DEEP U-NET REGRESSION AND HAND-CRAFTED FEATURE FUSION FOR ACCURATE BLOOD VESSEL SEGMENTATION

Yasmin M. Kassim[1,7], *O. V. Glinskii*[2,4,6], *V. V. Glinsky*[4,5,6], *V. H. Huxley*[2,6], *G. Guidoboni*[1,3], *K. Palaniappan*[1]

[1]Department of Electrical Engineering and Computer Science, Computational Imaging and VisAnalysis Lab
[2]Department of Medical Pharmacology and Physiology
[3] Department of Mathematics, [4] Dalton Cardiovascular Research Center
[5] Department of Pathology and Anatomical Sciences
[6]National Center for Gender Physiology
University of Missouri-Columbia, Columbia, MO 65211 USA
[7]Department of Computer Engineering, Al-Nahrain University, Baghdad, Iraq

ABSTRACT

Automated curvilinear image segmentation is a crucial step to characterize and quantify the morphology of blood vessels across scale. We propose a dual pipeline RF_OFB+U-NET that fuses U-Net deep learning features with a low level image feature filter bank using the random forests classifier for vessel segmentation. We modify the U-Net CNN architecture to provide a foreground vessel regression likelihood map that is used to segment both arteriole and venule blood vessels in mice dura mater tissues. The hybrid approach combining both hand-crafted and learned features was tested on 60 epifluorescence microscopy images and improved the segmentation of thin vessel structures by nearly 5% using the Dice similarity coefficient compared to U-Net.

Index Terms— Semantic vessel segmentation, deep learning, histogram equalization, random forests, U-Net.

1. INTRODUCTION

Automatic segmentation and quantification of blood vessels is important to characterize changes in flow dynamics during vascular network remodeling [1]. Veins are very much understudied with respect to arteries, mainly due to the difficulty of collecting data and imagery, despite the fact that they play a crucial role in maintaining tissue perfusion and homeostasis [2, 3]. This paper focuses on segmenting both arterioles and venules making an important contribution to the field. Such patient specific approaches can improve the diagnosis and treatment of chronic conditions such as glaucoma, hypertension, diabetes or estrogen deprivation. Changing response state of the vasculature to different physiological conditions can be extracted from biomedical imagery for quantitative research. Many automated analysis methods have been proposed to segment blood vessels especially retinal vessels in brightfield fundoscopy images [4] and larger arteries and veins in angiograms [5]. However, there has not been as much study of image analysis methods for segmenting the smallest vessels known as capillaries or microvasculature often imaged using fluorescence microscopy. Further, the majority of existing methods for segmenting vessel structures use hand-crafted image features [6, 7, 8, 9]. More recently deep learning based methods are being applied to biomedical imagery for accurate image segmentation of vessels [10, 11, 12, 13, 14, 15, 16]. We present a new methodology that combines the strength of both hand-crafted features and learned features to produce a robust state-of-the-art vessel segmentation algorithm, as shown in Figure 1 for the feature fusion pipeline. In our previous work we focused on segmenting only the arteriole part of the vasculature for a small set of epiflourescent imagery [14, 9];

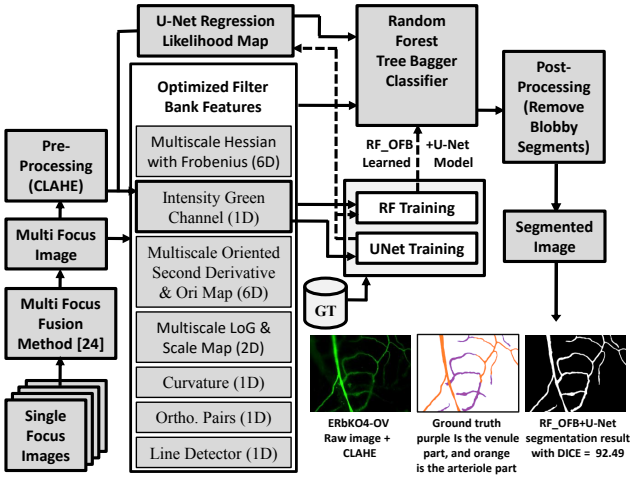
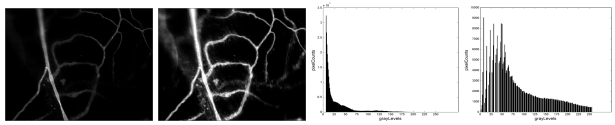


Fig. 1. Our proposed RF_OFB+U-NET dual pipeline. The grayed area is the main inference path composed of optimized filter bank (convolutional features) and U-Net regression likelihood map (sample result shown at bottom).



