GLOBAL SCIENCE LIBRARY

RESEARCH

Integrative Clinical Medicine & Therapeutics

Optimized Contrast Limited Histogram Equalization and Fully Convolutional Neural Networks for Segmentation of Retinal Blood Vessels

Dewei Hu[1], Han Shen[1] and Jason T. Smith[2*]

[1]Department of Electrical Engineering, Rensselaer Polytechnic Institute, USA

[2]Department of Biomedical Engineering, Rensselaer Polytechnic Institute, USA

Abstract

Retinal imaging for purposes of ophthalmologic diagnosis is conventionally used for ailments ranging from diabetic retinopathy to glaucoma. The regions of main interest include blood vessels, where abnormalities present in the vascular tree are used as major indicators of unwanted anomalies. Thus, the use of techniques in computer vision for the automated segmentation of blood vessels in retinal images has the potential to greatly aid ophthalmologic practitioners in the assessment phase – particularly in early detection of fundal ailments leading to better prognosis for the patients involved. Herein, an approach based around the use of an optimized, dual-pass contrast limited adaptive histogram equalization (CLAHE) pre-processing methodology coupled with a fully-convolutional neural network is presented. The author's results are compared with an unsupervised Gabor Wavelet-based segmentation along with further benchmarking with results recently presented in literature. Notably, the authors submission to the "DRIVE grand challenge", which resulted in an average dice coefficient of ~0.80, ranked within the top 17% of all submissions to date..

Keywords: *retinal vessel segmentation; ophthalmology; supervised; automated*

Introduction

The objective in ordinary ophthalmological vessel segmentation is to return a binary image of size n*x*m*x*1 (n*x*m being equal to the x and y dimensions of the input) for a given retinal image, where each non-zero pixel corresponds to a predicted vessel location. Figure 1 is given to further illustrate.

The DRIVE (Digital Retinal Images for Vessel Extraction) [1] and STARE (Structed Analysis of the Retina) [2] open-sourced datasets are currently the two major retinal image datasets used for benchmarking vessel segmentation results.

Since the inception of AlexNet in 2015 [3], an unprecedented degree of progress in analytic workflow optimization has taken place. The exponential growth regarding novel developments in Convolutional Neural Network (CNNs) -based processing is beginning to aid clinical practitioners in not only diagnosis [4], but multi- modal registration [5], increasing acquisition efficiency [6] and bringing previously speed-bottlenecked applications into clinical relevancy [7].

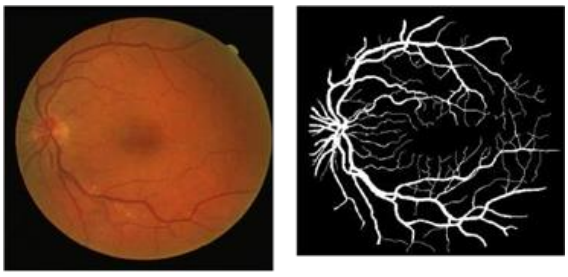

Figure 1. Raw retinal image (a) and corresponding, manually drawn vessel segmentation (b).

Due in large part to the development of the highly adaptable U-Net architecture [8], remarkably successful automated end-to-end segmentation in biomedical imaging has blossomed across

innumerable modalities and applications of interest. Logically it follows that the applications of CNNs for blood vessel segmentation in fundal images would prove worthy of investigation. Thus, a large number of groups across the globe have taken to working on the problem of vessel segmentation in fundal imaging via CNNs. A myriad of unique approaches can be observed, of which the next section will go into further depth.

Related Work

There are lots of work concentrating on the traditional approaches for image segmentation. The motivation of these traditional work is they don't need a lot of data used as training and testing data set and manual labels as ground truth. Also, in terms of processing time, the filtering methods have some advantages compared with machine learning algorithms. The major portion among this kind follow a pattern of enhancement coupled with thresholding. The aim of enhancement is to strengthen the contrast of the vessel with regard of the background to make it more separable. Then, the use of a thresholding method binarizes the image. Main ideas of enhancement share a common sense of utilizing some properties of the tubular structure. For example, the orientation. In [18], Hanung Adi Nurgroho, et al. applied the Gabor wavelet function as a filter to enhance the vessel structure in a specific orientation. Similarly, a matched Gaussian filter can play the same role. By considering the special properties in the neighborhood of the tubular shape, a line operator can also be applied to find the vessel [19].

