BLOOD VESSEL CHARACTERIZATION USING VIRTUAL 3D MODELS AND CONVOLUTIONAL NEURAL NETWORKS IN FLUORESCENCE MICROSCOPY

Aritra Chowdhury, Dmitry V. Dylov, Qing Li, Michael MacDonald, Dan E. Meyer, Michael Marino, Alberto Santamaria-Pang

GE Global Research Center, Niskayuna NY

ABSTRACT

We report an automated method for characterization of microvessel morphology in micrographs of brain tissue sections to enable the facile, quantitative analysis of vascular differences across large datasets consisting of hundreds of images with thousands of blood vessel objects. Our objective is to show that virtual 3D parametric models of vasculature are adequately capable of representing the morphology of naturally acquired data in neuropathology. In this work, we focus on three distinct morphologies that are most frequently observed in formalin-fixed, paraffin-embedded (FFPE) human brain tissue samples: single blood vessels showing no (or collapsed) significant lumen ("*RoundLumen-"*); single blood vessels with distinct lumen ("*RoundLumen+"*); two blood vessels bundled together in close proximity ("*Twins"*). The analysis involves extraction of features using pre-trained convolutional neural networks. A hierarchical classification is performed to distinguish single blood vessels (*RoundLumen*) from *Twins*; followed by a more granular classification between *RoundLumen*- and *RoundLumen*+. A side-by-side comparison of the virtual and natural data models is presented. We observed that classification models built on the virtual data perform well achieving accuracies of 92.8% and 98.3% for the two aforementioned classification tasks respectively.

Index Terms— Vasculature Analysis, Neuropathology, Convolutional Neural Networks, Parametric 3D models.

1. INTRODUCTION

Characterizing the morphology of vasculature in digital pathology is an important step in defining the microenvironment within brain tissue samples. In particular, understanding the geometry of vessel configuration and its changes during a disease may provide deeper insight into the progression of neuropathological degenerative diseases such as Alzheimer's disease. Images acquired using 20x object magnification from immunofluorescent (IF) stained 6 um tissue sections with collagen IV antibody are used in this work. We attempt to characterize three different types of blood vessel morphologies which are found in relative abundance in our image data set. They are singular blood vessels with no visible lumen, singular blood vessels with a distinct lumen and blood vessels appearing as a pair; which we have named *RoundLumen-*, *RoundLumen+*, and *Twins* correspondingly. In this work, we show that it is possible to characterize blood vessels using convolutional neural networks (CNN) as opposed to traditional image processing techniques which involve segmentation and hand-crafted feature extraction. Instead, here we use a pre-trained CNN to extract features from the images. This technique of "deep transfer learning" is compared to the visual bag of words (VBW) method for feature extraction [4].

Acquisition of natural training samples is a time consuming and labor intensive process. Deep learning requires abundant training data for tuning the large number of parameters of the various inherent models. If a certain class is imbalanced then the classification models could become prone to biased outcomes. The construction of 3D parametric models, presented here tackles these issues and creates a balanced high-fidelity classification model. In this study, we built a basic 3D vasculature model using our prior knowledge of blood vessel geometry, as guided by a pathologist. The 3D vasculature was repeatedly sliced at various angles and orientations to obtain 2D samples for training the machine learning model, thereby mimicking the physical sectioning of tissue during sample preparation for microscopy. In addition, a filtering technique was then used to fine-tune the virtual data to reflect the variability present in the naturally acquired samples. We train three models based on: virtual data, natural data and a mixture of both. The models are then tested on a reserved, independent portion of the naturally occurring data, with a hierarchical classification being performed to demonstrate a proof of concept.

The first classification task involves distinguishing between singular blood vessels (*RoundLumen*) and pair of blood vessels (*Twins*). The second task of finer granularity is the classification between *RoundLumen*- and *RoundLumen*+. We report classification metrics for both the classification tasks and observe that the artificial data improves upon the model trained from only the natural data. As far as we know, this is the first attempt to model vasculature using parametric 3D geometric methods exclusively to generate

virtual data. Statistical 2D and 3D shape models have been used extensively in medical image segmentation as detailed in [3]. However, these models use the training samples to statistically find a model of the object of interest. We do not generate virtual 2D models because we believe that we can obtain more variability in the training data by sectioning from the 3D models. Additionally, the 3D models may be extended to various other modalities in medical imaging.

2. DATA

This section provides a detailed explanation of the different morphologies of blood vessels that are explored in this study. The first subsection is a description of the natural data curated from actual postmortem human tissue samples. The second subsection involves the description of the virtual 3D model of blood vessels which are used for generating samples for training the convolutional neural networks.