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

Fig. 2. Visualization of applying contrast-limited adaptive histogram equalization: (a) Original image, (b) CLAHE enhanced, (c) Histogram of (a), and (d) Histogram of (b).

arterioles transport oxygenated blood. We extended our previous work on segmenting arterioles which have more distinct boundaries to also segment venule segments with diffuse boundaries [9] by adapting the widely used U-Net deep convolutional neural network architecture [17] to work with epifluorescence blood vessel images. In this paper, we study the automatic segmentation of *both* arteriole and venule anatomical structures and use a larger dura mater dataset with 60 images. The proposed approach in this paper extends our previous approach using random forests and filter bank features combined with a U-Net based likelihood score map for the foreground vessel structure. The U-Net feature map provides an additional feature input to the random forest learning algorithm to support more precise and robust pixel wise segmentation of the complete microvasculature in extremely challenging dura mater epifluorescence microscopy images.

2. OUR PROPOSED DUAL PIPELINE ARCHITECTURE RF_OFB+U-NET

The proposed dual CNN model provides a new deep learning architecture to exploit the strengths of learned feature extraction uding convolutional neural networks with hand-crafted features based on computer vision and image processing expert knowledge. Hand-crafted features show the benefit of improved performance with lower computational cost, less training data, applicability across domains without retraining and optimization for realtime systems [18]. Our proposed dual CNN pipeline shown in Figure 1 fuses vessel specific features using an engineered optimized filter bank with the foreground likelihood map from the modified U-Net CNN architecture. This aggregation along with a random forests (RF) classifier produces robust pixel-wise segmentation on challenging microscopy images of dura mater capillaries.

2.1. Pre-processing using contrast-limited adaptive histogram equalization (CLAHE)

The green channel pixel grayscale intensity values were enhanced before including them as a feature in the RF or as input to the U-Net deep learning network. Intensity is an important feature to characterize venule regions. Since venule structures are more diffuse with less distinctive boundaries compared to the arteriole regions, most derivative-based approaches fail to accurately segment venules. Our epifluorescence images are characterized by excessive stain in the background, along with non-homogeneous staining within the lumen for which histogram equalization and other stretching methods result in amplification of noise artifacts in the relatively homogeneous regions. Histogram-based enhancement using CLAHE was used to increase image contrast without increasing noise artifacts [19]. The CLAHE tile size was set to 25×25 pixels and 0.02 for the clip limit. Figure 2 shows the result of CLAHE-based histogram stretching to increase dynamic range and bring out faint vessel structures that would otherwise be missed during segmentation and reduce the likelihood of false negatives.

Table 1. Description and dimensionality of our optimized filter bank (OFB) for random forest (RF) classifier.

Feature	Dim.
$\tanh$ of Frobenius norm over 5 scales using Gaussian filters for five scales $\sigma = \{1, 2, 3, 4, 5\}$ with filtering window size of [-3σ, 3σ]	5
Z-score sigmoid max Frobenius norm over 5 scales	1
Intensity of green channel	1
Maximum response of oriented 2nd derivatives over 3 scales & 3 orientation maps using oriented second derivative filters of Leung-Malik (LM) [20] with $\sigma = 3 * \{\sqrt{2}, 2, 2\sqrt{2}\}$	6
Maximum response over eight scales of LoG filter & sigma map using LoG filters of LM [20] with $\sigma = \{\sqrt{2}, 2, 2\sqrt{2}, 4\}$ and 3σ	2
Curvature using oriented first and second derivatives of LM filters [20] with scale equal to $\sigma = 6$	1
Orthogonal pairs first derivative using three counterpart angles with scale equal to $\sigma = 6$, max-pooling is used then to obtain the response over all angles	1
Multiscale Line Detector [7]	1
Dim. of all features together	18

2.2. Vessel segmentation using random forests classifier with optimized filter bank (RF_OFB)

We use a random forests bagging classifier to predict the class label of each pixel as vessel or non-vessel. Bagging means bootstrap aggregation in which an ensemble of trees is grown for either regression or classification. Bagging is designed to decrease variance and avoid over-fitting since each tree is constructed using a different set of samples and variables. Averaging is used in the case of regression trees and voting in the case of classification to make function learning more accurate and with less over-fitting. We have improved our work in [9] to also segment the venule capillaries, which is more challenging compared to arteriole segmentation. Table 1 summarizes the hand-crafted features used and §3.2 has more details. Additional epifluorescence microscopy images, more features, more trees, improved pre- and post-processing are incorporated to obtain improved micro-vessel segmentations using the RF_OFB classifier.