As deep learning algorithms continuously mature in classification problem performance, entire image plane-image classification, which can also be considered as a classification problem, has steadily evolved in recent years. Given the large success and generalizability of U-Net's segmentation results, groups have taken to using the U-Net architecture directly, or with slight modifications, to perform their ophthalmological vessel segmentation. For instance, in 2017 Gao et al. [9] utilized U-Net for just this purpose, utilizing a pre- processing technique based around the Gaussian Matched Filter in order to enhance vessel contrast with background instead of the raw RGB image data. The group's evaluation on the DRIVE test dataset obtained the highest AUROC (area under the receiver operator characteristic) score out of any technique at the time.

In 2018, Alom, Md et al. utilized a modified U- Net architecture deemed "R2U-Net" (Recurrent Residual Convolutional Neural Network based on U-Net) to further enhance segmentation AUROC scores on the DRIVE dataset [10]. The group extracted 180,000 48x48x3 image patches from the limited number of retinal images available in the DRIVE training dataset (20), along with an extra 75,000 from the STARE dataset. Notably, the group did not utilize any image pre-processing before patch-based augmentation – deciding to work with the raw RGB data directly. The resulting AUROC and dice coefficient (DC) values obtained (0.9784 and 0.8175, respectively) top all other values obtained on the DRIVE test dataset at the time of the work's dissemination.

To recognize one more notable modified U-Net based retinal vessel segmentation work, Jin, Qiangguo et al. developed "DUNet" (Deformable U-Net) [11]. By utilized deformable convolutions [12], or, convolutions that allow for free-form deformation of the sampling procedure during convolution, the group's modified U-Net segmentation obtained dice coefficient results on the test dataset of DRIVE that rank among the top submitted to date (0.8154). The group claims the addition of these deformable convolutions avoided problems encountered by unsupervised techniques that only exploit vessel features through purely linear combinations of response, where "various shapes and scales can be captured via deformable receptive fields" (pg. 5, [12]).

Data Pre-Processing

The given raw data is color images shown in figure 2, as stated before, the pigment makes the vessel have similar color with the background. Thus, for whatever method we want to apply, we need to do some preprocessing so that the feature we want can stand out in some way. Although deep learning provides us a strong tool (CNNs) which can learn the binary classifier if we have enough data, the ill posed data can cause some problem in convergence and low classification accuracy. A straightforward idea is to augment the contrast so that the two classes, vessel and background, have larger distance. So here we applied the Contrast Limited Adaptive Histogram Equalization (CLAHE) method to redistribute the histogram.

This approach can give us a visually satisfying result as it is shown in figure 3. However, the downside is it can simultaneously over amplify the noise as well. Although the contrast limited is an inbuilt mechanism in this algorithm to limit the enhancement of the noise, this effect is still obvious especially when it is applied twice.

In the third image, we can see a lot of grainy noise within the eyeball. The authors tried using these images directly during training the neural network, resulting in a severe sub-optimal convergence of the loss function. But since the color difference is significantly better when we apply CLAHE twice, a median filter was added to filter out the grainy noise. The kernel is set to be 3x3 in order to protect the fragile small vessel branches. The fourth image seems to be more smooth and the convergence problem is solved. In conclusion, the output of the first stage is the smoothen histogram equalization image which is shown as the fourth image in figure 3.

Gabor wavelet approach

The traditional filtering approach can only deal with grayscale images. Thus, the first step is to extract the green channel from the color image. The reason of choosing the green channel is obvious, in figure 4, we can see that the green channel has a stronger contrast.

Figure 4 shows the three channels of the original image, the green channel extracted from the preprocessed image is strongly augmented in figure 5.

Though the result of the first stage is promising, there are still a lot of unimportant features in the image. For example, the optic disk with high intensity and the dark part near the peripheral eyeball. In this case, we cannot apply the thresholding before removing these features.

