Probabilistic Labeling of Cerebral Vasculature on MR Angiography

Benjamin Quachtran, Sunil Sheth, Jeffrey L. Saver,
David S. Liebeskind, and Fabien Scalzo[(✉)]

Stroke Program, Department of Neurology,
University of California, Los Angeles (UCLA), Los Angeles, USA
fscalzo@mednet.ucla.edu

Abstract. Automatic labeling of cerebrovascular territories would greatly advance our ability to systematically study large datasets and also provide rapid decision support during the assessment of stroke patients, for example. Previous attempts have been challenged by the wide inter-subject variation in vascular topography. We investigate the use of a probabilistic model that learns the configurational characteristics of vascular territories to better annotate the cerebrovasculature. In the George Mason Brain Vasculature database, we identified patients with MRA reconstructions segmented into seven major regions (left and right MCA, PCA, and ACA and Circle of Willis). We then augmented these labels by manually segmenting the MCA territory into an additional eight regions. Among 54 patients that met the inclusion criteria, 39 reconstructions were used as training input to the MCA, ACA, and PCA model among the 61 digital reconstructions of human brain arterial structures available. The model was then validated on an independent cohort of 15 patients. The MCA segmentation algorithm was trained and tested using leave-one-out crossvalidation. The algorithm was found to be $94 \pm 5.2\,\%$ accurate in annotating the seven major regions and $88 \pm 9.3\,\%$ accurate in annotating the MCA subterritories.

1 Introduction

Advancements in medical imaging and computational methods have led to a new wave of algorithmic-based techniques to analyze patient data. In the past, cerebral scans have been studied in a more case-by-case basis, leaving the interpretations subject to debate and reducing analysis efficiency. The inherent lack of quantification limits the recognition of correlations and patterns, obscuring meaningful clinical findings within the noise. Digitizing data and automating interpretation paves the way for deeper and more objective analysis, and in the case of cerebral angiography, moves towards more unbiased assessments that can directly improve clinical decision-making.

Previous attempts at automated 3D vessel segmentation [14] and labeling [1,3,15] have been challenged by the difficulties related to the quality of the images and the wide inter-subject variability that normally occurs. Image processing

and computer vision methods can play a major role in solving these issues if the algorithms are well trained to handle highly variable data. A study [2] conducted at the University of Toronto has investigated the use of Bayesian-based inference to label cerebral vasculature of mice. An arterial set is initialized with random labels and then a relaxation algorithm iteratively re-labels segments in manner that minimizes the data's energy function, a negated summation of each label's posterior probability. Despite being $\geq 75\,\%$ accurate in recognizing each of the sample's 54 territories, the method took an average of $100 \pm 18\,$h to segment each image. Recently, promising advances [13] have also been achieved for the automatic segmentation and labeling of the arterial tree from whole-body MRA data. Labeling of the vascular tree was performed by using a combination of graph-based and atlas-based approaches.

The characteristics chosen as inputs in the training data have proven to be a critical element in determining the robustness of an algorithm. Shape recognition, through use of principal component analysis, for example, provides a reasonably effective and computationally efficient way to quantify cerebral territories in segmenting the majority of vasculature sets [5]. The dependence on consistent arterial structure made by this technique is not resilient to the wide variation in morphological structure present in cerebral anatomy. Instead, selecting spatial features as inputs seem to lead to better results when performing vascular segmentation because of the visual nature of territorial labeling [2]. Labeling studies based on other modalities such as CT [6] have successfully used various features such as diameter, curvature, direction, and running vectors of a branch to infer the labels on abdominal arteries.

In this paper, we propose and test a method to automatically label the major vascular territories of cerebral reconstructions through use of kernel density estimation (KDE) and Bayesian inference. Through this form of probabilistic estimation, we are able to create accurate approximations of multi-dimensional density functions describing the cartesian coordinates of the right and left MCA, PCA, and ACA regions. From these probabilistic distributions, each vascular segment is evaluated as a function of the likelihoods of each sub-segment. A maximum a posteriori algorithm decides the appropriate label based on the optimal characterization of the evaluation.

