performs the segmentation and anatomical labeling of the cerebral vasculature jointly so as to guarantee a result that is biologically plausible. This is achieved by taking into account both the image evidence and the anatomical knowledge about the vascular connectivity simultaneously using probabilistic classifiers that capture both appearance and geometry features, and impose learned connectivity rules.

Not only does this approach yield better results than state-of-the-art methods, such as Türetken et al. (2013), it is also generic and could easily be applied to other curvilinear structures.

This paper extends our earlier work (Robben et al., 2014). We present here an improved derivation for the objective function, which has both theoretical and practical advantages. The objective function has an extra term, taking the directionality of the edges into account. This speeds up the optimization, but more importantly, yields improved segmentation accuracy. Additionally, we propose and compare three different approaches to handling the labeling term introduced previously in Robben et al. (2014). Also, the label classifier uses different features. Finally, we also provide a more thorough experimental evaluation of the approach and its individual components. We propose a new segmentation quality metric that not only looks at overlap of the segmentation but also at its topological correctness. Using this metric, we demonstrate the advantage of the proposed method over existing ones.

2. Method

Starting from the original image (Fig. 2a), we first compute a directed graph G_I , which is an overcomplete segmentation of the vasculature: it is assumed that it includes all the vessels in the image, but it may contain spurious branches that are not part of the vasculature (Fig. 2c). In a second step, we select a subgraph in G_I and anatomically label its edges such that it most likely represents the true vasculature. This is done jointly, by optimizing a global objective function that captures both the image evidence and the prior knowledge about the geometry and connectivity of the labeled arteries (Fig. 2d).

The method assumes that all training and testing images are in a common coordinate system and cover the same field of view. If this is not the case, the images need to be registered and cropped to their common field of view.

2.1. Creation of the overcomplete graph G_I

We first compute a 4D scale-space *tubularity* volume, in which the first three dimensions refer to the position and the last dimension refers to the vessel radius, or scale. The values represent the probability that there is a vessel of that radius centered at that voxel in the image (Fig. 2b). The probabilities are computed by a supervised cascaded classifier with three levels. A cascaded classifier (Viola and Jones, 2001) is a concatenation of classifiers of increasing complexity, where the first classifiers serve as a filter to save computational effort: if a classifier on a lower level estimates with very high confidence that a voxel is not part of the vasculature, those probability estimates are final. If not, the voxel is fed to the next layer that has a more complex classifier. Each level uses an Extremely Randomized Trees classifier (Geurts et al., 2006), as provided by Scikit-learn (Pedregosa et al., 2011). The first level has as only feature the intensity of the voxel itself, the second level uses the intensities of a $3 \times 3 \times 3$ neighborhood around the voxel and the third level uses the responses of the Frangi (Frangi et al., 1998) and OOFA (Law and Chung, 2010) vesselness filters at 16 different scales (uniformly from 0.6 mm to 3.5 mm) in addition to the neighborhood intensities. The cascaded classifier returns 16 probabilities, one for each scale. The probabilities for a single voxel do not sum up to one, as there is also the possibility that the voxel is not located on a vessel centerline. Finally, the non maxima suppressed volume is also computed (Türetken et al., 2013).

We sample local maxima of this suppressed tubularity volume by iteratively picking the current maximum value and then suppressing its connected neighborhood (radius 2.5 mm or about 5 voxels). These samples act as vertices V_I in the directed graph

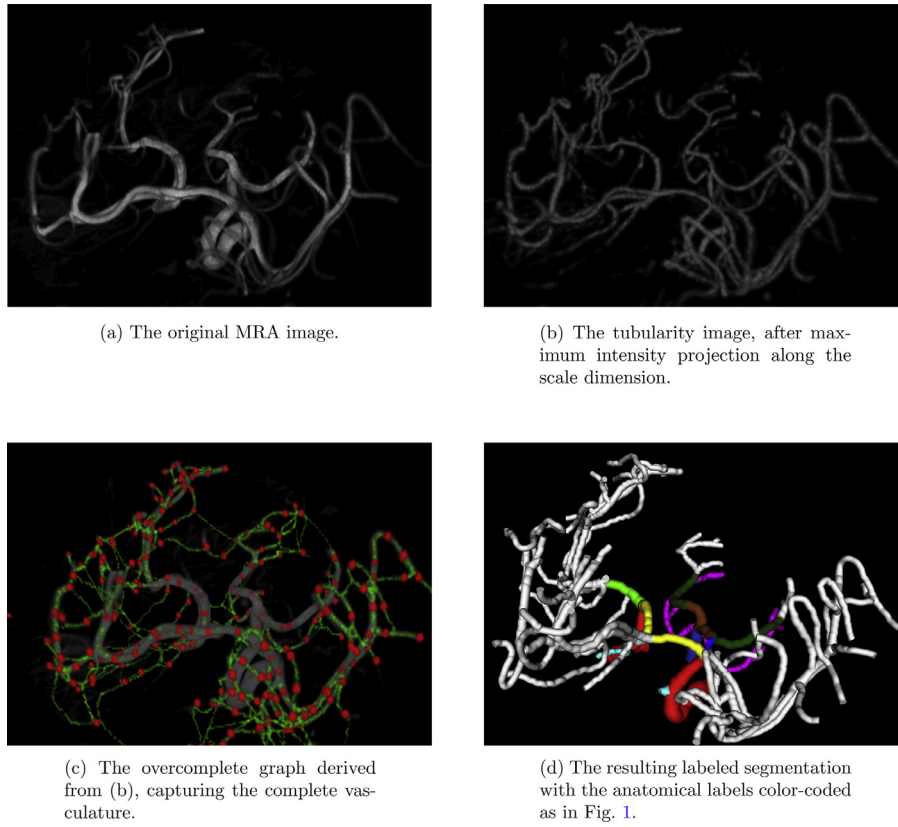


Fig. 2. Overview of the pipeline of the method.

$G_I = (V_I, E_I)$, which contains the desired segmentation of the vasculature (Fig. 2c). The edges E_I represent tubular segments. They are obtained by connecting pairs of samples that are within a certain distance from each other, along trajectories with maximal vessel centerline probability, as determined using the Fast Marching algorithm in the 4D tubularity volume (Hua and Yezzi, 2007). Since these trajectories are in scale space, they describe both the centerline positions and centerline radii. We assume that G_I is overcomplete such that its edges cover all the vessels in the image, but it may contain spurious edges that are not part of the vasculature.

In contrast to earlier approaches (Türetken et al., 2011) that attempt to sample the vertices from the true furcations of the vasculature, which are usually very hard to obtain accurately, the vertices in our graph do not need to coincide with the furcations. As a consequence, an edge does not necessarily correspond to a single vessel segment with a single anatomical label, but can cover (parts of) different vessel segments with different anatomical labels.

Note that the length of an edge is not limited to the suppression radius. On a straight line, two samples can be located up to 5 mm from each other and hence an edge pair can be up to 1 cm long. In the presence of bifurcations and curved segments, edges can become even longer.

2.2. Terminology

Given an image I , let $G = (V, E)$ be the directed graph obtained by adding a virtual vertex v_v to G_I such that $V = \{v_i\} = V_I \cup \{v_v\}$ and $E = \{e_{ij} = (v_i, v_j)\} = E_I \cup \{(v_v, v_i) | v_i \in V_I\}$. The virtual vertex will be used later to ensure the connectedness of our solution by forcing the existence of a directed path from the virtual vertex to every vertex in the solution. An illustration of G in a toy example is given in Fig. 3a.

The use of directed edges is preferred over undirected ones as this simplifies the mathematical formulation of this connectedness

constraint and other constraints on the topology of the solution (see Section 2.4).

While in earlier works the virtual vertex and directed edges are only used to formulate the connectedness constraint (Türetken et al., 2013), we propose to also learn the direction of the edges in the solution (see Section 2.3.2). Hereto, we indicate in the training set the centerlines' direction and connections with the virtual vertex. We indicate the direction according to the blood flow direction (see Fig. 1d). In some variants of the Circle of Willis, the direction of blood flow in the PComA and AComA is not known. In that case, we indicate the PComA from anterior to posterior, and the AComA from left to right. Since the training centerlines should meet the earlier mentioned connectedness constraint (there is a directed path from the virtual vertex to every piece of centerline), the virtual vertex will be connected to the points where blood enters the imaged region: the left and right ICA, the VBA (see Fig. 1d) and sometimes some small, unnamed segments that, due to the limited field of view, are not connected to the ICA or VBA.

Let $S = \{s_i\}$ be the set of the anatomical labels for distinct segments of the vasculature, extended with a void label (NA) for unnamed vessel segments (as illustrated in Fig. 1) and a virtual label for the edges between the virtual vertex and vertices in the image. The void label is necessary since we are interested in segmenting the whole cerebral vasculature rather than only the named segments of the CoW. The virtual label allows the method to learn and label where the directed vascular structure begins.

We formulate our problem in terms of pairs of consecutive directed edges in G , since it allows us to capture more global appearance and geometry information, and it gives rise to a linear objective function and constraints. Let $F = \{e_{ijk} = (e_{ij}, e_{jk})\}$ be the set of such edge pairs and \hat{L} the set of the edge pair labels. Since an edge pair can span multiple anatomical segments, an edge pair label is a tuple of potentially multiple segment labels: $\hat{L} \subset \bigcup_{n=1}^N \{(s_1, \dots, s_n) | \forall i : s_i \in S\}$, with N determined from the train-

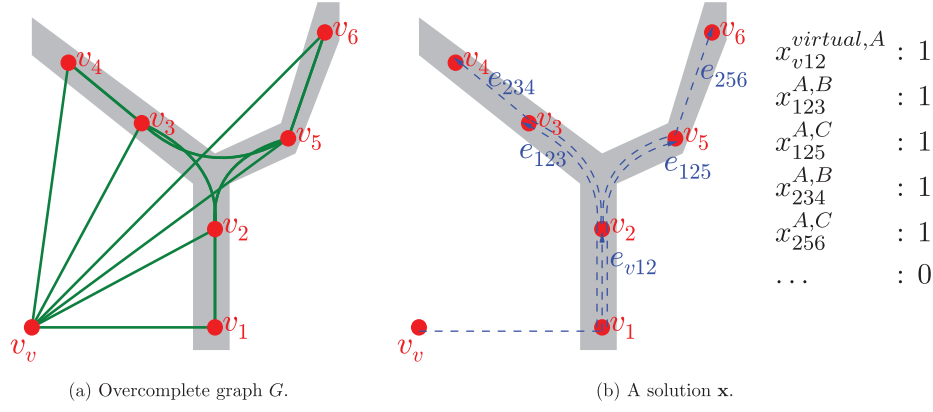


Fig. 3. Illustration of the used terminology on a toy example of a vascular structure (in gray). The overcomplete graph $G = (V, E)$ has vertices $V = \{v_i\}$, shown as red dots, and edges $E = \{e_{ij} = (v_i, v_j)\}$, shown as green lines. A subset of the set of all edge pairs F is shown as dashed blue lines. The shown edge pairs form a valid solution: the edge pairs are overlapping and connected to the virtual vertex v_v . (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Terminology and symbols regarding the graph and labels.

