Segmentation of Vasculature From Fluorescently Labeled Endothelial Cells in Multi-Photon Microscopy Images

Russell Bates, Benjamin Irving, Bostjan Markelc, Jakob Kaeppler, Graham Brown, Ruth J. Muschel, Sir Michael Brady, Vicente Grau, and Julia A. Schnabel

***Abstract*— Vasculature is known to be of key biological significance, especially in the study of tumors. As such, considerable effort has been focused on the automated segmentation of vasculature in medical and pre-clinical images. The majority of vascular segmentation methods focus on bloodpool labeling methods; however, particularly, in the study of tumors, it is of particular interest to be able to visualize both the perfused and the non-perfused vasculature. Imaging vasculature by highlighting the endothelium provides a way to separate the morphology of vasculature from the potentially confounding factor of perfusion. Here, we present a method for the segmentation of tumor vasculature in 3D fluorescence microscopic images using signals from the endothelial and surrounding cells. We show that our method can provide complete and semantically meaningful segmentations of complex vasculature using a supervoxel-Markov random field approach. We show that in terms of extracting meaningful segmentations of the vasculature, our method outperforms both state-of-the-art method, specific to these data, as well as more classical vasculature segmentation methods.**

***Index Terms*— Image segmentation, Markov random fields, machine learning, microscopy.**

I. Introduction

SEGMENTATION of vasculature continues to be an important and challenging topic of research in biomedical image analysis. An area where the investigation of vasculature is of particular interest is in the study of tumor angiogenesis. Neovascularization is known to be one of the key mechanisms in the growth and development of tumors [1]. While the vasculature present in tumors can provide valuable information for investigating the properties of a tumor, there are various biological and technological issues surrounding the imaging of this vasculature that any automated segmentation method must address. Most methods for angiography use bloodpool based contrast agents, which provide contrast against the surrounding tissue for blood vessels imaged with modalities such as Computed Tomography (CT), Magnetic Resonance Imaging (MRI) or Ultrasound (US). Bloodpool based contrast agents typically enhance vessels as bright, tubular objects and many methods for vessel segmentation exploit this tubular structure through analysis of the local Hessian eigensystem [2]. However, in tumors, we know that the vasculature system can be poorly or irregularly perfused due to the inherent leakage and chaotic organization [3]. By imaging vessels using only bloodpool based contrast agents, we would implicitly limit our analysis to only the perfused vasculature. Therefore, as we wish to perform an analysis of the fractions of perfused vasculature, any such contrast enhancement-based segmentation method will be insufficient.

Fluorescence microscopy provides an alternative way to image vasculature, by imaging the vascular endothelium directly. This can be done pre-clinically by genetically modifying the endothelial cells in mice to express a fluorescent protein, which can then be directly imaged. However, particularly in the case of tumor vasculature, the endothelium alone does not provide the tubular geometry that most standard methods of vascular segmentation rely upon, as only the outer wall is visible. In the imaging paradigm of bloodpool based contrast agents methods typically attempt to infer the presence of vasculature from the perfusion. This is fundamentally flawed in the case of leaky and poorly perfused vasculature and therefore we propose a much more robust paradigm to infer the perfused as well as non-perfused lumen of the vessel from the visible vascular walls, which will always be present, without imaging the lumen directly. For our proposed segmentation algorithm we use signals from both the endothelium and tumor cells fluorescence channels, avoiding any kind of bloodpool labeling. A whole tumor view of these channels can be seen in Figure 2 and single images tiles can be seen in Figure 1. The images shown in Figure 2 are composed of approximately 225 tiles of equivalent size to those shown in Figure 1.

Manuscript received April 28, 2017; revised June 30, 2017; accepted July 3, 2017. Date of publication August 9, 2017; date of current version December 28, 2018. This work was supported in part by the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007–2013) under REA Grant 625631, in part by Cancer Research U.K. (CRUK) under Grant C5255/A18085, in part by the CRUK Oxford Centre under Grant C5255/A15935, in part by the CRUK/EPSRC Oxford Cancer Imaging Centre under Grant C5255/A16466, and in part by the EPSRC Systems Biology Doctoral Training Centre, Oxford, under Grant EP/G03706X/1. *(Corresponding author: Russell Bates.)*

R. Bates, B. Irving, and V. Grau are with the Department of Engineering Science, Institute of Biomedical Engineering, University of Oxford, Oxford OX3 7DQ, U.K. (e-mail: russell.bates@eng.ox.ac.uk).

B. Markelc, J. Kaeppler, G. Brown, R. J. Muschel, and S. M. Brady are with the CRUK/MRC Oxford Centre for Radiation Oncology, University of Oxford, Oxford OX3 7DQ, U.K.

J. A. Schnabel is with the School of Biomedical Engineering and Imaging Sciences, King's College London, London SE1 7EH, U.K.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMI.2017.2725639

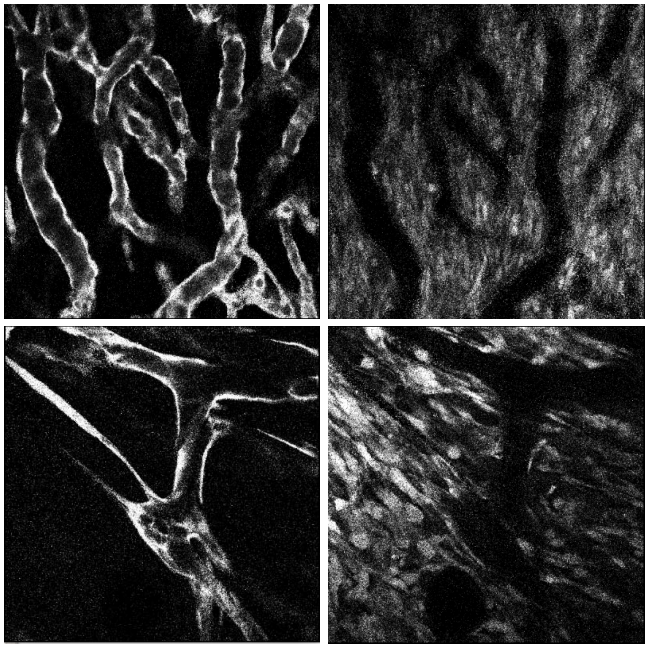

Fig. 1. Fluorescence channels showing a single tile (512 × 512 pixels) of endothelial cells (left) and the corresponding tumor cells (right).