2   Methods

2.1   Dataset Acquisition and Properties

The 3D arterial models were collected from the George Mason Brain Vasculature database [17] which is freely available at http://cng.gmu.edu/brava/ and contains 61 digital reconstructions created from magnetic resonance angiography (MRA) scans of healthy adult subjects (mean age, 31 years; age range, 19 to 64; 36 women). The dataset was created using Neuron_Morpho, a plug-in of ImageJ [5]. Through Neuron_Morpho, users are able to trace neurons contained in image stacks and export them in swc format, representing the cerebral vasculature as a series of interconnected cylindrical segments (segments have

a specified x, y, and z coordinate, type, radius, and parent segment). Following the use of this software, the reconstructions were additionally verified for accuracy through juxtaposition with 3D renderings of the MRA scans. In the Brain Vasculature database, each reconstruction is available both unlabeled and labeled, with the labeled models identifying seven major regions (Fig. 1): Circle of Willis and the left and right middle cerebral artery (MCA), posterior cerebral artery (PCA), and anterior cerebral artery (ACA). Vaa3D [9], an open source tool for 3D bio-image visualization and editing, was used to render and view these models.

While the arterial data from BraVa is pre-segmented into the seven aforementioned regions, we are interested in further annotating the MCA territory, a site of interest in stroke patients. To create the necessary training data, we used Vaa3D's built-in neuron utilities to expand the MCA territory into eight additional regions [7] (Fig. 2): (1) Posterior Temporal, (2) Temporo-Occipital, (3) Angular, (4) Posterior Parietal, (5) Rolandic, (6) Precental, (7) Prefrontal, and (8) Orbitofrontal. A neurologist from UCLA manually labeled the MCA regions of training images. Vasculatures in the database were co-registered using landmark points placed manually so that the relative location of the vessel segments could be used. Data registration is necessary to properly characterize the vasculatures anatomical locations with the respective territories.

2.2   Bayesian Labeling Framework

We assume that the labeling framework is presented with a set of unlabeled vessel segments $S_{i=1...n}$ that can be represented as a tree-structured graph. Each segment S_i is characterized by a state x_i in the model that represents the label probabilities of vascular territories. Because of the natural variations occurring in the vasculature across subjects, the number of states N for a specific model varies $[x_1, \ldots, x_N]$. A state x_i is a M-dimensional discrete vector that associates a probability to each of the M possible labels of the segment.

To each state x_i is associated an observation y_i, directly obtained from the normalized location of the segment and its radius. It differs from the state x_i in the sense that it comes vessel detectors that can be affected by noise and transient artifacts present in the image, whereas the value x_i is obtained through inference, thus believed to be more robust.

The labeling model is assumed to follow the general properties of a tree-structured Markov model where each state x_i is connected to at most one parent state x_i^p and can have several children states $x_{i,1...N_c}^c$ where $N_c \geq 0$. This means that the probability of a segment x_i given all the states available $x_{1...N}$ depends only its parent and children states,

$$p(x_i|x_{\{1...N\}}) = p(x_i) \, p(x_i|x_i^p) \, p(x_i|x_{i,1...N_c}^c) \tag{1}$$

