Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

Statistical modeling and knowledge-based segmentation of cerebral artery based on TOF-MRA and MR-T1

Na Li [a,b], Shoujun Zhou [a,*], Zonghan Wu [a,b], Baochang Zhang [a,b], Gang Zhao [c,*]

[a] *Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China*
[b] *Shenzhen Colleges of Advanced Technology, University of Chinese Academy of Sciences, Beijing, China*
[c] *Neurosurgery Department, General Hospital of Southern Theater Command, PLA, Guangzhou, China*

ARTICLE INFO

ABSTRACT

Background and objective: For cerebrovascular segmentation from time-of-flight (TOF) magnetic resonance angiography (MRA), the focused issues are segmentation accuracy, vascular network coverage ratio, and cerebral artery and vein (CA/CV) separation. Therefore, cerebral artery segmentation is a challenging work, while a complete solution is lacking so far.

Methods: The preprocessing of skull-stripping and Hessian-based feature extraction is first implemented to acquire an indirect prior knowledge of vascular distribution and shape. Then, a novel intensity- and shape-based Markov statistical modeling is proposed for complete cerebrovascular segmentation, where our low-level process employs a Gaussian mixture model to fit the intensity histogram of the skull-stripped TOF-MRA data, while our high-level process employs the vascular shape prior to construct the energy function. To regularize the individual data processes, Markov regularization parameter is automatically estimated by using a machine-learning algorithm. Further, cerebral artery and vein (CA/CV) separation is explored with a series of morphological logic operations, which are based on a direct priori knowledge on the relationship of arteriovenous topology and brain tissues in between TOF-MRA and MR-T1.

Results: We employed 109 sets of public datasets from MIDAS for qualitative and quantitative assessment. The Dice similarity coefficient, false negative rate (FNR), and false positive rate (FPR) of 0.933, 0.158, and 0.091% on average, as well as CA/CV separation results with the agreement, FNR, and FPR of 0.976, 0.041, and 0.022 on average. For clinical visual assessment, our methods can segment various sizes of the vessel in different contrast region, especially performs better on vessels of small size in low contrast region.

Conclusion: Our methods obtained satisfying results in visual and quantitative evaluation. The proposed method is capable of accurate cerebrovascular segmentation and efficient CA/CV separation. Further, it can stimulate valuable clinical applications on the computer-assisted cerebrovascular intervention according to the neurosurgeon's recommendation.

1. Introduction

Medical image computing strives for essential schemes of computer-assisted diagnosis and interventional therapy, e.g., the location and quantification of the cerebrovascular structures and focuses [1]. For whole cerebral artery imaging, TOF-MRA is widely used in the practices of clinical neurosurgery for its well-performing with non-invasive and non-radiant characteristics [2], by which accurate extraction of artery structures is very significant.

TOF-MRA is a popular imaging approach of cerebral artery, the resultant image acquisition inevitably contains local low contrast and/or low Signal-to-Noise-Ratio, especially partial venous sinus structures, which could affect the computer-assisted procedure. As a result, signal loss of the curve or distal vessels [3] often occurs in the TOF-MRA data. To extract rich vascular structures and remove venous structures, a galling fact is that the arteries, veins and dural sinus present the similar brightness features in the local complex context, which makes our work more challenging.

* Corresponding authors.
 E-mail addresses: sj.zhou@siat.ac.cn (S. Zhou), zg215@tom.com (G. Zhao).

For the segmentation of blood vessel from volumetric data, existing model-driven methods (as reviewed in [4, 5]) generally focus on scale analysis, deformation model, region growing, tracking, and statistical modeling. By vascular scale analysis, thresholding the multi-scale filtering results could acquire vascular structures with large size [6, 7]. With intensity and geometrical information, deformation model-based methods produced vascular segmentation with topological flexibility [8, 9]. Based on seed points and growth criterion, region growing methods have steadily realized vascular segmentation through a simple and reliable local operator [10, 11]. Tracking-based methods locally ensure the spatial continuity of vascular structure [12]. So far, these methods are inevitably affected by artifacts, noise, or highly depend on seed points or initial position deviation. Despite statistical modeling with Maximum a Posterior (MAP) and Markov random field (MRF) has received special attention on cerebrovascular segmentation [13–16], its performance is still affected by lack of effective exploratory of finite mixture models, model parameters, prior knowledge and energy constraints. Zhou et al. [14] improved the MAP-MRF model by incorporating multi-pattern neighborhood system (NBS) in the Markov high-level process, while the energy function of the high-level model is not enough to extract small-sized vessels or even the main vascular branch at low signal to noise ratio (SNR). In recent years, deep-learning methods with convolutional neural network [17, 18] were applied to segment cerebral vessels, while the high time cost of manual annotation still needs to be broken through. The model-driven method is helpful to solve these bottlenecks of deep-learning methods [18, 19], which is also one of our motivations in this research work.

It is necessary to separate arterial structure from venous ones. While few works are associated with CA/CV separation so far, some researchers focused on the arteries and veins (A/V) separation in the angiographic data of the organs other than the brain, such as pulmonary [20, 21], abdominal [22, 23], retinal [24, 25], and cerebral angiographic data [26–30]. Because of respective priori knowledge on arteriovenous information in different organs, the existing methods in these works are difficult to be applied to CA/CV separation.

This paper is organized as follows. In Section 2, the related works are introduced. Section 3 describes our methods of vascular segmentation and CA/CV separation in detail. The experimental setup, the dataset, and the results are presented in Section 4. In Section 5, some limitations of our method are analyzed and discussed in detail and followed by our conclusions in Section 6.

2. Related work

2.1. Statistical model for cerebrovascular segmentation from TOF-MRA data

Statistical methods have been used for vascular segmentation from brain TOF-MRA data. Vascular segmentation can be considered as a labeling problem, which is specified in terms of a set of sites $\Omega = \{s_n | n = 1, ..., N\}$ and the binarization label set $\mathcal{L} = \{L_V, L_B\}$, where L_V and L_B represent the vessel and background classes, respectively, and s_n represents the index number of a site in the image. In the MAP-MRF model, assume that label variable set $X = \{x_s | s \in \Omega, x_s \in \{L_V, L_B\}\}$ is an MRF, observed variable set $Y = \{y_s | s \in \Omega, y_s \in \mathcal{R}\}$ is conditionally independent. According to the Bayesian rule [31], the posterior probability $\rho(X = x_s | Y = y_s)$ meets the expression:

$$\rho(X = x_s | Y = y_s) \propto \rho(Y = y_s | X = x_s) P(X = x_s) \quad (1)$$

The likelihood probability $\rho(Y = y_s | X = x_s)$ is modeled as a finite mixture model (FMM) in the Markov low-level process; $\rho(X = x_s)$ is the prior probability of label variables and can be modeled as the Gibbs distribution (i.e., high-level model) in the Markov high-level process according to Hammersley-Clifford Theorem [32] as follows

$$\rho(X = x_s) = e^{-U(x_s)}/Z, \ Z = \sum_{x_s \in \mathcal{L}} e^{-U(x_s)} \quad (2)$$

where $U(x_s)$ is an energy function, and is computed in neighborhood $\mathcal{N}_s$ of s. According to [33], the energy function can be expressed as:

$$U(x_s) = V_1(x_s) + \sum_{s' \in \mathcal{N}_s} V_2(x_s, x_{s'}) \quad (3)$$

where $V_1(x_s)$ and $V_2(x_s, x_{s'})$ are the potentials respectively from single sites and pair-wise sites (PWS, i.e., the pairs of sites s and s'). By MAP criterion, the optimal label x_s^* is the one that maximizes the posteriori probability $\rho(x_s | y_s)$, given its the prior probability $\rho(x_s)$ and likelihood probability $\rho(y_s | x_s)$, and is expressed as

$$x_s^* = \arg\max_{x_s \in \mathcal{L}} \rho(x_s | y_s) \propto \arg\max_{x_s \in \mathcal{L}} (\rho(y_s | x_s)\rho(x_s)) \quad (4)$$

For the earlier MAP-MRF based methods, especially that of Hassouna et al. [13] and Zhou et al. [14], a Rayleigh and three Gaussian functions were used to construct a finite mixture model and fit the intensity distribution of a set of TOF-MRA data, which is easy to form the first order energy function (i.e., $U_1(x_s)$) in the Markov low-level process. To calculate the second order energy function in the Markov high-level process, the multilevel logistic (MLL) model [34] was used, which can be expressed as

$$U_2(x_s) = \sum_{s' \in \mathcal{N}_s} V_2(x_s, x_{s'}) = \sum_{s' \in \mathcal{N}_s} \beta_c E(x_s, x_{s'}) \quad (5)$$

where the regularization parameter β_c indicates the strength of the interaction between the label of neighboring sites s and s'; $E(x_s, x_{s'})$ is a PWS-potential and was generally denoted by a Kronecker delta function as

$$E(x_s, x_{s'}) = 1 - \delta(x_s - x_{s'}), \ s' \in \mathcal{N}_s \quad (6)$$

While the simple spatial constraint of above PWS-potential was only used to remove the label-noise and keep the neighboring label-continuity [13]. To improve the segmentation, Zhou et al. [14] calculated the energy function by minimizing the summation of PWS-potentials within multi-pattern NBS, i.e.,

$$U_2(x_s) = \min_m \left[\min_j \sum_{c_{m,j}} \beta_c E_{c_{m,j}}(x_s, x_{s'}) \right] \quad (7)$$

where $m = 1, ..., M, \ j = 1, ..., N_{C_m}$ are respectively the serial numbers of clique-classes and the patterns of each clique-class. As mentioned in [14], the multi-pattern NBS involves 45 neighborhood patterns. The energy expression (i.e., Eq. (7)) of multi-pattern NBS generalizes NBSs to the construction of clique potential V_2, where the 6- and 26-NBS are two particular cases of multi-pattern NBS.

Some limitations still exist in the state-of-the-art statistical model. On the one hand, the Markov low-level process inevitably suffers from large error of FMM-based histogram fitting, because (i) the background region occupies a great percentage of the volume of original data (more than 95%); (ii) the FMM-based likelihood estimation only depends on the intensity distribution and may be impacted at the region of low contrast and/or low SNR. On the other hand, the existing Markov high-level processes overly depend on the low-level process for initialization and iteration, thus are insufficient to regain the vessels at low SNR and/or low contrast regions. Besides, their high-level models depend on the summation of the Kronecker delta function, i.e., Eq. (6). The great difference between the segmentation results of the traditional methods [13,14] and ours is compared in Fig. 1.