Name	Description
I	The image
$G_I = (V_I, E_I)$	The directed tubular graph constructed from the image.
v_v	The virtual root vertex. This vertex is virtual, as it lies outside the image. In the solution, every edge should be indirectly connected to it.
$G = (V, E)$	The directed tubular graph used in the reconstruction.
$V = \{v_i\} = V_I + \{v_v\}$	The vertices of the graph.
$E = \{e_{ij} = (v_i, v_j)\}$	The directed edges of the graph.
$E_I = \{(v_v, v_i) v_i \in V_I\}$	
$F = \{e_{ijk} = (e_{ij}, e_{jk})\}$	The set of consecutive directed edge pairs in the graph.
e_{ijk} follows e_{gh}	This implies that $e_{ij} = e_{gh}$.
e_{ijk} is active	Edge pair e_{ijk} is present in the solution.
incoming edge pair of e_{ij}	Edge pair whose second edge is e_{ij} . E.g. e_{hij} .
$S = \{s_i\}$	Set of the segment labels, these correspond to anatomical names for arteries.
$\hat{L} \subset \cup_{n=1}^N \{(s_1, \dots, s_n) \forall i : s_i \in S\}$	Set of the edge pair labels.

ing set. To be a valid solution, the edge pairs must be overlapping and indirectly connected to the virtual vertex. We impose every edge e_{ij} to have no more than one incoming edge pair e_{hij} in the solution. As such, every edge in the solution will have exactly one incoming edge pair. Note that this does not impose a constraint on the topology of the solution, while it simplifies the mathematical formulation.

We define $\mathbf{X} = \{X_{ijk}^l\}$ to be the vector of binary random variables, each representing the (non-)existence of a vessel segment with edge pair label l along the directed edge pair e_{ijk} . Its realization is denoted by the binary vector $\mathbf{x} = \{x_{ijk}^l\}$, which we will call a solution. An illustration on a toy example is given in Fig. 3b. We call an edge pair e_{ijk} active if it is present in the solution: $\exists l : x_{ijk}^l = 1$. Equivalently, an edge e_{jk} is active when it is present in the solution: $\exists l, v_i : x_{ijk}^l = 1$. All the symbols and their meaning are presented in Table 1.

In the following, we pose the joint segmentation and labeling problem as an integer program (IP) over \mathbf{x} , subject to a set of constraints.

2.3. Objective function

We formulate the problem as a maximum a posteriori (MAP) inference over the binary variables \mathbf{x} :

$$\mathbf{x}^* = \arg \max_{\mathbf{x} \in \mathcal{X}} P(\mathbf{X} = \mathbf{x} | I, G) \quad (1)$$

where \mathcal{X} denotes the set of feasible solutions that satisfy the constraints described in Section 2.4. We decompose $\mathbf{X} = \{X_{ijk}^l\}$ into two sets of random variables: $\mathbf{T} = \{T_{ijk}\}$ and $\mathbf{L} = \{L_{ijk}\}$. The binary variable T_{ijk} indicates whether the edge pair e_{ijk} is part of the solution. The variable L_{ijk} represents the edge pair label of the edge pair e_{ijk} and its value is an element of \hat{L} . The relations are:

$$\forall X_{ijk}^l : X_{ijk}^l = 1 \iff T_{ijk} = 1 \wedge L_{ijk} = l. \quad (2)$$

Their realizations are respectively $\mathbf{t} = \{t_{ijk}\}$ and $\mathbf{l} = \{l_{ijk}\}$. The objective function contains now two terms:

$$P(\mathbf{X} = \mathbf{x} | I, G) = P(\mathbf{T} = \mathbf{t}, \mathbf{L} = \mathbf{l} | I, G) \quad (3)$$

$$= P(\mathbf{L} = \mathbf{l} | \mathbf{T} = \mathbf{t}, I, G) P(\mathbf{T} = \mathbf{t} | I, G). \quad (4)$$

One can think of $P(\mathbf{T} = \mathbf{t} | I, G)$ as the segmentation term, giving the probability of a certain segmentation \mathbf{t} conditioned on the image I and the overcomplete graph G . The term $P(\mathbf{L} = \mathbf{l} | \mathbf{T} = \mathbf{t}, I, G)$ can be considered as the labeling term, giving the probability of a certain labeling \mathbf{l} conditioned on the segmentation \mathbf{t} , the image I and overcomplete graph G . In the following, both will be elaborated.

2.3.1. Labeling term $P(\mathbf{L} = \mathbf{l} | \mathbf{T} = \mathbf{t}, I, G)$

To derive $P(\mathbf{L} = \mathbf{l} | \mathbf{T} = \mathbf{t}, I, G)$ we assume conditional independence of the label over the edge pairs given the image and the graph:

$$P(\mathbf{L} = \mathbf{l} | \mathbf{T} = \mathbf{t}, I, G) = \prod_{e_{ijk} \in F} P(L_{ijk} = l_{ijk} | T_{ijk} = t_{ijk}, I, G) \quad (5)$$

$$= \prod_{e_{ijk} \in F} [P(L_{ijk} = l_{ijk} | T_{ijk} = 1, I, G)^{t_{ijk}} \quad (6)$$

$$\times P(L_{ijk} = l_{ijk} | T_{ijk} = 0, I, G)^{1-t_{ijk}}]. \quad (7)$$

The probabilities of $P(L_{ijk} = l | T_{ijk} = 1, I, G)$ are obtained from a probabilistic classifier trained on a set of labeled edge pairs, as explained in Section 2.5. The estimation of $P(L_{ijk} = l | T_{ijk} = 0, I, G)$ is more involved. Since edge pairs that are not part of the solution do not have a real label, this distribution cannot be learned from annotated examples. Instead, we have to specifically define a suitable distribution. The most correct way would be stating that an edge pair that is not part of the solution has no label. However,

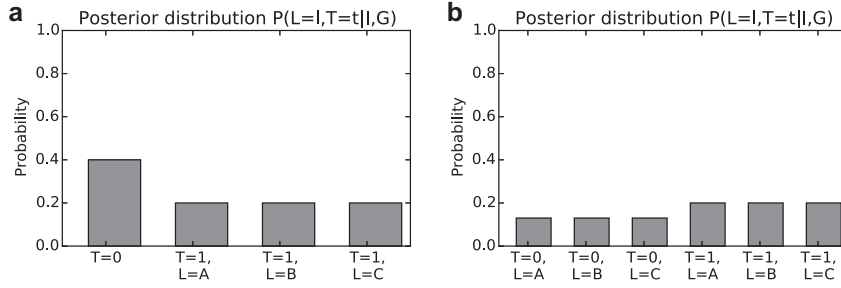


Fig. 4. Imagine a scenario with a single edge pair that has $P(T=1|I, G)=0.6$ and three as likely labels ($P(L=A|T=1, I, G)=P(L=B|T=1, I, G)=P(L=C|T=1, I, G)=0.33$). (a) If we say that edge pairs with $T=0$ do not have a label, then the MAP solution would not contain the edge pair, although it would make more sense to include the edge pair in the solution. (b) Using $P(L_{ijk}=l|T_{ijk}=0, I, G)=P(L_{ijk}=l|T_{ijk}=1, I, G)$ gives a posterior distribution with the modulus corresponding to the intended solution.

this can result in a scenario where the MAP solution of the simultaneous segmentation and labeling is not the solution we want, as demonstrated in Fig. 4a.

We propose three approaches to this problem. The first is to define $P(L_{ijk}=l|T_{ijk}=0, I, G)=P(L_{ijk}=l|T_{ijk}=1, I, G)$, which behaves better (as demonstrated in Fig. 4b). An edge pair e_{ijk} that is not part of the solution (i.e. $t_{ijk}=0$) has no real label and thus its label l_{ijk} in the solution is arbitrary and of no interest to us. Since we look for the most likely solution, an edge pair that is not part of the solution will be assigned the label with the highest probability. Thus, under optimization over x , the following equation holds:

$$P(\mathbf{L}=\mathbf{l}|\mathbf{T}=\mathbf{t}, I, G) = \prod_{e_{ijk} \in F} \left[P(L_{ijk}=l_{ijk}|T_{ijk}=1, I, G)^{t_{ijk}} \times \left(\max_{l'} P(L_{ijk}=l'|T_{ijk}=1, I, G) \right)^{1-t_{ijk}} \right]. \quad (8)$$

Allowing us to further derive:

$$P(\mathbf{L}=\mathbf{l}|\mathbf{T}=\mathbf{t}, I, G) = \prod_{e_{ijk} \in F} \left[\frac{P(L_{ijk}=l_{ijk}|T_{ijk}=1, I, G)}{\max_{l'} P(L_{ijk}=l'|T_{ijk}=1, I, G)} \right]^{t_{ijk}} \times \prod_{e_{ijk} \in F} \max_{l'} P(L_{ijk}=l'|T_{ijk}=1, I, G) \quad (9)$$

$$\propto \prod_{e_{ijk} \in F} \left[\frac{P(L_{ijk}=l_{ijk}|T_{ijk}=1, I, G)}{\max_{l'} P(L_{ijk}=l'|T_{ijk}=1, I, G)} \right]^{t_{ijk}} \quad (10)$$

$$= \prod_{e_{ijk} \in F} \prod_{l \in \hat{L}} \left[\frac{P(L_{ijk}=l|T_{ijk}=1, I, G)}{\max_{l'} P(L_{ijk}=l'|T_{ijk}=1, I, G)} \right]^{x_{ijk}^l}. \quad (11)$$

A second approach is to define $P(L_{ijk}=l|T_{ijk}=0, I, G)=1/|\hat{L}|$, i.e. that for edges not belonging to the solution all labels are equally likely. Following a similar derivation, we end up with:

$$P(\mathbf{L}=\mathbf{l}|\mathbf{T}=\mathbf{t}, I, G) \propto \prod_{e_{ijk} \in F} \prod_{l \in \hat{L}} \left[\frac{P(L_{ijk}=l|T_{ijk}=1, I, G)}{1/|\hat{L}|} \right]^{x_{ijk}^l}. \quad (12)$$

A third approach recognizes that not every label has an equally large prior probability to be found, especially since there is considerable variation in the length of the different vascular segments. By defining $P(L_{ijk}=l|T_{ijk}=0, I, G)=P(L_{ijk}=l)$, more importance is attached to the smaller segments:

$$P(\mathbf{L}=\mathbf{l}|\mathbf{T}=\mathbf{t}, I, G) \approx \prod_{e_{ijk} \in F} \prod_{l \in \hat{L}} \left[\frac{P(L_{ijk}=l|T_{ijk}=1, I, G)}{P(L_{ijk}=l)} \right]^{x_{ijk}^l}. \quad (13)$$

