2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Pixel Distribution Learning for Vessel Segmentation under Multiple Scales

Chenqiu Zhao[1] and Anup Basu[1]

Abstract— In this work we try to address if there is a better way to classify two distributions, rather than using histograms; and answer if we can make a deep learning network learn and classify distributions automatically. These improvements can have wide ranging applications in computer vision and medical image processing. More specifically, we propose a new vessel segmentation method based on pixel distribution learning under multiple scales. In particular, a spatial distribution descriptor named Random Permutation of Spatial Pixels (RPoSP) is derived from vessel images and used as the input to a convolutional neural network for distribution learning. Based on our preliminary experiments we currently believe that a wide network, rather than a deep one, is better for distribution learning. There is only one convolutional layer, one rectified linear layer and one fully connected layer followed by a softmax loss in our network. Furthermore, in order to improve the accuracy of the proposed approach, the RPoSP features are captured at multiple scales and combined together to form the input of the network. Evaluations using standard benchmark datasets demonstrate that the proposed approach achieves promising results compared to the state-of-the-art.

I. INTRODUCTION

Vessel segmentation is a fundamental problem of medical image processing with a wide range of applications, such as oncology [1], ophthalmology [2] and neurosurgery [3]. Previous approaches usually devised an artificial model to analyze the distribution of pixels for vessel segmentation. However, because of the diversity of medical images, which comes from the profiles of different patients, or machines, vessel segmentation is still a challenging problem in computer vision. Vessel segmentation is essentially a binary pixel-wise classification problem. A pixel is classified as a vessel or background based on comparison with its neighborhood. In this work, the distributions of spatial pixels are used for vessel segmentation, and a novel method based on distribution learning is proposed.

In our previous work [4], we demonstrated that a convolutional neural network can be guided to learn a statistical distribution by randomly permutating the temporal pixels. In this work, our previous technique is extended for vessel segmentation, since vessels in an image can be segmented by classifying the distributions of spatial pixels. The pixels are subtracted from their neighborhoods, and the distributions of the subtraction results are input into the network for vessel segmentation, as shown in Fig. I.

*This work was not supported by any organization

[1]Chenqiu Zhao and Anup Basu are with Department of Computing Science, University of Alberta, Canada zhao.chenqiu@ualberta.ca and basu@ualberta.ca

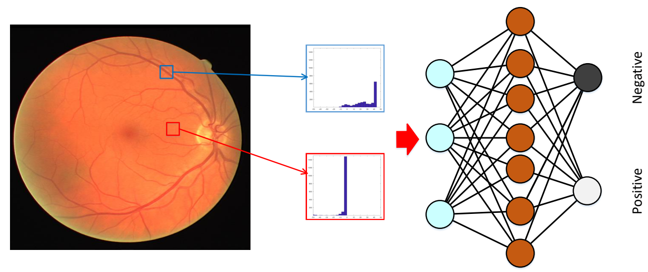

Fig. 1. Pixels Distribution Learning for Vessel Segmentation.

In the proposed approach, a spatial distribution descriptor named the Random Permutation of Spatial Pixels (RPoSP) feature is proposed. In particular, the spatial pixels are randomly permutated to guarantee that only the statistical information is retained. The RPoSP features are dynamically generated as the input to a convolutional neural network (CNN) for every training epoch. It indirectly forces the network to rely solely on the statistics of the distribution of spatial pixels. Moreover, several RPoSP features captured under different scales [5] are combined and used as the input of the network, to improve the accuracy of the proposed approach. The main differences between this paper and our previous work [4] are:

- Random Permutation of Spatial Pixels: In this paper, the distribution of spatial pixels rather than temporal pixels is learned by the network. Compared to the temporal pixels, the variation of spatial pixels includes higher complexity and diversity, which is one of the motivations behind capturing the distribution information under multiple scales.
- Multiple Scales: We capture the distribution information at multiple scales rather than only one scale. This strategy provides the network with better information for learning the distribution and improves the accuracy of the proposed approach.
- Network architecture: We simplify the network architecture considering the computational cost. Our architecture is quite simple; thus, it should probably not be considered “deep.” It only includes one convolutional layer, one rectified linear layer, and one fully connected layer.

The remainder of this paper is organized as follows. The next section describes related work. Spatial pixel distribution learning under multiple scales is discussed in Section 3. Experimental results are outlined in Section 4, before the work is concluded.