where $p(x_i)$ is the prior distribution and $p(x_i|x_j)$ represents the conditional dependency between two connected vessel segments. By introducing observations
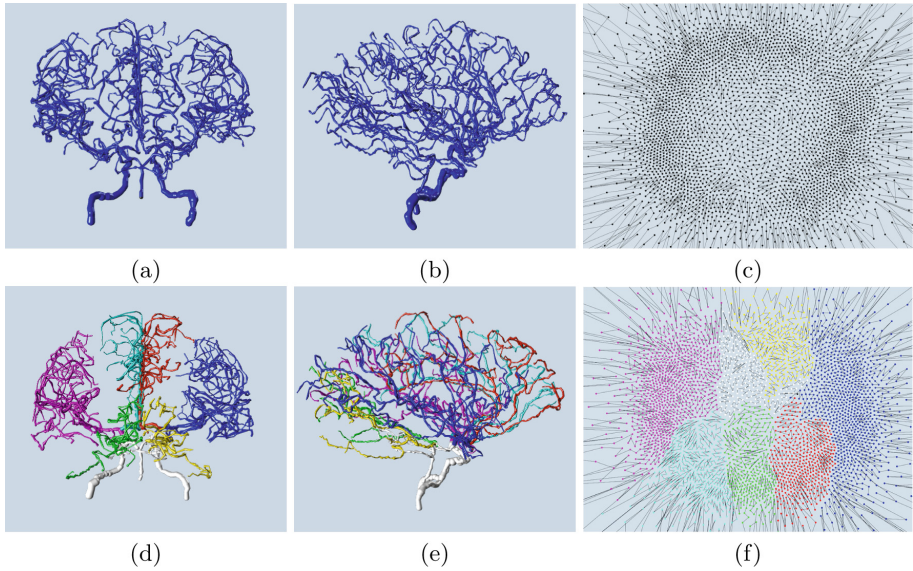
Fig. 1. Figures (a) and (b) show the unlabeled reconstructions of a cerebral vasculature. Figures (d) and (e) show their respective labels using color mapping. The annotations were chosen as Pink = LMCA, Blue = RMCA, Cyan = LACA, Red = RACA, Green = LPCA, Yellow = RPCA, White = Circle of Willis and ICA. The corresponding graphical representations of the unlabeled and labeled vessel segments are shown in (c) and (f). The vascular model shown in this figure was composed of 3514 vessel segments (Color figure online).

y_i in the model, the posterior marginal of the state is defined as follows,

$$p(x_i|y_{\{1...N\}}) \propto p(y_i|x_i)\, p(x_i) \int p(x_i|x_i^p)\, p(x_i^p|y_{\{1...N\},1})dx_i^p \; \ldots$$

$$\prod_j^{N_c} \int p(x_i|x_{i,j}^c)\, p(x_{i,j}^c|y_{\{1...N\}})\; dx_{i,j}^c \qquad (2)$$

where $p(y_i|x_i)$ is the likelihood. We propose to use a graphical model to represent this recursive problem, and Belief Propagation to perform the labeling process.

Graphical Model. The graphical model used in our labeling framework defines relations between pairs of nodes only. It is usually referred to as Pairwise Markov Random Field (PMRF) in the literature. As illustrated in Fig. 3, states $x_i \in x$ and observations $y_i \in y$ are represented in the graphical model by white, and shaded nodes, respectively. Edges represent dependencies between states by two types of functions: observation potentials $\phi(x_i, y_i)$ that are the equivalent of the likelihood part $p(y_i|x_i)$, and compatibility potentials $\psi_{ij}(x_i, x_j)$ that embed the conditional parts $p(x_i|x_j)$, $p(x_j|x_i)$ of the Bayesian formulation and can be used
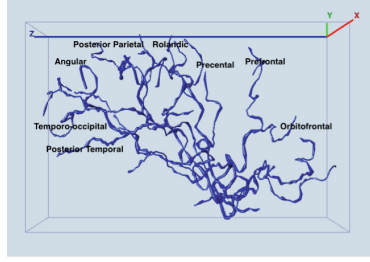
Fig. 2. Illustration of the additional annotations made by a neurologist on the middle cerebral artery (MCA) territory indicating the eight sub-territories of interest.
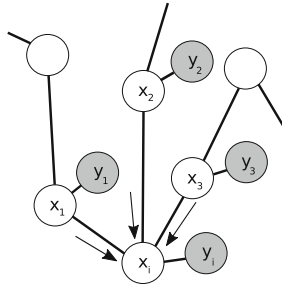
Fig. 3. In this graphical representation, y_1, y_2, y_3 indicate each node's observational features and x_1, x_2, x_3 represent their respective states. Messages received by state x_i are denoted by directed arrows.

by conditioning them in either directions during inference. In addition, the prior distribution over the labels is denoted $\psi_i(x_i)$.

