Distinguishing pericyte and smooth muscle cells in serial 2-photon tomography imaging of mouse brain blood vessels using convolutional neural networks

*Note: Sub-titles are not captured in Xplore and should not be used

Tyler Duane Ruch
Math and Computer Science
The Pennsylvania State University, Harrisburg
777 W Harrisburg Pike, Middletown Pa, 17057
ruch@psu.edu

***Abstract*—This document is a model and instructions for LaTeX. This and the IEEEtran.cls file define the components of your paper [title, text, heads, etc.]. *CRITICAL: Do Not Use Symbols, Special Characters, Footnotes, or Math in Paper Title or Abstract.**

***Index Terms*—component, formatting, style, styling, insert**

I. Introduction

Medical image segmentation and feature detection is arguably the most cutting-edge and important area in computerized image analysis. In fact, the first use of computerized image analysis was implemented to analyze cell microscopy images [1]. There is constant innovation in the field of medical imaging, whether it be in imaging for research, pathology, diagnosis, or one of many other possibilities. However, unlike many other image feature detection tasks, medical image data very rarely is labelled. In other tasks, such as identifying faces or animals in images, the manual labelling can be outsourced to the general population. However, in medical imaging, manual labelling must be done be a field expert, all of which have very valuable time and, often, better things to do with it.

This brings medical image segmentation in general to a predicament at this moment: there exists a massive amount of data, whose size and complexity increases more and more every day. Therefore, there is a large amount of room for automated segmentation and feature detection in medical images, on a large scale. Clearly, the algorithms used would be on a case-by-case basis, but techniques exist in the 2-dimensional plane and 3-dimensional plane to intelligently segment medical images with non-negligible results. Research in automatic image analysis has the potential to bridge the gap between professional fields, unlock insights about diseases [1], and save lives [4].

This paper outlines experiments performed on a dataset acquired by using PDGFR-CRE mice, which have an altered genome that fluoresces neurovascular cells. This allows serial 2-photon tomography to be used to generate extremely microscopic images of the neurovascular cell structure; in particular, the pericytes and smooth muscle cells. If these cells could be distinguished effectively and efficiently, a large amount of data could be gained from these images. Experiments will be carried out in the following order: 2-dimensional automated image segmentation, 3-dimensional automated image segmentation, 2-dimensional CNN segmentation, 3-dimensional CNN segmentation. The 2-dimensional CNN architecture will be U-Net [2] and the 3-dimensional CNN architecture will be DeepMedic [10].

It is reasonable to hypothesize that using 2D or 3D automated segmentation techniques will yield a dataset that is 80% as accurate as human ground truth labelled data, because human ground truth labeling has a bit of variance due to differeing expert opinion. This can then be used to train or augment a deep learning model that will make the segmentation accuracy comparable to an expert ground truth.

Initially, automated image segmentation techniques will be implemented in the 2D and 3D plane. Techniques including gaussian blur, mean-shift filtering, histogram thresholding, connected component analysis, shape analysis, and region growing will be experimented with in both the 2D and 3D space. These techniques are among the leading choices in automatic medical image segmentation [3] and combinations of them have had success in medical image segmentation in the past [4].To benchmark, a small dataset of expert generated ground truth data will be compared with the output of the various algorithm combination strategies. Based on the results of these tests, the CNNs will ideally be trained using training data generated with these algorithms.

Because these medical images lack labelling, if the automatic labelling of the 2D and 3D algorithms is sufficiently close to human ground truth labelling, (because even experts have labelling variance), theoretically, the algorithms could be used to create training sets for deep learning. Thus, this paper

Identify applicable funding agency here. If none, delete this.

proposes the use of the aforementioned techniques to generate training data for deep learning. Further experiments will be conducted, using the aforementioned human ground-truth data as a test set, and a massive automatically labelled dataset as a training set. Ideally, these CNN models can bring the segmentation performance from slightly/moderately below human ground truth labelling up to comparable/indistinguishable from human ground truth labelling.

TODO: Insert results and model structure later

II. Literature Review

This section will review related literature that aims to solve a similar medical image segmentation problem. In particular, this related literature either aims to solve a cellular image segmentation problem or provides techniques and knowledge concerning medical image segmentation that can be applied to cellular image segmentation. This section contains literature overviews concerning automated, image-based techniques that can be applied to cellular image segmentation, followed by overviews concerning specific automated cell segmentation pipelines and implementations, some of which include the use of deep learning. After this, this literature review provides literature concerning the biological context of the segmentation technique performed in this paper, followed by literature containing neurovascular segmentation, which the biological review suggests will be useful. Finally, this literature review concludes with literature concerning deep-learning approaches to medical image segmentation, which will be used at the end of this papers approach and also as a benchmark technique. Ultimately, the goal of this section is to inform the reader of current, related works and build a context for the combined and new approaches explored in this paper.