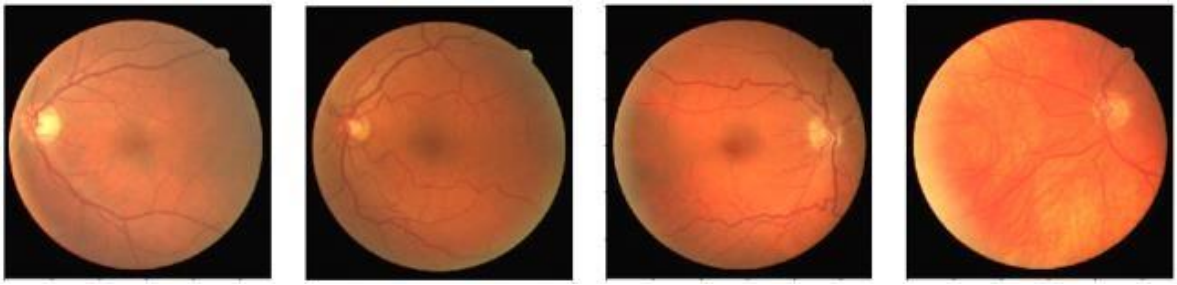

Figure 2. 4 input images are shown here as example

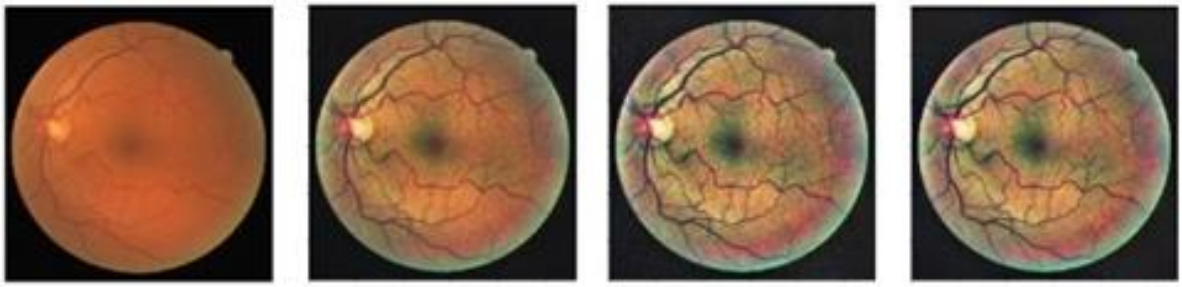

Figure 3. From left to right: 1) raw input, 2) first CLAHE, 3) second CLAHE, 4) after a median filter.

Figure 4. From left to right: 1) red channel, 2) green channel, 3) blue channel

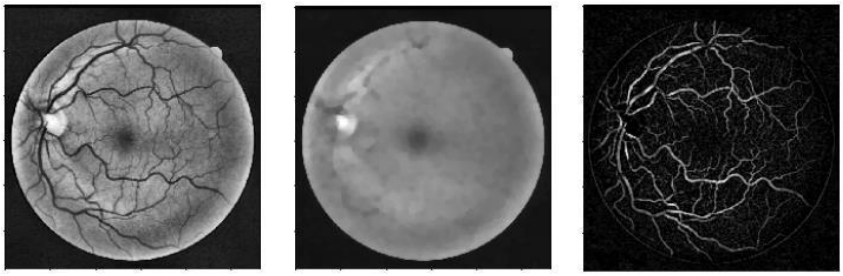

Figure 5. From left to right: 1) green channel, 2) background, 3) result.

Another property of the vessel structure is easy to remove though they are hard to stand out. The method is called top-hat morphological filter which can remove the unimportant feature by using the image subtract the background. The idea is straightforward. By doing an opening operation with a large ellipse kernel (in this case 11x11), all the vessels will be removed since most of the vessel have width smaller than 11 pixels. The result of the opening is shown in the second image of figure 5.

$$I = I_0 - I_0 \circ kernel$$

Then by subtracting this background image, we can get the pure vessel image which is shown in the third image in figure 6. The result is pretty good, but we want to further enhance the vessel so that we can capture more after the thresholding, then here comes the Gabor wavelet filtering.

