Cortical Vessel Segmentation for Neuronavigation using Vesselness-enforced Deep Neural Networks

Nazim Haouchine, Michael Nercessian, Parikshit Juvekar, Alexandra Golby, Sarah Frisken

Cortical Vessel Segmentation for Neuronavigation using Vesselness-enforced Deep Neural Networks

Nazim Haouchine, Michael Nercessian, Parikshit Juvekar, Alexandra Golby and Sarah Frisken

***Abstract*—We propose in this paper an efficient method to segment cortical vessels in craniotomy images acquired by the surgical microscope. Our method uses a vesselness-enforced convolutional neural network to classify each pixel of a craniotomy image as a vessel or surrounding tissue. This permits training the network not only on appearance-based features but also on geometrical-based constraints that will ensure the continuity of the vascular trees. Our solution uses neural style transfer to generate new instances of images from manually labeled data leading to augment the training dataset in an anatomically semantic manner. The generated images improve the generalization of our model to various types of cortical surface appearances and vascular geometries. We conducted experiments on real images from human patients that demonstrate that accurate intraoperative cortical vessel segmentation can be achieved.**

***Index Terms*—Semantic Segmentation, Vesselness Filter, Craniotomy, Image Analogy, Image-guided Neurosurgery**

I. Introduction

Most brain tumors are initially treated with surgical resection through craniotomy. It involves removing a part of the bone from the skull in order to expose the brain surface and gives surgeons direct access to the brain [1]. After opening the skull, the brain surface is visible and can be used for pre-operative to intra-operative registration [2], particularly vessels located at the brain surface (cortical vessels). This registration can be used to help locate tumors' positions and other anatomical brain features [1] [3].

Blood vessels play important roles in medical diagnosis and various vascular segmentation approaches have been proposed in a different contexts, including fetoscopy [4], retinal surgery [5], colonoscopy [6] and intra-abdominal endoscopy [7]. Nevertheless, in neurosurgery, very few studies investigated the segmentation of cortical vessels from craniotomy images acquired from a surgical microscope. The segmentation of blood vessels can help in registering pre-operative and intra-operative images to highlight the tumor's position, ensure surgical safety and improve the localisation of vessel malformations. Manual detection can be sufficient in certain cases where surgeons are involved using interactive tools. In [8] brain displacements are estimated by manually defining vessel centerlines and then using them to drive a Coherent Point Drift matching algorithm to register preoperative MRI scans. This method was later improved to extract vessel centerlines from multiview cameras and reconstruct their 3D shape [9]. Using an optically tracked stylus Luo *et al.* [10] manually identified cortical vessel features to estimate brain shift and corrected it for after dural opening using a model-based registration. In [11], cortical vessels were also manually selected by an operator and used for registration with the preoperative MRI. Vessels' centerlines were used to constrain a biomechanical model and estimate brain deformation. Although manual, these methods show the value of using cortical vessels as features to correctly drive a non-rigid registration.

Estimating the 3D shape of the visible part of the brain brings important information on the amount of deformation undergone because of the craniotomy. Although vessels are not specifically extracted, automatic methods for 3D reconstruction and tracking of cortical surfaces exist [12] [13]. In order to automatise vessel extraction, Jiang *et al.* [14] used phase-shifted 3D measurement with Frangi filter to detect cortical features from texture-based images of the brain surface. A two-step approach was proposed by Ding *et al.* [15] using a vesselness filtering enhancement then edge detection to extract centerlines from craniotomy images. The centerlines are later used to estimate cortical displacement in video sequences based on frame-by-frame vessel tracking [16]. This method however, does not generalize well due to the large amount of variability in craniotomy sizes, shapes and visual appearance. In [2] a deep learning approach is used to classify the vessels from the parenchyma. Centerlines are then extracted and an elastic minimization is performed between preoperative and intraoperative centerlines to update the underlying tumor location. This method was limited by the scarcity of real data available and used small image patches instead of the whole images which led to fragmentation and discontinuities in the vessel segmentation. To overcome this issue, Nercessian *et al.* [17] proposed to augment the dataset using a style-enforce neural image analogy method. Nevertheless, despite an improved classification, the resulting binary images showed vessel discontinuities.