2.3. U-Net deep learning CNN architecture adapted for blood vessel segmentation (U-Net)

Semantic segmentation has been widely adopted for biomedical segmentation after Long et al. [21] described an end-to-end fully CNN-based segmentation approach for natural images in the Pascal VOC 2012 dataset. Semantic segmentation performs pixel-wise grouping and labeling so that each pixel

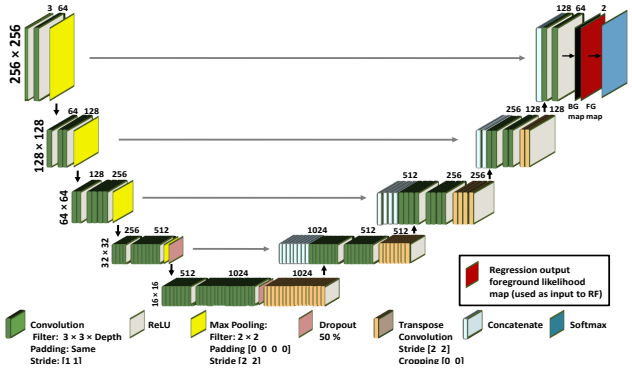


Fig. 3. Adapted U-Net architecture for capillary vessel segmentation in microscopy images of brain dura mater.

in the image has a unique class label or category that is typically used to classify objects such as people, vehicles, objects, animals, etc. Several deep architectures have been proposed for semantic segmentation [22]. U-Net[17] was designed for biomedical image segmentation and has become very popular after winning two challenges at ISBI 2015 (Grand Challenge and Cell Tracking Challenge). U-Net's name is inspired from its architecture layout which consists of a contraction path followed by an expansion path, each composed of several blocks; blocks maybe convolution or transpose convolution, ReLU and pooling. U-Net preserves spatial information even through max-pooling layers along the contraction path by concatenating intermediate layers with their corresponding feature maps along the opposite expansion path (see Fig. 3). As a result, larger structural context is captured along the contraction (feature) path without loss of spatial semantic details along the expansion (pixel labeling) path.

Figure 3 shows the modified U-Net architecture deep network that uses input image patches or tiles of size 256×256 pixels with filter sizes, padding, and strides shown. Tiles are used for training to decrease the number of layer (convolution) parameters needed and increase the training speed. Our architecture consists of four stages along both contraction and expansion paths; each block has 2 convolutions, 2 ReLU's and one max pooling, and transpose convolution (upsampling and interpolation) replaces max pooling along the expansion path. After each transpose convolution feature maps are concatenated from the encoder side to have higher resolution for pixel labeling. We incorporated two dropout layers, with a probability of 0.5, to reduce over-fitting.

3. EXPERIMENTAL RESULTS

3.1. Dataset and setup

The experiments were performed using high resolution epifluorescence microscopy images of mice dura mater for understanding changes to the vasculature network remodeling morphology under estrogen deprivation versus a control. A set of 40 epifluorescence ovary intact (OV), also referred to as intact female, images consisting of 20 ER-β wild-type (WT), and 20 ER-β knock-out (KO) intact mice were used. The treated post-ovariectomy (OVX) class includes a set of 20 epifluorescence images consisting of 10 WT and another 10 KO. These 60 images are selected and marked carefully by three experienced physiologists, using interactive methods [23], to obtain segmentations of the arterioles and venules. We use a Hessian feature-based multi-focus fusion algorithm [24] to obtain a set of higher quality all-focus epifluorescence images and our proposed framework was applied on these fused images[1]

3.2. Implementation details and comparison

In our experiments, we used four fold cross-validation using 40 OV/IF images from WT and KO intact mice. OVX images were set aside just for testing as they have different statistical characteristics that may adversely effect the classifier. In each experiment, we utilized 30 images for training and the remaining 10 for testing in addition to the 20 OVX images. The overall performance is estimated by averaging the cross-validation results. Table 2 compares the average performance of 5 evaluation metrics using the set of 60 images.