2.3.2. Segmentation term $P(\mathbf{T}=\mathbf{t}|I, G)$

We first introduce the set of undirected edge pairs: $\bar{F} = \{\bar{e}_{ijk} | e_{ijk} \in F \wedge i < k\}$. The random variables $\mathbf{T} = \{T_{ijk}\}$ are decomposed in random variables $\mathbf{Y} = \{Y_{ijk}\}$ and $\mathbf{D} = \{D_{ijk}\}$ with their realizations $\mathbf{y} = \{y_{ijk}\}$ and $\mathbf{d} = \{d_{ijk}\}$. The binary variable Y_{ijk} indicates whether the undirected edge pair \bar{e}_{ijk} is part of the solution: $Y_{ijk} = T_{ijk} + T_{kji}$. The binary variable D_{ijk} represents in which direction undirected edge pair \bar{e}_{ijk} appears in the solution: if D_{ijk} is true, the direction is from v_i to v_k , otherwise from v_k to v_i . The relations are:

$$\forall T_{ijk} \in \mathbf{T} : T_{ijk} = 1 \iff Y_{ijk} = 1 \wedge D_{ijk} = 1 \quad (14)$$

$$\forall e_{ijk} \in F : Y_{ijk} = Y_{kji} \quad (15)$$

$$\forall e_{ijk} \in F : D_{ijk} + D_{kji} = 1. \quad (16)$$

Assuming conditional independence over the undirected edge pairs, we get:

$$P(\mathbf{T}=\mathbf{t}|I, G) \quad (17)$$

$$= \prod_{\bar{e}_{ijk} \in \bar{F}} [P(D_{ijk}=d_{ijk}|Y_{ijk}=y_{ijk}, I, G)P(Y_{ijk}=y_{ijk}|I, G)] \quad (18)$$

$$= \prod_{\bar{e}_{ijk} \in \bar{F}} [(P(D_{ijk}=d_{ijk}|Y_{ijk}=1, I, G)P(Y_{ijk}=1|I, G))^{y_{ijk}} \times (P(D_{ijk}=d_{ijk}|Y_{ijk}=0, I, G)P(Y_{ijk}=0|I, G))^{1-y_{ijk}}]. \quad (19)$$

The probabilities of $P(D_{ijk}=d|Y_{ijk}=1, I, G)$ are obtained from a probabilistic classifier trained on a set of edge pairs and their directions, as explained in Section 2.5. Since undirected edge pairs \bar{e}_{ijk} that are not part of the solution do not have a real direction, the distribution of $P(D_{ijk}=d|Y_{ijk}=0, I, G)$ cannot be learned from annotated examples. Following the same reasoning as in Section 2.3.1, we define $P(D_{ijk}=d_{ijk}|Y_{ijk}=0, I, G)=P(D_{ijk}=d_{ijk}|Y_{ijk}=1, I, G)$.

Since we look for the most likely solution, an undirected edge pair that is not part of the solution will be assigned the direction with the highest probability. Thus, under optimization over x , the following equation holds:

$$P(\mathbf{T}=\mathbf{t}|I, G) \quad (20)$$

$$= \prod_{\bar{e}_{ijk} \in \bar{F}} \left[\left(P(D_{ijk}=d_{ijk}|Y_{ijk}=1, I, G)P(Y_{ijk}=1|I, G) \right)^{y_{ijk}} \times \left(\max_{d'} P(D_{ijk}=d'|Y_{ijk}=1, I, G)P(Y_{ijk}=0|I, G) \right)^{1-y_{ijk}} \right]. \quad (21)$$

This can be rewritten as:

$$P(\mathbf{T}=\mathbf{t}|I, G) \quad (22)$$

$$\propto \prod_{e_{ijk} \in \tilde{F}} \left[\frac{P(D_{ijk} = d_{ijk} | Y_{ijk} = 1, I, G) P(Y_{ijk} = 1 | I, G)}{\max_{d'} P(D_{ijk} = d' | Y_{ijk} = 1, I, G) P(Y_{ijk} = 0 | I, G)} \right]^{y_{ijk}} \quad (23)$$

$$= \prod_{e_{ijk} \in \tilde{F}} \left[\frac{P(D_{ijk} = 1 | Y_{ijk} = 1, I, G) P(Y_{ijk} = 1 | I, G)}{\max_{d'} P(D_{ijk} = d' | Y_{ijk} = 1, I, G) P(Y_{ijk} = 0 | I, G)} \right]^{y_{ijk} d_{ijk}} \quad (24)$$

$$= \prod_{e_{ijk} \in \tilde{F}} \prod_{l \in \tilde{L}} \left[\frac{P(D_{ijk} = 1 | Y_{ijk} = 1, I, G) P(Y_{ijk} = 1 | I, G)}{\max_{d'} P(D_{ijk} = d' | Y_{ijk} = 1, I, G) (1 - P(Y_{ijk} = 1 | I, G))} \right]^{x_{ijk}^l} \quad (25)$$

2.3.3. Objective function

Taking the logarithm of Eq. 4 results in an objective function that is linear in the x_{ijk}^l variables. Depending on which assumption we made in the derivation of the labeling term, we get a different objective function. We will refer to them as *Proposed 1*, *Proposed 2* and *Proposed 3*:

$$\sum_{e_{ijk} \in \tilde{F}} \sum_{l \in \tilde{L}} \left[w_{ijk}^Y + w_{ijk}^D + w_{ijk,l}^{L1} \right] x_{ijk}^l \quad (26)$$

$$\sum_{e_{ijk} \in \tilde{F}} \sum_{l \in \tilde{L}} \left[w_{ijk}^Y + w_{ijk}^D + w_{ijk,l}^{L2} \right] x_{ijk}^l \quad (27)$$

$$\sum_{e_{ijk} \in \tilde{F}} \sum_{l \in \tilde{L}} \left[w_{ijk}^Y + w_{ijk}^D + w_{ijk,l}^{L3} \right] x_{ijk}^l \quad (28)$$

with

$$w_{ijk}^Y = \log \frac{P(Y_{ijk} = 1 | I, G)}{1 - P(Y_{ijk} = 1 | I, G)} \quad (29)$$

$$w_{ijk}^D = \log \frac{P(D_{ijk} = 1 | Y_{ijk} = 1, I, G)}{\max_{d'} P(D_{ijk} = d' | Y_{ijk} = 1, I, G)} \quad (30)$$

$$w_{ijk,l}^{L1} = \log \frac{P(L_{ijk} = l | T_{ijk} = 1, I, G)}{\max_{l'} P(L_{ijk} = l' | T_{ijk} = 1, I, G)} \quad (31)$$

$$w_{ijk,l}^{L2} = \log \frac{P(L_{ijk} = l | T_{ijk} = 1, I, G)}{1/|\tilde{L}|} \quad (32)$$

$$w_{ijk,l}^{L3} = \log \frac{P(L_{ijk} = l | T_{ijk} = 1, I, G)}{P(L_{ijk} = l)} \quad (33)$$

For comparison, the objective function derived in our previous work (Robben et al., 2014) can be rewritten using the same notation as above:

$$\sum_{e_{ijk} \in \tilde{F}} \sum_{l \in \tilde{L}} \left[w_{ijk}^Y + w_{ijk,l}^{L3} \right] x_{ijk}^l \quad (34)$$

This corresponds to *Proposed 3* without the direction term.

Similarly, the objective function of Türetken et al. (2013) is given by:

$$\sum_{e_{ijk} \in \tilde{F}} w_{ijk}^Y t_{ijk}, \quad (35)$$

i.e. as a segmentation problem in the variables t_{ijk} without labeling component.

2.4. Constraints

Not every \mathbf{x} gives rise to a biologically plausible and feasible solution. We formulate constraints to enforce both the consistency and the connectedness of the solution. Furthermore, our algorithm learns from the annotated training data which edge pair labels and, more importantly, which configurations of labels are possible in the final solution. All these constraints are expressed by linear inequalities $W\mathbf{x} \leq \mathbf{b}$, where $W_{ij} \in \{-1, 0, 1\}$ and \mathbf{b} is a binary vector. In the following we give an overview of the constraints. The detailed mathematical formulation is given in Appendix A.

2.4.1. Consistency and connectedness

First, an active edge pair, i.e. an edge pair that is in the solution, can have only one edge pair label. Second, an edge e_{ij} can have at most one active incoming edge pair e_{hij} , as also illustrated in Fig. 3. Third, edges represent tubular segments and it is impossible that two edges that not share a vertex occupy the same physical 3D space. Hence, the solution cannot contain edges that spatially overlap but not have a common vertex. Fourth, all segment labels (except for the void label NA) correspond to anatomical names of arteries, which are coherent anatomical structures. By anatomical naming convention, two different anatomical segments are either not connected, or are connected at a single location. Now, recall that an edge pair label l is a tuple of segment labels: $l = (s_1, \dots, s_n)$, $s_i \in S$. Thus, if an edge pair e_{ijk} in the solution has edge pair label l that contains (s_a, s_b) , then the connection between s_a and s_b is somewhere on the edge e_{ij} or on the edge e_{jk} , and not on any other edge. Finally, to ensure the connectedness of the vascular segments, we enforce the active edge pairs to be connected to the virtual vertex v_v . An edge pair e_{ijk} is connected to v_v if there is a path, containing only active edge pairs, starting from v_v and containing e_{ijk} .

2.4.2. Learned label configurations

Although there is considerable inter-subject variation in the cerebral vasculature, some patterns hold invariably. These patterns can be expressed as rules: they state the edge pair labels that are forbidden for an edge pair, conditioned on the edge pair labels of certain neighboring edge pairs. The algorithm infers these rules from the training data and enforces them when segmenting and labeling unseen images.

When learning the allowed edge pairs labels of an edge pair e_{hij} , the algorithm conditions on various (combinations of) neighboring edge pairs (e.g. the preceding edge pair e_{ghi} , the following edge pair e_{ijk} , ...). In the following, an overview is given of the considered combinations and to which situations they refer to.

First, consider a single edge pair. The algorithm observes in the training data which edge pair labels can appear on a single edge pair, and restricts \tilde{L} to these edge pair labels.

Second, consider two edge pairs that follow each other, say e_{hij} and e_{ijk} . The algorithm observes in the training data which edge pair labels never appear on an edge pair e_{ijk} , given the edge pair label of its preceding edge pair e_{hij} .

Third, consider two edge pairs forming a bifurcation, say e_{hij} and e_{hik} . The algorithm observes in the training data which edge pair labels never appear on edge pairs forming a bifurcation. Additionally, it learns which edge pair labels never appear on an edge pair that does not form a bifurcation.

Fourth, consider an edge pair that forms a terminating branch, i.e. that is not followed by another edge pair. The algorithm observes in the training data which edge pair labels cannot be on a terminating branch.

Finally, since edge pairs are directed, it is possible to detect when they form a non-tree like topology. The algorithm observes

in the training data which edge pair labels cannot be on a non-tree like topology.