2.2. CA/CV separation

For separation of cerebral arteries and veins, several angiography imaging technologies with post-processing were employed,

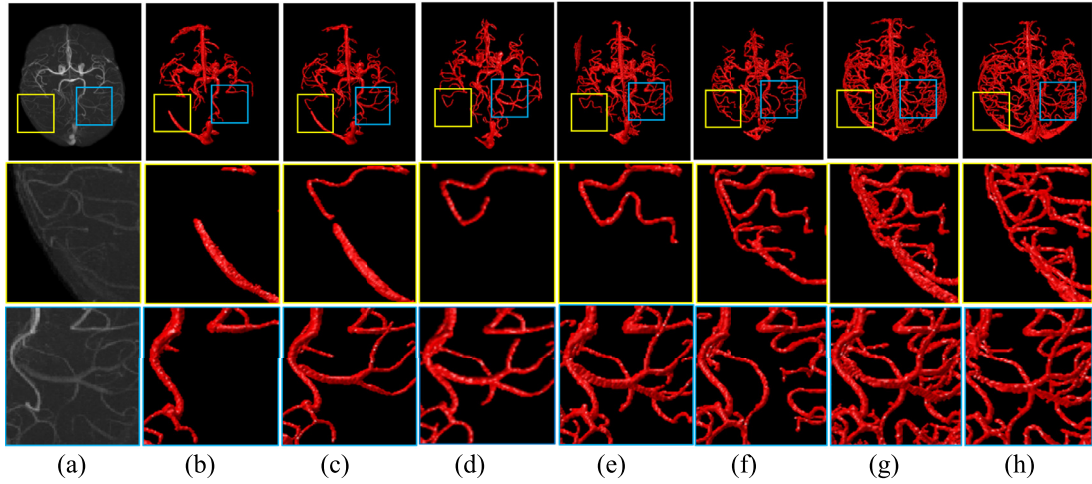| (a) | (b) | (c) | (d) | (e) | (f) | (g) | (h) |

Fig. 1. Comparisons of cerebrovascular segmentation results from TOF-MRA data between the proposed and the existing statistical methods, where 2nd and 3rd rows correspond to the yellow and blue boxes in the first row, respectively. (a) the maximum intensity projections (MIP) of the raw data; (b)-(e) the results by Lu et al. [15], Wen et al. [16], Hassouna et al. [13] and Zhou et al. [14]; (f) results by our Markov low-level model; (g) results by our Markov low- and high-level model; (h) the corresponding ground truth. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

such as 4D MR [26], 4D CT perfusion scans [27], single-wavelength laser speckle imaging [28], Dual-Wavelength Optical Imaging [29], and 4D CT [30]. Among these acquisitions, the time-intensity profiles provided by the 4D data or the spectral features of imaging modality enabled the CA/CV to be extracted independently. For the TOF-MRA data, there exists a similar intensity characteristic in between CA/CV. To the best of our knowledge, no one has performed an integrated work from cerebrovascular segmentation and CA/CV separation on the popular 3D TOF-MRA data, which is critical for the implementation of computer-aided diagnosis and intervention.

So far, the most related work of the CA/CV separation is [35]. In that work, incomplete TOF-MRA data were acquired only from the small part of the head, and they used the distance map based on cerebrospinal fluid and grey matter to recover the topology of cerebral superficial venous, which is time-consuming to compute the distance of each vascular voxels to the cerebrospinal fluid and grey matter. However, their work inspires our intention on the multi-modality and anatomical priori.

2.3. Our contributions

For extracting the cerebral arterial structures, a novel method is proposed to address several bottlenecks of the state-of-the-art statistical modeling and CA/CV separation. In our work, the main contributions are summarized as: (i) a new FMM is constructed by using a Gaussian mixture model (GMM) to fit the intensity histogram of the skull-stripped original data, thus to improve the adaptivity and validity of Markov low-level process. (ii) The vascular multi-scale filtering is employed to acquire the indirect prior,[1] i.e., the vascular feature map (VFM) and the vascular direction field (VDF). Meanwhile, the regularized parameter estimation is automated by the maximum pseudo likelihood estimation with a machine-learning process, thus avoids repeated trial-error-test estimation. (iii) The direct prior knowledge of cerebral anatomy is explored to locate cerebral artery, vein, and dural sinus, based on their relationship in between TOF-MRA and MR-T1 data. Then, a series of morphological logic operations are proposed to automate CA/CV separation from the cerebrovascular segmentation results.

---

[1] Note: To distinguish the two classes of prior knowledges, we call the VFM and VDF as indirect prior, while the direct prior is corresponding to the relationship of anatomical information acquired from TOF-MRA and MR-T1.

3. Methods

For extraction of the cerebral arterial tree, the main steps are as follows: firstly, we strip the skull from the original TOF-MRA data and use GMM to fit the resultant histogram. Secondly, multi-scale filtering is implemented to enhance the vascular structure and acquire the local characteristics of vascular shape. Thirdly, we construct a new MAP-MRF model and estimate the regularization parameter. Finally, we automate CA/CV separation according to the anatomical relation of the artery and dural sinus. Our framework is shown in Fig. 2.

3.1. Vascular hessian characteristics extraction

According to [36], we use the multiscale filtering algorithm to acquire two vascular characteristics, i.e., VFM and VDF. Given a site $s(x_1, x_2, x_3)$, its Hessian matrix can be calculated by

$$H_\sigma(s) = \sigma^2 (\partial^2 I_\sigma(s)/\partial x_i \partial x_j) \qquad (8)$$

where $i, j = 1, 2, 3$ and $I_\sigma(s)$ is the filter response with a Gaussian kernel at scale σ. By eigen decomposition, each Hessian matrix has three eigenvalues meeting $|\lambda_{1,\sigma}| \leq |\lambda_{2,\sigma}| \leq |\lambda_{3,\sigma}|$ and three corresponding eigenvectors, i.e., $\vec{v}_{1,\sigma}, \vec{v}_{2,\sigma}, \vec{v}_{3,\sigma}$. For the eigenvectors, $\vec{v}_{1,\sigma}$ indicates the direction along the vessel (minimum intensity variation), $\vec{v}_{2,\sigma}$ and $\vec{v}_{3,\sigma}$ form a base for the orthogonal plane. The vascular filter response at site s can be obtained by

$$V_\sigma = \begin{cases} 0 & if\, \lambda_{2,\sigma} \leq 0 \vee \lambda_{t,\sigma} \leq 0 \\ 1 & if\, \lambda_{2,\sigma} \geq \frac{\lambda_{t,\sigma}}{2} > 0 \\ \lambda_{2,\sigma}{}^2 (\lambda_{t,\sigma} - \lambda_{2,\sigma}) \left[\frac{3}{\lambda_{t,\sigma}+\lambda_{2,\sigma}}\right]^3 & otherwise \end{cases} \qquad (9)$$

where $\lambda_{t,\sigma}$ is computed by

$$\lambda_{t,\sigma} = \begin{cases} \lambda_{3,\sigma} & if\, \lambda_{3,\sigma} > \mu\, max_s \lambda_{3,\sigma}(s) \\ \mu\, max_s \lambda_{3,\sigma}(s) & if\, 0 < \lambda_{3,\sigma} \leq \mu\, max_s \lambda_{3,\sigma}(s) \\ 0 & otherwise \end{cases} \qquad (10)$$

where μ is a cutoff threshold between 0 and 1, and an empirical value (i.e., $\mu = 0.7$) is employed in our work. By comparing the aforementioned filter responses of the same site at different scales, the maximized the response of each site is considered as the fi-
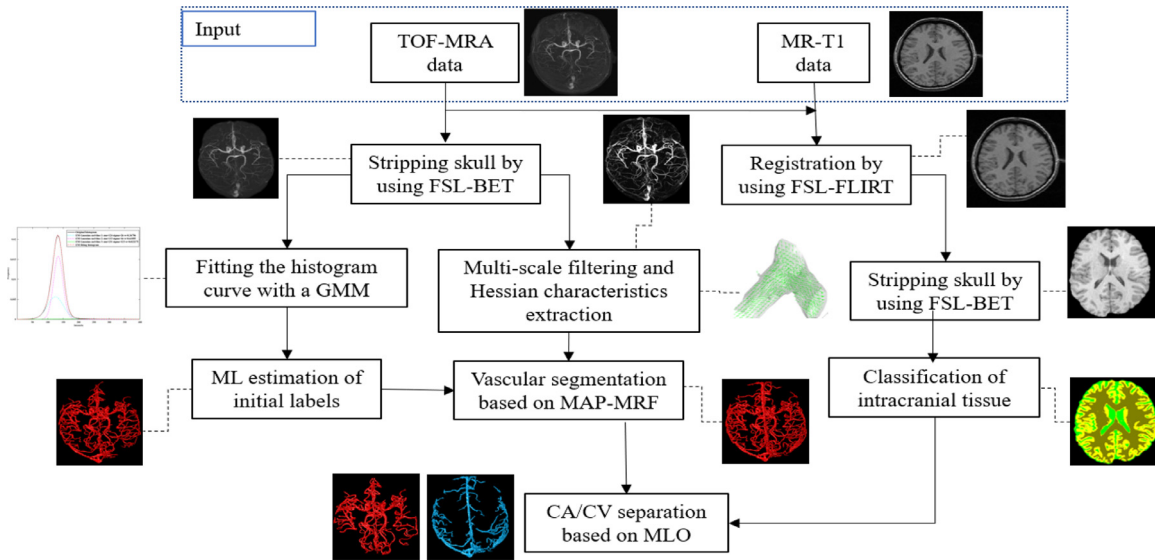
Fig.2. The framework of the proposed cerebral artery extraction from TOF-MRA and MR-T1 data.
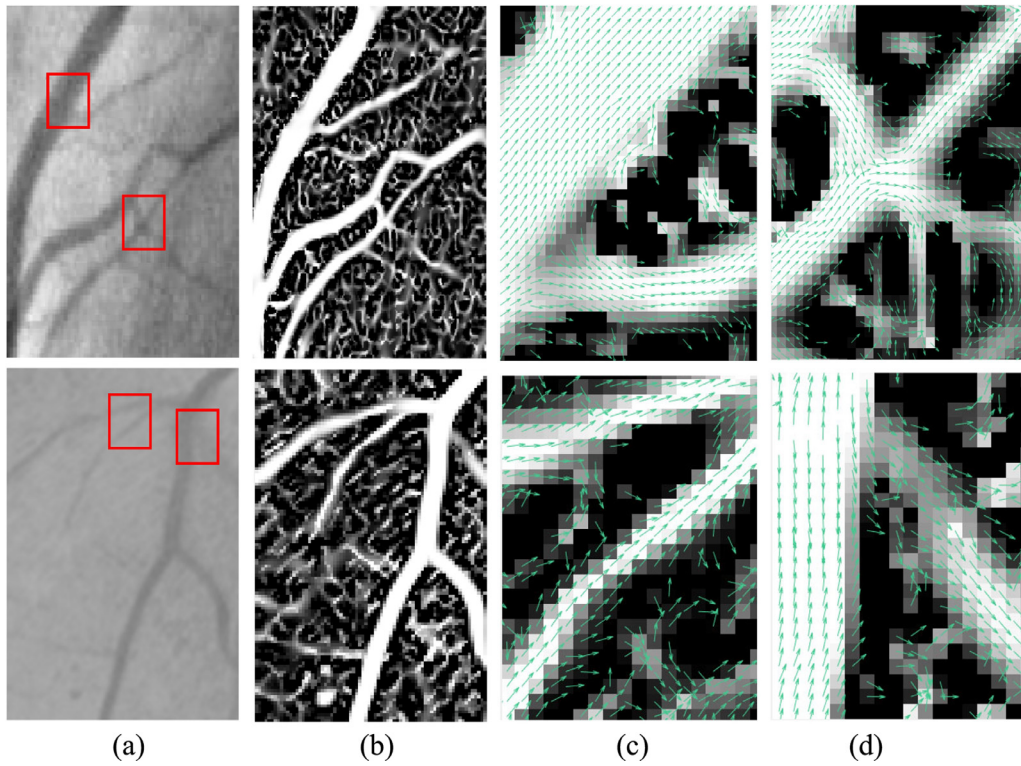