978-1-7281-1178-0/21/$31.00 ©2021 IEEE

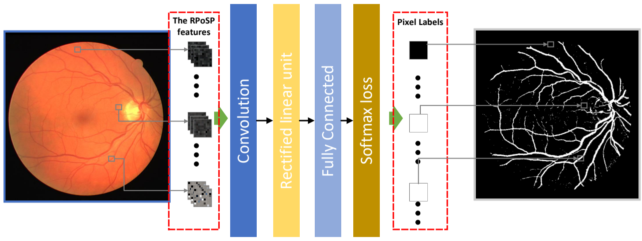

Fig. 2. Flowchart of the proposed approach.

TABLE I
Details of our network architecture, which consists of 4 convolutional layers, 3 batch normalization, 2 max pooling and a softmax operator.

Type	Filters	Layer size	Data size
Input Data			$15 \times 15 \times 15$
Convolution	10024	$15 \times 15 \times 15$	$1 \times 1 \times 10024$
Rectified linear unit			$1 \times 1 \times 10024$
Convolution	2	$1 \times 1 \times 10024$	$1 \times 1 \times 2$
Softmax			

II. Related Work

Medical image processing is an important problem in computer vision [6]. In particular, the segmentation of blood vessels is a challenging problem because of the extreme variations in the morphology of the vessels against a noisy background [7]–[10]. Many approaches have been developed on this topic, including several based on the deep learning networks. However, for brevity, only a few typical methods related to convolutional neural networks and distribution analysis are discussed here.

For vessel segmentation, several approaches based on the analysis of distributions have been proposed. For example, Hassouna et al. [11] utilized stochastic modeling to segment cerebrovascular structure from time-of-flight magnetic resonance angiography, and Evgin et al. [12] used k-means for rough liver vessel segmentation. In addition, Oliveira et al. [13] segmented liver vessels in CT images utilizing region-growing. A pixel was incorporated within a region if its intensity fell within a predefined range, which was defined by approximating the image histogram with a Gaussian Mixture Model (GMM) [14]. However, due to the complexity and diversity of different vessel structures, the Gaussian distribution may not be strong enough to handle all these cases. In contrast, we focus on how to make the network learn the distribution information automatically.

In addition, there are also several other methods related to deep learning networks. For example, Dasgupta et al. [15] formulated the segmentation task as a multi-label inference problem, which combined a convolutional neural network with structured prediction. Fu et al. [16] formulated the vessel segmentation problem as a boundary detection task. In particular, a multi-scale and multi-level convolutional neural network was utilized to learn a rich hierarchical representation. Similarly, Luo et al. [17] combined the prediction ability of CNNs and the segmentation ability of CRFs. Moreover, Vega et al. [18] utilized a Lattice neural network with Dendritic processing, which does not require parameters and can automatically construct its structure.

By contrast, we formulate vessel segmentation as a spatial distribution classification problem. Furthermore, our network architecture is much simpler than the others described in this section.

III. Spatial Pixel Distribution Learning under Multiple Scales

In this section, details of the proposed approach are discussed. The flowchart of our approach is shown in Fig. II, in which the Random Permutation of Spatial Pixel (RPoSP) features are captured first and then input into the network to label if a pixel belongs to a vessel or is part of the background.

For vessel segmentation in medical image processing, it is reasonable to segment vessels based on comparisons between the center pixel and its neighborhood. This is because the intensity of pixels in a vessel is significantly different compared to values in neighboring pixels. We focus on learning the distribution of these comparisons for segmenting vessels, utilizing a convolutional neural network. In addition, motivated by our previous work [4], it is possible to force a network to solely focus on the statistical distribution, by randomly permutating the temporal pixels which are used as the input to the network. However, in this work, the distribution information is derived from spatial pixels instead of temporal pixels. Thus, a new distribution descriptor named Random Permutation of Spatial pixels (RPoSP features), which is an extension of our previous work, is proposed. Since the complexity of the distribution captured from spatial pixels is higher than the one from temporal pixels, a multi-scale strategy is proposed to extract multiple RPoSP features to improve the robustness of the proposed approach. The combination of several RPoSP features captured from a particular pixel under different scales is input into the network for learning the distribution. The network architecture is devised as a classification network to classify pixels into the categories of a vessel or background.