2.5. Estimation of probability terms

The objective function contains three probability terms, which need to be estimated. For estimation of $P(Y_{ijk} = 1 | I, G)$, we use the classifier described in Türetken et al. (2013). The probabilities $P(D_{ijk} = 1 | Y_{ijk} = 1, I, G)$ and $P(L_{ijk} = l | T_{ijk} = 1, I, G)$ are estimated by extremely randomized trees classifiers (Geurts et al., 2006), available in Scikit-learn (Pedregosa et al., 2011). The features are the begin and end point coordinates, the direction (the normalized difference between the begin and end point coordinates) and the average radius of the edge pair. Since edge pairs are relatively long and more local geometric properties are also relevant, we split the centerline at its point of largest curvature in two parts and calculate all features also for each part separately, resulting in 3x10 features for the edge pair.

For each classifier, 100k training examples are randomly sampled from the centerlines of the manually labeled ground truth segmentations. Each example is a randomly sampled path (i.e. a piece of centerline) with its length sampled from the distribution of edge pairs lengths in the graphs G_l created on the images in the training set.

The extremely randomized trees classifier has several hyperparameters which needs to be set. In a grid search, the combination of hyperparameter values is found that minimizes the log loss, as measured in a five fold cross-validation on the training set. The classifier is subsequently retrained on the full training set using the optimal hypervalues. The hyperparameters (and the considered values in the grid search) are: the minimum number of samples per leaf (20, 35, 50, 100), the number of trees (100) and the maximum number of features considered at each split (40%, 80% and the square root of the number of features).

2.6. Optimization and post processing

We formulated the joint segmentation and labeling problem as an integer program (IP) over \mathbf{x} subject to a set of constraints. We solve the resulting IP to provable optimality¹ using the branch-and-cut procedure of the Gurobi Optimizer (Gurobi Optimization Inc., 2015). To reduce computation time, we only consider edge pairs e_{ijk} for which $P(Y_{ijk} = 1 | I, G) \geq 0.1$ and labels l for edge pair e_{ijk} for which $P(L_{ijk} = l | T_{ijk} = 1, I, G) \geq 0.05$.

The solution \mathbf{x} , which says which edge pairs belong to the vasculature and what their labels are, needs to be converted to labeled centerlines. We create a graph $G_{\mathbf{x}} = (V_{\mathbf{x}}, E_{\mathbf{x}})$ containing the selected edge pairs: $V_{\mathbf{x}} = \{v_i | \exists v_j, v_k, l : x_{ijk}^l \vee x_{jik}^l \vee x_{kji}^l\}$ and $E_{\mathbf{x}} = \{e_{ij} | \exists k, l : x_{ijk}^l \vee x_{kij}^l\}$. Every edge e_{ij} receives a set of segment labels: $\{s\} = \cap \{l | (\exists k : x_{ijk}^l) \vee (\exists k : x_{kij}^l)\}$. Since an edge can span multiple anatomical segments, some edges will be assigned more than one segment label. By merging the spatially overlapping parts of the edges and their segment label sets, we get the labeled centerlines.

3. Results

We first evaluate the labeling and the segmentation performance of our algorithm separately, each against a state-of-the-art

approach for the respective task, and then report our combined performance. All experiments are done with a leave-one-image-out cross-validation. We use 50 MRA images of the cerebral vasculature of healthy volunteers, originating from a publicly available dataset (Bullitt et al., 2005). The MRA images are acquired on a 3 T MR unit (Allegra, Siemens Medical Systems Inc., Germany) with a voxel size of 0.5mm \times 0.5mm \times 0.8mm and dimensions of 448 \times 448 \times 128. Ground truth segmentations were provided by Bogunović et al. (2013), manually labeled (by D.R.) and reviewed by an expert (G.W.). The dataset is representative for the topological variability in the Circle of Willis, as was demonstrated by Bogunović et al. (2013). The images are rigidly aligned and cropped to the region that covers the segmentations (Bogunović et al., 2013).

Segmentation quality metric. The performance of centerline segmentation algorithms is often quantified by the precision and recall of centerlines points. For every centerline point of the resulting segmentation, the distance to the closest ground truth centerline point is calculated. Points closer than tolerance ρ are considered true positives (TP) while the others are considered false positives (FP). Vice versa, for every centerline point of the ground truth, the distance to the closest centerline point of the result is calculated, where points closer than ρ are again considered true positives (TP) and the others false negatives (FN). This allows calculating the precision ($TP/(TP + FP)$) and the recall ($TP/(TP + FN)$). The F1 score combines these two metrics and is defined as their harmonic mean: $\frac{2 \cdot \text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$.

However, we are interested in evaluating not only the overlap but also the topological correctness of our segmentation. We propose a new *local-to-global segmentation quality metric* that measures the quality of a centerline segmentation on both a local and a global scale. Instead of considering the precision and recall of centerline points, we look at the precision and recall of centerline paths with varying lengths, where a path is a piece of centerline that does not branch (although it might pass through a furcation). For very short paths, the measure is equivalent to the overlap measure explained earlier. For longer paths, the measure gives an idea of the correctness of topology of the segmentation. To calculate it, we sample random paths with varying lengths from both our segmentation and the ground truth segmentation. For every random path of either segmentation, we look for a corresponding path in the other segmentation. A path is corresponding if it has a Hausdorff distance less than ρ to the sampled path. A such, both the F1 score, precision and recall of a segmentation can be calculated for different path lengths (i.e. scales). In all experiments, we set ρ to 1.5 mm and take at each scale 100,000 random paths per image.

3.1. Anatomical labeling of ground truth centerlines

We first evaluate the anatomical labeling component of our algorithm in isolation: we start from a ground truth segmentation and perform labeling. Hereto, we construct the overcomplete graph G from the ground truth centerlines (instead of from the image). This graph is unlabeled, but contains only edges that belong to the vasculature. By setting $P(T_{ijk} = 1 | I, G) = 1$, we can use our algorithm to label this graph. The result is an edge labeled graph, from which we can infer the position of several named bifurcations (ICA-OA, ICA-M1, ICA-PCoM, ACA1-ACoM, M1-M2, VBA-SCA, VBA-PCA1, PCA1-PCoM, PCA2-PCA3). The positions are compared with those in the ground truth annotation. Again we introduce some tolerance: if the Euclidean distance is smaller than 2 mm, it is considered a true positive. Since we use the same dataset as Bogunović et al. (2013), we can directly compare the performance.

¹ During optimization, the branch-and-cut solver looks for the highest scoring solution while simultaneously proving the theoretically exact upper bound on the objective value. Once the relative difference between the best found objective value and the upper bound is smaller than the demanded solution gap, the optimization stops. We use the default solution gap of $1e^{-4}$. Hence, we have the guarantee that the found solution is extremely close to the global optimal, as close as demanded.

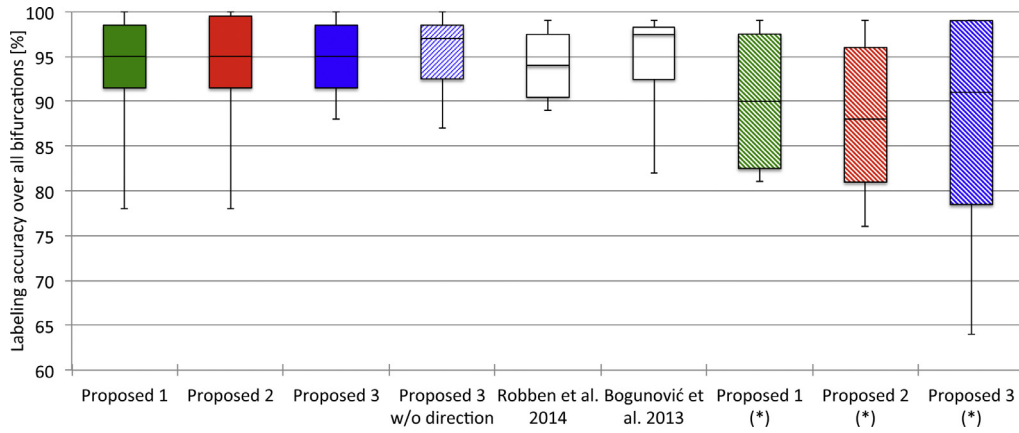


Fig. 5. Comparison of the labeling accuracy on the ground truth centerlines. For each method the distribution of the labeling accuracy of the different bifurcations is shown. An asterisk (*) indicates that randomly sampled vertices are used, see Section 3.1.2.

3.1.1. Vertices on furcations

In a first experiment we construct the overcomplete graph in a way that fully exploits the knowledge that it is constructed from the ground truth: the furcations of the ground truth centerlines are taken as the vertices of the overcomplete graph and additional vertices are added to ensure that edges of the overcomplete graph are about 5 mm long. Having edges of more or less equal length improves the classification accuracy. By sampling the ground truth furcations, we make sure that an edge pair label contains no more than two segment labels. The performance of anatomical labeling of ground truth centerlines was compared with the methods of Robben et al. (2014) and Bogunović et al. (2013). Fig. 5 shows the distribution of the labeling accuracy of the different bifurcations. Note that ‘Proposed 3 without direction’ uses the same objective function as Robben et al. (2014), but uses different features for the label classifier, hence the different results. Table B1 shows the accuracy, precision and recall for each method and for each bifurcation separately. For none of the bifurcations, the differences between the methods is significant, as determined by an unpaired permutation test ($P > 0.01$, without correction for multiple comparisons).

3.1.2. Randomly sampled vertices

When doing the simultaneous segmentation and labeling, the vertices are not placed on the exact positions of the furcations but on the positions with the highest tubularity values. To quantify the influence of this, we perform a second experiment. Now the vertices are randomly sampled from the centerlines in such a way that the distribution of the edge lengths is similar to the distributions found in overcomplete graphs constructed on real images. Results are shown in Fig. 5. Table B.2 shows the accuracy, precision and recall for each method and for each bifurcation separately. The labeling accuracy with randomly sampled vertices is significantly lower than with vertices on furcations ($P < 0.003$, without correction for multiple comparisons).

3.1.3. Stability of the constraints

In each fold of the leave-one-image-out cross-validation, the allowed label configurations are learned from a different subset of the dataset. We verified to which extent the allowed label configurations are stable throughout the cross-validation. The allowed label configurations are determined by the different sets of edge pair labels that are learned for each label l (see Appendix A: $L_l^{\text{incompatible}}$, B_l , $L_l^{\text{nobifurcation}}$, $L_l^{\text{non-terminal}}$, $L_l^{\text{notincoming}}$, and $L_l^{\text{nottouching}}$). Hence we can evaluate the stability of the constraints by evaluating the stability of these sets as obtained over all 50 folds. We quantify the

similarity between two sets A and B with the Dice similarity coefficient: $DSC(A, B) = \frac{|A \cap B|}{|A \cup B|}$. Then, for each label configuration separately (e.g. $L_l^{\text{nobifurcation}}$), we calculate the overall similarity as the mean DSC averaged over all edge pair labels l and over all pairwise comparisons between the 50 folds. Similarly, we calculate the minimal pairwise similarity: the minimal similarity of all pairwise comparisons where the similarity is the DSC after averaging over all edge pair labels. For each label configuration, the overall similarity was above 0.99 and the minimal pairwise dice was above 0.9.