Evaluating our dual RF_OFB+U-Net deep architecture followed these three steps: i) Training the U-Net architecture to provide an informative likelihood feature map to be used with other hand-crafted features, ii) Training the random forest bagging classifier combining the 18 filter bank features described in Table 1, with the U-Net likelihood map to form a 19-D feature vector, and, iii) Using the resulting trained model for inference to assess delineation of blood vessel networks in the dura mater epifluorescence microscopy images.

U-Net was trained using non-overlapping tiles. Our images are 1036×1360, with images padded to 1280×1536 and cropped into non-overlapping tiles of size 256×256 pixels. This data augmentation produced 30 tiles from each individual image to construct training samples equal to 900 tiles in each fold. However, this resulted in poor segmentation results, due to insufficient training sample size; as seen in Table 2 Row 4 (U-Net), a sensitivity of 78.35% indicates vessel segments have missed detections. In the second experiment, we used additional data augmentation to increase the training sample size by randomly cropping 50 tiles from each image, and each tile is augmented 8 times by adding random rotations, reflections, and scaling. As a result, our new training set consists of 12,000 samples $(30 \times 50 \times 8)$ for each fold. Our optimized U-Net has a notable improvement in performance, Row 5 in Table 2, with a Dice value of 86.88%, almost 4% higher than the first experiment with 83.03%.

The random forest was trained using the same cross validation approach as for U-Net. The likelihood map is obtained using the trained model of U-Net for each individual image, simultaneously, feature maps described in Table 1 are gener-

[1]Each of the 60 epifluorescence fused images used consist of a variable number of single-focus images at the same tissue location.

Table 2. Experimental results for six approaches to extract *both* arteriole and venule blood vessel segments in our challenging dura mater epifluorescence microscopy blood vessel images. We show the overall average sensitivity, precision, specificity, accuracy, and Dice coefficient for 60 images with **Bold** highlighting the best performance.

#	Machine Learning Approaches	Sensitivity	Precision	Specificity	Accuracy	Dice
1	RF + no preprocessing	86.61 ± 7.95	87.44 ± 5.45	99.16 ± 0.39	98.21 ± 0.93	86.68 ± 4.41
2	RF + preprocessing all features	89.03 ± 7.1	84.05 ± 8.09	98.89 ± 0.53	98.07 ± 0.89	85.93 ± 3.85
3	RF + preprocessing intensity only	87.66 ± 7.3	87.62 ± 5.3	99.15 ± 0.41	98.27 ± 0.93	87.35 ± 4.02
4	U-Net	78.35 ± 8.7	$\mathbf{89.54 \pm 7.7}$	$\mathbf{99.37 \pm 0.4}$	97.89 ± 1.1	83.03 ± 5.4
5	U-Net optimized	89.22 ± 6.2	85.81 ± 9.6	98.98 ± 0.76	98.24 ± 0.86	86.88 ± 4.9
6	Proposed method RF_OFB+U-NET	$\mathbf{89.68 \pm 6.1}$	86.96 ± 6	99.08 ± 0.53	$\mathbf{98.37 \pm 0.83}$	$\mathbf{88.00 \pm 3.6}$

ated to produce feature vectors for 42,268,800 pixel observations; with most of those observations for background pixels. For a more balanced sampling of the classes, we randomly removed background observation samples till the class distributions had almost an equal number of foreground and background pixels. The new 19-D feature vectors had 22,787,086 observations with their corresponding labels. After heuristic analysis we chose 70 trees to train the RF and 15 epochs to train U-Net with stochastic gradient descent and momentum optimization. The results in all experiments have been post-processed using morphological operations.

The first three rows of Table 2 shows the performance of hand-crafted features with RF classifier. It is interesting to note that Dice decreased by about 1% when the raw image was pre-processed before feature maps were generated (Row 1 vs 2). This decrease is likely due to increased noise (false positive pixels) in the segmentation. Applying pre-processing to the intensity feature only, produces the best Dice results among all segmentations using hand-crafted features. Rows 4 and 5 in Table 2 show U-Net experiments, and the last row shows our pipeline results. The overall average performance of our pipeline outperforms the stand-alone approaches in terms of Dice, accuracy and sensitivity. The combination of a set of strong hand-crafted features with a U-Net likelihood map adds more robustness and confidence for the random forests pixel-wise segmentation outputs as shown in Figure 4.