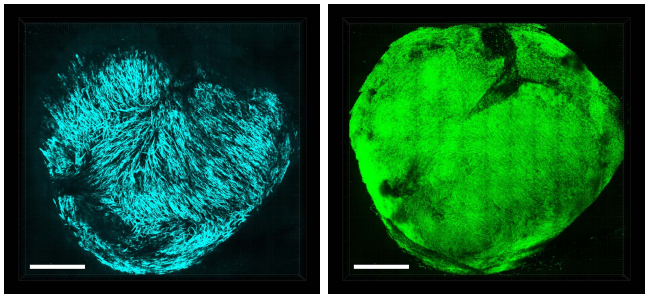

Fig. 2. Whole tumor images (225 tiles) showing endothelial cells (cyan) and the corresponding tumor cells (green).

In addition to these biological concerns, modern microscopy imaging techniques are capable of producing extremely large image volumes. While this allows us to image small structures with high resolution, there are challenges in terms of processing the entire volume in an efficient manner. Methods designed to operate on large image volumes such as these must have reasonable bounds on both computational time and memory usage.

To address these difficulties, we propose a method of segmentation in which we first perform a supervoxel segmentation of the image volume. We then extract image features from each supervoxel which are classified using a random forest classifier to determine the most likely structure present in the image. These likelihoods are then regularized by a vessel-oriented Markov Random Field (MRF) graphical model defined over the graph of supervoxels to generate our final segmentation of the vasculature. Central to our application is the ability to generate meaningful measurements of the underlying vasculature, rather than merely achieving a high volume overlap. We will show that not only does this method produce accurate and complete segmentations of the image, but that the semantic structure of the vasculature is maintained, allowing a straight forward skeletonization process to recover the full topology of the network.

II. Related Work

There has been very little work focused on the segmentation of vasculature labeled purely on the endothelium. Instead, the vast majority of vascular segmentation techniques depend on the use of bloodpool based contrast agents, providing a tubular structure [4]–[7]. A thorough review of methods applied to enhanced vasculature has been produced by Lesage *et al.* [8]. In fact, the lack of suitable segmentation techniques for data of this type means that even state-of-the-art imaging studies rely upon manual segmentations to acquire measurements [9], [10]. In the case of many biological studies, manual annotation becomes infeasible due to the scale of images acquired. A full exploration of the spatio-temporal behaviour of tumor vasculature requires that we analyze many large image volumes, both longitudinally and for a number of cases. Therefore, a suitable automated method must be developed for data of this type. The method of Narayanaswamy *et al.* [11] attempts to tackle a similar application by generating a bounding mesh from the endothelial cells. However, they focus on larger, non-pathological vasculature which does not suffer from the problems associated with the chaotic tumor neovasculature. Work by Almasi *et al.* on a similar domain avoids the general task of segmentation by enhancing 'key points' and utilizing graph considerations to extract the network connectivity [12], [13]. Deep Convolutional Neural Networks have been applied to a related problem also by Teikari *et al.* [14].

In the case of such large image volumes it becomes impractical to deal with each pixel individually, and a suitable reduced representation may be required. Often there exist large amounts of local redundancy in an image so we can consider a piecewise constant representation of the image in some feature space. We refer to these regions as supervoxels, which have been used to great effect in many imaging applications. A number of supervoxel clustering algorithms have been proposed. Felzenszwalb poses it as a graph partitioning problem [15], the Normalized Cut method of Shi and Malik solves a generalized eigenvalue problem [16] and the classical watersheds algorithm has also been used. Achanta *et al.* proposed the use of a localized k-means clustering for their supervoxel algorithm, Simple Linear Iterative Clustering (SLIC) [17] and a number of variants have been explored by transforming the clustering domain. Irving *et al.* performed SLIC clustering on the PCA modes of DCE-MRI signals [18] [19], Oda *et al.* suggest the inclusion of Hessian based signals into a SLIC algorithm [20] and Holzer *et al.* proposed a signal transform using the monogenic phase to provide a more suitable clustering domain for medical images [21].

The use of supervoxel-based features in graphical models has been successfully applied to object recognition tasks by Fulkerson *et al.* [22]. While Fulkerson *et al.* enforces supervoxel label continuity as a hard constraint, the use of supervoxels as a soft label constraint was explored by Kohli *et al.* in their Robust P^n method [23]. In biomedical applications, supervoxel graphical models have

been explored by Lucchi *et al.* [24] for the segmentation of mitochondria from SSEM volumes. In their work both shape and textural features are extracted from each supervoxel and are used to train an SVM classifier, which is then used in a CRF framework. The work of Zhu *et al.* [25], [26] has applied supervoxels to the segmentation of 3-dimensional vasculature. In this work they consider a semi-supervised graph transduction approach using Gaussian affinities. However, their use of bloodpool labeling meant that simple intensity and orientation features were sufficient for segmentation.

III. Materials

A. Data

For the particular focus of this work we segment vasculature from images acquired using high resolution fluorescent multi-photon microscopy, achieving a theoretical lateral resolution of 0.4μm and an axial resolution of 1.3μm. Voxels are sized 5μm in the z direction and 0.83μm in the x-y direction. We used an abdominal window chamber model in mice, which allowed for intravital imaging of the tumors [27]. The abdominal window chamber was surgically implanted in transgenic mice on C57Bl/6 background that had expression of red fluorescent protein tdTomato only in endothelial cells. The murine colon adenocarcinoma MC38 tumors with expression of green fluorescent protein (GFP) in the cytoplasm were induced by injecting 5μl of dense cell suspension in a 50/50 mixture of saline and matrigel (Corning, NY, USA). Prior to imaging we intravenously injected 100μl of Qtracker 705 Vascular Labels (Thermo Fisher Scientific, MA, USA) which is a blood-pool based labeling agent thus allowing us to determine whether vessel are perfused or not. The images of tumors were acquired 9–14 days after tumor induction with Zeiss LSM 880 microscope (Carl Zeiss AG), connected to a Mai-Tai tuneable laser (Newport Spectra Physics). We used an excitation wavelength of 940 nm and the emitted light was collected with Gallium Arsenide Phosphide (GaAsP) detectors through a 524–546 nm bandpass filter for GFP and a 562.5–587.5 nm bandpass filter for tdTomato and with a multi-alkali PMT detector through a 670–760 bandpass filter for Qtracker 705. A $20\times$ water immersion objective with NA of 1.0 was used to acquire a Zstacks-TileScan with dimensions of 512×512 pixels in x and y, and approximately 70 planes in z, with a z step of 5μm. Each tumor is covered by approximately 100–200 tiles, depending on the size. All animal studies were performed in accordance with the Animals Scientific Procedures Act of 1986 (UK) and Committee on the Ethics of Animal Experiments of the University of Oxford. The advantage of using both a labeled blood-pool based agent (Qtracker 705) and transgenic mouse model with fluorescently labeled endothelium is that it allows us to assess the functional behaviour of the tumor vasculature. Ground truth segmentations were acquired by a manual segmentation of each slice using the MITK software package [28]. Ground truth skeletons were derived from these segmentations using the automatic tracking and manual pruning tools available in the NeuTube software package [29].