3.2. Simultaneous segmentation and labeling: segmentation quality

We perform the simultaneous segmentation and labeling as described earlier. The resulting overcomplete graphs have on average 3563 edge pairs. The edge length is on average 4.3 mm and the 95-percentile is 7.9 mm. Likewise, an edge pair is on average 9.2 mm and has a 95-percentile of 15.1 mm. The integer problems have on average 19,599 binary variables and 237,405 constraints. Running single threaded on an AMD Opteron 6128 processor, the optimization takes on average 88 s for Proposed 1, 253 s for Proposed 2 and 510 s for Proposed 3.

We compare the segmentation quality of our method and the method of Türetken et al. (2013), which can be thought of as our algorithm without the simultaneous labeling. Both methods use the same overcomplete graph and the same estimations for $P(Y_{ijk} = 1 | I, G)$. To put these results into perspective, a second observer (D.R.) delineated three randomly selected images again. This gives an idea of the inter-observer variability. Finally, we also report the segmentation quality of the overcomplete graph. The local-to-global segmentation quality metric is calculated and results are shown in Fig. 6. While the differences between the three proposed methods are small, a paired permutation test shows that the F1 scores of Proposed 1 are significantly higher than the scores of Proposed 2 and Proposed 3 for the lengths between 0 and 10 mm ($P < 0.0001$, without correction for multiple comparisons).

To appreciate the improved objective function over the one presented in Robben et al. (2014), we show in Fig. 7 the results for the objective function without the direction term (or equivalently, with $P(D_{ijk} = d_{ijk} | Y_{ijk} = 1, I, G) = 0.5$). A paired permutation test shows that the F1 score is significantly lower without the direction term for lengths between 5 and 30 mm ($P < 0.0001$, without correction for multiple comparisons). Without the direction term, optimization takes on average 320 s for Proposed 1, 420 s for Proposed 2 and 808 s for Proposed 3.

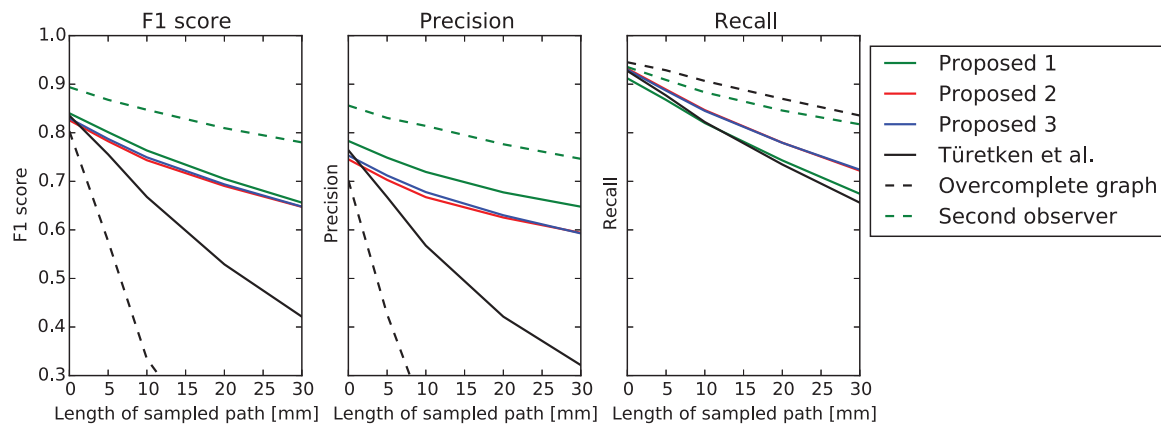


Fig. 6. The local-to-global reconstruction quality metric for both our method and Türetken et al. (2013).

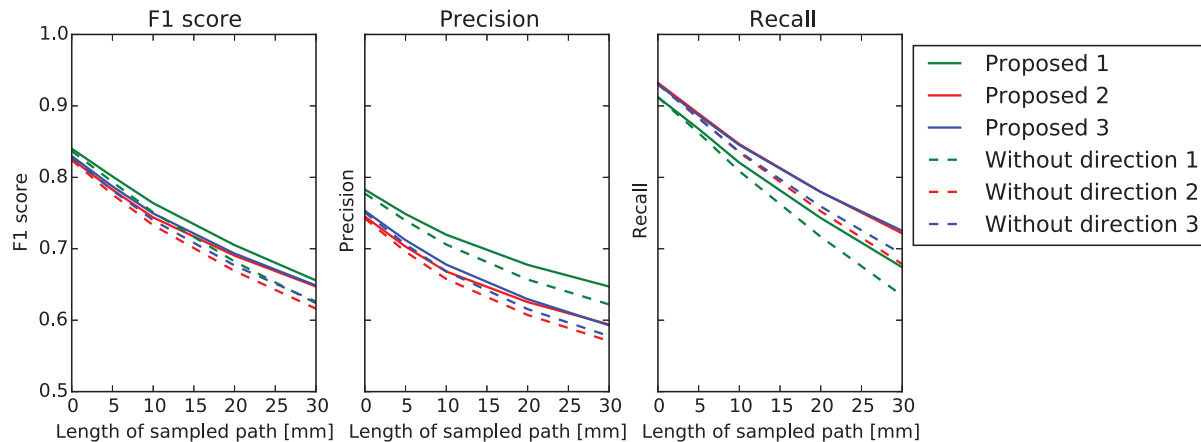


Fig. 7. The local-to-global reconstruction quality metric for our method both with and without the direction term. Note that *Without direction 3* corresponds to the objective function proposed in Robben et al. (2014).

3.3. Simultaneous segmentation and labeling: segmentation quality of the labeled segments

We report the precision and recall of the centerline points for each named vascular segment. The use of the local-to-global segmentation quality metric is not required, since named vascular segments always have the same topology: a non-branching curve. Results are given in Table 2. Finally, Fig. 8 shows automated segmentation and labeling results of our algorithm for two complete vasculatures.

4. Discussion and conclusion

We first assessed the performance of anatomical labeling of ground truth centerlines (shown in Table B1). The differences between the three proposed objective functions (Eq. 26, 27 and 28) are small, with the third one performing slightly, but not significantly, better than the other two. In comparison with Bogunović et al. (2013), a state of art method for anatomical labeling of the cerebral vasculature, we tend to have a higher precision but a lower recall, resulting in an accuracy that is on average about the same.

It should be noted that the method of Bogunović et al. (2013) requires topologically correct segmentations, and uses reference graphs explicitly stating bifurcation connectivity for the entire vasculature. Extending it to a larger number of bifurcations requires a steep increase in the number of reference graphs. For example, in their approach, inclusion of bifurcation of the left and right SCA (the VBA-SCA bifurcation), which can lie proximal or dis-

Table 2

Results of the simultaneous segmentation and labeling: precision (P) and recall (R) of the centerline points for every vascular segment.

Vascular segment	Proposed 1 (Eq. 26)		Proposed 2 (Eq. 27)		Proposed 3 (Eq. 28)	
	P	R	P	R	P	R
ICA-L	96	82	94	93	95	92
ICA-R	93	88	91	91	91	91
OA-L	86	48	67	77	65	77
OA-R	88	43	78	73	71	75
MCA1-L	94	81	94	80	84	99
MCA1-R	96	78	97	76	84	99
ACA1-L	97	94	97	96	96	98
ACA1-R	98	88	97	94	96	93
PComA-L	98	65	97	74	94	84
PComA-R	98	70	96	75	94	78
AComA	96	35	93	62	90	75
PCA1-L	95	80	94	85	87	87
PCA1-R	98	85	97	84	95	89
PCA2-L	95	88	95	85	91	87
PCA2-R	96	88	96	86	96	88

tal to the VBA-PCA1 bifurcation and not necessarily left-right symmetric, would result in a sixfold increase in the number of reference graphs from 8 to 48. In general, assuming that the topology of distant furcations is independent, their number of reference graphs increases exponentially with the number of furcations. In our approach, the constraints on the connections between certain labels

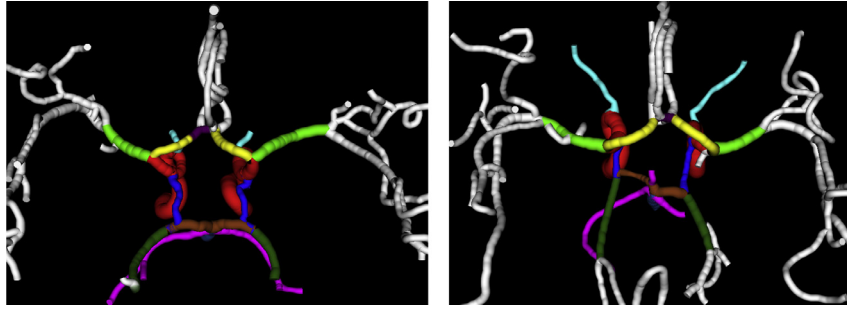


Fig. 8. Two automatically segmented and labeled vasculatures. As in Fig. 1, the colors indicate the anatomical labels of the vessel segments.

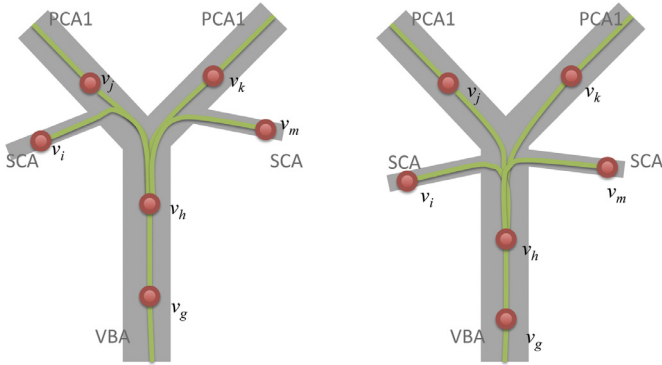


Fig. 9. The VBA-PCA1 and VBA-SCA bifurcations lie very close to each other, making it hard to consistently classify their composing edges. The edge pair e_{ghi} has in both cases very similar geometrical features but has different edge pair labels (left: VBA-PCA1-SCA, right: VBA-SCA). Because of the random sampling, the constraints cannot enforce a consistent labeling.

are both learned and enforced locally. This scales better to larger regions of interest while—as demonstrated—performing equally well.

When the ground truth segmentations are labeled with the vertices randomly sampled, we see a decrease in accuracy (Fig. 5). This is especially the case around the VBA-PCA1 bifurcation. The reason is that the VBA-PCA1 and the VBA-SCA bifurcation lie very close to each other (or sometimes even form a quadrifurcation) and are often represented by four partially overlapping edges for which the labels may not be consistent between subjects if the vertices are not coinciding with the actual furcations as illustrated in Fig. 9.