Since there is no way that we can make sure the orientation of the vessel within specific window, we need to prepare a stack of filters with diverse orientation. The for each pixel, we apply all of them and only retain the maximum value as result (Figure 7). In this project, we used the angular resolution of *15*° which means we have 12 filters to cover all the possible directions. Then we can get a much stronger vessel structure that is ready for thresholding.

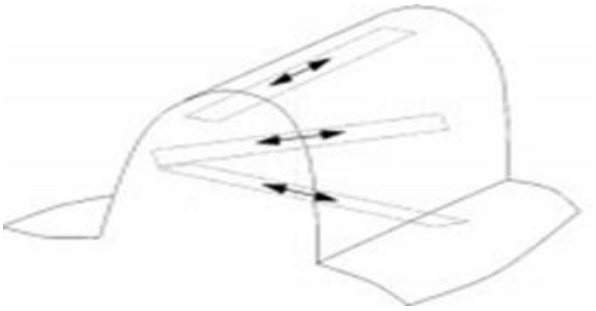

Figure 6. Relationship between filter and vessels

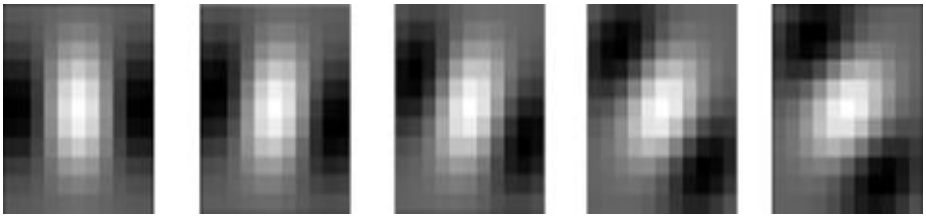

Figure 7. Filters with different orientations.

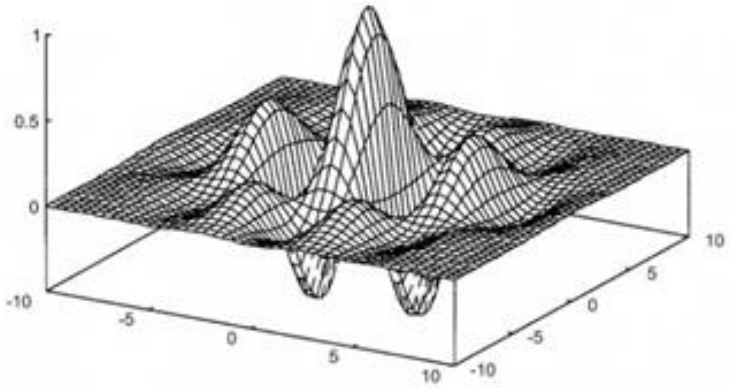

Figure 8. 2D Gabor wavelet.

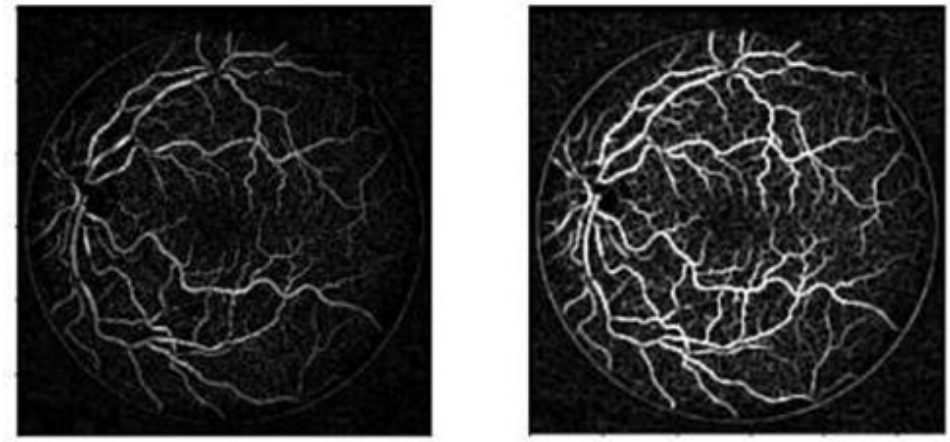

Figure 9. A comparison with the previous result.