Observation Model. An observation y_i represents the label information about the ith vessel segment that is directly extracted from the image. An observation $y_i \in R^M$ assigns a likelihood to each possible label with respect to the position of the segment a, b, c and its radius r using a previously learned kernel density estimation (KDE) $\hat{f}(S_i; \Theta_l)$ (Eq. 3) that is constructed by collecting a total of N labeled vessel segments $\{a_k, b_k, c_k, r_k\}_{1...N}$ across the training set,

$$\hat{f}(S_i; \Theta_l) = \frac{1}{N} \sum_{k=1}^{N} \mathcal{G}(S_i; \{a_k, b_k, c_k, r_k\}, \Sigma_i) \tag{3}$$

where $\mathcal{G}(S_i; \{a_k, b_k, c_k, r_k\}, \Sigma_i)$ is a Gaussian kernel centered at $\{a_k, b_k, c_k, r_k\}$ with standard deviation $\Sigma_{i,l}$ which is common to all the components k of a given label l. $\Theta_l = \{\{a_i^k, b_j^k, c_j^k, r_j^k\}_{1...N_c}, \Sigma_i\}$ denotes the parameter set of the KDE associated with a given label l.

Observations y_i are linked probabilistically to their state x_i through a Gaussian observation potential $\phi(x_i, y_i)$,

$$\phi(x_i, y_i) = \exp(-|y_i - x_i|^2/\sigma_o^2) \tag{4}$$

where σ_o is a smoothing parameter.

Compatibility Potentials. Compatibility potentials $\psi(x_i, x_j)$ define the relationship between two connected states. They are defined as a Gaussian difference between their arguments,

$$\psi(x_i, x_j) = \exp(-|x_i - x_j|^2/\sigma_t^2) \tag{5}$$

where the standard deviation σ_t of the model can be estimated using maximum likelihood (ML) on training data.

2.3 Labeling Using Belief Propagation

Labeling vessel segments in the 3D vasculature amounts to estimating $p(x_i|y_{\{1...N\}})$, the posterior belief associated with the state x_i given all observations $y_{\{1...N\}}$ accumulated. Thus, labeling is achieved through inference in our graphical model. One way to do this efficiently is to use Belief Propagation (BP) [8], a method implemented successfully in numerous computer vision applications such as for image segmentation [16] and object recognition [10–12].

It is a message passing algorithm for graphical models where messages are repeatedly exchanged between nodes to perform inference. Following the notation of BP, a message m_{ij} sent from node i to j is written,

$$m_{i,j}(x_j) \leftarrow \int \psi_{i,j}(x_i, x_j) \, \psi_i(x_i) \, \phi_i(x_i, y_i) \prod_{k \in \mathcal{N}_i \setminus j} m_{k,i}(x_i) \, \mathrm{d}x_i \tag{6}$$

where $\mathcal{N}_{i \setminus j}$ is the set of neighbors of state i where j is excluded, $\psi_{i,j}(x_i, x_j)$ is the pairwise potential between nodes i, j, and $\phi_i(x_i, y_i)$ is the observation potential.

After any iteration of message exchanges, each state can compute an approximation $\hat{p}(x_i|y_{\{1...N\}})$, called belief, to the marginal distribution $p(x_i|y_{\{1...N\}})$ by combining the incoming messages with the local observation:

$$\hat{p}(x_i|y_{\{1...N\}}) \leftarrow \psi_i(x_i) \, \phi_i(x_i, y_i) \prod_{k \in \mathcal{N}_i} m_{k,i}(x_i) \tag{7}$$

For tree-structured graphs like ours, the beliefs (Eq. 7) will converge to the exact solution $p(x_i|y_{\{1...N\}})$ [4].

3 Experiments

The goal of the experiments presented in this section is to demonstrate the effectiveness of the developed probabilistic labeling algorithm. To evaluate the model,

Table 1. Results of the probabilistic labeling model on 15 test subjects.

| Cerebral territory	Accuracy
LACA	$88.9 \pm 14.3\,\%$
RACA	$93.8 \pm 10.3\,\%$
LPCA	$93.6 \pm 10.3\,\%$
RPCA	$97.3 \pm 3.2\,\%$
LMCA	$98.7 \pm 1.7\,\%$
RMCA	$98.7 \pm 1.4\,\%$
Circle of Willis	$84.1 \pm 14.4\,\%$
Average	$\mathbf{94 \pm 5.2\,\%}$

two sets of experiments were ran; the first one using the 7 major vascular terri-
tories provided as part of BraVa, and the second one based on the 8 additional
MCA sub-territories established in this study (as described in Sect. 2.1). In both
experiments, we evaluate the accuracy by finding the percentage of segment
labels in the testing data consistent with the manual annotation.