One approach to cellular image segmentation is using entirely image-based transformational/analytical methods to isolate, detect, and separate cellular features. The ultimate goal is to create a binary mask from an image that contains cellular features, in which the mask contains only the cellular features. Many of these conventional techniques are explained by Sharma et al. in Automated medical image segmentation techniques [3]. These techniques include edge-based, region-based, and textural-based segmentation, along with thresholding algorithms, which include global thresholding, dynamic thresholding, and local thresholding. This work also explains the concept of pipelining techniques, that is, using multiple segmentation techniques or sub-techniques in a row in order to achieve a satisfactory result. This brings the addition of Separating touching and overlapping objects in particle images A combined approach, [5]. This paper outlines an approach for separating cellular structures in binary masks, which uses the changes in pixel density over the image to detect areas where overlap is possible, then removes them using the mask later. I propose the use many of the techniques outlined by Sharma et al. to form my own pipeline to detect cellular features, along with Korath et al.s approach of region overlap removal, as an addition to my pipeline. This paper proposes that this combined pipeline approach will yield satisfactory results in completing the proposed task of segmenting pericyte and smooth muscle cells in neurovascular images.

In terms of designing an automated cellular segmentation algorithm pipeline, it is intuitively useful to understand the biological context surrounding the images that the pipeline is processing. For this reason, this paper explores Establishment and Dysfunction of the Blood-Brain Barrier, by Zhao et al. [11]. The particular areas of significance in this paper surround the diagrams and explanations of pericytes and smooth muscle cells in the context of the blood-brain barrier. The incorporation of the knowledge of the shape and location of these cells in relation to one another is crucial to the design of the automatic segmentation process outlined in this paper. Zhao et al. explain that the smooth muscle cells are located surrounding much larger (by a factor of 2-3) blood vessels than those of pericytes. Knowing this, along with knowing how to isolate the location of blood vessels in the data images, could potentially facilitate the differentiation between the pericytes and smooth muscle cells with quite high accuracy. With vessel segmentation being a potentially important goal in relation to the biological context, it is important that the segmentation of vessels is an achievable goal by analyzing recent literature on the topic. In Blood vessel segmentation algorithms Review of methods, datasets and evaluation metrics, Moccia et al. discuss automated segmentation methods specifically tuned to vessel detection [8]. This paper discusses which methods are effective for which type of imaging, and what some common pipelines for vessel detection are. In addition, not only are these techniques useful for vessel detection, but some may be leveraged for cell segmentation as well. Specifically, Moccia et al. discuss accuracy evaluation metrics, such as positive predictive value, that will be invaluable in evaluating the methods in this paper, both for vessel and cell segmentation. Additionally, one work that implements some of these strategies is Retinal Blood Vessel Segmentation by Means of Scale-Space Analysis and Region Growing, by Martnez-Prez et al. [9]. In this work, analysis techniques incorporating vessel width, size, orientation, and other geometrical features (also included in the paper by Moccia et al.) are used in a scale-space analysis and region growing algorithm that achieves very favorable results in the segmentation of retinal blood vessels. As a result, it is a very feasible strategy, given the biological context, to include an investigation into segmenting the data in this experiment into blood vessel features, and use these blood vessel features to differentiate between cell types for the ultimate cell segmentation goal.

Development of the approach tested in this paper involved exploring current, related works that also tested implementations of medical or cellular image segmentation. One possible approach is to use image-based transformational/analytical methods, as outlined in the previous paragraph. One of these pipeline approaches was used by Zhou et al. to segment breast ultrasound images [4]. This approach involved cropping, gaussian filtering, histogram equalization, pyramid mean-shift filtering, and graph cuts segmentation, and the experiment yielded useful results. Additionally, another purely analytical