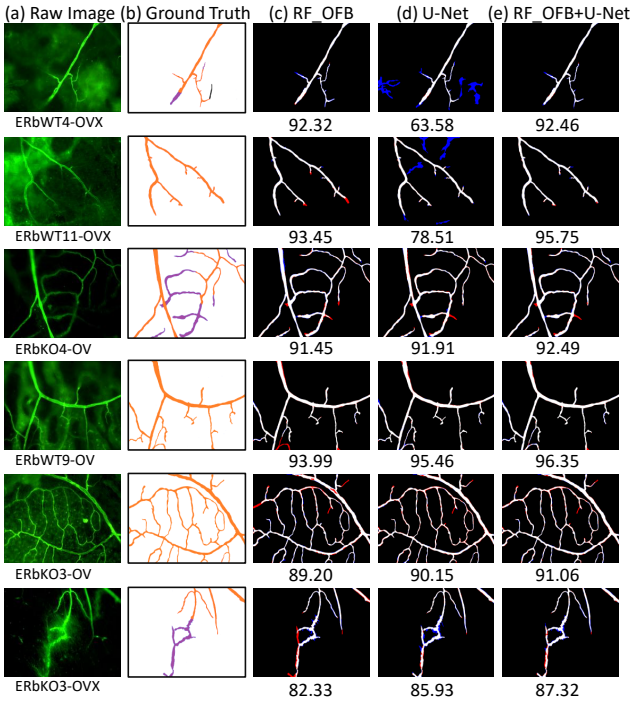


Fig. 4. Comparison of proposed pipeline for vessel segmentation using Dice values. (a) Input epifluorescence microscopy image, (b) Ground truth segmentation, (c) RF_OFB related to row 3 in Table 2, (d) U-Net for row 5 in Table 2, and (e) Our pipeline RF_OFB+U-Net. In the ground truth, **orange** segments correspond to the arteriole part, and, **purple** segments represent venules. In the automatic segmentation results, white regions represent correctly segmented pixels, and pixels colored **red** are missing (false negative) and **blue** are extra regions (false positive) compared to GT. The RF_OFB+U-Net (Col (e)) is better than either single approach method.

4. CONCLUSIONS

The proposed RF_OFB+U-Net deep learning architecture for automatic blood vessel segmentation combines an optimized convolution feature filter bank with U-Net learned vessel regression feature map using a random forest semantic segmentation classifier. Selective image enhancement using CLAHE has a positive impact on performance. The proposed hybrid approach outperforms either individual hand-crafted or deep learning U-Net feature groups for vessel segmentation in terms of accuracy (98.4%) and Dice coefficient (88%).

5. ACKNOWLEDGMENTS

This work was partially supported by awards from U.S National Institute of Health NINDS R01 NS110915, R01 DK095501 (VHH), NSF CNS-1429294 and DMS-1853222/1853303. Animals were graciously provided by Dennis Lubahn. YMK was partially supported by an HCED Government of Iraq doctoral scholarship. The content does not necessarily represent the official views of NIH or NSF.

6. REFERENCES

[1] G. Guidoboni, A. Harris, and R. Sacco, *Mathematical Modeling of Ocular Fluid Dynamics: From Theory to*

Clinical Applications, Springer-Birkhauser, 2019.

[2] G. Guidoboni et al., "Intraocular pressure, blood pressure, and retinal blood flow autoregulation: A mathematical model to clarify their relationship and clinical relevance," *Investigative Ophthalmology & Visual Science*, vol. 55, no. 7, pp. 4105–4118, 2014.

[3] M. Glucksberg and R. Dunn, "Direct measurement of retinal microvascular pressures in the live, anesthetized cat," *Microvascular Research*, vol. 45, no. 2, pp. 158–165, 1993.

[4] M. Fraz et al., "Blood vessel segmentation methodologies in retinal images – A survey," *Computer Methods and Programs in Biomedicine*, vol. 108, no. 1, pp. 407–433, 2012.

[5] R. M. Tayebi et al., "Coronary artery segmentation in angiograms with pattern recognition techniques – A survey," in *Int. Conf. Advanced Computer Science Applications and Technologies*, 2013, pp. 321–326.