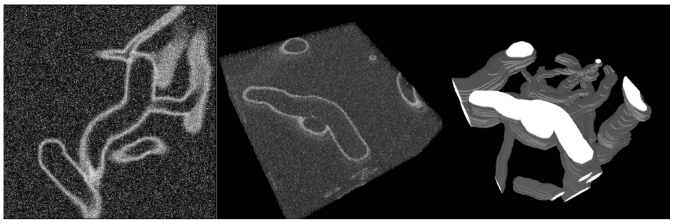

Fig. 3. 2D view of synthetic endothelium (left), 3D rendering of synthetic image volume (center), 3D rendering of synthetic ground truth segmentation (right).

B. Synthetic Data

We also tested on a synthetic dataset that aimed to replicate a number of the difficulties present in our real data. The data is sampled anisotropically to the same aspect ratio as the real data, and models a number of vessels of varying diameter. Gaussian and Poisson noise are added to simulate the conditions of the original images, with noise levels tuned to match those present. An example of this data, along with the ground truth segmentation, is shown in Figure 3.

IV. Methods

Here we will give an overview of the proposed pipeline, followed by a detailed description of each stage. We begin by performing a transformation of our input image into our proposed hesSLIC domain. This domain is then used for SLIC local clustering, to produce a superpixel over-segmentation. Features are then extracted from the entirety of the original image and assigned to their respective superpixel. Aggregations and averages of the features are taken for each superpixel. These are then passed to a Random Forest decision tree ensemble, which defines a unary potential for our MRF graphical model. The pairwise terms are also defined between each pair of neighboring superpixels according to various constraints. This system is then solved to find the most likely labeling of our image and to produce the desired segmentation.

A. Supervoxels

The size of the images that may nowadays be routinely acquired with microscopy methods provides a challenge for image processing, and for graphical models in particular. To address this, we first perform a pre-processing step by which we subdivide the image into a set of congruent regions known as supervoxels [17] (superpixels in 2D). These supervoxels provide a, piecewise constant, reduced representation of the image in some feature space. Defining a graph of the adjacency of these supervoxel regions is a much less demanding task and opens up the possibility of efficiently using probabilistic graphical models on our large image volumes.

In this work we consider the Simple Linear Iterative Clustering (SLIC) supervoxel method of Achanta *et al.* [17]. SLIC supervoxels are both computationally fast and memory efficient to calculate. They are known to provide regularly sized, uniform, supervoxels. The number and compactness of these supervoxels is controlled by the user. The SLIC method is an extension of k-means clustering from a feature space to a combined spatial+feature space. We begin by initializing a

regular grid of cluster centers with step size S. Then we assign each pixel to its closest center according to some weighted Euclidean metric:

$$Dist = \sqrt{Dist_F^2 + \left(\frac{Dist_S}{S}\right)^2 m^2} \quad (1)$$

where $Dist_S$ is the Euclidean distance between the center and the pixel, and $Dist_F$ is the Euclidean distance in feature space. The SLIC algorithm can be considered to exist on a sliding scale between k-means clustering and a regular grid. The position on this sliding scale is determined by the compactness parameter m. By tuning m, a balance can be found between edge localization and regularity.

However, we find that the concept of simply clustering pixel intensities is not well suited to images such as those in our primary application, where there is not a clear distinctive edge between many of the different structures, and edges which do exist may be soft. This has also been observed by Holzer *et al.* , inspiring their monoSLIC method [21], which utilizes the monogenic phase of the image.

In this implementation we introduce a modified version of the SLIC algorithm, which we title hesSLIC. SLIC relies on strong image gradients to create image boundaries; however in the case of medical images, specifically of vasculature, strong boundaries are not present and instead, soft gradients show us the edge of objects, and curvature cues tell us about possible vascular structures. When attempting to cluster on the smoothed signal, there is no clear boundary between the regions of high and low signal, this can result in a 'banding' of supervoxels along this transitional region. In the transformed hesSLIC signal on the bottom, the boundaries are made clear and cluster boundaries will align with the signal transition boundaries. We observe that a step function, convolved with a smoothing kernel, creates a smooth, sigmoidal function. However, the original stepping point is recoverable by examining the sign of the second derivative. We perform a transformation of the pixel domain by first computing the eigenvalues $\lambda = \{\lambda_1, \lambda_2, \lambda_3\}$ of the Hessian matrix H at each pixel in the image I, at scale σ. For these experiments we take $\sigma = 2$.

We define,

$$m = \arg\max_i(|\lambda_i|) \quad (2)$$

and

$$\phi(\lambda_m) = \text{sgn}(\lambda_m)(1 - \exp(-|\lambda_m \nabla I|)) \quad (3)$$

such that this switching behavior is emphasized in the low-curvature transitional areas, rather than noisy regions in which the sign of the second derivative may oscillate wildly. While the monoSLIC method [21] analyses the sign of the monogenic signal, our method considers the local curvature.

As the Hessian matrix is symmetric, our set of eigenvectors will form an orthonormal basis and the infinity norm of the eigenvalues gives the magnitude and direction of the principal curvature. If we consider the eigenvalues to vary smoothly over the image domain, then in order for the largest magnitude eigenvalue to change sign we must transition through a region of zero curvature. This means we should be transitioning from 'convex' to 'concave' regions. Thus, by examining only the dominant eigenvalue we can partition the space into regions of consistent curvature, concave or convex. Here we create a strong boundary at points at which the dominant curvature changes direction. This means that the enforced supervoxel boundaries will exist along a subset of the so-called 'parabolic lines' of the surface, the set of points at which at least one of the principal curvatures is equal to zero. In the case of vasculature, we are sure that the principal direction, corresponding to the greatest curvature, is perpendicular to the direction of the vessel and we therefore create strong boundaries along the length of the vessel, as desired. In addition to this we normalize according to the local image gradient and mean curvature, which helps to prevent rapid oscillations of the signal in low signal regions. The transformed signal, ϕ in Equation 3, is to be used in the original SLIC algorithm with the $Dist_F$ term from Equation 1 being calculated on this signal rather than the image pixel values directly.