To quantify the performance of the ACA, MCA, and PCA labeling, we used
39 annotated reconstructions from BraVa as training to build and optimize the
graphical model, utilizing the cartesian data and radii as inputs. The algorithm
was then tested on the data of 15 remaining subjects. The split between the
training and testing set was done randomly. For the MCA subdivision, we tested
the algorithm using leave-one-out crossvalidation. In this method, we train the
model using all subjects except one, and test on the remaining subject, repeating
the process for each set.

Table 1 summarizes the accuracy of the ACA, MCA, and PCA segmentation
over the 15 tested data sets was $94 \pm 5.2\,\%$ (as illustrated in Fig. 5). For the
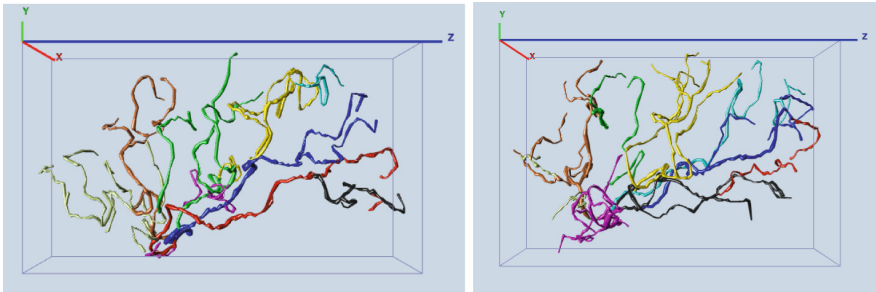
Fig. 4. Results of the MCA segmentation algorithm on two vascular models. Black
= Posterior Temporal, Red = Temporo-Occipital, Blue = Angular, Cyan = Poste-
rior Parietal, Yellow = Rolandic, Green = Precental, Brown = Prefrontal, Beige =
Orbifrontal (Color figure online).
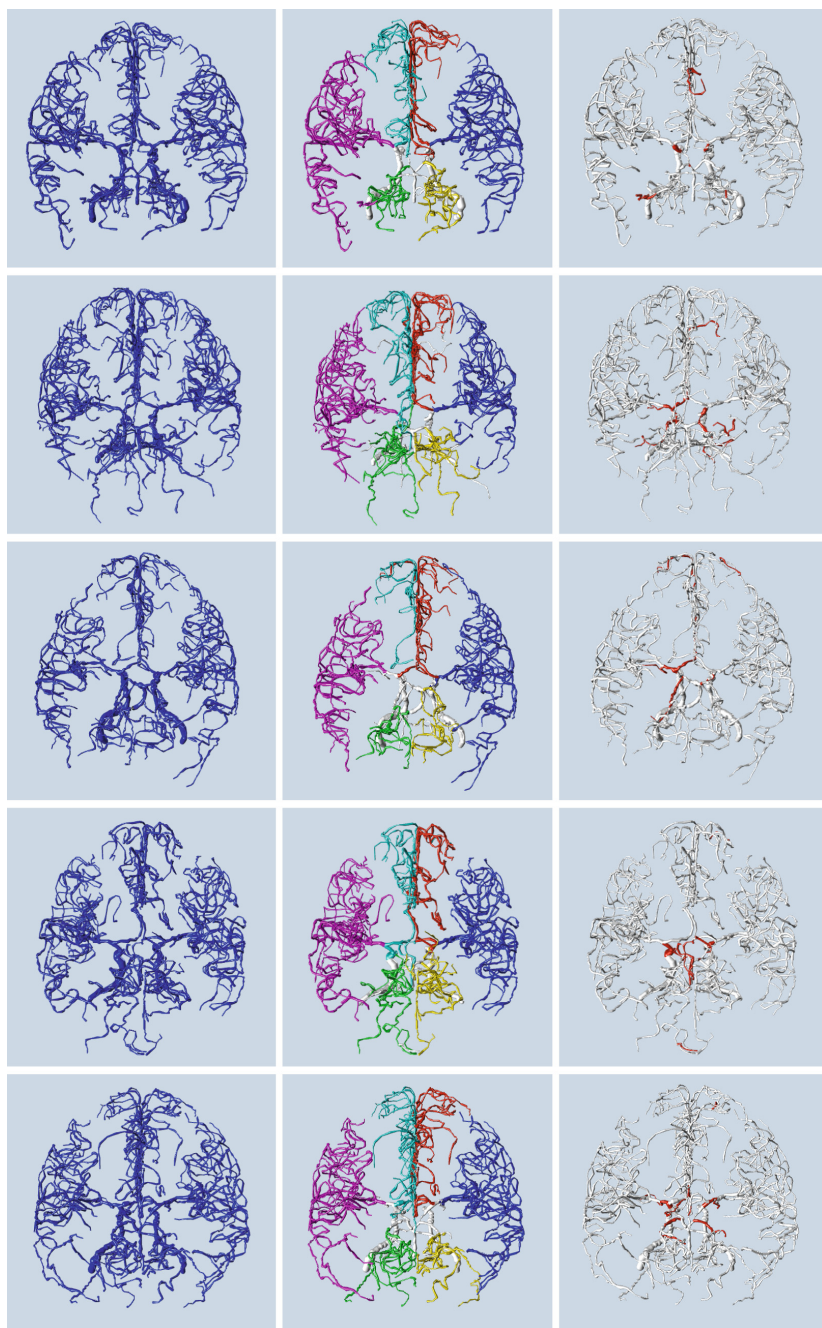
Fig. 5. The left and middle columns show the unlabeled models and their respective probabilistic segmentation results for five subjects obtained from BraVa [17]. The right column indicates the automated labeling errors in red (Color figure online).