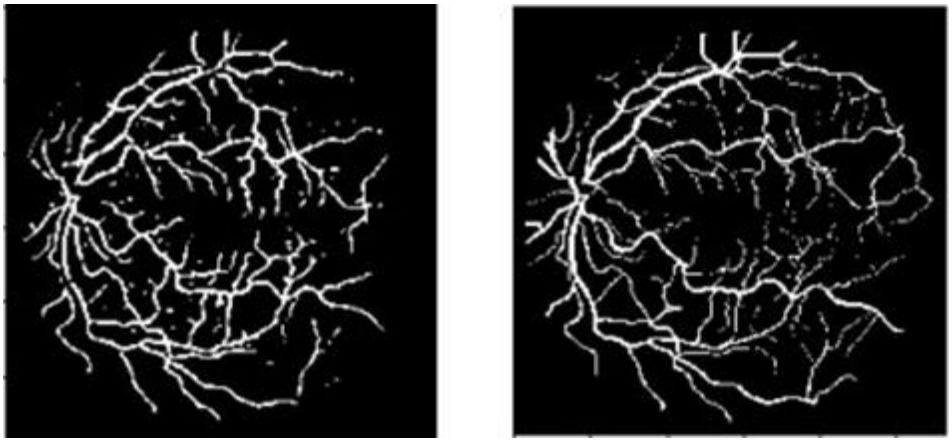

Figure 10. Left: output of the algorithm right: ground truth.

We want to make use of the orientation of the tubular structure. We have a lot of filters have a ripple shape, for example, the Gaussian filter. But the Gaussian filter also has an effect of smoothing which may be harmful to the sharp edges. So the Gabor wavelet is using the Gaussian filter multiplied by a sinusoid function (Figure 8).

$$ker(x, y;\ \lambda, \theta, \varphi, \sigma, \gamma)\ =\ \exp(-\ \frac{x^2+\gamma^2y^2}{2\sigma^2})\ ex(i(2\pi\frac{x}{y}+\ \varphi))$$

When we apply these Gabor wavelet filter, we do the 2D convolution with every pixel. If the window in which the pixel lie in has a vessel that has aligned orientation with the ripple, then the 2D convolution will result in a maximum response. The relationship can be visualized in figure 9.

Otsu's method. The regional thresholding may not work well in this case because there are still lots of grainy noise though a median filter is applied in the preprocessing stage. By comparing the binarized result with the ground truth, we can say by using Gabor wavelet filter and Otsu's method, we captured most of the main branches although some of the thinner vessels are lost (Figure 10).

CNN Approach

Given that the DRIVE challenge provides just 20 images for training, each of which possessing a size of 565x584 spatially, commonly used techniques of image augmentation was performed to scale up the overall data- size available for training. Fifty 28x28x3 patches were extracted at random for each of the twenty available datasets – each of which undergoing a 90° clockwise rotation and a vertical flip. This process resulted in an increase to an overall 3,000 images for use in training and validation of our model. Figure 11 is given for illustration.

Given the nature of the problem, the author's crafted the CNN architecture in a fully-convolutional fashion. The network structure is illustrated in detail in figure 12.

The network was written and trained using the machine-learning library Keras [16] with Tensorflow [17] backend in python. The chosen design was fashioned to reduce overall parameter count while retaining the powerful localization and segmentation capability inherent to U-Net. This was accomplished by reducing the number of maxpooling and deconvolution layers (to just two, respectively) along with reduction of filters used during most convolutional operations.

Immediately following the second deconvolution layer, which projected the spatial dimension back to that of the input, a fully-convolutional projection to 2D was performed (Figure 12, leftmost box). The use of dropout after both convolutions (50% in both cases) was necessary, as overfitting would result almost immediately after the beginning of training otherwise. Given that the ground- truth was binary (no vessel and vessel assigned to zero and one, respectively) a sigmoid activation was used as the final operation to map each spatial prediction between zero and one.