method was used by Liu et al. [6] to segment hematopoietic cells from blood smears. This approach introduced an iterative-based threshold algorithm, which used a variant of a simulated annealing algorithm. It also used median-filtering to remove noise and contour detection to isolate the cell shapes. Ultimately, the approach for the experiment outlined in this paper will use a combination ideas from Zhou et al. and Liu et al. to construct a pipeline, particularly median filtering, gaussian filtering, histogram equalization, and another thresholding algorithm variant. In addition, there are clearly uses for deep learning in medical image segmentation. The main drawback to this approach is that medical images lack a large enough ground-truth dataset to train models properly. However, in instances where there exists a large amount of ground-truth data, deep learning can have satisfactory results. In White blood cell classification and counting using convolutional neural network, by Macawile et al. [7], an approach known as transfer learning is used on a moderately sized classified dataset of different types of white blood cells. Macawile et al. used pretrained networks with a few extra training layers to train their data for a short time. In the majority of the pretrained networks tested, accuracy was above 95%. This paper also outlines preprocessing techniques and error metrics that will be useful in my ultimate approach, with or without deep learning. Ultimately, there is a plethora of literature containing implemented automatic medical image segmentation techniques, many of which have very satisfactory results.

As one of the goals for this research is to create training sets for deep learning automatically, it is paramount to explore current literature on deep learning for medical image segmentation. In fact, there are a few cutting-edge convolutional neural network architectures designed specifically for medical image segmentation. For the purposes of this experiment, a well-established architecture for a 3D convolutional neural network and a 2D convolutional neural network will be necessary for potential experimentation. In Invited Talk: U-Net Convolutional Networks for Biomedical Image Segmentation, [2] , Ronneberger discusses the application, structure, implementation, and performance of U-Net, a 2D CNN designed for medical image segmentation. U-Net is designed to work with small datasets and to isolate complex features with low computational time. Additionally, Ronneberger discusses some augmentation techniques for small ground truth sets that can create larger sets of training and test data from small datasets. Furthermore, in DeepMedic for Brain Tumor Segmentation, [10], Kamnitsas et al. discuss the structure, implementation, and performance of DeepMedic, a 3D CNN architecture designed for medical image segmentation. DeepMedic uses a small kernel approach in its 3D convolutional architecture, which allows for relatively low computational time requirements, along with parallel convolutional pathways that allow for the maximum amount of context extraction in the network. Kamnitsas et al. also discuss preprocessing and postprocessing techniques for 3D CNNs in this network, along with the fact that DeepMedic has desirable performance on medical image segmentation benchmarks. The experiments performed in this paper aim to generate training and test data to be fed to these networks. Ultimately, the hope is that the networks will achieve a greater performance than the automated segmentation algorithm can achieve alone, or alternatively that the performance achieved using DeepMedic and U-Net can be used as a benchmark for the automated segmentation techniques. Therefore, these two papers will be vital in understanding and implementing these networks for the experiments conducted for this research.

Thus, the current literature on cell image segmentation and medical image segmentation, both with deep learning and without, have yielded very promising results. In order to get the most out of these techniques, developing a pipeline of methods is a crucial step. In this paper, the biological context of the brain processes being imaged is crucial to the success of the approach. Therefore, the approach proposed in this paper will include a combination of techniques that has not been proposed in the past, particularly for the rare images that are being analyzed. Particularly, the approach in this paper will combine automatic cell image segmentation methods and vascular structure segmentation methods, and use them, respectively, to extract cell masks from images and differentiate between these cells. Then, this technique can be used to train a CNN to achieve even greater (3 dimensional) context learning.

III. Methodology

This section provides an in-depth description of the methodology and procedure used to segment and differentiate pericyte and smooth muscle cells in blood brain barrier neurovascular microscopy scans. It outlines the acquisition of the data, a baseline shallow-learning approach, and a deep neural network approach which uses a convolutional neural network, sliding window approach, as presented in [12].

A. Data

First, labelled data was obtained from Dr. Yongsoo Kim containing 120 512x512 images, along with segmented black and white masks containing separately labelled pericyte and smooth muscle cell features. The images were captured using serial two-photon tomography and are greyscale. They are all manually labelled by a neuroscience expert. 20 images were randomly chosen as test data, and 100 were kept for training data.

B. Baseline

To provide a baseline for the experiment, it was necessary for the sake of result evaluation to develop a baseline segmentation method that used no deep learning. For the sake of simplicity, techniques outlined in [3] were predominantly analyzed. Ultimately, a combination of thresholding and connected component area analysis was used. These techniques are outlined in [3] and [4]. Specifically, Otsu thresholding was used, followed by a connected component analysis removing features below 10 pixels in area and above 250 pixels in

area. These techniques yielded a set of preliminary images segmented to identify both smooth muscle cells and pericytes in the dataset. This segmentation served as a baseline of comparison for the deep methods described shortly.