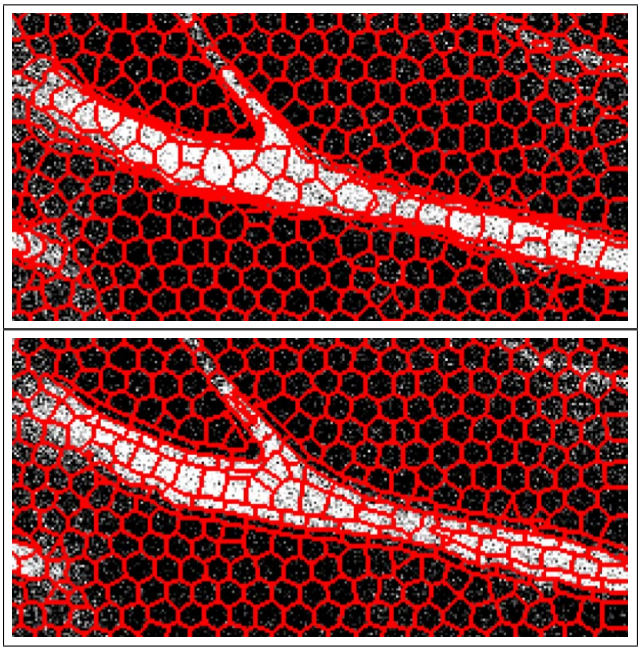

Fig. 4. Over-segmentation of vascular structure by two different supervoxel algorithms. Top: original SLIC algorithm, banding occurs along edges. Bottom: hesSLIC signal transform has significantly reduced banding.

The results of this method may be observed in Figure 4 where the SLIC algorithm (top) suffers from the 'banding' issues described earlier, whereas the hesSLIC (bottom) manages to deal with this boundary appropriately. We find that on our data the hesSLIC method provides an excellent semantic deconstruction of the image into structures such as lumen, endothelium, tumor and background.

An algorithmic overview of this method can be seen in Algorithm 1, and is further illustrated in Figure 5.

B. Markov Random Fields

Markov Random Fields (MRFs) provide a principled framework in which we are able to combine probabilistic

Algorithm 1 Overview of hesSLIC Algorithm

Require: Image I
1: Initialize output image ϕ
2: **for** pixel p in image I **do**
3: Compute local Hessian matrix at pixel p in image I
4: Compute magnitude of local image gradient at pixel p in image I
5: Compute eigenvalues λ of Hessian matrix
6: $m = \arg\max_i |\lambda|$
7: $\phi(p) = \text{sgn}(\lambda_m)(1 - \exp(-|\lambda_m \nabla I|))$
8: **end for**
9: Perform SLIC [17] on new image ϕ.

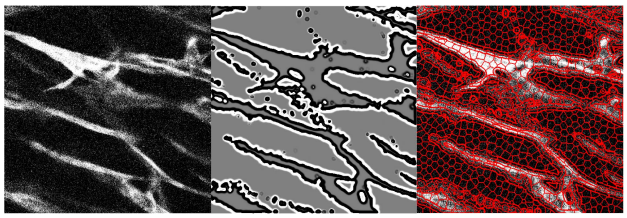

Fig. 5. Transformation from endothelium channel (left) to ϕ domain (center) to superpixel segmentation (right).

information provided by the image with a set of spatial prior beliefs. We consider this a semantic segmentation task where we attempt to assign a label to every part of the image, denoting what kind of structure is present at that particular location. In these images we identify 4 labels: tumor, endothelial cells, vessel lumen and background. This is how we motivate a 4-label pairwise Markov Random Field model.

In full generality, an MRF describes a joint probability structure over an un-directed graph. If we combine this with the positivity condition, $P(\mathbf{x}) > 0, \forall \mathbf{x}$. It may be shown, by the Hammersley-Clifford theorem, that this is equivalent to a Gibbs random field and as such can be factorized into conditional probabilities over the cliques c.

$$P(\mathbf{x}) = \frac{1}{Z}\exp\left(-E(\mathbf{x})\right) = \frac{1}{Z}\exp\left(-\sum_c E_c(\mathbf{x}_c)\right) \tag{4}$$

For tractability we consider only pairwise cliques,

$$E(\mathbf{x}) = \sum_{i\in N} U_i(x_i) + \sum_{(i,j)\in C} V_{ij}(x_i, x_j) \tag{5}$$

where E represents a potential induced by a given labeling x and Z is a partition function required for a correctly normalized distribution.

We can interpret this formulation as the function U_i defining the node i's inclination toward being assigned the label x_i. This function can be unique for each node i, since the function U_i is conditioned on the data itself. The function V_{ij} describes the interaction at the edge connecting node i with node j, taking labels x_i and x_j respectively. By penalizing boundaries between different labels, we can provide a form of spatial regularization. As our function V_{ij} may also be conditioned on the data it improves clarity to rewrite our energy as,

$$E(\mathbf{x}) = \sum_{i\in N} U_i(x_i) + \sum_{(i,j)\in C} G(f_i, f_j)H(x_i, x_j) \tag{6}$$

where, G is a function of the features at nodes i and j and H is a function of the labels alone. The functions G and H can then be defined globally. We can then solve this approximately using the Graph Cut method of Boykov and Jolly [30]

In our formulation we treat each superpixel as a node of the graph. Here, f_i, p_i and S_i correspond, respectively, to the feature vector, centroid location and set of constituent pixels for superpixel i. Our unary potentials are taken as the probabilistic outputs from a Random Forest classifier [31], the details of which will be addressed in Section IV-C. As we wish to define a potential rather than a probability, we denote the probabilistic output of the Random Forest U_{RF} and finally say that,

$$U = -\log(U_{RF}) \tag{7}$$

We define our pairwise function over the features, $G(f_i, f_j)$, to be the product of four functions. Through the function G_1 we enforce that the strength of interactions is modulated by the distance between two nodes. The function G_2 enforces that interactions are strongest between nodes with similar features. The difference in size between the supervoxels is accounted for by G_3 and the agreement of the local directionality of the two nodes is dealt with by G_4.

$$G_1 = \exp\left(-\frac{|p_i - p_j|^2}{\gamma_d}\right) \tag{8}$$

$$G_2 = \exp\left(-\frac{|f_i - f_j|^2}{\gamma_f}\right) \tag{9}$$

$$G_3 = \frac{|i|}{\min(|j|, N_{min})} \tag{10}$$

$$S_{i,j} = (v_i \cdot \tfrac{p_i - p_j}{||p_i - p_j||}) \tag{11}$$

$$G_4 = |S_{i,j} - (|S_{i,j} - 1)\exp\left(-\frac{f_{i,v}}{\gamma}\right) \tag{12}$$

$$G(i,j) = \lambda G_1(i,j)G_2(i,j)G_3(i,j)G_4(i,j) \tag{13}$$

Tuning parameters γ_f,γ_d and γ govern the strength of these interactions and can be either calibrated or estimated from the data. $f_{i,v}$ is given by the local Hessian strength (L_2-norm of the eigenvalues) however this could be replaced by a vesselness measure for suitable data. This term modulates the extent to which anisotropic regularization should occur. The general strength of regularization is modified by the parameter λ.

We define H to represent our biological understanding of the image, e.g. that lumen should not co-occur with tumor cells (because it should be surrounded by endothelial cells) or that endothelium should not share a boundary with background (because they should be embedded in the tumor tissue). As this paper compares binary segmentation methods, we represent H by a simple Potts model, taking a constant value for different labels and 0 if the labels are the same. However, more complex relationships between labels may be learned and encoded.