In this paper, we propose an automated approach to segment cortical vessels in craniotomy images seen through the surgical microscope. Our method uses a convolutional neural network (CNN) and improves upon previous approaches by enforcing the vessel's continuity avoiding fragmented segmentation. The fragmentation-free segmentation permits to efficiently extract centerlines that can be used to further perform pre-operative to intra-operative registration. Our approach outperforms the segmentation results of previous methods as demonstrated on both real patient and synthesized data.

This work was supported in part by the National Institute of Health under Grant R01 EB027134-01, R01 NS049251 and BWH Radiology Department Research Pilot Grant Award.

Nazim Haouchine, Parikshit Juvekar, Alexandra Golby and Sarah Frisken are with Harvard Medical School and the Brigham and Women's Hospital, Boston, MA, USA (e-mail: nhaouchine@bwh.harvard.edu).

Michael Nercessian is with Cornell University, Ithaca, NY, USA, and the Brigham and Women's Hospital, Boston, MA, USA

II. METHOD

A. Overview

Given an input image of the brain surface $\mathbf{x}$, our solution aims at segmenting the cortical vessels to produce a binary image $\mathbf{y}$. As illustrated in Figure 1, our method is a based on a *two-stage* model composed of: 1) a *segmenter* network $\mathcal{S}$ that estimate from the input image $\mathbf{x}$ a probability map $\mathbf{p}$ that outline the cortical vessels in the image; 2) a *vesselness-enforced* network $\mathcal{V}$, that, given the previously extracted probability map, corrects the geometrical continuities of the segmented images to produce a smooth and continuous vascular structure. To train the network $\mathcal{S}$, images of the brain surface are synthesized using a *style-enforced neural image analogy* network $\mathcal{A}$ and a recursive label generation algorithm Φ. Composing the networks $\mathcal{V}$ and $\mathcal{S}$ will result in a predicted binary image $\mathbf{y}$, as illustrated in Fig. 1.

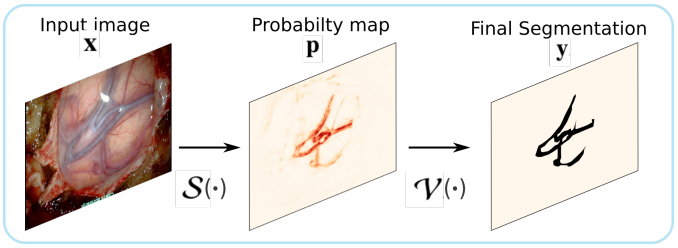


Fig. 1: Given an input image of the brain surface $\mathbf{x}$, our solution aims at segmenting the cortical vessels to produce a binary image $\mathbf{y}$ through a probability map $\mathbf{p}$ using the composition of two networks $\mathcal{V}$ and $\mathcal{S}$.

B. Semantic Segmentation Network

Let us define the training set $T^{\mathcal{S}} = \{(\mathbf{x}_i, \mathbf{y}_i)\}_i$ composed of images $\mathbf{x}_i$ of the brain surface and their corresponding labels $\mathbf{y}_i$, with $i \in |T^{\mathcal{S}}|$. The labels $\mathbf{y}_i$ correspond to the vessels, parenchyma and background. We first start by training the feature extractor $\mathcal{S}(\mathbf{x}; \theta_{\mathcal{S}})$ with $\theta_{\mathcal{S}}$ being the learned parameters for network $\mathcal{S}$. We use the feature extractor to generate a per-pixel probability map $\mathbf{p}$ that will associate with each pixel of an image $\mathbf{x}$ the probability of that pixel belonging to one of the labels.

The network $\mathcal{S}$ is a deep CNN that follows an original U-Net architecture [18] with repeated convolutional layers, each followed by a ReLU and a max pooling operation then a Softmax final layer. The classifier assigns a class label to each pixel corresponding to the background, parenchyma, and vessel. The background class is, however, excluded from the training. Indeed, we empirically found that, excluding this class (which highly varies from one patient to another) from the training, increases the segmentation accuracy. Thus, the background class is excluded from the loss that is only estimated between the two classes parenchyma and vessel. We use a *categorical cross-entropy* loss since it was empirically found that it results in improved accuracy with our data compared to Dice loss.