MCA subdivision, we observed an accuracy of $88 \pm 9.3\%$ for the 7 tested sets (Fig. 4). These results demonstrate that our model is effective in predicting vascular regions in arterial reconstructions. The larger MCA subdivision deviation is a result of one experiment producing an accuracy of 67%, indicating that our model is not totally impervious to the wide variation in vascular topography. The major artery segmentation algorithm proved to be more consistent in its results; the accuracy for any test never fell below 79%.

4   Discussion

The results obtained during our experiments have demonstrated that the use of a graphical model that combines kernel density estimation (KDE) in conjunction with belief propagation inference is capable to learn and label vascular territories automatically with high accuracy. We believe this performance can be attributed to both the robustness of nonparametric estimation of the likelihood model and the optimal sharing of information by the message passing algorithm. Training and labeling times allow for processing of a large cohort of subject data; reconstructions can be labeled in a matter of seconds, and complete training and optimization in a few minutes.

In order to fully evaluate how well the model performs in real clinical conditions, a larger, more diverse pool of subjects needs to be tested. Because all training and test sets were retrieved from a single database, we are exposed to sampling bias, perhaps overly estimating the accuracy of the results of the framework on a new dataset where the input features would be extracted with a vessel segmentation technique different from Neuron_Morpho.

More importantly, the ability of our framework to study brain vasculatures automatically is hindered by pre-processing tasks that require manual input, preventing large data sets from being analyzed. Being able to automatically label vascular territories solves a piece of that puzzle, however, more efforts should be invested in the automatic vessel segmentation process.

In terms of clinical relevance of the developed tool for stroke patients, the automatic labeling allows us to quantify arterial regions missing due to occlusion and assess brain health. The MCA territory is a common site of occlusion in ischemic stroke and, through automated labeling, we can identify areas of the MCA that are frequently targeted and recognize features symptomatic of stroke. The proposed algorithm provides a framework for the learning and labeling of other vascular territories as well; whole-body and abdominal vessel segmentation are also very relevant clinical applications that could benefit from the probabilistic framework presented in this paper.