The combination of G_1 and G_3 means that we can avoid the issues of a non-regular supervoxel lattice by taking a fully connected graph and eliminating sufficiently weak interactions. G_1 allows us to set a characteristic interaction distance by γ_d and G_3 prevents undue influence from small supervoxels. The interaction term G_2 provides us with the classical edge preserving regularization that provides the majority of the regularization. However, it is known that regularization of this form can suffer from the so-called 'shrinking problem' where larger, convex structures are encouraged in order to minimize boundary areas. Anisotropic regularization terms such as $S_{i,j}$ have been suggested in terms of MRF formulations previously [32], [33] however we incorporate a transition function G_4 to give highly anisotropic interactions only in regions of strong curvature ($f_{i,v}$) and isotropic interactions elsewhere. This takes a form similar to that found in the Vessel Enhancing Diffusion of Manniesing *et al.* [34].

C. Random Forests

Random Forests are an ensemble learning technique of random decision trees [31], [35]. The use of random, bagged, decision trees has been shown to reduce overfitting and improve generalization error for high dimensional classification. The decision forest is constructed by randomly sampling from the feature space as well as randomly sampling from the training set with replacement to produce an ensemble of binary decision trees. A key advantage of using ensembles of decision trees is that features are only ever compared among themselves, which avoids concerns about feature scaling. We can perform hard classification by taking the modal value of the decision tree outputs, or we can obtain a probabilistic output by considering the distribution of responses. We would like to return a probabilistic output so that we can feed it back into our MRF as a unary potential after transforming according to Eq. 7.

We perform Random Forest classification for each supervoxel by constructing a feature vector similar to that used by Kaynig *et al.* [32]. This feature set contains Gaussians taken at a number of scales, as well as differences between Gaussians and Hessian eigenvalues computed at multiple scales. For each supervoxel we then form a single feature vector by taking the median value achieved over the supervoxel for each feature. In addition to this, the feature vector for each supervoxel is concatenated with the average feature vector of its neighboring supervoxels, providing additional context for the classifier as has been explored by Fulkerson *et al.* [22]. Features are computed on both the endothelium and tumor channels, as visible in Figure 1. For these experiments we trained a decision forest consisting of 100 trees, with a maximum tree depth of 10 splits. Sigmas for the Gaussian kernels used for feature extraction were taken between 1 and 20.

V. Results

We compare the results of our segmentation algorithm against a gold-standard, manual-segmentation. We compared the manual segmentation against our method, a median-filtering approach, a commonly used Hessian-based, vesselness filtering method [4] and a state-of-the-art vessel segmentation method specifically designed for labeled vessel laminae [11]. In the case of median filtering and vesselness, the images are thresholded to produce a final segmentation. The method of Narayanaswamy *et al.* was made available as part of the Farsight Toolkit (*www.farsight-toolkit.org*) [11]. All the parameters and thresholds are optimized on a calibration image and held constant over the test set. The Random Forest model is trained over the entire data set in a leave-one-out framework, so for each test case, the model has been trained on all images, excluding the current test case.

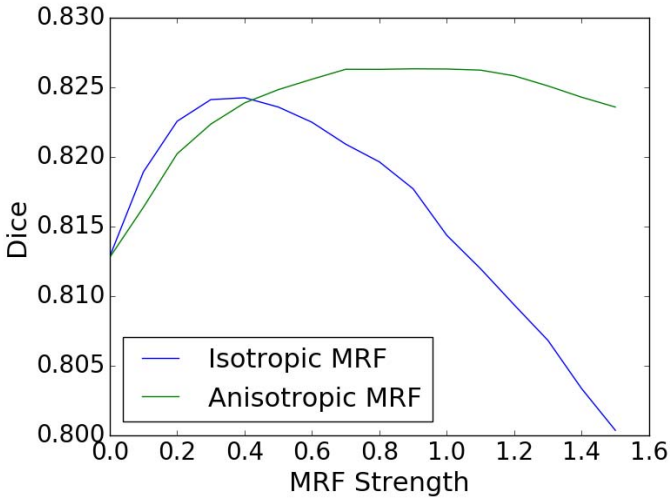

Fig. 6. Comparison of the effects of MRF strength on isotropic and anisotropic regularization.

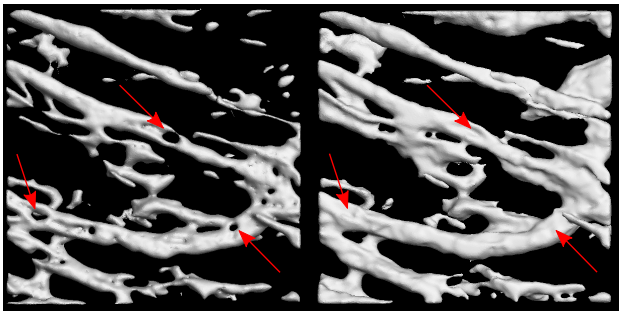

Fig. 7. Left: Method of Narayanaswamy [11]. Right: Proposed Method. Here the method of Naranayaswamy struggles to fill the vessel lumens, leaving gaps in the segmentation.

We found, in the case of the method of Narayanaswamy *et al.* , that although the method had been specifically designed for data of this type, the computational time meant that this method would not be well suited for our purposes (∼70,000s per tile). This can be compared to our method (∼100s per tile) and the median filtering method (∼10s per tile). We found that the method of Narayanaswamy performed well in regions of strong signal but failed to regularize in regions of low signal leaving incomplete sections of the vasculature as well as 'holes'. This is not unexpected as the method was designed for larger, more mature, vasculature where the endothelium would be regular and more clearly labeled. The results of this method, in comparison with our method, can be seen in Figure 7.

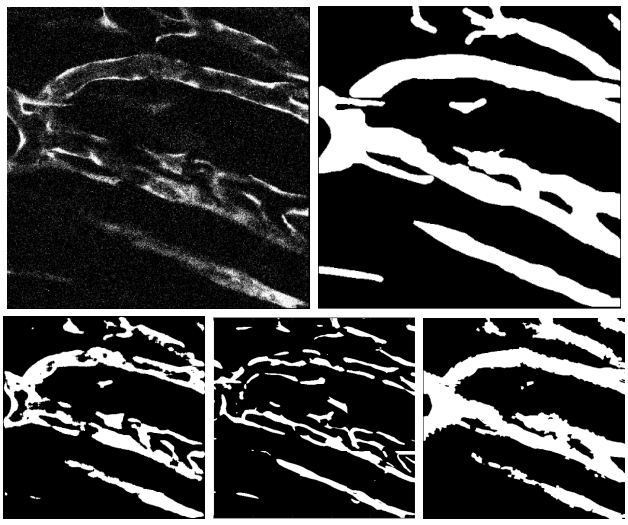

Fig. 8. Top left: Endothelial cell channel, top right: manual segmentation, bottom left: median filtering method, botton centre: vesselness filtering method, bottom right: proposed method.