[6] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever, "Multiscale vessel enhancement filtering," *Int. Conf. Medical Image Computing and Computer Assisted Intervention (MICCAI)*, pp. 130–137, 1998.

[7] U. T. V. Nguyen, A. Bhuiyan, L. A. F. Park, and K. Ramamohanarao, "An effective retinal blood vessel segmentation method using multi-scale line detection," *Pattern Recognition*, vol. 46, no. 3, pp. 703–715, 2013.

[8] D. Marín, A. Aquino, M.E. Gegúndez-Arias, and J.M. Bravo, "A new supervised method for blood vessel segmentation in retinal images by using gray-level and moment invariants-based features," *IEEE Transactions on Medical Imaging*, vol. 30, no. 1, pp. 146, 2011.

[9] Y. M. Kassim, V. B. S. Prasath, R. Pelapur, O. Glinskii, R. J. Maude, V. Glinsky, V. Huxley, and K. Palaniappan, "Random forests for dura mater microvasculature segmentation using epifluorescence images," in *IEEE Engineering in Medicine and Biology Society Conf. (EMBC)*, 2016, pp. 2901–2904.

[10] P. Teikari, M. Santos, C. Poon, and K. Hynynen, "Deep learning convolutional networks for multiphoton microscopy vasculature segmentation," *arXiv: 1606.02382*, 2016.

[11] H. Fu et al., "DeepVessel: Retinal vessel segmentation via deep learning and conditional random field," in *Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2016, pp. 132–139.

[12] Q. Li et al., "A cross-modality learning approach for vessel segmentation in retinal images," *Transactions on Medical Imaging*, vol. 35, no. 1, pp. 109–118, 2016.

[13] T. J. Jebaseeli, C. A. Durai, and J. D. Peter, "Segmentation of retinal blood vessels from ophthalmologic diabetic retinopathy images," *Computers & Electrical Engineering*, vol. 73, pp. 245–258, 2019.

[14] Y. M. Kassim, V B. S. Prasath, O Glinskii, V Glinsky, V Huxley, and K. Palaniappan, "Microvasculature segmentation of arterioles using deep CNN," in *IEEE Int. Conf. on Image Processing*, 2017, pp. 580–584.

[15] Y. M. Kassim and K. Palaniappan, "Extracting retinal vascular networks using deep learning architecture," in *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2017, pp. 1170–1174.

[16] Y. M. Kassim, R. J. Maude, and K. Palaniappan, "Sensitivity of cross-trained deep CNNs for retinal vessel extraction," in *IEEE Engineering in Medicine and Biology Society Conf. (EMBC)*, 2018, pp. 2736–2739.

[17] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Int. Conf. Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2015, pp. 234–241.

[18] J. L. Schönberger, H. Hardmeier, T. Sattler, and M. Pollefeys, "Comparative evaluation of hand-crafted and learned local features," in *IEEE Conf. Computer Vision and Pattern Recognition*, 2017, pp. 6959–6968.

[19] S. M. Pizer et al., "Adaptive histogram equalization and its variations," *Computer Vision, Graphics, and Image Processing*, vol. 39, no. 3, pp. 355–368, 1987.

[20] T. Leung and J. Malik., "Representing and recognizing the visual appearance of materials using three-dimensional textons," *International Journal of Computer Vision*, vol. 43, no. 1, pp. 29–44, 2001.

[21] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 3431–3440.

[22] A. Garcia-Garcia et al., "A review on deep learning techniques applied to semantic segmentation," *arXiv:1704.06857*, 2017.

[23] S. Meena, V. B. S. Prasath, Y. M. Kassim, R. J. Maude, O. Glinskii, V. Glinsky, V. Huxley, and K. Palaniappan, "Multiquadric spline-based interactive segmentation of vascular networks," in *IEEE Eng. in Medicine and Biology Society Conf. (EMBC)*, 2016, pp. 5913–5916.

[24] R. Pelapur, V. B. S. Prasath, F. Bunyak, O. Glinskii, V. Glinsky, V. Huxley, and K. Palaniappan, "Multi-focus image fusion on epifluorescence microscopy for robust vascular segmentation," in *IEEE Eng. in Medicine and Biology Society Conf.*, 2014, pp. 4735–4738.