(a)   (b)   (c)   (d)

Fig. 3. The VDF of a 2D image. For each row, Subfigure (a) represents original data; subfigure (b) shows the VFM of subfigure (a); subfigure (c) and (d) show the VDF of the target in the red box in subfigure (a), which show the complementary characteristics of irregular vascular structures (at low contrast and bifurcation areas) in VFM and VDF.

nal filter output, i.e., $\mathcal{V}_{\sigma'}$, where σ' is the corresponding optimal scale. We consider the multi-scale filter response as the VFM and the field, which consists of all the $\bar{v}_{1,\sigma'}$ of the sites in data, as VDF. For simplicity, $\bar{v}_1(s)$ is used to represent the vascular direction vector of a site s. The above-mentioned characteristics are shown in Fig. 3. It is worth mentioning that we assume that the direction of $\bar{v}_1(s')$ around a site s is consistent within the neighborhood $s' \in \mathcal{N}_s$. Besides, there are still some uneven characteristic in VFM, e.g., contrast, bifurcation connection, and tissue-edge response. As a supplement, the VDF reveals the local direction of image-ridges, while its directional consistency would be weakened at bifurcation.

3.2. Intensity- and shape-based statistical modeling

Cerebral vessels have wide overlapping intensity distribution with that of non-vascular tissues in original data, and often lead to over- or under-segmentation if we directly use FMMs fitting its histogram. As a fact, the large part of cerebrovascular voxels distributes in the middle and high intensity, and the proportion of cerebrovascular voxels in intracranial volume is relatively small (about no more than 5% in anatomy). Meanwhile, numerous experiential validations demonstrate that the histogram-fitting accuracy barely rests on the middle- and high-intensity range and especially impacted by the fitting error at the low-intensity range (generally
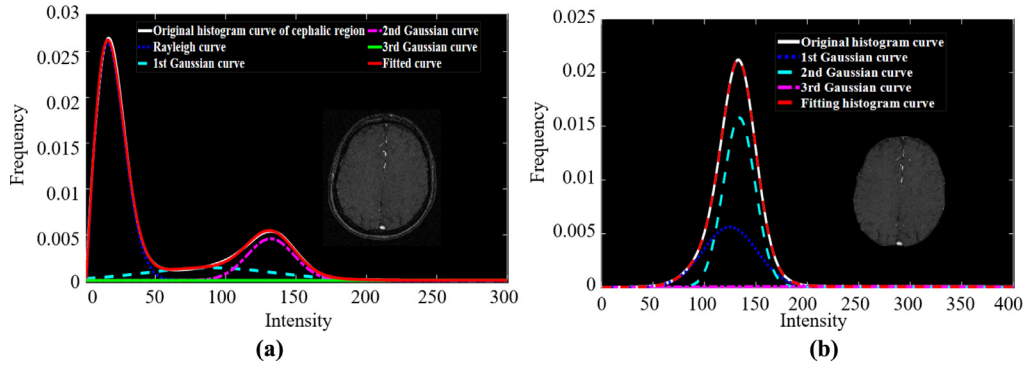
Fig. 4. Comparison of original histogram curves and the histogram-fitting results in different object regions of TOF-MRA data. (a) histogram-fitting of cephalic region with one Rayleigh and three Gaussian distributions; (b) histogram-fitting of intracranial region with three Gaussian distributions.

the background region in TOF-MRA), which results in much error and computational redundancy. While using the Brain Extraction Tool (BET) in FSL [37] to strip the skull from original data, the resultant ROI stabilizes a histogram-fitting with mixture Gaussian distributions and facilitate the following steps.

3.2.1. Low-level model

By analyzing the skull-stripped TOF-MRA data, we divide the volume voxels into three classes according to their intensity: voxels with high intensity are considered as the cerebral vessel, while voxels with medium- and low- intensity are judged as brain tissues and cerebrospinal fluid, respectively. The histogram-fittings of cephalic and intracranial regions are compared in Fig. 4. Different from traditional FMMs (e.g., Fig. 4(a)) using one Rayleigh and three Gaussian functions to fit the cephalic region-histogram, the intracranial region-histogram is fitted by using two Gaussian functions (i.e., $f_{g1}(y_s)$ and $f_{g2}(y_s)$) for the background class and another one (i.e., $f_{g3}(y_s)$) for the vascular class (e.g., Fig. 4(b)). At each site s, the GMM of intensity value y_s is expressed as:

$$f_{GMM}(y_s) = \sum_{l=1}^{3} \frac{w_l}{\sqrt{2\pi}\sigma_l} exp\left(\frac{-(y_s - \mu_l)}{2\sigma_l^2}\right) \tag{11}$$

where μ_l, σ_l and w_l are the mean, variance and proportionality coefficient of the *l-th* Gaussian function. All the nine values (i.e., μ_l, σ_l, w_l with $l = 1,2,3$) of GMM parameters should be estimated efficiently. According to [14], the authors automated a Rayleigh-Gaussian mixture distribution by using the K-means based pre-classification and parameter initialization following with expectation-maximum (EM) algorithm, we deduce a GMM-EM (referred in Appendix A) to automate the above-mentioned parameters estimation. Note, the GMM, as with other FMMs, still has the common characteristic of lacking geometric shape of vascular prior, for which shape-based modeling is explored as follows.

3.2.2. High-level model

Since only relying on intensity information, the initial segmentation result (label field) obtained by the low-level model generally do not contain the vascular structures in the low-contrast image area, which also makes it difficult to obtain the complete vascular structures even by the Markov high-level process that exert a traditional NBS-energy constraint (Eq. (7)) to label field. Aiming to regain the underlying vessels from the low-contrast or background regions, VFM and VDF are employed as the vascular shape priors and are used to construct the Markov high-level model. For the new NBS energy constraint, the PWS-potential is first analyzed and regularized as follows.

In the VDF, given vectors $\vec{v}_1(s)$ and $\vec{v}_1(s')$ at the PWS of s and s', the cosine similarity of the two vector is

$$Similarity = \frac{\vec{v}_1(s) \cdot \vec{v}_1(s')}{|\vec{v}_1(s)||\vec{v}_1(s')|} \tag{12}$$

Note, to avoid denominators being zero, a very small value (e.g., 10^{-15}) is added to the denominator for effective calculation. Then, a normalized phase continuity measure of the two sites is denoted by

$$S_{s,s'} = 1 - Similarity \tag{13}$$

In the VFM, given $\nabla \mathcal{V}_{s,s'}$ is the? numerical intensity difference of the PWS, and $\nabla \mathcal{V}_{max}$ is the maximum value within $s' \in \mathcal{N}_s$, a normalized measure for the intensity smoothness of the two sites is expressed as

$$\nabla \mathcal{I}_{s,s'} = \frac{|\nabla \mathcal{V}_{s,s'}|}{|\nabla \mathcal{V}_{max}|} \tag{14}$$

Combining the two measures, we can construct a compound measure, i.e.,

$$\rho_{s,s'} = \epsilon_1 \nabla \mathcal{I}_{s,s'} + \epsilon_2 \mathcal{S}_{s,s'} \tag{15}$$

where two weight coefficients meet $\varepsilon_1 + \varepsilon_2 = 1$ and average assignment will balance the two measures, i.e., $\varepsilon_1 = \varepsilon_2 = 0.5$. The traditional PWS-potential of Eq. (6) is only to regulate the label values at the position of $x_s \neq x_{s'}$ (mainly situated in the edges of the vessels or the block-shaped structures). Applying the measure $\rho_{s,s'}$ to either edge (where $x_s \neq x_{s'}$) or background (where $x_s = x_{s'} = L_B$), we can derive a new PWS-potential, namely

$$\mathcal{P}(x_s, x_{s'}) = \begin{cases} 0 & x_s = x_{s'} = L_V \\ 1 - \rho_{s,s'} & x_s = x_{s'} = L_B \\ \rho_{s,s'} & x_s \neq x_{s'} \end{cases} \tag{16}$$

where $x_s, x_{s'} \in \{L_V, L_B\}$ and $s' \in \mathcal{N}_s$. Different from the traditional U_1 and U_2 Eqs. (5)–((7)) in Section 2.1, the GMM $f_{GMM}(y_s)$ is used to construct the new energy function (i.e., $U_I(x_s)$) of the Markov low-level model, while the potential $\mathcal{P}(x_s, x_{s'})$ would be used on NBSs to derive a new energy function (i.e., $U_{II}(x_s, x_{s'})$) of the Markov high-level model, namely,

$$U_{II}(x_s, x_{s'}) = \min_{m} \left[\min_{j} \sum_{c_{m,j}} \beta_c \mathcal{P}_{c_{m,j}}(x_s, x_{s'}) \right] \tag{17}$$

3.2.3. MAP-MRF model

Combining the above energy functions of low- and high-level models, i.e., Eqs. (11) and (17), with the MP-NBS, the new posterior probability (MAP-MRF model) is derived as

$$\rho(x_s|y_s) \propto exp\left[-U_I(x_s)\right]exp\left[-U_{II}(x_s, x_{s'})\right] \tag{18}$$

where the two terms on the right can be specifically expressed as

$$\exp\left[-U_I(x_S)\right] = \begin{cases} f_{g3}(y_s), & x_s = L_V \\ \sum_{l=1}^{2} w_l f_{gl}(y_s) / \sum_{l=1}^{2} w_l, & x_s = L_B \end{cases} \tag{19}$$

and

$$\exp\left[-U_{II}(x_s, x_{s'})\right] = \exp\left\{ - \min_m \left[\min_j \sum_{c_{m,j}} \beta_c \mathcal{P}_{c_{m,j}}(x_s, x_{s'}) \right] \right\} \tag{20}$$

A steady label field can be obtained through iteration conditional mode (ICM) [13, 33], where the initial label field is acquired by the maximum likelihood estimate based on the low-level model. In each iteration step, according to MAP criterion, if the posterior probability of $x_s = L_V$ no less than that of $x_s = L_B$, we can infer the site to be a vessel site. The above Markov process (Eq. (18)) with ICM will make the global convergence within few iterations. The resultant region with label $x_s = L_V$ ($s \in \Omega$) covers the cerebrovascular structures, however, it includes cerebral artery, vein, and dural sinus.