TABLE I
RESULTS FOR hesSLIC VS SLIC COMPARISON

Method	Dice
SLIC-MRF00	0.77 ± 0.027
hesSLICMRF00	0.81 ± 0.025
SLIC-MRF05	0.78 ± 0.029
hesSLIC-MRF05	$\mathbf{0.83 \pm 0.027}$

We found the median filtering method to be by far the most effective of the other methods tested, due to its ability to 'fill' the lumen of the vessel by filtering the strong signal from the vessel walls. However, it is not a context aware method and we found that it was unable to fill the larger vessels found in our images. This can be observed in Figure 8.

We also investigate the role played by our proposed hesSLIC algorithm in improving segmentation. We perform this analysis on a synthetic dataset designed to replicate a number of the complexities of our real data. We compare the Dice overlap using both the traditional SLIC algorithm and our hesSLIC algorithm for two different values of the MRF regularization. The results of this can be seen in Table I, where we show the mean Dice score over our synthetic dataset, along with standard deviations. Here the hesSLIC variant out-performs the regular SLIC algorithm both with and without MRF regularization.

To explore the impact of anisotropic regularization in the MRF, as imposed by Eq. 12, we test the effectiveness on our synthetic data. Experiments were run over a range of MRF strengths, represented by the parameter λ in Eq. 13, both with and without the contribution from Eq. 12. This provides a directional anisotropic regularization, encouraging regularity in the direction of 'vessel structures'. We present an averaged Dice score over 10 synthetic image volumes. The results of this analysis are shown in Figure 6.

Although prevalent in the literature, we find the Dice overlap score to be poorly suited to the analysis of vascular segmentation. The Dice overlap relies on an assumption of local convexity such that change in the score should be small with respect to small deviations from the optimal solution. This is not the case in vasculature where small lateral changes perpendicular to the direction of vasculature may have a magnified effect on the Dice score. In addition to this, the Dice score does not assess the success with which a method has performed a true semantic segmentation, which, in this case means its ability to recover the true topology of the vasculature. Although we report comparable results to the median filtering method in terms of Dice overlap, we direct our main focus toward analysis of the extracted vascular topology, via skeletonization. In order to do this, we present the results of a Hausdorff analysis as well as the error in reported vessel diameter, this can be seen in Tables II, III and IV. We derive a skeletal representation from each segmentation by applying a simple iterative thinning algorithm [36]. We refine each skeleton by pruning artifactual branches that did not exceed some minimum length, or the diameter of their parent vessel. For these experiments we chose a minimum branch length of 10 voxels.

We define the Hausdorff Average [37], [38] as

$$d_H(A, B) = \frac{1}{|A|} \sum_{a \in A} d(a, B)$$

where $d(a, B)$ is the minimum Euclidean distance between a point a and a set B, given by $d(a, B) = \min_{b \in B} |a - b|$.

In this case the sets A and B denote the set of skeletal points of the segmented image and the ground truth image respectively. We propose this as a method to quantify the semantic value of the segmentation. While a Dice overlap shows that a significant pixel co-occurrence is achieved, the pixels that are not included in the segmentation may fundamentally change the meaning of the segmented structure. By quantifying the segmentation according to this distance metric rather than a simple segmentation overlap, we reward a method which captures the semantic meaning of the image, the vascular network.

Although we observe no significant difference in the values of the Dice overlap, we test the significance of the Mean Hausdorff Distance and Mean Vessel Diameter Error using a paired Wilcoxon Signed-Rank Test. We report a significant (tested at $p = 0.05$) improvement of the proposed method over the others in both the Hausdorff average distance (Table III) and the mean vessel diameter error (Table IV).

VI. DISCUSSION

Here we have presented a method for the multi-label segmentation of vasculature in 3D, endothelium-labeled, multi-photon microscopy images. We build up a primitive structure of the image using a novel supervoxel algorithm and generate likelihoods over the labels for each supervoxel using a supervised Random Forest algorithm. These likelihoods are then regularized using a vessel-oriented MRF, which encodes ideas about the anisotropic nature of vasculature as well as the relative likelihood of co-occurrence for different labels. We have shown that this method out-performs a number of other methods, most notably in its ability to produce full

TABLE II
DICE OVERLAP COMPARISON

Case #	1	2	3	4	5	6	7	8
Proposed	0.789	**0.761**	**0.673**	0.679	**0.888**	**0.878**	**0.761**	**0.710**
Median Filter	**0.816**	0.665	0.599	**0.779**	0.841	0.739	0.575	0.368
Frangi Filter	0.353	0.326	0.310	0.297	0.295	0.250	0.232	0.134
Narayanaswamy [11]	0.719	0.606	0.537	0.649	0.625	0.525	0.468	0.250

TABLE III
MEAN HAUSDORFF DISTANCE

Case #	1	2	3	4	5	6	7	8
Proposed	**5.02**	**11.66**	**11.94**	9.66	**4.96**	**4.61**	**4.85**	**4.01**
Median Filter	12.32	45.23	75.64	12.61	11.66	17.39	24.14	26.07
Frangi Filter	9.52	26.90	23.44	12.70	8.38	11.94	12.52	10.12
Narayanaswamy [11]	7.48	12.69	13.89	**7.81**	6.78	6.78	8.06	9.22

TABLE IV
MEAN VESSEL DIAMETER ERROR

Case #	1	2	3	4	5	6	7	8
Proposed	**2.61**	**4.37**	**4.51**	4.72	**1.94**	**1.65**	**1.61**	**1.20**
Median Filter	3.72	5.69	6.34	**4.31**	2.55	2.91	2.61	2.00
Frangi Filter	3.43	5.07	5.63	4.58	4.58	4.28	3.16	3.66
Narayanaswamy [11]	2.99	5.12	5.16	4.89	4.82	5.84	3.70	5.96