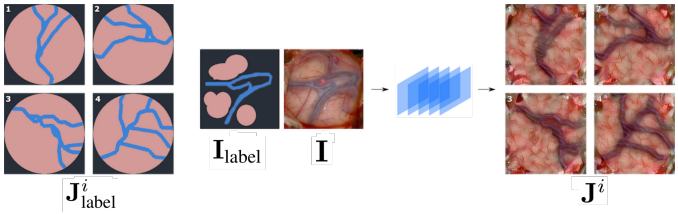


Fig. 2: Dataset augmentation: using the image labels $\mathbf{J}^i_{\text{label}}$ generated recursively using Φ, a style image $\mathbf{I}$ and its label $\mathbf{I}_{\text{label}}$ (partially labelled), we can synthesize a large number of target images $\mathbf{J}^i$ using style-enforced neural image analogy with varying appearances and geometries.

C. Vesselness Network

Using the output probability maps we can generate a thresholded binary image that will represent the segmentation of the vessels. The thresholding results in a discontinuous segmented vascular network. In order to obtain a continuous and smooth segmentation, we enforce a vesselness constrain using another convolutional network $\mathcal{V}(\mathbf{p}; \theta_{\mathcal{V}})$ with parameters $\theta_{\mathcal{V}}$ to transform probability maps to binary images, preventing the segmentation discontinuities. We use probability maps instead of thresholded images since probability maps retain the structural reasoning of the anatomical structure while keeping uncertainty information at the frontiers of classes. In order to train this network, we use the training set $T^{\mathcal{V}} = \{(\mathbf{p}_j; \mathbf{y}_j)\}_j$. This set make use of the already available binary images $\mathbf{y}_j$ used by the segmenter and their corresponding inferences $\mathbf{p}_j = \mathcal{S}(\mathbf{x}_j; \widehat{\theta_{\mathcal{S}}})$; with $j \in |T^{\mathcal{V}}|$.

The network architecture of $\mathcal{V}$ consists of 6 blocks in the encoder, composed of two convolutional and one ReLU layers. Each block except the last one is followed by an average pooling layer with a stride of 2 to decrease the spatial dimension. There are 6 blocks in the decoder part. Filter sizes and padding are similar to the original U-Net [18]. A bilinear upsampling layer is used at the beginning of each block, increasing the spatial dimension by a factor of 2, followed by two convolutional and ReLU layers. We finally optimize the following loss function over the parameters $\theta_{\mathcal{V}}$ of the network $\mathcal{V}$:

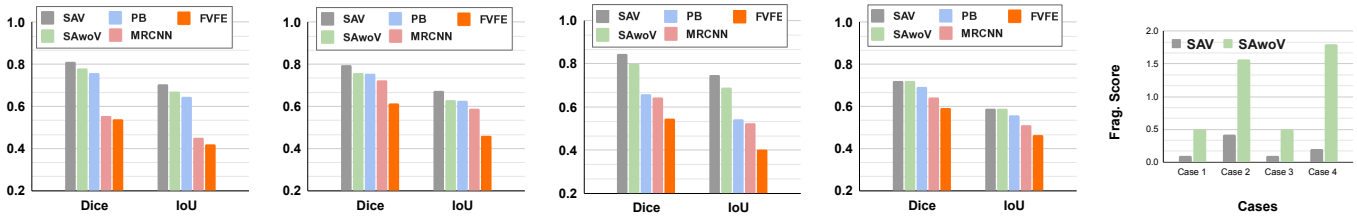
$$L(\theta_{\mathcal{S}}, \theta_{\mathcal{V}}, T^{\mathcal{S}}, T^{\mathcal{V}}) = \sum_j \left\| \mathcal{V}(\mathcal{S}(\mathbf{x}_j; \theta_{\mathcal{S}}); \theta_{\mathcal{V}}) - \mathbf{y}_j \right\|_2 \quad (1)$$

Note that we compose the network $\mathcal{S}$ that extracts the probability map from image $\mathbf{x}$ and network $\mathcal{V}$ that transform the probability maps to continuous binary images enforcing a smooth vessel geometry. These networks are trained separately.

D. Style Sources and Recursive Label Generation

Following on [17], the set $T^{\mathcal{S}}$ is augmented using a style-enforced neural image analogy. We used the feed-forward generation networks method [19], which can generate high-quality textures hundreds of times faster than traditional neural style transfer models proposed by Gatys et al. [20].