2.1. Natural data

FFPE Postmortem brain tissue samples from ten subjects with neurological disorders underwent sequential IF-multiplexing and fluorescent imaging. For each subject, approximately 25 images were acquired. This involves a cycling process of tissue section staining with 2-3 dye-labeled antibodies, imaging, dye inactivation and repeated staining with a new set of dye-labeled antibodies. Images underwent illumination correction, registration, stitching and auto-fluorescence subtraction. Collagen IV was used as a marker to detect all blood vessels. An image overlay is shown in Fig. 1.

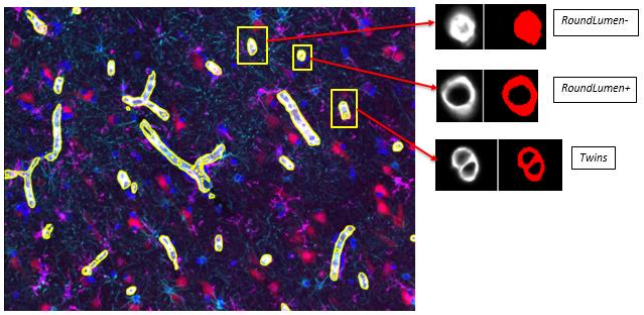


Figure 1: Depiction of the different morphologies in the natural data with respect to a multichannel image, overlaid with different protein markers. The three types of morphologies analyzed in this study is represented on the right.

The number of instances of the three different types of morphologies are depicted in Table 1.

Table 1. Number of instances in the three classes of vascular morphology

Morphology	Frequency
RoundLumen-	689
RoundLumen+	3427
Twins	266
Total	4382

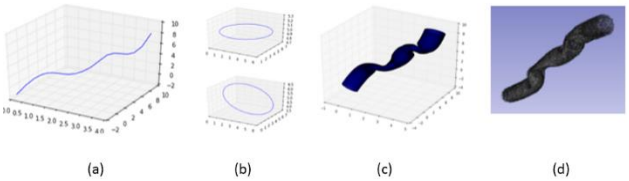


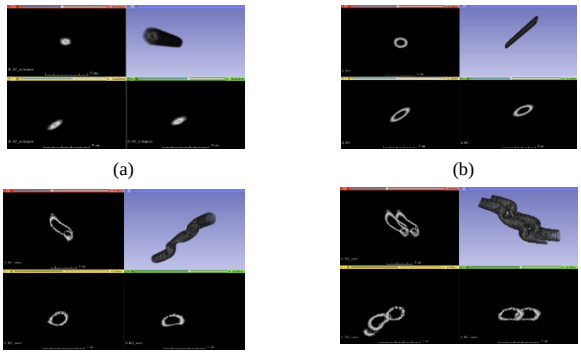
Figure 2: Development of the 3D virtual model.

2.2. Virtual data

The construction of the artificial model starts with defining a set of control points in three dimensional Cartesian coordinates. The control points reflect the basic structure that the blood vessel is supposed to represent. This is followed by interpolating between the points using a 3D cubic spline interpolator. This forms the skeleton or the center line that represents the center or the lumen of the blood vessel and is shown in Fig. 2(a). The 3D volume of the blood vessel is constructed after this step. We first define a number of parameters; the inner radius of the blood vessel (r); the outer radius (R). The number of sampling points along the spline (N); the number of sampling points in the radial direction (Nr). At each sampling point; we define a circular disk along the z-axis; by randomly perturbing the values of r and R. We also define an intensity model for the blood vessels depicted in Eq. 1. From the natural images, it seems that the intensity is high in the periphery of the blood vessel and decays towards the lumen and as we move away from the periphery. We model this using an exponential decay in the following form:

$$I(d) = I_{max}\exp(-\alpha\,|r' - d|) \qquad (1)$$

where, I_{max} is the maximum intensity, α is the calibration coefficient (in mm^{-1} units), $r'=(R+r)/2$. d is the distance from the center of the lumen. At each point on the disc, we define the voxel density as a normal distribution with mean $I(d)$ and standard deviation 0.01. This is followed by formulating the rotation matrix by calculating the angle between the tangent to the spline at that sampling point and the z-axis. The points corresponding to each point on the disc are therefore mapped or rotated along the curve by multiplying the coordinates with the rotation matrix. An example of this rotation is depicted in Fig. 2(b). We then discretize the coordinates such that we obtain an actual 3D image in the form of an array. This is depicted in Fig. 2(c). The intensity values are normalized and assigned to the corresponding discretized points in the 3-dimensional array. The volume rendered version of the 3D image is depicted in Fig.2(d). Therefore, by changing the parameters of the model we can build several different 3D images and slice it at various angles to mimic the natural tissue cross sections at various depths and angles. Some examples the various models are depicted in Fig. 3.