C. CNN

The deep neural network implemented to identify pericyte and smooth muscle cells in the dataset will be structured in the following ways. The model is written using Python, with the Keras library and the Tensorflow backend. The model has the following layers:

- 2D convolution, kernel size: 3, stride: 2, activation function: relu, input size: 65x65
- 2D max pooling, size (2x2)
- 2D convolution, kernel size: 3, stride: 2, activation function: relu
- 20% dropout
- 2D convolution, kernel size: 3, stride: 2, activation function: relu
- 20% dropout
- Dense layer, activation function: Sigmoid, output: probability that a cell exists in the center pixel of the input image (between 0 and 1)
- Loss function: Binary Crossentropy

The methodology of the model can be deciphered from its structure. The model takes as input an 65x65 greyscale image, in the form of a tensor (numpy array). The model uses dropout to avoid overfitting [2]. The model uses max-pooling to reduce data size but still retain important features from a large initial data size [12]. The final layer is a dense layer with a sigmoid activation function, along with a binary crossentropy loss function, because the model is being trained on a binary decision. The objective of the model is to determine whether or not the central pixel in the image is part of a region of interest in the image, as determined by the ground truth segmented images. This is the approach taken by [12] and implemented with success. In other words, the model determines whether or not the central pixel of the input image is a pericyte or smooth muscle cell. The model is trained using the same set of images, but different sets of masks, and thus can identify pericytes and smooth muscle cells together, and distinguish between them. The idea behind this model is that the small image size provides the model with the capability to process a large amount of data, along with allowing it to process many images quickly. Additionally, the center-pixel identification strategy allows the CNN to learn to use the context around the center pixel to determine if that center pixel contains a cell, just as the human eye would [12]. This model allows the small amount of data to be increased by only training on a small portion of each image, and still allows the CNN to distinguish patterns and features by learning them.

D. Data Preprocessing

The source images are 512x512 greyscale, and the ground truth masks are very sparse in terms of the amount of identified, segmented cells.

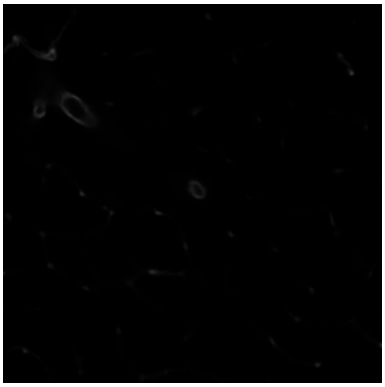

Fig. 1. A raw brain microscopy image scan from the dataset.

Fig. 2. A corresponding ground truth mask for the previous image.

Due to the sparse nature of the data, the data generation and preprocessing must be done in such a way that maximum features can be extracted and learned. Thus, choosing a random 65x65 patch and training on it is not optimal, due to the fact that there is a low probability that this fragment will contain a cell. Thus, an approach was taken akin to [12]. Every pixel in every cell in the ground truth image set was used to generate a 65x65 image for training in the positive category. Subsequently, random sections of the training dataset portion not containing cells in the center pixel were chosen for the negative category such that the positive and negative examples were equal. In instances where the center pixel lies within the window size (65x65) from the edge of the image, the image is mirrored in that area to create the slice, as outlined in [12].

The model was trained for each cell type over 1000 epocs with a batch size of 25 images. The model was evaluated using a custom loss function, which is described in the following subsection. This loss function is designed to maximize cellular feature extraction, and not emphasize exact cellular accuracy.

E. Data Augmentation

In [12], Ciresan et al. had success with random 90 rotations on images when generating training data. The data in this experiment were augmented accordingly. Additionally, [2] added directly to this idea and proposed various additional augmentation techniques in their CNN paper. These augmen-

tations are vital to generating a large amount of training data from a small training dataset. Thus, to generate an even larger amount of potential data, these techniques, including shifting and random elastic deformations were also used to augment the training data.