3.2.4. Estimation of regularization parameter

In the Markov process, the parameter β_c in Eq. (17) will balance the actions between low- and high-level processes for a specific original data. However, the acquisitions of different individual subjects have respective features on contrast, SNR, and vasculature distribution, which are non-negligible for possible impact to the regularization parameter estimation. The maximum pseudo-likelihood [33] based machine-learning algorithm is used to estimate the parameter, considering that the homogeneous distribution accepts the assumption of a constant $\beta_c = \beta$ for all image voxels. Given the initial label field, the pseudo-likelihood is equivalent to the product of conditional likelihood probabilities, and is expressed as

$$\text{PL}(x) = \prod_{s \in \Omega, s' \in \mathcal{N}_s} P(x_s | x_{s'}) = \prod_{s \in \Omega, s' \in \mathcal{N}_s} \frac{exp[-U_s(x_s = L_V, \ x_{s'})]}{\sum_{x_s} exp[-U_s(x_s, \ x_{s'})]} \tag{21}$$

Here, the total energy is determined by its first- and second-order energy functions, i.e.

$$U_s(x_s, x_{s'}) = U_I(x_s) + U_{II}(x_s, x_{s'}) \tag{22}$$

where the first-order energy function $U_I(x_s)$ meets the following expressions:

$$\begin{cases} f_{GMM}(y_s | x_s = L_V) = exp(-U_I(y_s | x_s = L_V)) \\ f_{GMM}(y_s | x_s = L_B) = exp(-U_I(y_s | x_s = L_B)) \end{cases} \tag{23}$$

Optimal parameter $\hat{\beta}_c$ can be obtained by maximizing the pseudo-likelihood as well as minimizing the negative logarithm of pseudo-likelihood, namely,

$$-\ln\text{PL}(x | \beta) = \sum_{s \in \Omega, s' \in \mathcal{N}_s} \left[U_I(x_s) + U_{II}(x_s, x_{s'}) \right.$$

$$\left. + \ln \sum_{x_s \in \{L_V, L_B\}} \exp(-U_I(x_s) - U_{II}(x_s, x_{s'})) \right] \tag{24}$$

Considering that it is difficult to find an analytical solution, we use the gradient descent strategy to iteratively find the optimal $\hat{\beta}_c$. The iteration equation is expressed as

$$\beta_t = \beta_{t-1} + \gamma \nabla(\ln\text{PL}(x | \beta_{t-1})), \quad \text{with } t = 1, 2, \dots \tag{25}$$

where γ is the iteration step size, then we derive the following

equation (referred in the Appendix B).

$$-\nabla(\ln\text{PL}(x | \beta)) = -\frac{\partial(\ln\text{PL}(x | \beta))}{\partial \beta}$$

$$= \sum_{s \in \Omega} \left(\delta(x_s - L_V)N_{L_V} + \delta(x_s - L_B)N_{L_B} \right.$$

$$\left. - \frac{f_{L_V} \exp(-\beta N_{L_V})N_{L_V} + f_{L_B} \exp(-\beta N_{L_B})N_{L_B}}{f_{L_V} \exp(-\beta N_{L_V}) + f_{L_B} \exp(-\beta N_{L_B})} \right) \tag{26}$$

where the delta function $\delta(x_s - L_V)$ produces 1 and 0 at $x_s = L_V$ and $x_s = L_B$, respectively; The GMMs f_{L_V} and f_{L_B} produce the conditional probability $\rho(y_s | L_V)$ and $\rho(y_s | L_B)$ for a voxel belonging to vascular and background classes, respectively. The likelihood potential N_{L_V} and N_{L_B} are the derivative of $U_{II}(x_s, x_{s'})$ and expressed as:

$$\frac{\partial(U_{II}(x_s))}{\partial \beta} = \begin{cases} N_{L_V} = \min_m \left(\min_j \sum_{c_{m,j}} \mathcal{P}(x_s = L_V, x_{s'}) \right) \\ N_{L_B} = \min_m \left(\min_j \sum_{c_{m,j}} \mathcal{P}(x_s = L_B, x_{s'}) \right) \end{cases} \tag{27}$$

Given an initial β_0, Eq. (25) will be iteratively updated until β_t and $\nabla(-\ln\text{PL}(x | \beta_t))$ no longer change. The resultant $\hat{\beta}_c$ approaches to a global parameter on the MRF space.

3.3. CA/CV separation

The cerebral arteries and veins present themselves with similar luminance information in the original data, meanwhile, there are also some kissing vessels in between the segmented cerebral arteries, superficial veins (include dural sinus), and intracranial non-vascular structures, which presents an interwoven structure. The resultant vascular structures should be separated with each other. As follows, we differentiate these cerebrovascular structures by using their anatomic relationships. The cerebral TOF-MRA presents vasculature imaging, while the MR-T1 presents the structure of cerebrospinal fluid, gray matter, and white matter. Their anatomic relationships are shown in Fig. 5, where the white, light gray, dark gray, and black regions in each slice denote white matter, grey matter, cerebrospinal fluid, and tentative region, respectively, while the red and green regions denote arteries and several superficial veins, respectively; and the green voxels in the black region denote dural sinuses. Furthermore, CA/CV separation are explored cerebral anatomy knowledge and detailed as follows.

Generally, our cerebrovascular segmentation result includes complete cerebral arteries and part of the superficial venous system. Thus, it is necessary to extract the whole arterial structures and remove the venous structures, while the MAP-MRF based statistical model cannot differentiate the two vascular classes merely through the intensity information or the label fields generated during the iteration processes. Therefore, we try to distinguish the arteries from another vascularity via anatomic difference. Note that dural sinuses (locate in the dura mater and include the superior sagittal sinus, transverse sinus, straight sinus, etc.) in TOF-MRA constitute the terminal of superficial venous system and distribute without closely accompanied arteries [38], and is far from the external edge of CSF, see Fig. 5. Therefore, the dural sinuses are firstly isolated to obtain part of the venous structures to acquire a tentative region (that locates outside the external edge of CSF) by using the MR-T1 data registering on the same space size. Then, we perform skull-stripping and further divide the voxels in MR-T1 into four categories of regions, i.e. CSF, GM, WM and tentative region (R_{ten}). In the above steps, the morphological process with logic operation is used to realize CA/CV separation, and the details are as follows. All regions involved in this process are shown intuitively in Fig. 6.
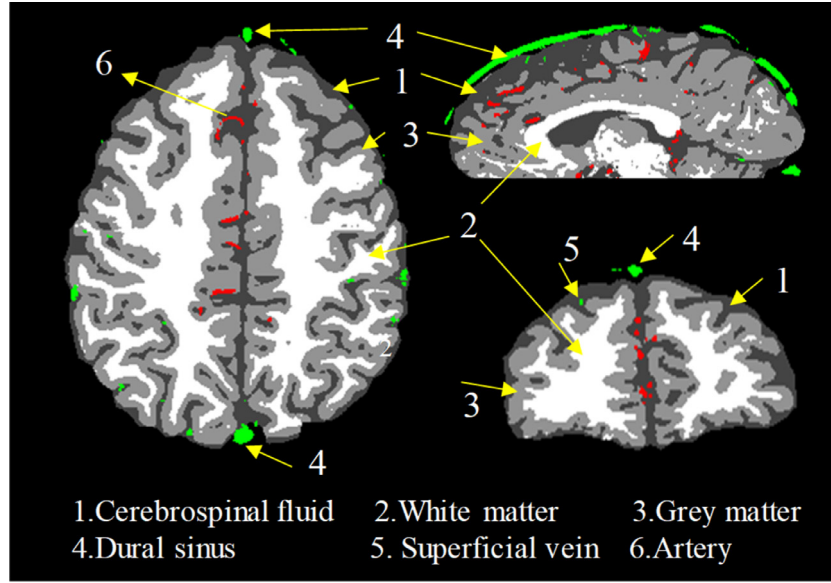
Fig. 5. The spatial relationship in between cerebral arteries, veins, dural sinuses, and intracranial tissue, where the classified tissue and vascularity from MR-T1 and TOF-MRA are fused together and shown in the axial, sagittal and coronary slices.
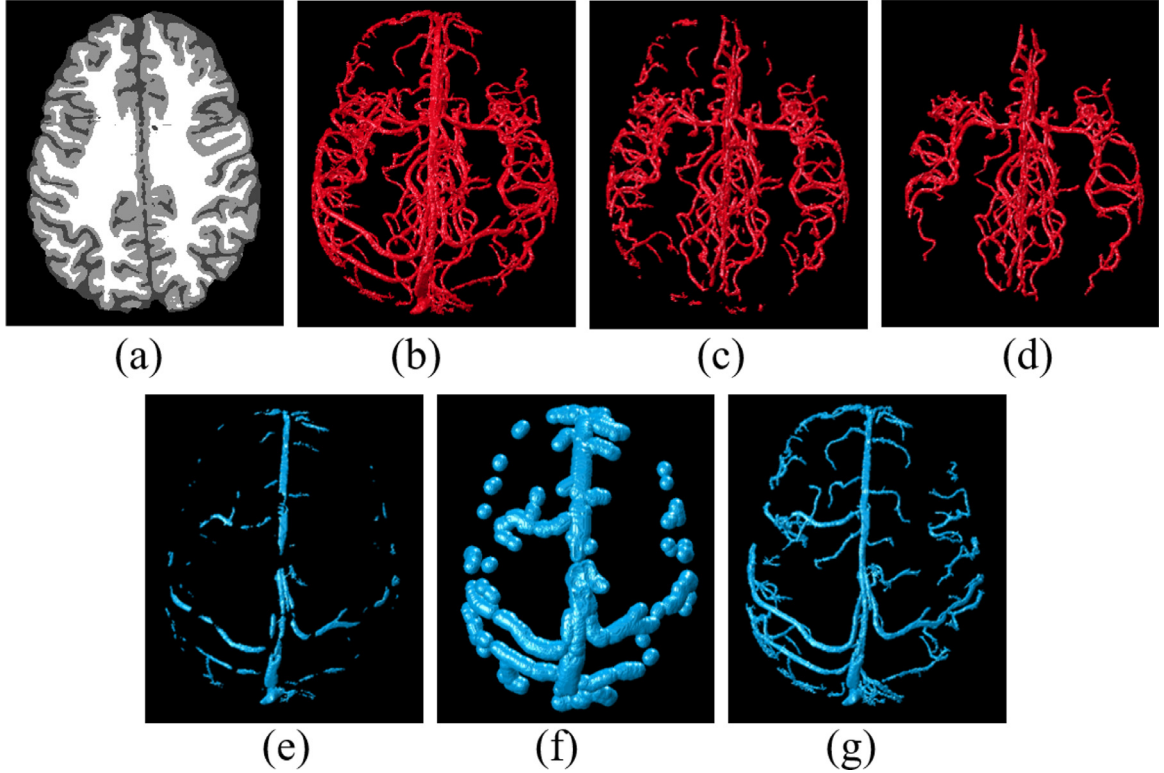

Fig. 6. The image representation of the regions involved in morphology and logic operations. Subfigure (a) shows a slice of MRA-T1, where the black region is R_{ten}; Subfigures (b)-(g) show the region of V_{seg}, V_{artery}^{ini}, V_{artery}, V_{durals}, $V_{ven\_candidate}$ and $V_{venious}$, respectively.