Binary cross-entropy was used as the loss during training. Figure 13 is given to illustrate the loss training and validation behavior whilst using the raw and dual- CLAHE processed data. As one can observe from figure 14, the use of the processed data resulted in a much faster convergence than that obtained during training with the unprocessed image patches. Since the ground-truth is exactly the same in both cases, this can be interpreted as an indication of the network's ability to distinguish vessel from non-vessel more effectively using the processed data. Thus, the overall computational expense needed during training the model with the pre-processed data is decreased significantly by the use of the presented pre-processing methodology.

Because we only have the ground truth for the training data, we do analysis here base on the training data set. From the result shown in figure 15, there are three main drawback of the algorithm. The first one is there are lots of small particles that are false positive due to the noise created in the CLAHE preprocessing stage. We can certainly remove them but not all of them are noise, so there will be trade off since we can hardly verify if it is noise or not. The second one is the missing of thin vessel structure. And the last one is, the difference in width is seriously weakened. We can see that although some of the vessels are detected, the width of it is not exactly the same with the original image. This problem is caused by the enhancement method itself.

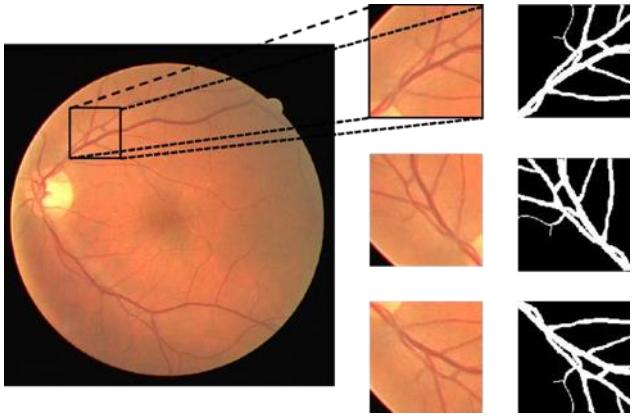

Figure 11. Example unprocessed training image (left) and a randomly extracted patch trio – the centermost and bottom rows obtained via rotation and vertical flip (right).

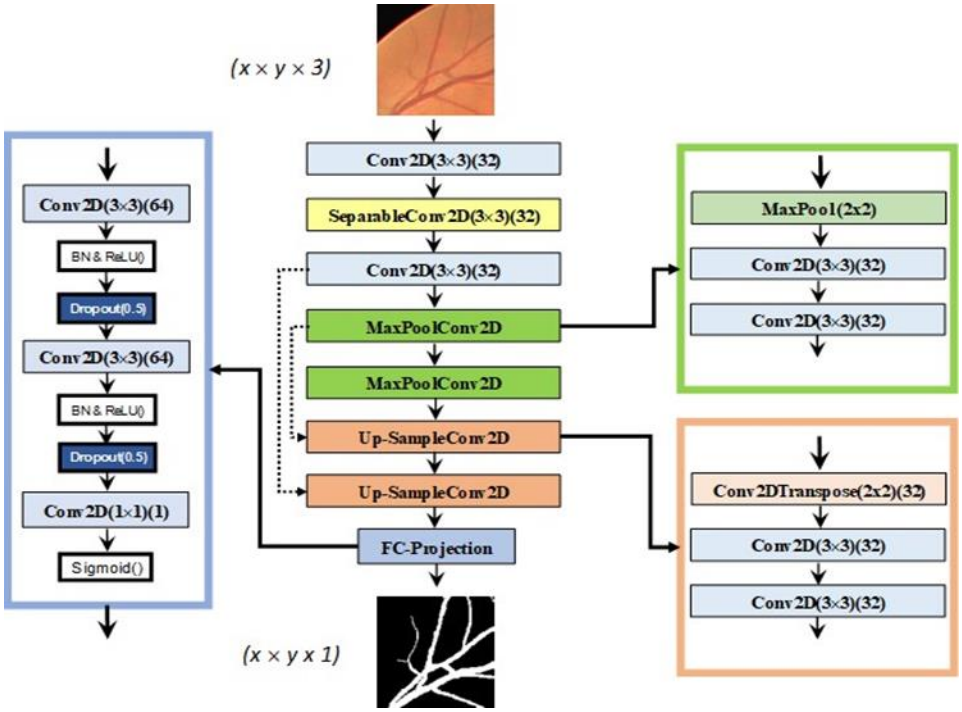

Figure 12. CNN architecture used for the final competition submission.