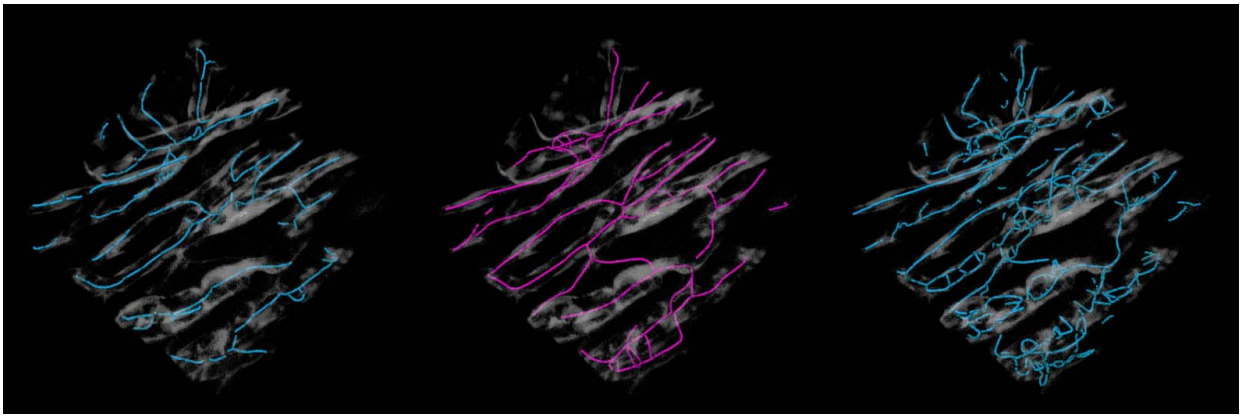

Fig. 9. Skeletons extracted from segmentation. Left: Proposed method, Middle: Ground Truth, Right: Median Filtering.

segmentations of the vasculature, and thus accurately extract the underlying topology of the vasculature.

A key advantage of the method proposed here, in comparison to other segmentation methods, is that although here we compared against binary segmentations, the method extends to as many classes as are present in the training annotations. An example in this case would be to separate the image into tumor cells, lumen, endothelial cells and background, as mentioned previously. This can be done to an arbitrary degree with any biological understanding of the relationship between labels encoded in the label affinity function H from Eq. 6. In motivating this method, we placed great importance on the ability to fully capture the topology of the vascular network in an efficient manner. The ability of this method to achieve both of these goals allows the analysis of large, irregular, vascular structures such as those found in tumors during angiogenesis. Although many segmentation methods value a high Dice overlap, we emphasize the fact that any region identified by the segmentation method should correspond to a complete section of the vasculature. It is this desire that leads us to prioritize the Hausdorff distance. One of the key strengths of the proposed method is the ability to fully segment sections of the vasculature. However, this context awareness can cause the method to reject sections of vasculature that are only partially visible. The ability of the thresholding method to segment only partially visible sections of the vasculature produces a favorable Dice score but makes the results wholly unsuitable for further analysis. This can be observed in Figure 9 where, although the median filtering method detects vasculature in approximately the correct parts of the image, the skeleton that is extracted is disjointed and contains many artifactual branches due to the incorrect topology of the segmentation.

From our synthetic experimental results, shown in Table I, we see an improved ability to segment the structures when using the hesSLIC formulation. We intuitively attribute this to an ability to perform better clustering in the presence of soft gradients in the image. In our synthetic experiments into the effects of modifying MRF strength λ, shown

in Figure 6, we also see an improved result by the application of anisotropic regularization. In addition to a higher peak, we note that the results in the anisotropic case appear less sensitive to the value of λ. This could potentially be due to a reduction in the impact of the 'shrinking problem' discussed earlier.

Another key advantage of our method is the relative speed with which we are able to produce segmentations, in comparison to methods such as that of Narayanaswamy *et al.* As the desired application of this method is to segment longitudinal data sets of large image volumes, each with upward of billions of voxels, it is imperative that we are able to process them efficiently.

A potential area for future improvement is the limitations imposed by our edge potentials in the MRF which are tuned rather than learned. Although methods exist that are able to learn fully connected pairwise CRFs [24], [39], they require densely annotated training sets to capture the full extent of spatial interactions. The expectation of the existence of fully annotated training sets for many applications is unrealistic. Future work will focus on the suitability of semi-supervised methods to achieve fully supervised levels of performance on sparse annotations. It is possible that this may be done in the current framework using label-transduction methods [40]. Interesting work in the transduction and interactive learning for sparsely labeled superpixel microscopy images has also been undertaken by Su *et al.* [41]. A method that can take sparse image annotations and use them to leverage information from large set of unlabeled parts of the image to create high quality segmentations would be an extremely powerful tool. This would have very broad applications in novel imaging experiments where large training sets are not readily available and where there is a high time-cost in producing such a training set.

REFERENCES

[1] D. Hanahan and R. A. Weinberg, "The hallmarks of cancer," *Cell*, vol. 100, no. 1, pp. 57–70, 2000.

[2] R. D. Rudyanto *et al.*, "Comparing algorithms for automated vessel segmentation in computed tomography scans of the lung: The VESSEL12 study," *Med. Image Anal.*, vol. 18, pp. 1217–1232, Oct. 2014.

[3] R. A. Weinberg, *The Biology of Cancer*. New York, NY, USA: Taylor & Francis, 2013.

[4] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever, "Multiscale vessel enhancement filtering," in *Medical Image Computing and Computer-Assisted Intervention* (Lecture Notes in Computer Science). Berlin, Germany: Springer, 1998, pp. 130–137.

[5] C. Xiao, M. Staring, D. Shamonin, J. H. C. Reiber, J. Stolk, and B. C. Stoel, "A strain energy filter for 3D vessel enhancement with application to pulmonary CT images," *Med. Image Anal.*, vol. 15, no. 1, pp. 112–124, Feb. 2011.

[6] K. Krissian, G. Maladain, N. Ayache, R. Vaillant, and Y. Trousset, "Model-based multiscale detection of 3D vessels," in *Proc. Workshop Biomed. Image Anal.*, Jun. 1998, pp. 202–210.

[7] C. Lorenz, I.-C. Carlsen, T. M. Buzug, C. Fassnacht, and J. Weese, "Multi-scale line segmentation with automatic estimation of width, contrast and tangential direction in 2D and 3D medical images," in *CVRMed-MRCAS*. Berlin, Germany: Springer, 1997, pp. 233–242.

[8] D. Lesage, E. D. Angelini, I. Bloch, and G. Funka-Lea, "A review of 3D vessel lumen segmentation techniques: Models, features and extraction schemes," *Med. Image Anal.*, vol. 13, pp. 819–845, Dec. 2009.

[9] J.-Y. Huang *et al.*, "*In vivo* two-photon fluorescence microscopy reveals disturbed cerebral capillary blood flow and increased susceptibility to ischemic insults in diabetic mice," *CNS Neurosci. Therapeutics*, vol. 20, pp. 816–822, Sep. 2014.

[10] P. R. Barber, B. Vojnovic, S. M. Ameer-Beg, R. J. Hodgkiss, G. M. Tozer, and J. Wilson, "Semi-automated software for the three-dimensional delineation of complex vascular networks," *J. Microscopy*, vol. 211, pp. 54–62, Jul. 2003.