Vascular voxels in the tentative region are considered as dural sinuses (i.e., V_{durals}). This is obtained from cerebrovascular segmentation result (i.e., V_{seg}), which reflects a relationship of

$$V_{durals} = V_{seg} \cap R_{ten} \qquad (28)$$

Considering that the possible errors occur from registration or brain tissue classification and the TOF-MRA presents incomplete dural sinuses, a venous candidate space (i.e., $V_{ven\_candidate}$) should be explored on basis of the obtained dural sinuses. We use a sphere-like structure element (named *strel*) with radius $r_V = 8$ (an

empirical maximal sinuses-radius) to dilate the obtained dural sinuses, i.e.,

$$V_{ven\_candidate} = V_{durals} \oplus strel \qquad (29)$$

Then, an initial arterial structure V_{artery}^{ini} is attained by

$$V_{artery}^{ini} = V_{seg} - (V_{seg} \cap V_{ven\_candidate}) \qquad (30)$$

For description of arterial structure with some connected regions (i.e., V_{artery}), the above-mentioned operations remove the discontinuous structures (they mainly consist of vein structures, and

Table 1
Models for the comparison in low-level process.

Model	Preprocess of datasets	The FMM
Lu et al. [15]	Multiscale filtering	FMM 1: $f(i) = \sum\limits_{l=1}^{2} w_{El} f_{El}(i) + w_G f_G(i)$
Wen et al. [16]	Removing non-vascular structures	FMM 2: $f(i) = w_R f_R(i) + \sum\limits_{l=1}^{2} w_{Gl} f_{Gl}(i)$
Hassouna et al. [13] and Zhou et al. [14]	None	FMM 3: $f(i) = w_R f_R(i) + \sum\limits_{l=1}^{3} w_{Gl} f_{Gl}(i)$
The proposed	Skull-stripping	FMM 4: $f(i) = \sum\limits_{l=1}^{3} w_{Gl} f_{Gl}(i)$

are believed as insignificant ones). The resultant vein structure (i.e., $V_{venious}$) can be obtained by

$$V_{venious} = V_{seg} - V_{artery} \tag{31}$$

4. Experiments and results

To validate our algorithms of vascular segmentation and CA/CV separation, the corresponding experiments were explored in detail and arranged as follows. The data used were acquired from a public database (http://www.insight-journal.org/midas/community/view/21) which contained 109 sets of MR images of the brain of healthy subjects. These datasets were acquired on a 3T unit under standardized protocols, including TOF-MRA and MR-T1 data. The size of each T1 data was 256 slices of 170×256 voxels, with voxel size $1 \times 1 \times 1$ mm^3; while the size of each TOF-MRA was 128 slices of 448×448 voxels, with voxel size $0.5 \times 0.5 \times 0.8$ mm^3. The main algorithms were coded by MATLAB 2018a. These experiments were carried out on a PC Inter(R) Core (TM) i7-4790 CPU @ 3.60 GHz, 24 GB-RAM.

Since there were 20 datasets without T1 data, only 89 datasets had been used to verify the CA/CV separation method. All the 109 datasets were used in cerebrovascular segmentation. We randomly selected 10 datasets as the training ones from the public databases to produce corresponding ground truths (GTs) for quantitative validation, while the remaining datasets were the testing datasets used for clinical visual evaluation. Each set of GTs included two parts. One was the manually segmented cerebrovascular structure on the TOF-MRA data under the supervision of three experienced neurosurgeons, which is used to evaluate our statistical method. The other was the manually labeled arteries and veins by the neurosurgeons from the GTs, which was used to evaluate the CA/CV separation results. To avoid confusion, the reference standard of arteries and veins in each GT were annotated and denoted respectively as A_{ref} and V_{ref} (with $A_{ref} \cup V_{ref} = $ GT). All the manual operations were implemented by using MIMICS (i.e., the medical image processing software). To determine a vascular voxel as arterial or venous class, the GTs were overlaid with the classified intracerebral tissue image for the neurosurgeons' annotation. Note, in our experiments, all quantitative and qualitive results uses the largest connected regions of cerebrovascular segmentation.

4.1. Quantitative evaluation of cerebrovascular segmentation

For current statistical models on TOF-MRA data, the FMMs are shown in Table 1. Our strategy optimized traditional models in their Markov low- and high-level processes. To evaluate their segmentation processes, a cross comparison experiments consisted of 24 composite modes by arranging low- and high-level models (see Table 2) with respective PWS-potential on NBSs. The composite modes include the existing statistical models used for vascular segmentation. Dice similarity coefficient (DSC) was employed to validate the segmentation effects of each method and was expressed as

$$\text{DSC} = \frac{2TP}{2TP + FP + FN} \tag{32}$$

where TP, FP and FN represented the number of true positives, false positive, and false negative, respectively. The evaluation results are recorded in Table 2. According to the values in the second column of the table, it can be seen that the results obtained by our low-level model have the highest average DSC of 0.887 and the lowest variance of 0.027, comparing with that by the other three low-level models. Furthermore, each row in Table 2 evaluates the low-level model combining with different PWS-potentials and NBSs, where the multi-pattern NBS improves the DSCs but with longer computation time than other NBSs; meanwhile, the combination of Markov low- and high-level models achieve higher DSC than mere low-level models, and the proposed modeling strategy obtains higher DSCs than the traditional MRF models. For example, in the 4-th column of Table 2, the proposed high-level model optimizes existing ones and obtains the DSCs of 0.913, 0.912, and 0.933, respectively with 6-, 26-, and multi-pattern NBSs. Overall consideration of DSC and running time recorded in Table 2, the 6-NBS based Markov model has a better cost performance.

In addition, four existing methods [13–16] and the proposed method were compared by using the average FPR and false FNR in between their 10 GTs and corresponding segmentation results. The FPR and FNR were expressed as

$$\text{FPR} = \frac{FP}{FP + TN} \tag{33}$$

$$\text{FNR} = \frac{FN}{FN + TP} \tag{34}$$

The average values of FNR and FPR of different methods were recorded in Table 3. As it was expected, the proposed model generated the least FNR and the standard deviation. Comparing with the other three methods in Table 3, the average FNR by ours is greatly reduced, while there is a little increment of FPR.

4.2. Clinical evaluation

The clinical evaluation was proceeded by three neurosurgeons using observational method. We divide the evaluation results into three levels namely, *Good case*: there is over 85% network coverage of the cerebrovascular structure, while covered vessels have unobvious over-segmentation. *General case*: over 60% network coverage, where or a small quantity of over-segmentation and pseudo-vascular structures are acceptable. *Poor case*: less 40% network coverage, only the large-sized vessels were obtained, or there exist too many over-segmentation, and pseudo-vascular structures. The evaluation results of the four different methods were recorded in Table 4. According to Table 4, the proposed model performs well on most of the 99 testing datasets with 86 good cases, 10 general cases, and only 3 bad cases, in comparing with the results in Zhou et al. [14].

Table 2
DSC and processing time comparison of segmentation results by composite modes of statistical modelling.

Low-level models	Results of low-level process	Optimized results by traditional high-level model with Eq. (6) and NBS			Optimized results by proposed high-level model with Eq. (16) and NBS		
		6-NBS	26-NBS	MP-NBS	6-NBS	26-NBS	MP-NBS
FMM 1	0.662±0.062	0.512±0.053	0.385±0.057	0.519±0.053	0.684±0.059	0.678±0.055	0.773±0.079
	23.62(s)	34.84(s)	33.45(s)	48.54(s)	36.36(s)	46.21(s)	108.43(s)
FMM 2	0.494±0.096	0.482±0.093	0.476±0.091	0.488±0.093	0.495±0.096	0.493±0.095	0.498±0.096
	18.94(s)	29.92(s)	30.56(s)	43.91(s)	32.97(s)	41.85(s)	104,78(s)
FMM 3	0.769±0.072	0.741±0.066	0.704±0.049	0.754±0.068	0.768±0.079	0.761±0.075	0.774±0.074
	22.38(s)	33.25(s)	33.61(s)	47.39(s)	36.03(s)	45.76(s)	107.21(s)
FMM 4	0.887±0.027	0.843±0.027	0.791±0.033	0.848±0.027	**0.913±0.022**	**0.912±0.025**	**0.933±0.017**
	17.23(s)	27.86(s)	27.20(s)	41.30(s)	30.83(s)	39.93(s)	102.91(s)

Note: The DSCs are denoted with mean value ± standard deviation and the process times are denoted by average value in seconds.

Table 3
The quantitative evaluation of segmentation methods.

Methods & Ref.	FNR		FPR (%)	
	Average	Std	Average	Std
Lu et al. [15]	0.502	0.066	0.072	0.031
Wen et al. [16]	0.673	0.080	0.011	0.006
Hassouna et al. [13]	0.453	0.082	0.069	0.038
Zhou et al. [14]	0.406	0.076	0.079	0.043
The proposed	0.158	0.041	0.091	0.019

Note: 'std' represents standard deviation.

Table 4
Visual evaluation of the results of different segmentation methods.

Method & Ref.	Good	General	Bad
Lu et al. [15]	0	54	45
Wen et al. [16]	0	49	50
Hassouna et al. [13]	7	78	14
Zhou et al. [14]	10	80	9
The proposed	86	10	3

Table 5
The quantitative evaluation for CA/CV separation.

	FPR	FNR	agreement
Average	0.022	0.041	0.976
Std	0.017	0.022	0.021

Note: 'std' represents standard deviation.