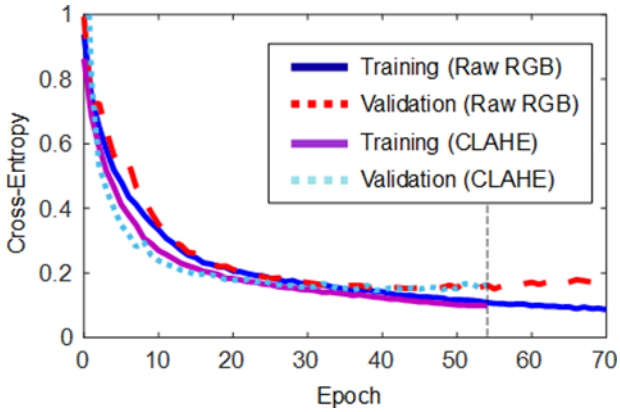

Figure 13. Training and validation loss (cross-entropy) for two training cycles: 1) using the raw RGB retinal image patches and 2) using the dual-CLAHE processed image patches.

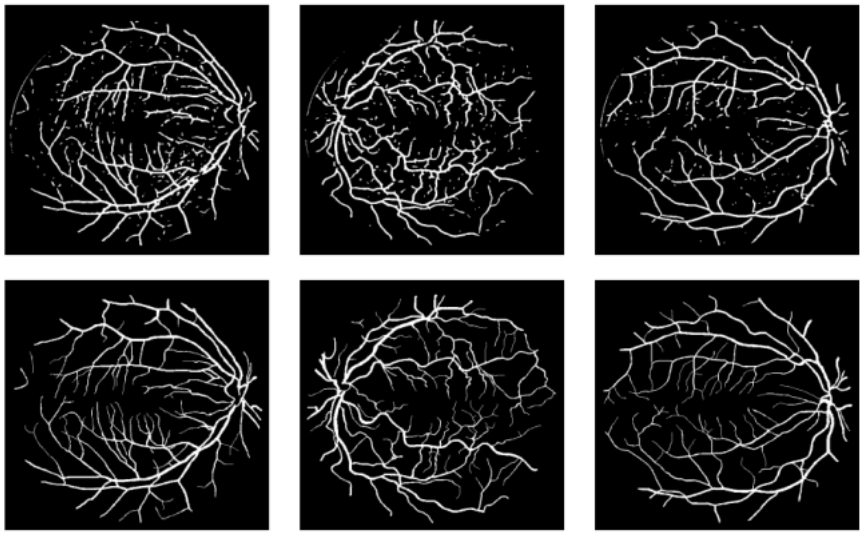

Figure 14. Training data result (first row is the result of algorithm, second row is ground truth).

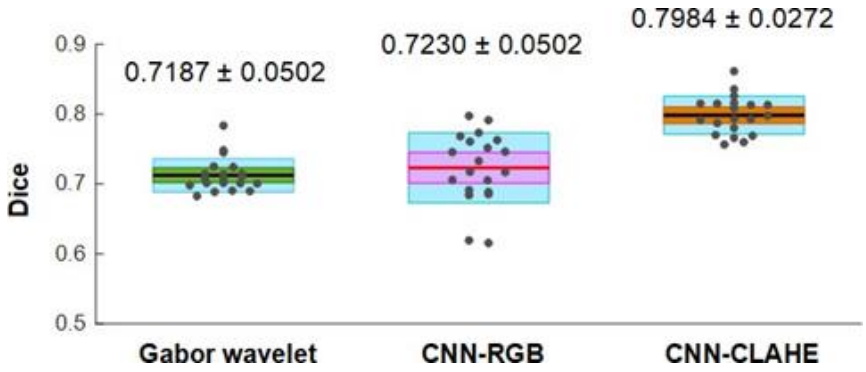

Figure 15. Dice coefficient performance for all algorithms.