Specifically, we define $\mathcal{A}(\{\mathbf{I}, \mathbf{I}_{\text{label}}, \mathbf{J}_{\text{label}}\}; \theta_{\mathcal{A}})$, with $\theta_{\mathcal{A}}$ being the learned parameters for network $\mathcal{A}$. This network



(a) Case 1 (b) Case 2 (c) Case 3 (d) Case 4 (e) Frag. Score

Fig. 3: Comparison with state-of-the-art methods. Our method **SAV** produces more accurate results than other approaches and produces less fragments.

generates a target image $\mathbf{J}$ from the style of a source image $\mathbf{I}$ following a labelled input image $\mathbf{J}_{\text{label}}$ while keeping the high-level content of $\mathbf{I}_{\text{label}}$ (See Figure 2).

The source images $\mathbf{I}$ and their corresponding sources labels $\mathbf{I}_{\text{label}}$ were selected from the original dataset on the basis of lack of obstructions, high quality, and representative anatomy. As illustrated in Figure 2, we used sparse labeling in our style sources since it was found to lead to higher quality results. This is possibly caused by the fact that sparse labeling produces less variability and noise.

We randomly generated image labels with vessel-like structures to mimic real craniotomy image labels using a procedural algorithm Φ. This algorithm takes as inputs the desired image size s, a progression step parameter Δp that represent the vessel branches length, a bifurcation angle $\alpha \in [-90, 90]$ to direct the bifurcation, the branches width b_{width} and depth $b_{\text{depth}} \in \{0, 1, 2, 3\}$ to limit the number of branches.

Algorithm 1: Recursive Label Generation: Φ

Input: s, Δp, α, b_{width}, b_{depth}
Output: $\mathbf{J}_{\text{label}}$

Given an initial pixel p_0 and a threshold $\mathcal{T} \in [0, 1]$
Given a Random operator $\Omega(\cdot)$
Given a Bezier Curve function $\mathcal{B}(\cdot)$
if $b_{depth} > 0$ **then**
  sample Angle α
  $p_1 = p_0 + \cos(\alpha) \times \Delta p$
  $\mathcal{B}(p_0, p_1, b_{\text{width}}, \mathbf{J}_{\text{label}})$
  if $\Omega(cdot) < \mathcal{T}$ **then**
    sample Angle α
    $\Phi(s, \Delta p, \alpha, b_{\text{width}}, b_{\text{depth}} - 1)$
  $p0 = p1$

This algorithm will recursively build new vessels paths with variable widths b_{width} and lengths Δp, branching out in either side at random angle between α, while limiting the bifurcation to a consistent number b_{depth} generating a randomly sampled multi-class label image and ultimately build the set $\{\mathbf{J}^i_{\text{label}}\}_i$.

E. Inference

Finally, at run-time, given an image $\mathbf{x}$ we predict the corresponding segmentation $\mathbf{y}$ by composing the networks $\mathcal{V}$ and $\mathcal{S}$ together:

$$\mathbf{y} \leftarrow \mathcal{V}(\mathcal{S}(\mathbf{x}; \widehat{\theta_{\mathcal{S}}}); \widehat{\theta_{\mathcal{V}}}) \tag{2}$$

where $\widehat{\theta_{\mathcal{S}}}$ and $\widehat{\theta_{\mathcal{V}}}$ are the resulting parameters from the training. It should be noted that the network $\mathcal{A}$ is not used at run-time.

III. Results

A. Dataset

The dataset for the segmenter $T^{\mathcal{S}}$ was composed of 4200 images generated using neural image analogy from 10 craniotomy style source images and 250 manually annotated inputs. This dataset was splitted into a training set and a validation set (75%/25%) by making sure that the validation set does not include augmentation from neural image analogy. The training set was augmented using geometric transformations. The model was trained using mini-batches of size 8 over 94 epochs with an Adam optimizer and a learning rate of 0.001.

On the other hand, the dataset $T^{\mathcal{V}}$ was composed of 1160 probability maps and annotated labels. We also used geometric transformations to expand the dataset. The model was trained using mini-batches of size 16 over 68 epochs with an Adam optimizer and a learning rate of 0.001.

We used the Tensorflow framework on an NVidia GeForce GTX 1070 (www.tensorflow.org) .