The clinical evaluation also involved intuitive comparison from neurosurgeons' visual experience and was carried out by visualizing the segmentation results from the testing datasets. Ten sets of the segmentation results were shown in Fig. 7, which included four, three, and three sets of high-, middle-, and low-contrast datasets, respectively. From Fig. 7 we can see that all methods show similar performance on the segmentation of the vessel with large size or in high contrast region. Seeing in detail, it is obvious that the proposed model performs better on the segmentation of the vessel with small size or in low contrast region, e.g., Fig. 7(f).

4.3. Validation of the regularization parameter estimation

It is necessary to verify the effectiveness of regularization parameter estimation, we applied Eqs. (25)–(27) to the datasets with GTs, and differentiate the Markov high-level processes with different NBSs. We verified the effectiveness of the parameter estimation as follows, then apply the automatic estimation for all the datasets.

Like the trial-error-test used by Zhou et al. [14], firstly, we took 9 test values in between 0 and 1 at the equal interval (i.e. 0.1) as the beta-values and computed the DSC of each segmentation result with different NBSs. Then, a parameter range of length 0.1 was determined around the optimal beta-value (as the center) corresponding to the maximum DSC. Through this trial-error-test, we can validate whether the estimated regularization parameter is significant and close to the optimal beta-value. The validation was illustrated and compared in Fig. 8. For each training dataset, the automatically estimated parameter was compared with the optimal beta-value interval corresponding to the maximum DSC, as the re-

sult, the estimated parameter falls in the test beta-values interval on corresponding maximum DSC. For example, on the left side of Fig. 8, the high-level model with 6-, 26-, and multi-pattern NBSs produce the optimal test beta-values of 0.4, 0.6, and 0.4, respectively on the dataset of Normal001, while the estimated parameters all fall into the corresponding test intervals of 0.35–0.45, 0.55–0.65, and 0.35–0.45. Similar effects emerged in 9 other training datasets, see the right side of Fig. 8.

4.4. Evaluation of CA/CV separation

To validate the results by the proposed CA/CV separation method, both quantitative and qualitative comparisons were used in this part. The quantitative strategy was undertaken in [21], where a fair scheme of this classified assessment avoided the impact of cerebrovascular over-segmentation. Namely, only the voxels presented both in the CA/CV separation result and the reference standard of arteries and veins were considered. Similar to [21], a voxel-based agreement was used to provide the quantitative evaluation and defined as

$$\text{agreement} = \frac{length\left(A_{ref} \cap A_{cal}\right) + length\left(V_{ref} \cap V_{cal}\right)}{length\left(\left(A_{ref} \cup V_{ref}\right) \cap \left(A_{cal} \cup V_{cal}\right)\right)} \quad (35)$$

where the function $length()$ is defined to calculate the total number of voxels in a region. A_{ref} and V_{ref} were the sets of arteries and veins of the reference standard, and A_{cal} and V_{cal} were the corresponding results of the CA/CV separation. Based on the reference standard, FNR and FPR were also calculated referring with the following expressions

$$\text{FNR} = \frac{length\left(A_{ref} \cap V_{cal}\right)}{length\left(A_{ref}\right)} \quad (36)$$

$$\text{FPR} = \frac{length\left(V_{ref} \cap A_{cal}\right)}{length\left(V_{ref}\right)} \quad (37)$$

The results were recorded in Table 5. which shows the average agreement of CA/CV separation achieve 0.976, while the FPR and FNR reach to 0.022 and 0.041, respectively. The CA/CV separation results over the testing datasets (mentioned in Section 4.2) were shown in the last two rows of Fig. 7 from which we can see the superficial venous structures (including a part of dural sinuses and
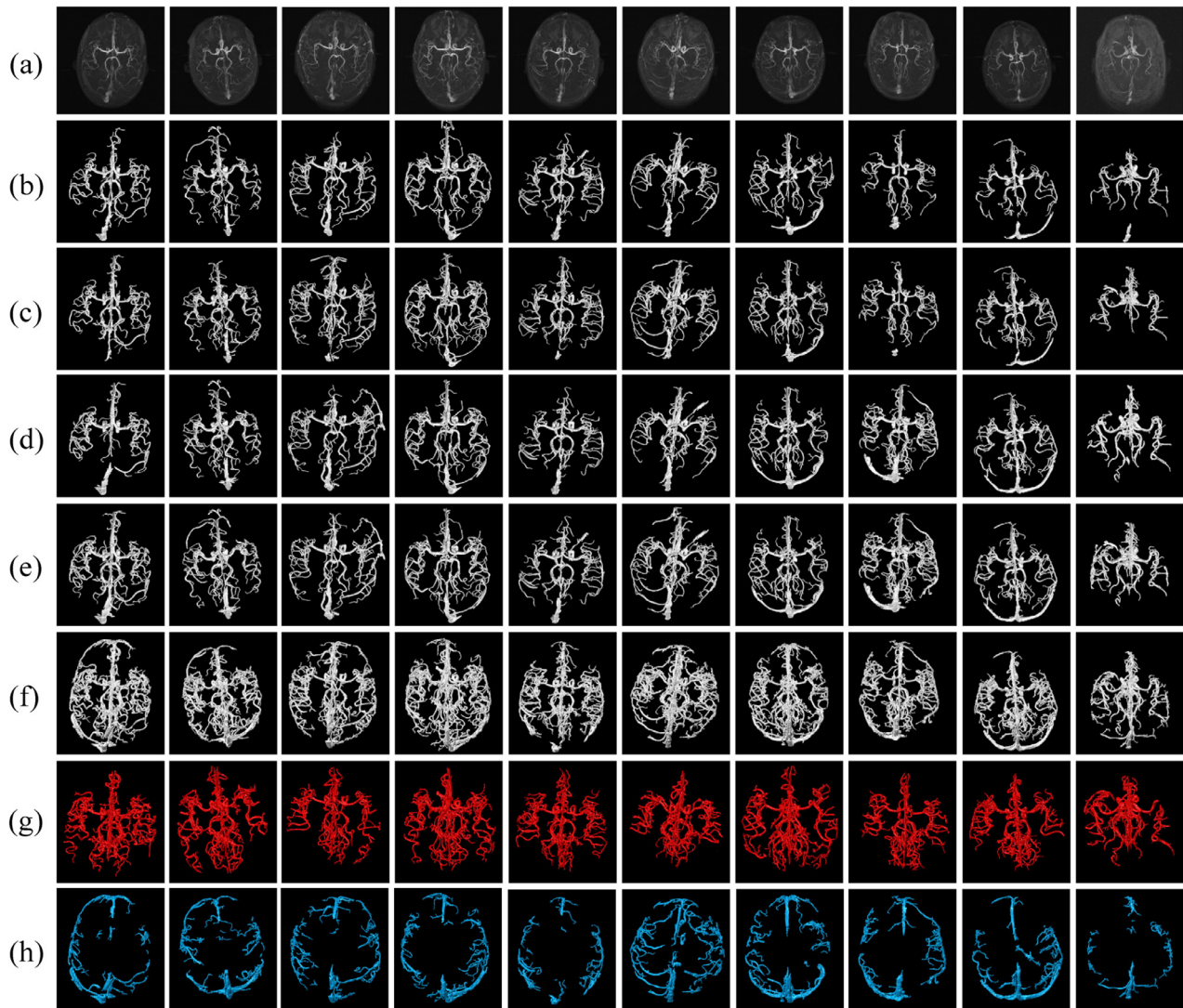
Fig. 7. Comparisons of the segmentation results from ten TOF-MRA datasets by five different methods. Each column corresponds a dataset. From left to right, the datasets are Normal-010, −017, −019, −021, −012, −013, −020, −022, −011, −029. The subfigures in each row represent (a) the MIP of the raw data and the resultant segmentation by (b) Lu et al. [15], (c) Wen et al. [16], (d) Hassouna et al. [13], (e) Zhou et al. [14], and (f) the proposed method's, respectively. Subfigures in (g) and (h) show the separated arterial and venous structures corresponding to that in (f).
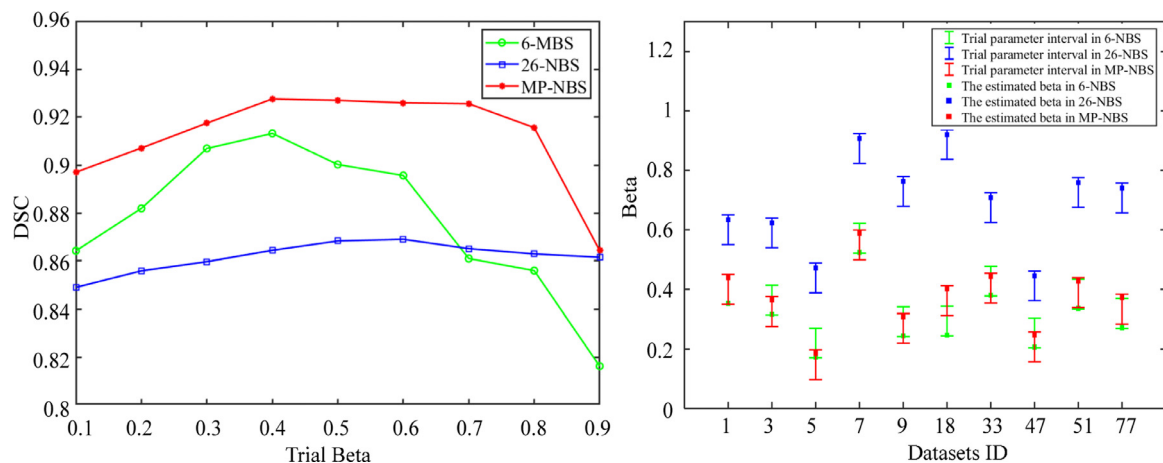
Fig. 8. Verification of estimated regularization parameter with trial-error-test. The left: DSCs obtained with the trial beta-values in 6-, 26-, and MP-NBS given an example with dataset Normal001. The right: comparison of trial parameter intervals and estimated parameter values in 6- 26- and MP-NBS over 10 normal datasets.

cortical veins) can be completely separated from the arterial structures, while lots of fine arterial structures are not faultily divided into venous ones, especially in the posterior arterial area.

5. Discussion

For quantitative evaluation of cerebrovascular segmentation, the accuracy for each segmentation approach was firstly presented in Table 2. It can be found that our statistical modeling strategy leads to more sensitive and accurate cerebrovascular segmentation than the existing methods, e.g., higher DSCs and lower variance by ours, which partly benefits from the high SNR produced by the skull-stripping operation. In theory, the traditional energy function (i.e., Eq. (7)) of the Markov high-level model is effective in terms of removing noise, but it cannot recover the vascular structures in low-contrast region. To be more specific, vascular voxels in the low-contrast region with initial neighborhood labels of background class (resulting from the low-level process) cannot be reclassified to vascular class no matter how many iterations, which results in loss of small-sized vessels or disconnection of the curved vessels. Overall, the traditional Markov high-level processes over depend on spatial continuity of the label field generated from the low-level model, and have nothing to do with the original image information; therefore, their roles are only to denoise label field while impossible to restore the potential vascular structure during the Markov iteration process. In contrast, the proposed high-level process relies closely on vascular shape characteristic (i.e., VFM and VDF) to explore and restore the blood vessels in the low-contrast image context, thus improves the sensitivity of vessel detection.