F. Accuracy Function

A custom loss function was written to evaluate and train the CNN model, and also to evaluate the baseline thresholding performance. This loss function is also used to compare the performance of the various segmentation techniques used to segment the data images into pericytes and smooth muscle cells. The loss function calculates the F-score of a segmentation image and its ground truth mask by calculating the false positive, true positive, and true negative rate as follows. The algorithm first locates the connected pixel areas in both images. Each of these connected pixel areas is considered a cell, with a location. True Positive: Cell in generated segmentation image with at least 1 pixel from this cell touching a cell in the ground truth image. False Positive: Cell in generated segmentation image that does not have at least 1 pixel from this cell touching a cell in the ground truth image. False Negative: Cell in ground truth image that does not have at least 1 pixel from this cell touching a cell in the generated segmentation image. F-score is calculated as usual, by calculating and using precision and recall [7]:

$$F1 = \frac{2 \cdot precision \cdot recall}{precision + recall} \tag{1}$$

The idea of this accuracy/loss function is that, rather than calculating pixelwise accuracy, this function calculates cell-wise accuracy. As long as the segmentation method generates a feature that is overlapping with the desired feature it is locating a cell that is, at the very least, extremely close to the right classification area. This is useful for two reasons: 1) Ultimately, the application of this model is in the counting of cells in certain regions of the brain, and 2) Ultimately, even when trying to locate specific cell locations, it is better to classify all cells very close to their proper location than to classify only some cells with exact pixel-by-pixel accuracy, because exact pixel-by-pixel accuracy is impossible due to variance in expert opinion and, thus, variance in human ground truth. Therefore, a very close approximation is always desirable in cell segmentation, so this loss function was crafted and implemented to evaluate performance based on this fact.

In conclusion, a shallow and deep learning approach was taken to identify cellular features in brain microscopy scans, and a custom F-score cell-wise loss function was used to evaluate the accuracy and usefulness of these features.

REFERENCES

[1] A. Madabhushi and G. Lee, Image analysis and machine learning in digital pathology: Challenges and opportunities, Medical Image Analysis, vol. 33, pp. 170175, 2016.

[2] O. Ronneberger, Invited Talk: U-Net Convolutional Networks for Biomedical Image Segmentation, Informatik aktuell Bildverarbeitung fr die Medizin 2017, pp. 33, 2017.

[3] N. Sharma, A. Ray, K. Shukla, S. Sharma, S. Pradhan, A. Srivastva, and L. Aggarwal, Automated medical image segmentation techniques, Journal of Medical Physics, vol. 35, no. 1, p. 3, 2010.

[4] Z. Zhou, W. Wu, S. Wu, P.-H. Tsui, C.-C. Lin, L. Zhang, and T. Wang, Semi-automatic Breast Ultrasound Image Segmentation Based on Mean Shift and Graph Cuts, Ultrasonic Imaging, vol. 36, no. 4, pp. 256276, 2014.

[5] J. M. Korath, A. Abbas, and J. A. Romagnoli, Separating touching and overlapping objects in particle images A combined approach, p. 6.

[6] B. Liu, C. Yin, Z. Liu, and Y. Zhang, Automatic Segmentation on Cell Image Fusing Gray and Gradient Information, in 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Lyon, France, 2007, pp. 56245627.

[7] M. J. Macawile, V. V. Quinones, A. Ballado, J. D. Cruz, and M. V. Caya, White blood cell classification and counting using convolutional neural network, in 2018 3rd International Conference on Control and Robotics Engineering (ICCRE), Nagoya, 2018, pp. 259263.

[8] S. Moccia, E. De Momi, S. El Hadji, and L. S. Mattos, Blood vessel segmentation algorithms Review of methods, datasets and evaluation metrics, Computer Methods and Programs in Biomedicine, vol. 158, pp. 7191, May 2018.

[9] M. E. Martnez-Prez, A. D. Hughes, A. V. Stanton, S. A. Thom, A. A. Bharath, and K. H. Parker, Retinal Blood Vessel Segmentation by Means of Scale-Space Analysis and Region Growing, in Medical Image Computing and Computer-Assisted Intervention MICCAI99, vol. 1679, C. Taylor and A. Colchester, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 1999, pp. 9097.

[10] K. Kamnitsas et al., DeepMedic for Brain Tumor Segmentation, in Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, vol. 10154, A. Crimi, B. Menze, O. Maier, M. Reyes, S. Winzeck, and H. Handels, Eds. Cham: Springer International Publishing, 2016, pp. 138149.

[11] Z. Zhao, A. R. Nelson, C. Betsholtz, and B. V. Zlokovic, Establishment and Dysfunction of the Blood-Brain Barrier, Cell, vol. 163, no. 5, pp. 10641078, Nov. 2015.

[12] D. C. Ciresan, L. M. Gambardella, and A. Giusti, Deep Neural Networks Segment Neuronal Membranes in Electron Microscopy Images, p. 9.

IEEE conference templates contain guidance text for composing and formatting conference papers. Please ensure that all template text is removed from your conference paper prior to submission to the conference. Failure to remove the template text from your paper may result in your paper not being published.