Method	Metric	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Overall
SwoAV	**IoU**	0.646	0.645	0.647	0.642	0.651	0.647
	Dice	0.785	0.787	0.786	0.782	0.789	0.786
SAwoV	**IoU**	0.713	0.711	0.705	0.716	0.701	0.709
	Dice	0.825	0.823	0.820	0.827	0.816	0.822
SAV	**IoU**	0.745	0.741	0.744	0.746	0.742	0.744
	Dice	0.853	0.851	0.853	0.854	0.852	0.852

TABLE I: Effectiveness of different components design on a Five-fold cross-validation.

B. Effectiveness of different components design

We first perform an ablation study to analyze the contribution of each component of our approach. We test the impact of augmenting the data with a neural style transfer model and the impact of using a vesselness-enforced network. To this end, we train three variants of our model: **SwoAV** which is a traditional U-Net without neural style transfer and vesselness enforcement, **SAwoV** which is the segmenter only (similar to

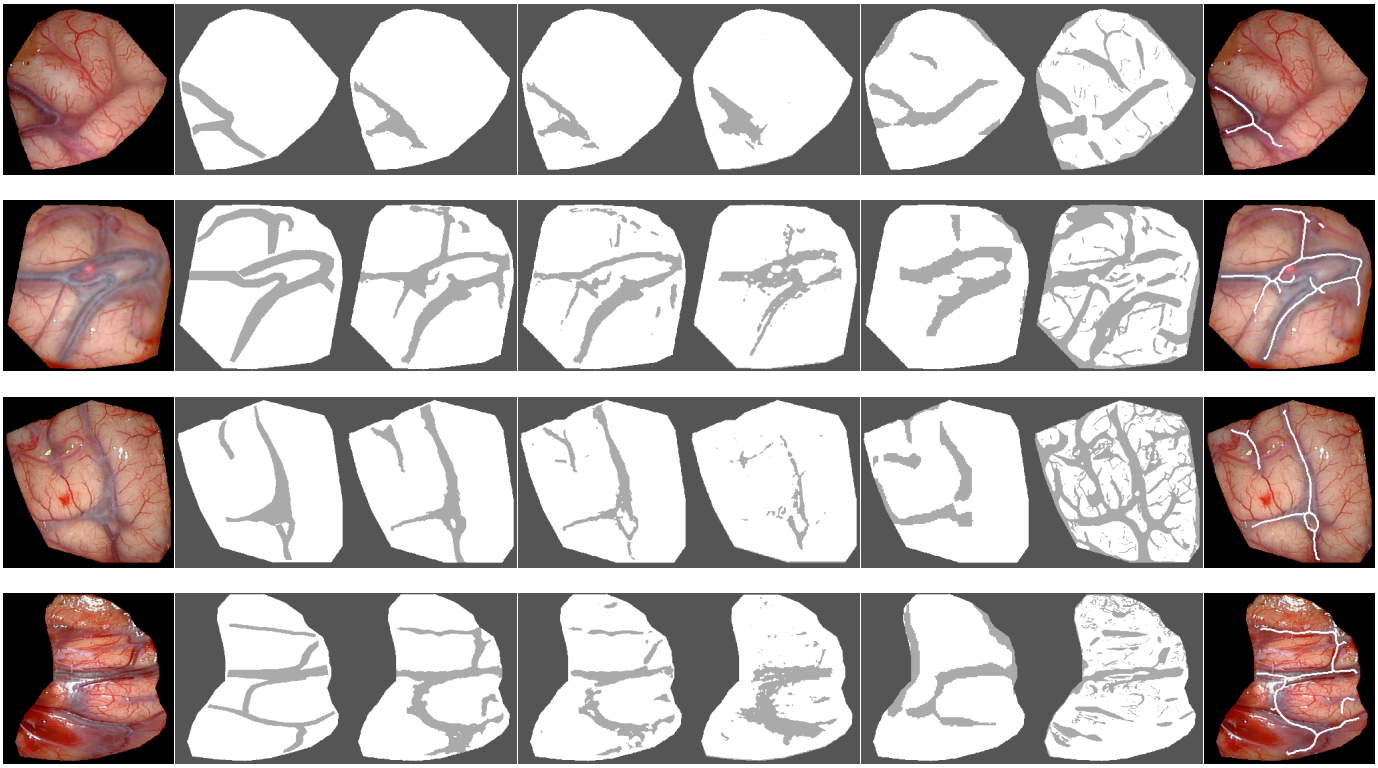


Fig. 4: Qualitative comparison of our deep learning approach against related approaches for cortical vessel segmentation. First row is case 1, second is case 2, third is case 3 and fourth is case 4. The last column shows centerlines generated from our segmented images (**SAV**).