The procedure of extracting RPoSP features under multiple scales is shown in Fig. 3. We introduce the extraction of RPoSP features for one pixel, but the procedure is identical for each pixel. Let us denote a given vessel image as $I(x, y)$, where x and y represent locations of pixels. The patches with the center location of (x, y) under multiple scales are extracted, and the intensity of the center pixel is subtracted from them. In particular, every pixel of patches is subtracted with the center pixel to generate a subtracted patches. Following this, the RPoSP features are captured by randomly permutating entries of subtracted patches at a particular scale. Mathematically, this can be described as follows:

$$
\begin{aligned}
&RPoSP_{x,y}(R_i, R_o) = I(x,y) - I(x + r(m), y + r(n)),\\
&m, n \in [1 \ R_i], \quad r(m), r(n) \in [1 \ R_o]
\end{aligned}
\tag{1}
$$

where $RPoSP_{x,y}(R_i, R_o)$ denotes the RPoSP feature extracted from the pixel located at (x, y). m, n are the indices of an entry in a patch and $r()$ is the random permutation to generate a random position according to the input indices. R_i and R_o are the parameters to control the size of RPoSP features under multiple scales. In particular, R_o is the radius of patches under different scales, and R_i is the radius of the RPoSP features. The reason we use two parameters is that the RPoSP features captured from different scales need to be linked together to input into the network. First, several patches with different radius of R_o are extracted. These patches are extracted from different scales thus have different radius of R_o In order to link them together, we used down-sampling techniques to guarantee than the number of pixels is $R_i \times R_i$. Therefore, these patches from different scales can be reshaped into the same size of $R_i \times R_i$, and can be linked together and input into the network.

The network architecture in the proposed approach is devised for classification, as shown in Table I. The input of the network is a combination of RPoSP features captured at multiple scales; and, the output is the label corresponding to the pixel where these RPoSP features are extracted. Mathematically, the steps above can be shown as follows:

$$
\begin{aligned}
&\ell_{x,y} = D(\mathcal{L}^{\theta}(\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_N)),\\
&\mathcal{F}_n = RPoSP_{x,y}(m, n : R_n^i, R_n^o)
\end{aligned}
\tag{2}
$$

where $\ell_{x,y}$ is a binary label of the pixel at location (x, y) identifying it as a vessel or the background, $\mathcal{L}$ is the learning block and $\mathcal{D}$ is the decision block. θ denotes the parameters of the learning block. The learning block $\mathcal{L}$ consists of convolutional, and rectified linear layers. The decision block $\mathcal{D}$ includes a fully connected layer linked with a Softmax loss.

There are several differences between the proposed approach and our previous work during network training. In our previous work, the data input into the network for training was only generated once. Thus, the input of the network was the same for every training epoch. Under this condition, it is possible that the network overfits the pattern implied in random permutations rather than learn the statistical information included in RPoSP features. In order to address this issue, a dynamic training strategy is proposed as a compensation. In this approach, the entries of RPoSP features are randomly re-permutated by new permutations for every training epoch. This strategy effectively prevents our network from overfitting, and improves the accuracy of the proposed approach for complex scenes.

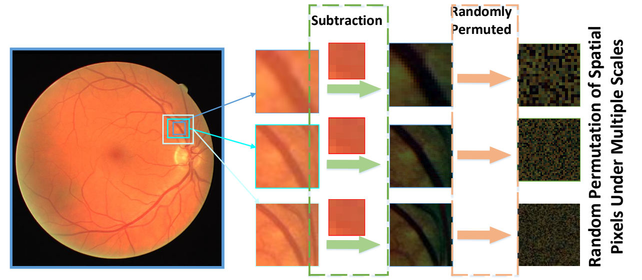

Fig. 3. The extraction of Random Permutation of Spatial Pixel (RPoSP) features under multiple scales.

IV. EXPERIMENTAL RESULTS

In this section, we evaluate the proposed approach. Our approach is compared with several state-of-the-art methods [15]–[20] on the DRIVE [21] dataset. In particular, both of these methods are based on a deep learning network. It should be noted that the training data is important for supervised methods and has a direct contribution in their performance. Methods with more training data are expected to achieve better results. This is especially true when deep learning networks are utilized. In the DRIVE [21] dataset, there are 20 images included in the training set, with another 20 images contained in the testing set. For the methods compared, all the 20 images in the training set are used for training the network. In contrast, since the proposed approach extracts training instances at a pixel-level, many training instances can be captured within one image. Considering this, our network is trained with 15 images, accounting for the limitations on our computational resources.