5   Conclusion

The results indicate that the developed framework, which is based on Bayesian formulation of the labeling problem, is effective in labeling vascular features in 3D reconstructions using belief propagation algorithm. We anticipate that this

approach will prove useful for learning arterial characteristics in any setting, particularly neurovascular, and circumventing the need for manual annotation. In order to fully automate the labeling process, an effective registration method should be considered. This requirement will be addressed in future work.

References

1. Dufour, A., Ronse, C., Baruthio, J., Tankyevych, O., Talbot, H., Passat, N.: Morphology-based cerebrovascular atlas. In: ISBI, pp. 1210–1214 (2013)
2. Ghanavati, S., Lerch, J.P., Sled, J.G.: Automatic anatomical labeling of the complete cerebral vasculature in mouse models. Neuroimage **95**, 117–128 (2014)
3. Ghanavati, S., Lerch, J.P., Sled, J.G.: Improved method for automatic cerebrovascular labelling using stochastic tunnelling. In: International Workshop on Pattern Recognition in Neuroimaging, pp. 1–4 (2014)
4. Ihler, A.T., Fisher, J.W., Willsky, A.S.: Message errors in belief propagation. In NIPS, no. 17, pp. 609–616. MIT Press (2005)
5. Mut, F., Wright, S., Ascoli, G.A., Cebral, J.R.: Morphometric, geographic, and territorial characterization of brain arterial trees. Int. J. Numer. Method Biomed. Eng. **30**(7), 755–766 (2014)
6. Oda, M., Hoang, B.H., Kitasaka, T., Misawa, K., Fujiwara, M., Mori, K.: Automated anatomical labeling method for abdominal arteries extracted from 3D abdominal CT images (2012)
7. Osborn, A.: Osborn's Brain: Imaging, Pathology, and Anatomy. Amirsys Pub., Salt Lake City (2013)
8. Pearl, J.: Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann Publishers Inc., San Francisco (1988)
9. Peng, H., Ruan, Z., Long, F., Simpson, J.H., Myers, E.W.: V3D enables real-time 3D visualization and quantitative analysis of large-scale biological image data sets. Nat. Biotechnol. **28**(4), 348–353 (2010)
10. Scalzo, F., Piater, J.H.: Statistical learning of visual feature hierarchies. In: CVPR, p. 44 (2005)
11. Scalzo, F., Piater, J.H.: Unsupervised learning of visual feature hierarchies. In: Perner, P., Imiya, A. (eds.) MLDM 2005. LNCS (LNAI), vol. 3587, pp. 243–252. Springer, Heidelberg (2005)
12. Scalzo, F., Piater, J.H.: Adaptive patch features for object class recognition with learned hierarchical models. In: CVPR, pp. 1–8 (2007)
13. Shahzad, R., Dzyubachyk, O., Staring, M., Kullberg, J., Johansson, L., Ahlstrom, H., Lelieveldt, B.P., van der Geest, R.J.: Automated extraction and labelling of the arterial tree from whole-body MRA data. Med. Image Anal. **24**(1), 28–40 (2015)
14. Stefancik, R., Sonka, M.: Highly automated segmentation of arterial and venous trees from three-dimensional magnetic resonance angiography (MRA). Int. J. Cardiovasc. Imaging **17**(1), 37–47 (2001)
15. Uchiyama, Y., Yamauchi, M., Ando, H., Yokoyama, R., Hara, T., Fujita, H., Iwama, T., Hoshi, H.: Automated classification of cerebral arteries in MRA images and its application to maximum intensity projection. IEEE Eng. Med. Biol. Soc. **1**, 4865–4868 (2006)
16. Wang, J., Cohen, M.F.: An iterative optimization approach for unified image segmentation and matting. In: ICCV, vol. 2, pp. 936–943 (2005)

17. Wright, S.N., Kochunov, P., Mut, F., Bergamino, M., Brown, K.M., Mazziotta, J.C., Toga, A.W., Cebral, J.R., Ascoli, G.A.: Digital reconstruction and morphometric analysis of human brain arterial vasculature from magnetic resonance angiography. Neuroimage **82**, 170–181 (2013)