Table 3 shows the values of FNR and FPR (reflected the degree of under- and over-segmentation, respectively) for five different statistical models. Comparing with the other three methods in Table 3, the average FNR by ours is greatly reduced, that means our methods can obtain the better completeness of vascular segmentation. A minor flaw is that a little increment of FPR is produced by our methods. The reason is that the vessel enhancement filters respond to brain tissues boundaries (e.g., the brain sulci) to a lesser extent. However, the qualitative comparison in Fig. 7 further proves that our proposed strategy can obtain larger cerebrovascular network coverage, even in low contrast context.

Many experiments indicated that DSC cannot fully express the effect of vascular segmentation alone. Because the thicker vessels account for a large proportion of vascular volume, while its network length only occupies less part of the total network. See the last row of Table 4, the neurosurgeons verified the very higher network coverage rate by ours than that by others. The traditional methods produced poor vascular extraction results not only in DSC but also in the network coverage rate, according to the neurosurgeons.

The regularization parameter estimation was quantitively validated in each dataset that has GT. It is undeniable that there is a positive correlation between the pseudo-likelihood and the resultant DSC in an ideal condition. Due to the brain tissue and artifacts in the angiography, the resultant DSCs are inevitably affected. It can be seen from the trial and error process that the smaller the regularization parameter value would produce the more remarkable effect of the low-level model, and cause the more under-segmentation results during the Markov iteration process; vice versa. Fortunately, trial-and-error verification like experimental results in Fig. 8 proved that the regularization parameter estimation approaches the real solution. Therefore, our method automates the MAP-MRF based statistical modeling.

With acquired cerebrovascular complete topology, the CA/CV separation was automated by knowledge-guided morphological operation with Eqs. (28)–(31). By analyzing the results, we found that the main errors in the CA/CV separation occurred in the adjoining areas of the lateral cerebral veins and middle cerebral arteries. Overall, all the evaluations demonstrated high agreement on our statistical modeling advantages.

6. Conclusion

Brain arterial network extraction from TOF-MRA data is of great significance and also a rarely studied field. This method proposed a novel statistical modeling strategy and developed a MAP-MRF model for cerebrovascular segmentation. Existing statistical models suffer from over- or under-segmentation in the low-level process and fail to obtain small vessel in the high-level process; thus, we optimized the two processes. In addition, the proposed method of CA/CV separation based on spatial prior in between vascular and non-vascular structures can successfully separate arteries, veins and dural sinus, thus demonstrates great clinical value. Summarily, cerebrovascular segmentation and CA/CV separation were solved successfully in this paper. The experimental results proved that our method can generate more complete and accurate results from the quantitative and qualitative perspectives in comparing with other methods, thus indicates that the proposed method has a great capability of clinical application, especially for the computer-assisted cerebrovascular diagnosis and interventional therapy.

Declaration of Competing Interest

The authors declare that they have no conflicts of interest.

Acknowledgements

Appendices

Appendix A. EM parameters update equations

This section shows how to derive the parameters update equations of the proposed Gaussian Mixture Model (GMM) in its closed form. The posterior probability of the Gaussian distribution is given by

$$f(gl|y_i) = \frac{w_l f_{gl}(y_i)}{f(y_i)}, l = 1, 2, 3 \tag{A.1}$$

The fundamental equation of the EM expectation step for the finite mixture of densities is given by

$$Q\left(\Theta^{k+1}, \Theta^k\right) = \sum_{i=1}^{N} \sum_{l=1}^{M} \log\left(w_l^{k+1} f_l\left(y_i|\Theta^{k+1}\right)\right) f\left(gl|y_i, \Theta^k\right) \tag{A.2}$$

where Q is the expected value of the log-likelihood function of the Gaussian mixture distribution, Θ^k is the mixture parameters at iteration k, M is the total number of classes, which is the three in our case, and N is the total number of voxels in the data volume. Eq. (A.2) can be rewritten as the sum of two independent terms:

$$Q\left(\Theta^{k+1}, \Theta^k\right) = \sum_{i=1}^{N} \sum_{l=1}^{M} \log\left(w_l^{k+1}\right) f(gl|y_i, \Theta^k)$$
$$+ \sum_{i=1}^{N} \sum_{l=1}^{M} \log\left(f_l\left(y_i|\Theta^{k+1}\right)\right) f(gl|y_i, \Theta^k) \tag{A.3}$$

The maximization step of the EM finds the mixture parameters by maximizing each term of Eq. (A. 3) independently. Let us maximize the term containing w_l^{k+1} under the constraint $\sum_{l=1}^{M} w_l = 1$ using the Lagrange multiplier and solve the following equation:

$$\frac{\partial}{\partial w_l^{k+1}} \left[\sum_{i=1}^{N} \sum_{l=1}^{M} \log \left(w_l^{k+1} \right) f(gl|y_i, \Theta^k) + \lambda \left(\sum_{l=1}^{M} w_l^{k+1} - 1 \right) \right] = 0 \tag{A.4}$$

or

$$\sum_{i=1}^{N} \frac{1}{w_l^{k+1}} f(gl|y_i, \Theta^k) + \lambda = 0 \tag{A.5}$$

Summing both sides over l, we get that $\lambda = -N$ resulting in

$$w_l^{k+1} = \frac{1}{N} \sum_{i=1}^{N} f(gl|y_i, \Theta^k) \tag{A.6}$$

The mean of each Gaussian distribution can be derived by maximizing the term of Eq. (A.3) that contains μ_l^{k+1},

$$\frac{\partial}{\partial \mu_l^{k+1}} \left[\sum_{i=1}^{N} \log \left(\frac{1}{\sqrt{2\pi} \sigma_l^{k+1}} \exp \left(-\frac{(y_i - \mu_l^{k+1})^2}{2\sigma_l^{2k+1}} \right) \right) f(gl|y_i, \Theta^k) \right] = 0 \tag{A.7}$$

Hence, the update equation of the mean of the Gaussian distribution is given by

$$\mu_l^{k+1} = \frac{\sum_{i=1}^{N} y_i f(gl|y_i, \Theta^k)}{\sum_{i=1}^{N} f(gl|y_i, \Theta^k)} \tag{A.8}$$

The variance of each Gaussian distribution can be derived by maximizing the term of Eq. (A.3) that contains $(\sigma_l^{k+1})^2$,

$$\frac{\partial}{\partial (\sigma_l^{k+1})^2} \times \left[\sum_{i=1}^{N} \log \left(\frac{1}{\sqrt{2\pi} \sigma_l^{k+1}} \exp \left(-\frac{(y_i - \mu_l^{k+1})^2}{2(\sigma_l^{k+1})^2} \right) \right) f(gl|y_i, \Theta^k) \right] = 0 \tag{A.9}$$

Hence, the update equation of the variance of the Gaussian distribution is given by

$$(\sigma_l^2)^{k+1} = \frac{\sum_{i=1}^{N} (y_i - \mu_l^{k+1})^2 f(gl|y_i, \Theta^k)}{\sum_{i=1}^{N} f(gl|y_i, \Theta^k)} \tag{A.10}$$

Appendix B. regularization parameter estimation

According to Eqs. (16)–(20), the first- and second-order energy functions in Eqs. (18)–(20) could be written respectively as

$$\begin{cases} U_I(x_s) = -\ln f_{x_s} \\ U_{II}(x_s, x_r) = \min_m \left(\min_j \sum_{c_{m,j}} \beta \mathcal{P}(x_s, x_{s'}) \right) \end{cases} \tag{B.1}$$

We can deduce the partial derivative of $U_{II}(x_s, x_{s'})$ as

$$\frac{\partial (U_{II}(x_s))}{\partial \beta} = \begin{cases} N_{L_V} = \min_m \left(\min_j \sum_{c_{m,j}} \mathcal{P}(x_s = L_V, x_{s'}) \right) \\ N_{L_B} = \min_m \left(\min_j \sum_{c_{m,j}} \mathcal{P}(x_s = L_B, x_{s'}) \right) \end{cases} \tag{B.2}$$

For the negative logarithm of pseudo-likelihood, i.e. Eq. (24), we could deduce its partial derivative as

$$\frac{\partial (-\ln PL(x|\beta))}{\partial \beta}$$

$$= \sum_{s \in \Omega, s' \in \mathcal{N}_s} \left(\frac{\partial (U_{II}(x_s))}{\partial \beta} + \frac{\partial \left(\ln \sum_{x_s \in \{L_V, L_B\}} \exp \left(-U_I(x_s) - U_{II}(x_s, x_{s'}) \right) \right)}{\partial \beta} \right)$$

$$= \sum_{s \in \Omega, s' \in \mathcal{N}_s} \left(\frac{\partial (U_{II}(x_s))}{\partial \beta} + \frac{\begin{array}{c} -\exp \left(-U_I(L_V) - U_{II}(L_V, x_{s'}) \right) \frac{\partial (U_{II}(L_V))}{\partial \beta} \\ -\exp \left(-U_I(L_B) - U_{II}(L_B, x_{s'}) \right) \frac{\partial (U_{II}(L_B))}{\partial \beta} \end{array}}{\begin{array}{c} \exp \left(-U_I(L_V) - U_{II}(L_V, x_{s'}) \right) \\ + \exp \left(-U_I(L_B) - U_{II}(L_B, x_{s'}) \right) \end{array}} \right) \tag{B.3}$$

Since the label x_s of the site s in current label field is either L_V or L_B, which make $\frac{\partial (U_{II}(x_s))}{\partial \beta}$ can be N_{L_V} or N_{L_B}, so Eq. (B.3) could be rewritten as

$$\frac{\partial (-\ln PL(x|\beta))}{\partial \beta} = \sum_{s \in \Omega, s' \in \mathcal{N}_s} \left(\delta (x_s - L_V) N_{L_V} + \delta (x_s - L_B) N_{L_B} \right.$$

$$\left. - \frac{f_{L_V} \exp \left(-\beta N_{L_V} \right) N_{L_V} + f_{L_B} \exp \left(-\beta N_{L_B} \right) N_{L_B}}{f_{L_V} \exp \left(-\beta N_{L_V} \right) + f_{L_B} \exp \left(-\beta N_{L_B} \right)} \right) \tag{B.4}$$

References

[1] M. Ghazal, Y. Al Khalil, A. El-Baz, An unsupervised parametric mixture model for automatic cerebrovascular segmentation, in: cardiovasc, Imaging Image Anal. (2018) 95–108 CRC Press, doi:10.1201/9780429441493-5.