Thus, if the goal is to anatomically label ground truth centerlines, it is preferable to use vertices that are located on the furcations. However, when the centerlines are not given and segmentation has to be performed, it is not known upfront where the furcations are located. In this case, we have to use randomly sampled vertices, which still gives reasonably good labeling results.

Looking at the segmentation quality (Fig. 6), precision and recall on a local scale are relatively similar for the compared methods. On a larger scale, the precision of our approach clearly outperforms that of Türetken et al. (2013), while the recall stays about the same. This means that the segmentations of both methods overlap equally well with the ground truth, but that the topology of our solution is closer to that of the ground truth. The additional constraints regarding the connectivity of the different labels help inferring the correct topology in ambiguous cases such as kissing vessels, which is also illustrated in the close-ups of Fig. 10.

The direction term improves both the precision and the recall on every scale (Fig. 7), making the proposed method more accurate than the one of Robben et al. (2014). Additionally, it improves the calculation time. Without the direction term, the objective function

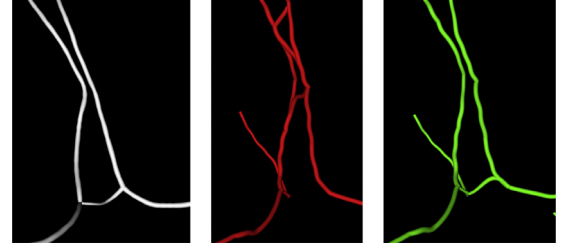


Fig. 10. The centerlines of the ACA1, AComA and ACA2 according to the ground truth (left), the method of Türetken et al. (middle), and the Proposed method 3 (right). Note the absence of spurious connections between the left and right ACA2 in our solution.

does not have a strong preference for the direction of the edges in the solution. On the other hand, the directions of the edges in the solution are strongly constrained by the connectedness requirement. As a result, there are invalid configurations that still have relatively high objective values. We think that the direction term penalizes such configurations, which makes the objective function sharper and the process of proving the optimality of the solution easier and thus faster.

In this work, we introduced three different cost functions. When comparing the segmentation quality of our three objective functions, we see that Proposed 2 and Proposed 3 perform similar. Proposed 1 has a higher precision, but a lower recall, resulting in an F1 score that is higher than the F1 score of the other two. From a theoretical point of view, we prefer Proposed 1: imagine a scenario where we would enlarge our field of view, and would segment and label additional vascular segments. If there is no interaction between these new segments and the existing ones, we would expect that the segmentation inside the original field of view remains the same. However, with Proposed 2 and 3, $w_{ijk,l}^{L2}$ (Eq. 32) and $w_{ijk,l}^{L3}$ (Eq. 33) will become larger, making it more favorable to include e_{ijk} in the solution and hence potentially altering the segmentation. In general, with increasing number of labels, Proposed 2 and 3 have the tendency to include more edges in the solution, which might explain the higher recall and lower precision observed in the experiments (see Fig. 6). In practice, we found no difference in labeling accuracy and a small but significant increase in the F1 score of the segmentation. Therefore, we suggest using the Proposed 1.

A crucial assumption for the method is that G_I is overcomplete. Fig. 6 shows that the overcomplete graph has—as expected—a very low precision and—as assumed—a very high recall. Even though the recall is not 100%, it is on all scales higher than the recall of the second human observer, suggesting that our assumption of covering all of the vasculature holds for practical purposes. Despite this, it might occur a vertex is placed on the position where two vessels kiss, which makes the correct reconstruction impossible.

In this work, we explore one particular way for creating the overcomplete graph. However, there are alternative approaches for obtaining an initial overcomplete segmentation of the vasculature. One could modify the current approach and use different tubularity values (Türetken et al., 2013; Schneider et al., 2014; Sironi et al., 2014) or a different cost function to connect the samples. Alternatively, one could use a greedy vessel tracking algorithm (Aylward and Bullitt, 2002; Wong and Chung, 2007; Gülsün and Tek, 2008; Friman et al., 2010) with a very high termination criterion to generate centerlines that serve as graph G_l . However, it is out of the scope of this work to evaluate which technique performs better at creating an overcomplete segmentation.

This work is based on the method of Türetken et al. (2013) but differs on several important points. First, the method now performs simultaneous anatomical labeling, which has ramifications on both the objective function and the constraints. Not only does this result in an anatomically labeled segmentation, it also improves the segmentation itself. Additionally, it allows the method to be fully automated: it removes the need for a manually indicated start point for each vascular structure. Second, we include the direction term in the objective function, which also improves the segmentation. Finally, we calculate the tubularity image differently, in a way that is better suited for the studied MRA images.

One could argue that this method is not truly simultaneously segmenting and labeling the vasculature, but rather first segmenting the structure (creating the overcomplete graph) and then removing unlikely connections while labeling. This would consider the step of creating the overcomplete graph as a segmentation. However, we think that our algorithm selects and labels the set of tubular segments that best explains the image data and we consider the overcomplete graph as a smartly selected subset from the space of all tubular segments. We do not expect the overcomplete graph to be a proper segmentation, just that it contains a good segmentation. This is illustrated by the fact that less than 30% of its 10 mm paths are correct (Fig. 6).

To the best of our knowledge, we presented the first algorithm for simultaneous segmentation and anatomical labeling of the vasculature. Our maximum a posteriori formulation results in an integer program and the globally optimal solution is found by a branch-and-cut algorithm. We evaluated our approach on a publicly available dataset of 50 cerebral MRA images. We showed the feasibility and the added value of incorporating knowledge about the anatomical structure into the segmentation process and demonstrated that our approach compares favorably against specialized state-of-the-art algorithms that address the segmentation and labeling problems separately.

However, there are still opportunities to improve this work. A first issue is the robustness against inter scanner variability. The classifiers that evaluate the probability that a centerline belongs to the vasculature depend on the absolute values of the image intensities. In MRA images, the intensities do not have a physical unit and depend on the specific scanner and protocol used. In this work, we used data from a single scanner; to handle data from a different scanner, these classifiers should be retrained. However, it would be interesting to explore to what extent a diverse training set in combination with intensity and feature standardization can overcome this need. Alternatively, one could resort to more advanced techniques such as transfer learning (Becker et al., 2013).

We focused on the cerebral vasculature in MRA images, but application of the algorithm on other modalities or organs is possible and interesting. In general, this would require a dataset of images of that organ and in that modality along with manually segmented and labeled centerlines. The dataset should be large enough to capture the natural variation of the anatomy. Depending on the modality, it might be beneficial to use different photometric features. In case only the modality changes and the organ of interest

remains the cerebral vasculature, it might be possible to reuse the anatomical knowledge learned from the MRA images. For example, for CT angiography (CTA) images in the arterial phase, it can be assumed that the geometric properties of the vasculature are similar to those in MRA. Hence, after registration of the CTA images to the same coordinate system as the MRA images, we could reuse the direction classifier, the label classifier and the rules on the labeled configurations. The photometric appearance of vessels in CTA is different than in MRA, such that we need to retrain all classifiers that use photometric properties.

Finally, in our method, the observed patterns in the vessel connectivity are learned and enforced as hard constraints. While proven to be efficient on our database of healthy patients, they might be too restrictive for diseased vasculature e.g. patients with severely stenosed arteries. Working with less labels or the use of soft constraints might be necessary here.

Acknowledgement

David Robben is supported by a Ph.D. fellowship of the Research Foundation - Flanders (FWO). The MRA brain images from healthy volunteers used in this paper were collected and made available by the CASILab at The University of North Carolina at Chapel Hill and were distributed by the MIDAS Data Server at Kitware Inc. We would also like to thank the authors of Bogunović et al. (2013) for sharing their centerline delineations. This work was supported by KU Leuven's Concerted Research Action GOA/11/006 and by the EU ERC project MicroNano.

Appendix A. Constraints

A1. Consistency and connectedness

An active edge pair, i.e. an edge pair that is in the solution, can have only one edge pair label:

$$\forall e_{ijk} \in F : \sum_{l \in \mathcal{L}} x_{ijk}^l \leq 1. \quad (\text{A.1})$$

An edge e_{ij} can have at most one active incoming edge pair e_{hij} , as also illustrated in Fig. 3:

$$\forall e_{ij} \in E : \sum_{e_{hij} \in F, l \in \mathcal{L}} x_{hij}^l + \sum_{e_{hji} \in F, l \in \mathcal{L}} x_{hji}^l \leq 1. \quad (\text{A.2})$$

We want the segmentation to be geometrically consistent, that is: prohibit that the solution contains edges that occupy the same physical 3D space. Hereto, we use the disjoint edges constraint from Türetken et al. (2013), which lets at most one edge pair be active among all those that either contain e_{ij} or overlap with it.

Finally, all segment labels (except for 'NA' of course) correspond to anatomical names of arteries, which are coherent anatomical structures. By anatomical naming convention, two different segments are either not connected, or are connected on a single place—typically a bifurcation. Now, recall that an edge pair label l is a tuple of segment labels: $l = (s_1, \dots, s_n)$, $s_i \in S$. Thus, if an edge pair e_{ijk} in the solution has edge pair label l that contains (s_a, s_b) , then the connection between s_a and s_b is somewhere on the edge e_{ij} or on the edge e_{jk} , and not somewhere else.

By introducing the following notation:

$$(s_a, s_b) \in l \Rightarrow \exists i : l_i = s_a \wedge l_{i+1} = s_b, \quad (\text{A.3})$$

this knowledge can be enforced as:

$$\text{connections} = \{(s1, s2) | s1 \in S, s2 \in S\} \quad (\text{A.4})$$

$$\forall e_{ghi} \in F, e_{jkm} \in F, |\{v_g, v_h, v_i\} \cap \{v_j, v_k, v_m\}| \leq 1,$$

$$c \in \text{connections} : \sum_{\substack{l \in \hat{L} \\ c \in l}} x_{ghi}^l + \sum_{\substack{l \in \hat{L} \\ c \in l}} x_{jkm}^l \leq 1. \quad (\text{A.5})$$

This approach requires $O(|F|^2)$ constraints. We use an approximation, only considering edge pairs that share at least one vertex, since those are most likely to violate this constraint.

$$\forall e_{ghi} \in F, e_{jkm} \in F, |\{v_g, v_h, v_i\} \cap \{v_j, v_k, v_m\}| = 1,$$

$$c \in \text{connections} : \sum_{\substack{l \in \hat{L} \\ c \in l}} x_{ghi}^l + \sum_{\substack{l \in \hat{L} \\ c \in l}} x_{jkm}^l \leq 1. \quad (\text{A.6})$$

A1.1. Connectedness

To ensure the connectedness of the vascular segments, we enforce the active edge pairs to be connected to the virtual vertex v_v . An edge pair e_{ijk} is connected if there is a path, containing only active edge pairs, starting from v_v and containing e_{ijk} . Türetken et al. (2013) enforce this by introducing additional flow variables to explicitly count the number of distinct paths to each vertex and through every edge. However, this is computationally very demanding and therefore we propose using a local approximation, enforcing that an active edge pair needs another active incoming edge pair (except if the edge pair starts at the virtual root vertex):

$$\forall e_{ijk} \in F, v_i \neq v_v : \sum_{l \in \hat{L}} x_{ijk}^l \leq \sum_{e_{hij} \in F, l \in \hat{L}} x_{hij}^l. \quad (\text{A.7})$$

This constraint is a local approximation of the true connectedness constraint, since it allows to have a circle of edge pairs that is not connected to the virtual root, e.g. e_{ghi} , e_{hij} , e_{ijg} . However, we did not encounter this in practice.