During the experiments, Acc, Se, Sp, DSC and MCC metrics are used for evaluations. The definitions of these metrics are shown below:

$$
\begin{aligned}
&Acc = \frac{TP + TN}{n}, Se = \frac{TP}{TP + FN}, Sp = \frac{TN}{TN + FP},\\
&DSC = \frac{2TP}{FP + FN + 2TP},\\
&MCC = \frac{(TP \times TN) - (FP \times TN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}
\end{aligned}
$$

where TP and FP are True Positive and False Positive. Here, positive refers to a vessel, while negative represents background. True denotes that the result of this detection is correct, while False means otherwise. Thus, TP means that the result of the detection is a vessel as well as being the ground-truth.

Quantitative comparisons between the proposed approach and state-of-the-art methods are shown in Table II. In addition, due to the length of paper, only the qualitative results of the proposed approach are shown in Fig. 4. As shown in Table II, the proposed approach achieves promising results compared to other state-of-the-art methods. In particular, the proposed approach achieves the highest scores in Se and Acc, which are

TABLE II
Quantitative comparison between the proposed approach and other state-of-the-art methods on the DRIVE [21] dataset.

Methods	Acc	Se	Sp	DSC	MCC
Luo et al. [17]	0.95	0.75	-	-	-
Dasgupta et al. [15]	0.95	0.75	-	-	-
Fu et al. [16]	0.95	0.76	-	-	-
Vega et al. [18]	0.94	0.74	0.96	0.69	0.66
Wang et al. [19]	0.95	0.74	0.98	-	-
Fraz et al. [20]	0.95	0.74	0.98	-	-
Proposed approach	0.95	0.76	0.97	0.78	0.51

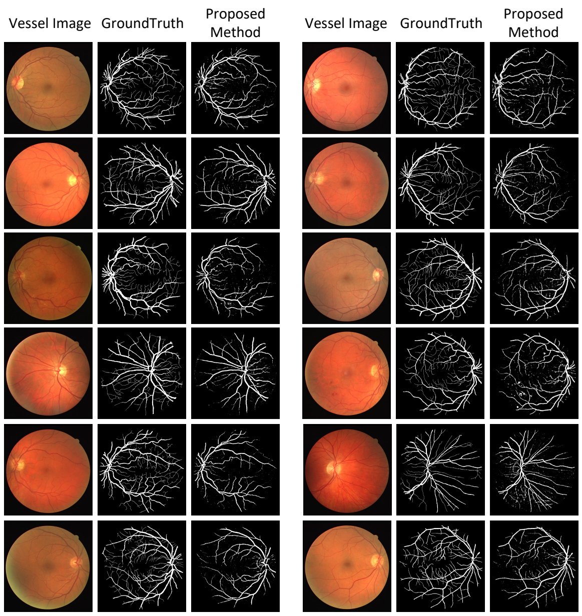

Fig. 4. Qualitative results of the proposed approach on the DRIVE [21] dataset.

considered as the completeness and the accuracy of the vessel mask generated.

There are some disadvantages to the proposed approach. Although it achieves good scores in the Acc metric, which is considered as the accuracy. Since the entries of the RPoSP features are randomly permutated during every training epoch, it is possible that the RPoSP features fed into the network during the testing cases are never shown in the training procedure. In this condition, the network may falsely generate some random noise. In addition, as shown in Fig. 4, the proposed approach work well for the big vessel segmentation but not for the fine structures. This is mainly because the down-sampling procedure which drop information of fine structures.

V. Conclusion

In this paper, we proposed a novel distribution learning approach for vessel segmentation. In particular, a spatial pixel distribution descriptor named Random Permutation of Spatial Pixel (RPoSP) feature was proposed to indirectly force the convolutional network to learn statistical distributions. The spatial pixels contained in the RPoSP features were randomly permutated to guarantee that only the statistical information was retained. Furthermore, several RPoSP features were captured at multiple scales and combined together to input into the network. This improved the robustness and accuracy of the proposed approach. Benefitting from the strong learning ability of deep learning networks, the proposed approach achieved promising results compared to others state-of-the-art methods on a standardized dataset. In the future, we will work on optimization of our code to reduce the computational cost of the proposed approach.

VI. Acknowledgments

The authors thank NSERC, Canada and UAHJIC for financial support of this research.

References

[1] Peter Carmeliet and Rakesh K Jain, "Angiogenesis in cancer and other diseases," nature, vol. 407, no. 6801, pp. 249, 2000.