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

Figure 3: 3D virtual models and their corresponding projections along different planes of view (a) Linear model of *RoundLumen-*, (b) Linear model of *RoundLumen+*, (c) Non-linear model of *RoundLumen+* and (d) Non-linear model of *Twins.*

The process of constructing different types of morphologies in blood vessels is the same as explained above. Fig. 3(a) is a blood vessel with a no lumen (*RoundLumen-*) and a linear skeleton. Figures 3(b) and 3(c) are blood vessels with a single lumen having linear and non-linear structures respectively. Fig. 3(d) is a model of a *Twin.* A simple way of creating these multi-vessel structures is by perturbing or shifting the sampling points of the skeleton along a random direction. As shown in Fig. 3; the different morphologies that commonly occur naturally can be generated. As explained in the introduction; this serves as a viable alternative to using natural data for training convolutional networks.

3. METHODS

Convolutional neural networks (CNNs) consist of multiple neuron collections which process portions of the input image called receptive fields. The outputs are then tiled so that the input regions overlap and this in turn produces a better representation of the original image. This is what makes CNNs translation invariant. The CNN is made up of four types of layers: the input layer, the convolutional layer, the pooling layer; and the fully connected layer. The input layer is where the networks accept the images. The images consist of raw pixel values depicted by width, height and the number of channels. The convolutional layer will compute the output of the neurons that are connected to local regions in the input, each computing a dot product between their weights and a receptive field. The pooling layer performs a down sampling operation. The high-level reasoning in the neural network is done by fully connected layers. Their activations can be performed by matrix multiplication

In this work, we use the pre-trained convolutional neural networks as a feature extractor. This network consists of weights trained on the ImageNet dataset. We extract the 6th layer in the network which is a 4096-dimensional vector as a representation of the image. This may be considered as a transfer learning model because we transfer the weights learnt from another domain to blood vessel recognition.

We use a pre-trained neural network called AlexNet [2] to extract features from the data and perform an experiment to show that pre-trained CNNs are efficient in representing the vascular morphology. The experiment is performed on the natural data where 33% of the data is held out as test data and the rest is used for training. Two models are developed, one of them uses the visual bag of words (VBW) [4], a feature extraction method to extract the features; the other uses the *AlexNet* architecture to extract the features. A three-class classification (one vs rest) is performed using the logistic regression classifier. The accuracy, f1-score, precision and recall calculated on the same test data are reported for comparison. The results in Table 2 show that pre-trained convolutional neural networks are a good choice for representation of vascular morphology.

Table 2. Comparison of feature extraction methodologies

Feature extractor	Accuracy	f1-score	Precision	Recall
AlexNet	**91.92**	**91.93**	**91.98**	**91.92**
VBW	78.38	77.38	76.71	78.38

3. EXPERIMENTS AND RESULTS

Features are extracted using the *AlexNet* architecture which is trained on the ImageNet [1] database. The patches obtained from the natural data in section 2.2 and the virtual data obtained from the parametric 3D model described in section 2.2 are resized to size 224x224. This is because the input size of the images is a fixed parameter for the AlexNet architecture. The weight parameters are used to extract the features in a feedforward manner. This is called transfer learning. 33% of the natural data is held out as test data. All the experiments are performed on this dataset for maintaining consistency in the results.

A filtering technique is introduced to appropriately extract slices from the 3D volumes. This is done by obtaining the probabilities of the artificial data using a model trained on the natural training data. The probabilities of the corresponding images are then sorted and the images with the highest probabilities are selected. This is a way to boost the artificial model. The filtered virtual data is then used to retrain the classifier. Examples of both the natural and artificial data are represented in the Fig. 4. A hierarchical classification is performed to first classify the single blood vessels from blood vessels that are bundled in pairs i.e *RoundLumen* vs *Twins*. The second classification task involves distinguishing between *RoundLumen-* and *RoundLumen+*. We perform three different types of training to demonstrate our proposed methods. The first type is with only the naturally occurring data. The number of training samples of each class is 66% of the numbers shown in Table 1. The second type consists only of virtual data that has been filtered by the natural model.