A2. Learned label configurations

Although there is considerable inter-subject variation in the cerebral vasculature, some patterns hold invariably. The algorithm learns these patterns from the training data and enforces them when segmenting and labeling unseen images. The patterns can be expressed as rules: they state the edge pair labels that are possible for a certain edge pair, given the edge pair labels of certain neighboring edge pairs.

Following, $L_l^{\text{incompatible}} \subseteq \hat{L}$ is the set of the edge pair labels that are in the training data never observed on an edge pair e_{ijk} that follows an edge pair e_{hij} with edge pair label l . This knowledge is expressed in the constraint:

$$\forall e_{ijk} \in F, l \in \hat{L} : \sum_{e_{hij} \in F} x_{hij}^l + \sum_{l' \in L_l^{\text{incompatible}}} x_{ijk}^{l'} \leq 1. \quad (\text{A.8})$$

Bifurcation 1. Some edge pair labels can occur only in a bifurcation or due to the sampling right after the bifurcation (illustrated for ICA – M1 in Fig. A1). We introduce $B_l = (\{l_1\}, \{l_2\})$, two sets of edge pair labels. A label l is only observed on an edge pair e_{fgh} if there is another edge pair e_{fgm} with a label l_1 or if there is an edge pair e_{mfg} with label l_2 .

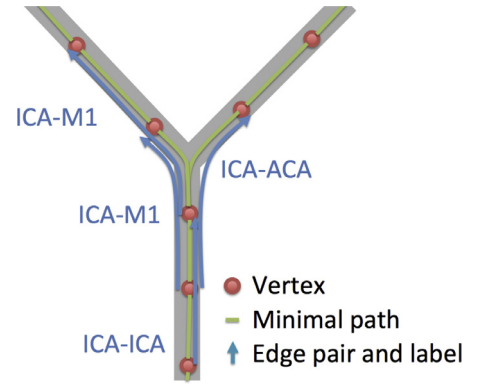
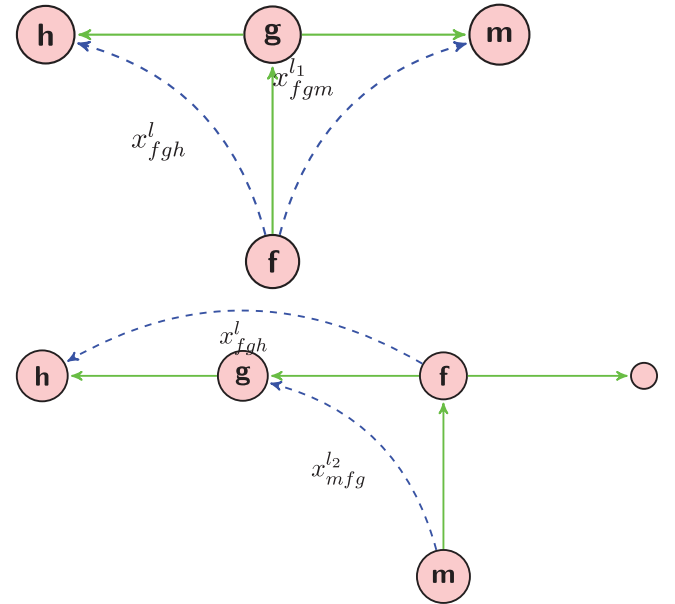


Fig. A1. Illustration of the configuration of the edge pair labels in the bifurcation of the ICA into the M1 and A1. ICA – A1 can only occur when there is a bifurcation (ICA followed by A1 and M1) or it's preceded by ICA – A1



This is enforced as:

$$\forall l \in \hat{L}, e_{fgh} \in F : x_{fgh}^l \Rightarrow \left(\sum_{e_{fgm} \in F, l_1 \in B_l[1]: m \neq h} x_{fgm}^{l_1} \geq 1 \right) \vee \left(\sum_{e_{mfg} \in F, l_2 \in B_l[2]} x_{mfg}^{l_2} = 1 \right). \quad (\text{A.9})$$

Rewritten as a linear inequality:

$$\forall l \in \hat{L}, e_{fgh} \in F : x_{fgh}^l \leq \sum_{e_{fgm} \in F, l_1 \in B_l[1]: m \neq h} x_{fgm}^{l_1} + \sum_{e_{mfg} \in F, l_2 \in B_l[2]} x_{mfg}^{l_2}. \quad (\text{A.10})$$

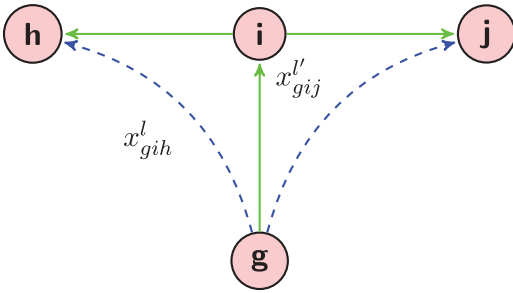
Bifurcation 2. Some edge pair labels can never occur in a bifurcation. Now, $L_l^{\text{nobifurcation}} \subseteq \hat{L}$ is the set of edge pair labels that are

Table B1

Comparison of the labeling performance on the ground truth centerlines. Accuracy (A), precision (P) and recall (R) are reported for the proposed methods with the vertices placed on the furcations and for the method of Bogunović et al. (2013). An unpaired permutation test between the proposed methods and Bogunović et al. (2013) shows no significant differences in labeling accuracy for any of the bifurcations.

Bifurcation	Proposed 1 (Eq. 26)			Proposed 2 (Eq. 27)			Proposed 3 (Eq. 28)			Proposed 3 w/o direction (Eq. 34)			Robben et al. 2014			Bogunović et al. 2013		
	A	P	R	A	P	R	A	P	R	A	P	R	A	P	R	A	P	R
ICA-OA	99	99	100	99	99	100	99	99	100	99	99	100	99	100	99	–	–	–
ICA-M1	100	100	100	100	100	100	100	100	100	100	100	100	99	99	100	99	100	99
ICA-PComA	98	100	98	100	97	97	98	100	97	98	100	97	93	94	96	97	98	98
ACA1-AComA	96	100	95	96	100	95	96	100	95	96	100	95	92	93	97	98	97	100
M1-M2	78	78	100	78	78	100	88	88	100	87	87	100	89	89	100	82	82	100
VBA-SCA	95	99	96	95	99	96	95	99	96	97	99	98	95	98	97	–	–	–
VBA-PCA1	90	98	91	90	98	91	90	98	91	92	98	93	94	100	93	96	96	100
PCA1-PComA	95	100	92	95	100	92	95	100	92	97	100	95	96	100	94	98	97	100
PCA2-PCA3	93	94	97	93	94	97	93	94	97	93	95	96	89	92	93	–	–	–

never observed to be on an edge pair e_{gij} that has the same first edge as an edge pair e_{gih} with edge pair label l :



$$\forall e_{gih} \in F, l \in \hat{L}, e_{gij} \in F, v_j \neq v_h : x_{gih}^l + \sum_{l' \in L_i^{nobi\text{furcation}}} x_{gij}^{l'} \leq 1 \quad (\text{A.11})$$

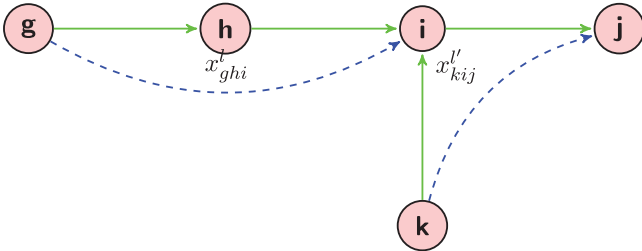
or:

$$\forall x_{gih}^l \in X, e_{gij} \in F, v_j \neq v_h : x_{gih}^l + \sum_{l' \in L_i^{nobi\text{furcation}}} x_{gij}^{l'} \leq 1. \quad (\text{A.12})$$

Terminating. $L^{non-terminal} \subseteq \hat{L}$ is the set of edge pair labels that are never observed on an edge pair that is not followed by another edge pair. Some edge pair labels cannot terminate but should be continued:

$$\forall l \in L^{non-terminal}, e_{hij} \in F : x_{hij}^l \leq \sum_{e_{kjm} \in F, l' \in \hat{L}} x_{kjm}^{l'}. \quad (\text{A.13})$$

Non-tree 1. An active edge pair cannot always end in the middle of another active edge pair. $L_i^{notincoming} \subseteq \hat{L}$ is the set of edge pair labels which are never observed on an active edge pair whose middle vertex is the last vertex of an active edge pair with edge pair label l .



We state:

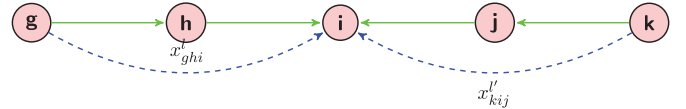
$$\forall e_{kij} \in F, e_{hi} \in E, l \in \hat{L} : \sum_{e_{ghi} \in F, h \neq k} x_{ghi}^l + \sum_{l' \in L_i^{notincoming}} x_{kij}^{l'} \leq 1. \quad (\text{A.14})$$

Table B2

Comparison of the labeling performance on the ground truth centerlines. Accuracy (A), precision (P) and recall (R) are reported for the proposed method with the vertices sampled randomly.

Bifurcation	Proposed 1 (Eq. 26)			Proposed 2 (Eq. 27)			Proposed 3 (Eq. 28)		
	A	P	R	A	P	R	A	P	R
ICA-OA	99	100	99	99	100	99	99	100	99
ICA-M1	95	96	99	96	96	100	99	99	100
ICA-PComA	96	97	97	95	97	96	95	96	97
ACA1-AComA	83	92	86	81	91	83	75	85	79
M1-M2	81	81	100	81	81	100	89	89	100
VBA-SCA	83	98	85	81	98	83	82	100	82
VBA-PCA1	82	100	80	76	92	79	64	88	67
PCA1-PComA	99	100	98	96	100	94	99	100	98
PCA2-PCA3	90	97	90	88	94	89	91	94	95

Non-tree 2. An active edge pair cannot always end at the end of another active edge pair. $L_i^{nottouching} \subseteq L$ is the set of edge pair labels which are never observed on an active edge pair whose last vertex is the last vertex of an active edge pair with edge pair label l .