[17]), and **SAV** (full model) which is the segmenter with the augmentation and the vesselness enforcement. To quantify the accuracy of our method we use the Intersection-over-Union (IoU) metric and the Dice score (F1 accuracy).

We can observe from Table I that removing the style augmentation from the model harms performance, which is expected, as measured by IoU 0.647 and Dice 0.786 for the variant **SwoAV** . We also verify that adding the continuity constraints to the segmentation improves the accuracy as measured by IoU 0.709 and Dice 0.822 for **SAwoV** and IoU 0.744 and Dice 0.852 for the full model **SAV**, which validate our hypothesis to enforce the vesselness in order to remove the fragmentation.

C. Comparison with state-of-the-art methods

In addition to artificially generated images, we tested our model's performance retrospectively on real craniotomy images from 4 patients. We compared the performance of our method **SAV** against the **SAwoV** method (similar to [17]), a baseline U-Net using patch-based data augmentation techniques [2] **PB**, the semantic segmentation method Mask R-CNN [21] **MRCNN** as well as the Frangi vesselness filter estimator **FVFE** used in the method proposed by Ding *et al.* [15] in the final step. These methods were chosen because of the relevance to the application or because of their popularity in the general task of multi-classe semantic segmentation. Adding the geometrical constraints yields better segmentations as noted in the third column of Figure 4 when comparing against **SAwoV**. The patch-based U-Net **PB** and Mask R-CNN **MRCNN** did not achieve consistent segmentation over all the images, particularly noted in the fifth and sixth column of Figure 4. Segmentation using the Frangi vesselness filter **FVFE** appeared to be less specific for the larger vessels and much noisier compared to our method. Performance of all models and methods on the patient images is illustrated in Figure 4. Our model achieved the highest Dice and IoU values across these images as reported in Figure 3.

We also define a fragmentation score to estimate quantitatively the impact of adding the geometric continuity enforcement. This score is computed as being the fraction of islands in the image (isolated set of pixels) w.r.t the number of vessel branches. We can notice from Figure 3-(e) that our method produces very few fragmentations.

IV. Conclusion

We presented in this paper a deep learning approach to cortical vessel segmentation in craniotomy images. We have shown that composing two networks, where the first one estimates the probability of a pixel being part of a vessel or not, and the second one enforces a geometry consistency constraints, leads to smoother segmentation with very few fragments. In addition, we have shown that training on data

generated artificially with neural style transfer performs well when evaluated on real patient data. This result circumvents a common problem of deep learning methods—lack of real labeled data. Our method can be used in the field of image-guided neurosurgery as it provides a robust and automatic method for extracting the vascular centerlines at cortical level that can be used in registration algorithms.

Future work will consist of assembling a larger data repository by improving our neural image analogy approach with the aim of generating more realistic images. In addition, we will extend the segmentation of cortical vessels to sulci. This will provide us with a larger graph-like structure (vessels and sulci) that can compensate for the potential vessel occlusions.

REFERENCES

[1] M. Miga et al., "Clinical evaluation of a model-updated image-guidance approach to brain shift compensation: experience in 16 cases," *International Journal of Computer Assisted Radiology and Surgery*, vol. 11, no. 8, pp. 1467–1474, Aug 2016.

[2] N. Haouchine, P. Juvekar, W. M. Wells III, S. Cotin, A. Golby, and S. Frisken, *Deformation Aware Augmented Reality for Craniotomy Using 3D/2D Non-rigid Registration of Cortical Vessels*, 09 2020, vol. 12264, pp. 735–744.

[3] J. F. Fraser et al., "Brainlab image guided system," in *Textbook of Stereotactic and Functional Neurosurgery*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 567–581.

[4] S. Bano, F. Vasconcelos, L. Shepherd, E. Vander Poorten, T. Vercauteren, S. Ourselin, A. David, J. Deprest, and D. Stoyanov, *Deep Placental Vessel Segmentation for Fetoscopic Mosaicking*, 09 2020, pp. 763–773.