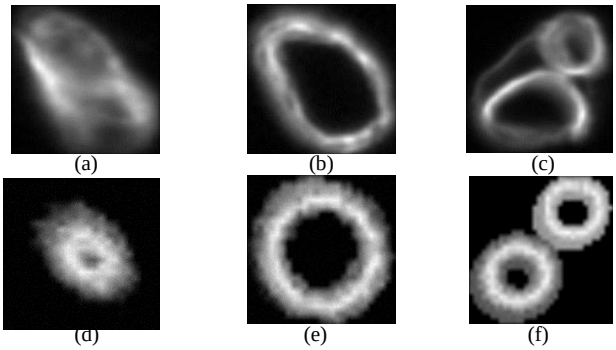


Fig. 4: Examples of vessel classes *RoundLumen-* (a,d), *RoundLumen+* (b,e) and *Twins* (c,f) for natural (a,b,c) and virtual data (d,e,f).

The total number of artificial samples is restricted to 10000 per class. Finally, the third type consists of both the artificial and natural training samples and we refer as *Mixed*. The 10000 examples generated from the artificial data are added to the natural training data. In addition, we use an oversampling technique called synthetic minority over sampling (SMOTE) [5] to equalize the number of instances for both classes in all experiments. All the results are reported on the held out 33% of the natural data depicted in Table 1. The accuracy, f1-score, precision, recall and receiver operating characteristic (ROC), and precision-recall (PR) curves are reported in the following tables and figures for each of the two classification tasks.

Table 3. Results of binary classification between *RoundLumen* and *Twin*

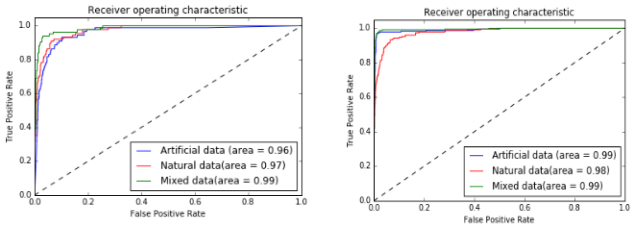
Data	Accuracy	f1-score	Precision	Recall
Artificial	92.81	59.36	45.24	**86.36**
Natural	96.34	71.03	68.42	73.86
Mixed	**97.71**	**81.76**	**79.57**	84.01

Table 4. Results of binary classification between *RoundLumen-* and *RoundLumen+*.

Data	Accuracy	f1-score	Precision	Recall
Artificial	98.38	99.02	99.38	98.67
Natural	94.55	96.72	96.89	96.55
Mixed	**98.60**	**99.16**	**99.29**	**99.03**

The ROC curves are calculated using the minority class in both the classification tasks, i.e. *Twins* for the first classification task and *RoundLumen-* for the second task.

From Table 3, we can see that the artificial data captures the differences between the two classes. It is also able to identify *Twins*, which is the minority class in Task 1, from the high recall. Therefore, the results are boosted when we combine both the artificial and natural data. In addition, the ROC curves in Figure 5 confirms our hypothesis that virtual data may be used for building the models. The naturally trained model performs better than the virtual trained model. However, as we can see from the ROC curves, the model built from the mixed data improves the performance.



(a) (b)

Fig. 5: ROC curves for classification: (a) *RoundLumen* and *Twins* and (b) *RoundLumen-* and *RoundLumen+*.

Table 4 presents the corresponding results for the second classification task. In this case, we see that the virtual data performs better than the natural data and boosts the performance when trained on its own or in union with the natural data. This task is relatively simple compared to the first task because the here the difference exists in the presence or absence of the lumen. Therefore, this is simpler to model using artificial data. In the first task, more variability is introduced due to the presence of more than one vessel. Therefore, it is the case that natural data performs better than its artificial counterpart.

4. CONCLUSIONS

We have shown that the use of deep learning algorithms trained with a mixture of virtual and natural data results in a more accurate prediction of vasculature morphologies than the use of standard feature extraction methods (such as visual bag of words). The methodology of complementing natural data with synthetic samples holds potential for becoming a standard approach in deep learning with unbalanced datasets. In neuroscience, it could be applied to help elucidate the underlying mechanisms of common neurological degenerative diseases.

5. REFERENCES

[1] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: *A large-scale hierarchical image database*. In IEEE CVPR, pp. 248–255, 2009.

[2] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. *Imagenet classification with deep convolutional neural networks.* In Advances in neural information processing systems, pp. 1097–1105, 2012.

[3] Heimann, Tobias, and Hans-Peter Meinzer. "Statistical shape models for 3D medical image segmentation: a review." *Medical image analysis* 13.4 (2009): 543-563.

[4] Yang, Jun, et al. "Evaluating bag-of-visual-words representations in scene classification." *Proceedings of the international workshop on Workshop on multimedia information retrieval*. ACM, 2007.

[5] Chawla, Nitesh V., et al. "SMOTE: synthetic minority over-sampling technique." *Journal of artificial intelligence research* 16 (2002): 321-357.