We state:

$$\forall e_{hi} \in E, e_{ji} \in E, l \in \hat{L}, j \neq h : \sum_{e_{ghi} \in F} x_{ghi}^l + \sum_{l' \in L_i^{nottouching}, e_{kji} \in F} x_{kji}^{l'} \leq 1. \quad (\text{A.15})$$

Appendix B

See Table B1 and B2 for detailed results of anatomical labeling of ground truth centerlines.

References

- Aylward, S.R., Bullitt, E., 2002. Initialization, noise, singularities, and scale in height ridge traversal for tubular object centerline extraction. *IEEE Trans. Med. Imaging* 21 (2), 61–75.
- Becker, C.J., Christoudias, C.M., Fua, P., 2013. Non-linear domain adaptation with boosting. In: Burges, C., Bottou, L., Welling, M., Ghahramani, Z., Weinberger, K. (Eds.), *Advances in Neural Information Processing Systems*. Curran Associates, Inc., Lake Tahoe, NV, pp. 485–493.
- Bilgel, M., Roy, S., Carass, A., Nyquist, P.A., Prince, J.L., 2013. Automated anatomical labeling of the cerebral arteries using relief propagation. In: Sebastien Ourselin, D.R.H. (Ed.), *SPIE Medical Imaging*, 8669, pp. 1–6.
- Bogunović, H., Pozo, J.M., Cárdenes, R., Frangi, A.F., 2011. Anatomical labeling of the anterior circulation of the circle of willis using maximum a posteriori classification. In: Fichtinger, G., Martel, A., Peters, T. (Eds.), *MICCAI*. Springer Berlin Heidelberg, Toronto, Canada, pp. 330–337.

- Bogunović, H., Pozo, J.M., Cardenes, R., Roman, L.S., Frangi, A.F., 2013. Anatomical labeling of the circle of willis using maximum a posteriori probability estimation. *IEEE Trans. Med. Imaging* 32 (9), 1587–1599.
- Brisman, J.L., Song, J.K., Newell, D.W., 2006. Cerebral aneurysms. *N. Engl. J. Med.* 355 (9), 928–939.
- Buelow, T., Lorenz, C., Wiemker, R., Honko, J., 2006. Point based methods for automatic bronchial tree matching and labelling. In: Manduca, A., Amini, A.A. (Eds.), *SPIE Medical Imaging*, 6143, p. O1430.
- Bullitt, E., Zeng, D., Gerig, G., Aylward, S., Joshi, S., Smith, J.K., Lin, W., Ewend, M.G., 2005. Vessel tortuosity and brain tumor malignancy: a blinded study. *Acad. Radiol.* 12 (10), 1232–1240.
- Deschamps, T., Cohen, L.D., 2001. Fast extraction of minimal paths in 3D images and applications to virtual endoscopy. *Med. Image Anal.* 5, 281–299.
- Feragen, A., Petersen, J., Owen, M., Lo, P., Thomsen, L.H., Wille, M.M.W., Dirksen, A., de Bruijne, M., 2012. A hierarchical scheme for geodesic anatomical labeling of airway trees. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (Eds.), *MICCAI*. Springer Berlin Heidelberg, Nice, pp. 147–155.
- Frangi, A., Niessen, W., Vincken, K., 1998. Multiscale vessel enhancement filtering. In: Wells, W.M., Colchester, A., Delp, S. (Eds.), *MICCAI*. Springer Berlin Heidelberg, Cambridge, MA, pp. 130–137.
- Friman, O., Hindennach, M., Kühnel, C., Peitgen, H.O., 2010. Multiple hypothesis template tracking of small 3D vessel structures. *Med. Image Anal.* 14 (2), 160–171.
- Geurts, P., Ernst, D., Wehenkel, L., 2006. Extremely randomized trees. *Mach. Learn.* 63 (1), 3–42.
- Ghanavati, S., Lerch, J.P., Sled, J.G., 2014. Automatic anatomical labeling of the complete cerebral vasculature in mouse models. *NeuroImage* 95, 117–128.
- van Gijn, J., Rinkel, G.J., 2001. Subarachnoid haemorrhage: diagnosis, causes and management. *Brain* 124 (Pt. 2), 249–278.
- van Ginneken, B., Baggerman, W., Van Rikxoort, E.M., 2008. Robust segmentation and anatomical labeling of the airway tree from thoracic CT scans. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (Eds.), *MICCAI*. Springer Berlin Heidelberg, New York, NY, pp. 219–226.
- González, G., Türetken, E., Fleuret, F., Fua, P., 2010. Delineating trees in noisy 2D images and 3D image-stacks. In: *CVPR*. San Francisco, CA, pp. 2799–2806.
- Graham, M., 2006. Robust graph-theoretic methods for matching and labeling anatomical trees. The Pennsylvania State University. Ph.D. thesis.
- Gülsün, M.A., Tek, H., 2008. Robust vessel tree modeling. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (Eds.), *MICCAI*. Springer Berlin Heidelberg, New York, NY, pp. 602–611.
- Gurobi Optimization Inc., 2015. *Gurobi Optimizer Reference Manual*.
- Hendrikse, J., Hartkamp, M.J., Hillen, B., Mali, W.P., van der Grond, J., 2001. Collateral ability of the circle of Willis in patients with unilateral internal carotid artery occlusion: border zone infarcts and clinical symptoms. *Stroke* 32 (12), 2768–2773.
- Hua, L., Yezzi, A., 2007. Vessels as 4-D curves: global minimal 4-D paths to extract 3-D tubular surfaces and centerlines. *IEEE Trans. Med. Imaging* 26 (9), 1213–1223.
- Krabbe-Hartkamp, M.J., van der Grond, J., de Leeuw, F.E., de Groot, J.C., Algra, A., Hillen, B., Breteler, M.M., Mali, W.P., 1998. Circle of willis: morphologic variation on three-dimensional time-of-flight MR angiograms. *Radiol.* 207 (1), 103–111.
- Law, M.W.K., Chung, A.C.S., 2010. An oriented flux symmetry based active contour model for three dimensional vessel segmentation. In: Daniilidis, K., Maragos, P., Paragios, N. (Eds.), *ECCV*. Springer Berlin Heidelberg, Heraklion, Greece, pp. 720–734.
- Lesage, D., Angelini, E.D., Bloch, I., Funka-Lea, G., 2009. A review of 3D vessel lumen segmentation techniques: models, features and extraction schemes. *Med. Image Anal.* 13 (6), 819–845.
- Lu, L., Bi, J., Yu, S., Peng, Z., Krishnan, A., Zhou, X.S., 2009. Hierarchical learning for tubular structure parsing in medical imaging: a study on coronary arteries using 3D CT Angiography. In: *IEEE 12th International Conference on Computer Vision*, 2009. Kyoto, pp. 2021–2028. doi:10.1109/ICCV.2009.5459445.
- Matsuzaki, T., Oda, M., Kitasaka, T., Hayashi, Y., Misawa, K., Mori, K., 2015. Automated anatomical labeling of abdominal arteries and hepatic portal system extracted from abdominal CT volumes. *Med. Image Anal.* 20 (1), 152–161.
- Mori, K., Oda, M., Egusa, T., Jiang, Z., Kitasaka, T., Fujiwara, M., Misawa, K., 2010. Automated nomenclature of upper abdominal arteries for displaying anatomical names on virtual laparoscopic images. In: Liao, H., Edwards, P.J., Pan, X., Fan, Y., Yang, G.-Z. (Eds.), *MIAR*. Springer Berlin Heidelberg, Beijing, pp. 353–362.
- Mori, K., Ota, S., Deguchi, D., Kitasaka, T., Suenaga, Y., Iwano, S., Hasegawa, Y., Takabatake, H., Mori, M., Natori, H., 2009. Automated anatomical labeling of bronchial branches extracted from CT datasets based on machine learning and combination optimization and its application to bronchoscope guidance. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (Eds.), *MICCAI*. Springer Berlin Heidelberg, London, United Kingdom, pp. 707–714.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: machine learning in python. *J. Mach. Learn. Res.* 12, 2825–2830.
- Rempfler, M., Schneider, M., Ielacqua, G.D., Xiao, X., Stock, S.R., Klops, J., Székely, G., Andres, B., Menze, B.H., 2015. Reconstructing cerebrovascular networks under local physiological constraints by integer programming. *Med. Image Anal.* 1–9.
- Robben, D., Sunaert, S., Thijs, V., Wilms, G., Maes, F., Suetens, P., 2013. Anatomical labeling of the Circle of Willis using maximum a posteriori graph matching. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (Eds.), *MICCAI*. Springer Berlin Heidelberg, Nagoya, Japan, pp. 566–573.
- Robben, D., Menze, B.H., 2015. Simultaneous segmentation and anatomical labeling of the cerebral vasculature. In: Golland, P., Hata, N., Barillot, C., Hornegger, J., Howe, R. (Eds.), *MICCAI*. Springer Berlin Heidelberg, Cambridge, MA, pp. 307–314.
- Schneider, M., Hirsch, S., Weber, B., Székely, G., Menze, B.H., 2014. Joint 3-D vessel segmentation and centerline extraction using oblique Hough forests with steerable filters. *Med. Image Anal.* 19 (1), 220–249.
- Sironi, A., Lepetit, V., Fua, P., Epfl, D.L., 2014. Multiscale Centerline detection by learning a scale-space distance transform. In: *CVPR*. Columbus, Ohio, USA.
- Tschirren, J., McLennan, G., Palágyi, K., Hoffman, E.A., Sonka, M., 2005. Matching and anatomical labeling of human airway tree. *IEEE Trans. Med. Imaging* 24 (12), 1540–1547.
- Türetken, E., Becker, C., Glowacki, P., Benmansour, F., Fua, P., 2013. Detecting irregular curvilinear structures in gray scale and color imagery using multi-directional oriented flux. In: *ICCV*, pp. 1553–1560.
- Türetken, E., Benmansour, F., Andres, B., Pfister, H., Fua, P., 2013. Reconstructing loopy curvilinear structures using integer programming. In: *CVPR*. Portland, OR, pp. 1822–1829.
- Türetken, E., Blum, C., González, G., Fua, P., 2010. Reconstructing geometrically consistent tree structures from noisy images. In: Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (Eds.), *MICCAI*. Springer Berlin Heidelberg, Beijing, China, pp. 291–299.
- Türetken, E., González, G., Blum, C., Fua, P., 2011. Automated reconstruction of dendritic and axonal trees by global optimization with geometric priors. *Neuroinformatics* 9 (2–3), 279–302.
- Viola, P., Jones, M., 2001. Rapid object detection using a boosted cascade of simple features. In: *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. CVPR 2001, 1 1–511–518.
- Wong, W.C.K., Chung, A.C.S., 2007. Probabilistic vessel axis tracing and its application to vessel segmentation with stream surfaces and minimum cost paths. *Med. Image Anal.* 11 (6), 567–587.