[5] M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A. Rudnicka, C. Owen, and S. Barman, "Blood vessel segmentation methodologies in retinal images - a survey," *Computer methods and programs in biomedicine*, vol. 108, pp. 407–33, 04 2012.

[6] M. Golhar, Y. Iwahori, M. Bhuyan, K. Funahashi, and K. Kasugai, *Blood Vessel Delineation in Endoscopic Images with Deep Learning Based Scene Classification*, 06 2018, pp. 147–168.

[7] B. Lin, Y. Sun, J. E. Sanchez, and X. Qian, "Efficient vessel feature detection for endoscopic image analysis," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 4, pp. 1141–1150, 2015.

[8] F. Marreiros, S. Rossitti, C. Wang, and O. Smedby, "Non-rigid deformation pipeline for compensation of superficial brain shift," vol. 8150, 09 2013.

[9] F. M. M. Marreiros, S. Rossitti, P. M. Karlsson, C. Wang, T. Gustafsson, P. Carleberg, and Örjan Smedby, "Superficial vessel reconstruction with a Multiview camera system," *Journal of Medical Imaging*, vol. 3, no. 1, pp. 1 – 1, 2016.

[10] M. Luo et al., "A comprehensive model-assisted brain shift correction approach in image-guided neurosurgery: a case study in brain swelling and subsequent sag after craniotomy," in *Medical Imaging 2019: Image-Guided Procedures, Robotic Interventions, and Modeling*, vol. 10951, 2019, pp. 15 – 24.

[11] N. Haouchine, P. Juvekar, A. Golby, W. Wells, S. Cotin, and S. Frisken, "Alignment of cortical vessels viewed through the surgical microscope with preoperative imaging to compensate for brain shift," vol. 11315, 03 2020, p. 60.

[12] S. Ji, X. Fan, D. W. Roberts, A. Hartov, and K. D. Paulsen, "Cortical surface shift estimation using stereovision and optical flow motion tracking via projection image registration," *Medical Image Analysis*, vol. 18, no. 7, pp. 1169 – 1183, 2014.

[13] A. Kumar, M. Miga, T. Pheiffer, L. Chambless, R. Thompson, and B. Dawant, "Automatic tracking of intraoperative brain surface displacements in brain tumor surgery," 2014, pp. 1509–1512.

[14] J. Jiang et al., "Marker-less tracking of brain surface deformations by non-rigid registration integrating surface and vessel/sulci features," *International journal of computer assisted radiology and surgery*, vol. 11, 03 2016.

[15] S. Ding, M. I. Miga, R. C. Thompson, I. Garg, and B. M. Dawant, "Automatic segmentation of cortical vessels in pre- and post-tumor resection laser range scan images," in *Medical Imaging 2009: Visualization, Image-Guided Procedures, and Modeling*, M. I. Miga and K. H. Wong, Eds., vol. 7261, International Society for Optics and Photonics. SPIE, 2009, pp. 41 – 48.

[16] S. Ding, M. Miga, T. Pheiffer, A. Simpson, R. Thompson, and B. Dawant, "Tracking of vessels in intra-operative microscope video sequences for cortical displacement estimation," *IEEE transactions on bio-medical engineering*, vol. 58, pp. 1985–93, 02 2011.

[17] M. Nercessian, N. Haouchine, P. Juvekar, S. Frisken, and A. Golby, "Deep cortical vessel segmentation driven by data augmentation with neural image analogy," in *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, 2021, pp. 721–724.

[18] O. Ronneberger, P.Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *MICCAI*, vol. 9351, 2015, pp. 234–241.

[19] D. Ulyanov, V. Lebedev, Andrea, and V. Lempitsky, "Texture networks: Feed-forward synthesis of textures and stylized images," in *Proceedings of The 33rd International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, M. F. Balcan and K. Q. Weinberger, Eds., vol. 48. New York, New York, USA: PMLR, 20–22 Jun 2016, pp. 1349–1357.

[20] L. A. Gatys, A. S. Ecker, and M. Bethge, "A neural algorithm of artistic style," *arXiv*, Aug 2015. [Online]. Available: http://arxiv.org/abs/1508.06576

[21] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," 2017, cite arxiv:1703.06870Comment: open source; appendix on more results. [Online]. Available: http://arxiv.org/abs/1703.06870