[2] Peter A Campochiaro, "Molecular pathogenesis of retinal and choroidal vascular diseases," Progress in retinal and eye research, vol. 49, pp. 67–81, 2015.

[3] Sara Moccia, Elena De Momi, Sara El Hadji, and Leonardo S Mattos, "Blood vessel segmentation algorithms—review of methods, datasets and evaluation metrics," Computer methods and programs in biomedicine, vol. 158, pp. 71–91, 2018.

[4] C. Zhao and A. Basu, "Dynamic deep pixel distribution learning for background subtraction," IEEE Transactions on Circuits and Systems for Video Technology, pp. 1–1, 2019.

[5] Rui Shen, Irene Cheng, and Anup Basu, "Qoe-based multi-exposure fusion in hierarchical multivariate gaussian crf," IEEE Transactions on Image Processing, vol. 22, no. 6, pp. 2469–2478, 2012.

[6] Z. Liu, Y. Song, C. Maere, Q. Liu, Y. Zhu, H. Lu, and D. Yuan, "A method for pet-ct lung cancer segmentation based on improved random walk," in ICPR, 2018, pp. 1187–1192.

[7] Amos Sironi, Vincent Lepetit, and Pascal Fua, "Multiscale centerline detection by learning a scale-space distance transform," in CVPR, June 2014.

[8] V. Kaul, A. Yezzi, and Y. Tsai, "Detecting curves with unknown endpoints and arbitrary topology using minimal paths," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 34, no. 10, pp. 1952–1965, Oct 2012.

[9] Z. Shi, H. Xie, J. Zhang, J. Liu, and L. Gu, "Vessel enhancement based on length-constrained hessian information," in ICPR, 2018, pp. 2869–2874.

[10] N. Eladawi and M. Elmogy etc., "Early diagnosis of diabetic retinopathy in octa images based on local analysis of retinal blood vessels and foveal avascular zone," in ICPR, 2018.

[11] M. Sabry Hassouna, A.A. Farag, Stephen Hushek, and Thomas Moriarty, "Cerebrovascular segmentation from tof using stochastic models," Medical Image Analysis, vol. 10, no. 1, pp. 2 – 18, 2006.

[12] Evgin Goceri, Zarine K Shah, and Metin N Gurcan, "Vessel segmentation from abdominal magnetic resonance images: adaptive and reconstructive approach," International journal for numerical methods in biomedical engineering, vol. 33, no. 4, pp. e2811, 2017.

[13] Dário AB Oliveira, Raul Q Feitosa, and Mauro M Correia, "Segmentation of liver, its vessels and lesions from ct images for surgical planning," Biomedical engineering online, vol. 10, no. 1, pp. 30, 2011.

[14] Brian S Everitt, "Finite mixture distributions," Wiley StatsRef: Statistics Reference Online, 2014.

[15] A. Dasgupta and S. Singh, "A fully convolutional neural network based structured prediction approach towards the retinal vessel segmentation," in ISBI, April 2017, pp. 248–251.

[16] Huazhu Fu, Yanwu Xu, Stephen Lin, Damon Wing Kee Wong, and Jiang Liu, "Deepvessel: Retinal vessel segmentation via deep learning and conditional random field," in MICCAI, 2016, pp. 132–139.

[17] Yuansheng Luo, Lu Yang, Ling Wang, and Hong Cheng, "Efficient cnn-crf network for retinal image segmentation," in Cognitive Systems and Signal Processing, 2017, pp. 157–165.
[18] Roberto Vega, Gildardo Sanchez-Ante, Luis E. Falcon-Morales, Humberto Sossa, and Elizabeth Guevara, "Retinal vessel extraction using lattice neural networks with dendritic processing," Computers in Biology and Medicine, vol. 58, pp. 20 – 30, 2015.
[19] Shuangling Wang, Yilong Yin, Guibao Cao, Benzheng Wei, Yuanjie Zheng, and Gongping Yang, "Hierarchical retinal blood vessel segmentation based on feature and ensemble learning," Neurocomputing, vol. 149, pp. 708 – 717, 2015.
[20] Muhammad Moazam et al. Fraz, "An ensemble classification-based approach applied to retinal blood vessel segmentation," IEEE Transactions on Biomedical Engineering, vol. 59, no. 9, pp. 2538–2548, 2012.
[21] J. Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken, "Ridge-based vessel segmentation in color images of the retina," IEEE Transactions on Medical Imaging, vol. 23, no. 4, pp. 501–509, April 2004.