[11] A. Narayanaswamy, S. Dwarakapuram, C. S. Bjornsson, B. M. Cutler, W. Shain, and B. Roysam, "Robust adaptive 3-D segmentation of vessel laminae from fluorescence confocal microscope images and parallel GPU implementation," *IEEE Trans. Med. Imag.*, vol. 29, no. 3, pp. 583–597, Mar. 2010.

[12] S. Almasi, X. Xu, A. Ben-Zvi, B. Lacoste, C. Gu, and E. L. Miller, "A novel method for identifying a graph-based representation of 3-D microvascular networks from fluorescence microscopy image stacks," *Med. Image Anal.*, vol. 20, pp. 208–223, Feb. 2015.

[13] S. Almasi, A. Ben-Zvi, B. Lacoste, C. Gu, E. L. Miller, and X. Xu, "Joint volumetric extraction and enhancement of vasculature from low-SNR 3-D fluorescence microscopy images," *Pattern Recognit.*, vol. 63, pp. 710–718, Mar. 2016.

[14] P. Teikari, M. Santos, C. Poon, and K. Hynynen. (Jun. 2016). "Deep learning convolutional networks for multiphoton microscopy vasculature segmentation." [Online]. Available: https://arxiv.org/abs/1606.02382

[15] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based image segmentation," *Int. J. Comput. Vis.*, vol. 59, no. 2, pp. 167–181, Sep. 2004.

[16] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 8, pp. 888–905, Aug. 2000.

[17] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, Nov. 2012.

[18] B. Irving *et al.*, "Automated colorectal tumour segmentation in DCE-MRI using supervoxel neighbourhood contrast characteristics," in *Medical Image Computing and Computer-Assisted Intervention* (Lecture Notes in Computer Science), vol. 8673, P. Golland, N. Hata, C. Barillot, J. Hornegger, and R. Howe, Eds. Berlin, Germany: Springer, 2014, pp. 609–616.

[19] B. Irving *et al.*, "Pieces-of-parts for supervoxel segmentation with global context: Application to DCE-MRI tumour delineation," *Med. Image Anal.*, vol. 32, pp. 69–83, Aug. 2016.

[20] H. Oda *et al.*, "Hessian-assisted supervoxel: Structure-oriented voxel clustering and application to mediastinal lymph node detection from CT volumes," *Proc. SPIE*, vol. 10134, p. 101341D, Mar. 2017.

[21] M. Holzer and R. Donner, "Over-segmentation of 3D medical image volumes based on monogenic cues," in *Proc. CVWW*, 2014, pp. 1–8.

[22] B. Fulkerson, A. Vedaldi, and S. Soatto, "Class segmentation and object localization with superpixel neighborhoods," in *Proc. ICCV*, Sep./Oct. 2009, pp. 670–677.

[23] P. Kohli, L. Ladický, and P. H. S. Torr, "Robust higher order potentials for enforcing label consistency," *Int. J. Comput. Vis.*, vol. 82, pp. 302–324, May 2009.

[24] A. Lucchi, K. Smith, R. Achanta, G. Knott, and P. Fua, "Supervoxel-based segmentation of mitochondria in EM image stacks with learned shape features," *IEEE Trans. Med. Imag.*, vol. 31, no. 2, pp. 474–486, Feb. 2012.

[25] Y. Zhu *et al.*, "Three-dimensional vasculature reconstruction of tumour microenvironment via local clustering and classification," *Interface Focus*, vol. 3, no. 4, p. 20130015, 2013.

[26] Y. Zhu *et al.*, "Coupling oriented hidden Markov random field model with local clustering for segmenting blood vessels and measuring spatial structures in images of tumor microenvironment," in *Proc. IEEE Int. Conf. Bioinform. Biomed.*, Nov. 2011, pp. 352–357.

[27] L. Ritsma, E. J. A. Steller, S. I. J. Ellenbroek, O. Kranenburg, I. H. M. Borel Rinkes, and J. van Rheenen, "Surgical implantation of an abdominal imaging window for intravital microscopy," *Nature Protocols*, vol. 8, pp. 583–594, Mar. 2013.

[28] I. Wolf *et al.*, "The medical imaging interaction toolkit (MITK): A toolkit facilitating the creation of interactive software by extending VTK and ITK," *Proc. SPIE*, vol. 5367, pp. 16–27, May 2004.

[29] L. Feng, T. Zhao, and J. Kim, "neuTube 1.0: A new design for efficient neuron reconstruction software based on the SWC format," *eNeuro*, vol. 2, no. 1, pp. 1–10, 2015.

[30] Y. Y. Boykov and M.-P. Jolly, "Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images," in *Proc. ICCV*, 2001, pp. 105–112.

[31] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[32] V. Kaynig, T. Fuchs, and J. M. Buhmann, "Neuron geometry extraction by perceptual grouping in ssTEM images," in *Proc. CVPR*, 2010, pp. 2902–2909.

[33] V. Grau, J. C. Downs, and C. F. Burgoyne, "Segmentation of trabeculated structures using an anisotropic Markov random field: Application to the study of the optic nerve head in glaucoma," *IEEE Trans. Med. Imag.*, vol. 25, no. 3, pp. 245–255, Mar. 2006.

[34] R. Manniesing, M. A. Viergever, and W. J. Niessen, "Vessel enhancing diffusion: A scale space representation of vessel structures," *Med. Image Anal.*, vol. 10, no. 6, pp. 815–825, 2006.

[35] T. K. Ho, "The random subspace method for constructing decision forests," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 8, pp. 832–844, Aug. 1998.

[36] C. Pudney, "Distance-ordered homotopic thinning: A skeletonization algorithm for 3D digital images," *Comput. Vis. Image Understand.*, vol. 72, pp. 404–413, Dec. 1998.

[37] M.-P. Dubuisson and A. K. Jain, "A modified Hausdorff distance for object matching," in *Proc. 12th Int. Conf. Pattern Recognit.*, vol. 1. Oct. 1994, pp. 566–568.

[38] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge, "Comparing images using the Hausdorff distance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 9, pp. 850–863, Sep. 1993.

[39] P. Krähenbühl and V. Koltun, "Efficient inference in fully connected CRFs with Gaussian edge potentials," in *Proc. NIPS*, 2012, pp. 109–117.

[40] A. Criminisi, J. Shotton, and E. Konukoglu, "Decision forests: A unified framework for classification, regression, density estimation, manifold learning and semi-supervised learning," *Found. Trends Comput. Graph. Vis.*, vol. 7, nos. 2–3, pp. 81–227, Feb. 2012.

[41] H. Su, Z. Yin, S. Huh, T. Kanade, and J. Zhu, "Interactive cell segmentation based on active and semi-supervised learning," *IEEE Trans. Med. Imag.*, vol. 35, no. 3, pp. 762–777, Mar. 2016.