[2] P.D. Schellinger, G. Richter, M. Köhrmann, A. Dörfler, Noninvasive angiography (magnetic resonance and computed tomography) in the diagnosis of ischemic cerebrovascular disease, Cerebrovasc. Dis. 24 (2007) 16–23, doi:10.1159/000107375.

[3] Ö. Özsarlak, J.W.V. Goethem, M. Maes, P.M. Parizel, MR angiography of the intracranial vessels: technical aspects and clinical applications, Neuroradiology 46 (2004) 955–972, doi:10.1007/s00234-004-1297-9.

[4] D. Lesage, E.D. Angelini, I. Bloch, G. Funka-Lea, A review of 3D vessel lumen segmentation techniques: models, features and extraction schemes, Med. Image Anal. 13 (2009) 819–845, doi:10.1016/j.media.2009.07.011.

[5] S. Moccia, E. De Momi, S. El Hadji, L.S. Mattos, Blood vessel segmentation algorithms—Review of methods, datasets and evaluation metrics, Comput. Methods Programs Biomed. 158 (2018) 71–91, doi:10.1016/j.cmpb.2018.02.001.

[6] Q. Li, J. You, D. Zhang, Vessel segmentation and width estimation in retinal images using multiscale production of matched filter responses, Expert Syst. Appl. 39 (2012) 7600–7610, doi:10.1016/j.eswa.2011.12.046.

[7] U.T.V. Nguyen, A. Bhuiyan, L.A.F. Park, K. Ramamohanarao, An effective retinal blood vessel segmentation method using multi-scale line detection, Pattern Recognit. 46 (2013) 703–715, doi:10.1016/j.patcog.2012.08.009.

[8] Y. Cheng, X. Hu, J. Wang, Y. Wang, S. Tamura, Accurate vessel segmentation with constrained B-snake, IEEE Trans. Image Process. 24 (2015) 2440–2455, doi:10.1109/tip.2015.2417683.

[9] T. Lv, G. Yang, Y. Zhang, J. Yang, Y. Chen, H. Shu, L. Luo, Vessel segmentation using centerline constrained level set method, Multimed. Tools Appl. 78 (2019) 17051–17075, doi:10.1007/s11042-018-7087-x.

[10] Y. Zhao, X. Wang, X. Wang, F.Y. Shih, Retinal vessels segmentation based on level set and region growing, Pattern Recognit. 47 (2014) 2437–2446, doi:10.1016/j.patcog.2014.01.006.

[11] M.M. Almi'ani, B.D. Barkana, A modified region growing based algorithm to vessel segmentation in magnetic resonance angiography, in: 2015 Long Isl, Syst. Appl. Technol. IEEE (2015) 1–7, doi:10.1109/lisat.2015.7160191.

[12] Y. Chen, Y. Zhang, J. Yang, Q. Cao, G. Yang, J. Chen, H. Shu, L. Luo, J.-.L. Coatrieux, Q. Feng, Curve-like structure extraction using minimal path propagation with backtracking, IEEE Trans. Image Process. 25 (2015) 988–1003, doi:10.1109/tip.2015.2496279.

[13] M.S. Hassouna, A.A. Farag, S. Hushek, T. Moriarty, Cerebrovascular segmentation from ToF using stochastic models, Med. Image Anal. 10 (2006) 2–18, doi:10.1016/j.media.2004.11.009.

[14] S. Zhou, W. Chen, F. Jia, Q. Hu, Y. Xie, M. Chen, J. Wu, Segmentation of brain magnetic resonance angiography images based on MAP–MRF with multi-pattern neighborhood system and approximation of regularization coefficient, Med. Image Anal. 17 (2013) 1220–1235, doi:10.1016/j.media.2013.08.005.

[15] P. Lu, J. Xia, Z. Li, J. Xiong, J. Yang, S. Zhou, L. Wang, M. Chen, C. Wang, A vessel segmentation method for multi-modality angiographic images based on multi-scale filtering and statistical models, Biomed. Eng. Online 15 (2016) 120, doi:10.1186/s12938-016-0241-7.

[16] L. Wen, X. Wang, Z. Wu, M. Zhou, J.S. Jin, A novel statistical cerebrovascular segmentation algorithm with particle swarm optimization, Neurocomputing 148 (2015) 569–577, doi:10.1016/j.neucom.2014.07.006.

[17] R. Phellan, A. Peixinho, A. Falcão, N.D. Forkert, Vascular segmentation in tof mra images of the brain using a deep convolutional neural network, in: intravasc. Imaging Comput. Assist. Stenting, Large-Scale Annot. Biomed. Data Expert Label Synth., Springer, 2017: pp. 39–46, 10.1007/978-3-319-67534-3_5

[18] F. Zhao, Y. Chen, F. Chen, X. He, X. Cao, Y. Hou, H. Yi, X. He, J. Liang, Semi-supervised cerebrovascular segmentation by hierarchical convolutional neural network, IEEE Access 6 (2018) 67841–67852, doi:10.1109/access.2018.2879521.

[19] M.J. Cardoso, T. Arbel, S.L. Lee, V. Cheplygina, S. Balocco, D. Mateus, G. Zahnd, L. Maier-Hein, S. Demirci, E. Granger, Intravascular Imaging and Computer Assisted Stenting, and large-scale annotation of biomedical data and expert label synthesis, in: CVII-STENT Second Int. Work. LABELS, Springer, 2017, 10.1007/978-3-319-67534-3

[20] J.-P. Charbonnier, M. Brink, F. Ciompi, E.T. Scholten, C.M. Schaefer-Prokop, E.M. Van Rikxoort, Automatic pulmonary artery-vein separation and classification in computed tomography using tree partitioning and peripheral vessel matching, IEEE Trans. Med. Imaging 35 (2015) 882–892, doi:10.1109/tmi.2015.2500279.

[21] C. Payer, M. Pienn, Z. Bálint, A. Shekhovtsov, E. Talakic, E. Nagy, A. Olschewski, H. Olschewski, M. Urschler, Automated integer programming based separation of arteries and veins from thoracic CT images, Med. Image Anal. 34 (2016) 109–122, doi:10.1016/j.media.2016.05.002.

[22] R.M. Stefancik, M. Sonka, Highly automated segmentation of arterial and venous trees from three-dimensional magnetic resonance angiography (MRA), Int. J. Cardiovasc. Imaging 17 (2001) 37–47, doi:10.1023/A:1010656618835.

[23] T. Lei, J.K. Udupa, P.K. Saha, D. Odhner, Artery-vein separation via MRA-an image processing approach, IEEE Trans. Med. Imaging 20 (2001) 689–703, doi:10.1109/42.938238.

[24] B. Dashtbozorg, A.M. Mendonça, A. Campilho, An automatic graph-based approach for artery/vein classification in retinal images, IEEE Trans. Image Process. 23 (2013) 1073–1083, doi:10.1109/tip.2013.2263809.

[25] V.S. Joshi, J.M. Reinhardt, M.K. Garvin, M.D. Abramoff, Automated method for identification and artery-venous classification of vessel trees in retinal vessel networks, PLoS ONE 9 (2014) e88061, doi:10.1371/journal.pone.0088061.

[26] F. Santini, S. Patil, S. Meckel, K. Scheffler, S.G. Wetzel, Double-reference cross-correlation algorithm for separation of the arteries and veins from 3D MRA time series, J. Magn. Reson. Imaging An Off. J. Int. Soc. Magn. Reson. Med. 28 (2008) 646–654, doi:10.1002/jmri.21499.

[27] A. Mendrik, E. Vonken, B. van Ginneken, E. Smit, A. Waaijer, G. Bertolini, M.A. Viergever, M. Prokop, Automatic segmentation of intracranial arteries and veins in four-dimensional cerebral CT perfusion scans, Med. Phys. 37 (2010) 2956–2966, doi:10.1118/1.3397813.

[28] N. Feng, J. Qiu, P. Li, X. Sun, C. Yin, W. Luo, S. Chen, Q. Luo, Simultaneous automatic arteries-veins separation and cerebral blood flow imaging with single-wavelength laser speckle imaging, Opt. Express 19 (2011) 15777–15791, doi:10.1364/oe.19.015777.

[29] Y. Wang, D. Hu, Y. Liu, M. Li, Cerebral artery–vein separation using 0.1-Hz oscillation in dual-wavelength optical imaging, IEEE Trans. Med. Imaging 30 (2011) 2030–2043, doi:10.1109/tmi.2011.2160191.

[30] M. Meijs, A. Patel, S.C. van de Leemput, M. Prokop, E.J. van Dijk, F.-E. de Leeuw, F.J.A. Meijer, B. van Ginneken, R. Manniesing, Robust segmentation of the full cerebral vasculature in 4D CT of suspected stroke patients, Sci. Rep. 7 (2017) 15622, doi:10.1038/s41598-017-15617-w.

[31] C.W. Therrien, Decision, Estimation, and Classification: an Introduction to Pattern Recognition and Related Topics, Wiley, New York, 1989, doi:10.5860/choice.27-0356.

[32] J. Besag, Spatial interaction and the statistical analysis of lattice systems, J. R. Stat. Soc. Ser. B. 36 (1974) 192–225, doi:10.1111/j.2517-6161.1974.tb00999.x.

[33] S.Z. Li, Markov Random Field Modeling in Image Analysis, Springer, London, 2009, doi:10.1007/978-1-84800-279-1.

[34] H. Derin, H. Elliott, Modeling and segmentation of noisy and textured images using Gibbs random fields, IEEE Trans. Pattern Anal. Mach. Intell. (1987) 39–55, doi:10.1109/tpami.1987.4767871.

[35] N. Passat, C. Ronse, J. Baruthio, J.-P. Armspach, M. Bosc, J. Foucher, Using multimodal MR data for segmentation and topology recovery of the cerebral superficial venous tree, Int. Symp. Vis. Comput. (2005) 60–67 Springer, doi:10.1007/11595755_8.

[36] T. Jerman, F. Pernuš, B. Likar, Ž. Špiclin, Enhancement of vascular structures in 3D and 2D angiographic images, IEEE Trans. Med. Imaging 35 (2016) 2107–2118, doi:10.1109/tmi.2016.2550102.

[37] M. Jenkinson, C.F. Beckmann, T.E.J. Behrens, M.W. Woolrich, S.M. Smith, Fsl, Neuroimage 62 (2012) 782–790, doi:10.1016/j.neuroimage.2011.09.015.

[38] A.L. Rhoton Jr, The cerebral veins, Neurosurgery 51 (2002) S1–159, doi:10.1097/00006123-200210001-00005.