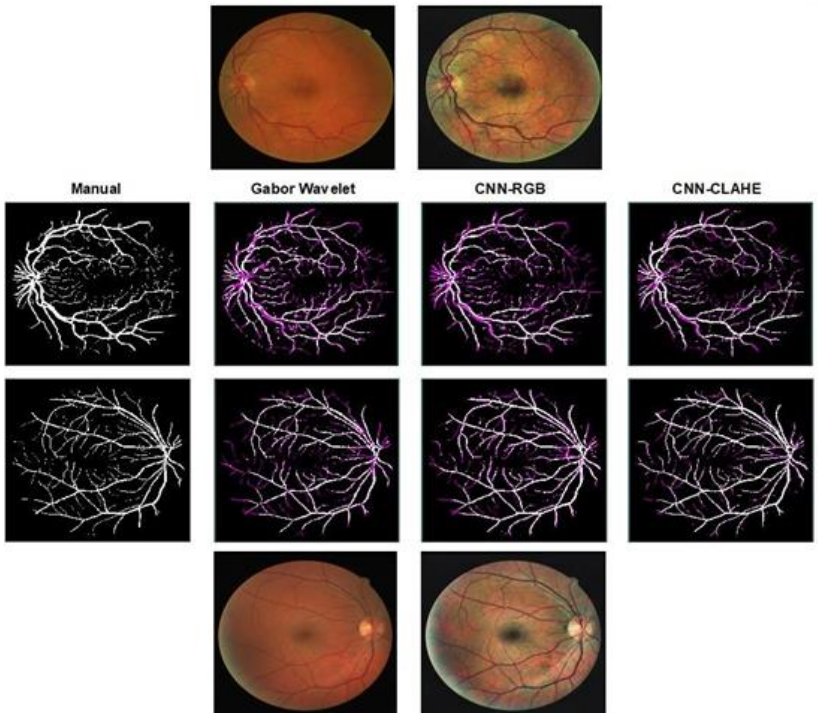

Figure 16. Result comparison

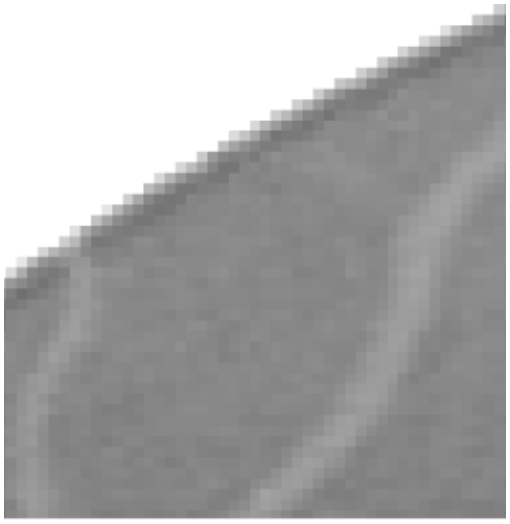

Figure 17. Retinal image border inside FOV.

And for the testing data set, we reached a 0.7187 for dice coefficient evaluation after we submit our result to the contest. And this score stands at the 67th rank.

Result of CNN Algorithm

By the CNN approach, we get a much better performance. We test two circumstances, one is using the raw data as input to train the convolution neural network, another is use preprocessed images to feed the model. The formal one reached 0.7230 in dice coefficient evaluation for the testing data. As we expected, the latter one improved dramatically to 0.7984 and ranked 21st out of 118 groups.

Results

Result of Gabor Wavelet algorithm

This In Figure 16, two segmentation cases are shown. The second row is the case with lowest dice-coefficient and the third row is the case with highest dice-coefficient. For each segmentation result, the white part is the segmented vessel and the purple part is the difference image which is given by subtracting the original image from the segmentation result and taking the absolute value. We can see from left to right, the difference is decreasing and our result is improving.

FOV border. For the DRIVE dataset, sometimes there is a thin line along the border inside the FOV. An example can be viewed in Fig. 16. The thin line has high intensity and resembles a vessel structure. Hence, it is likely to have large response when a gaussian filter is applied. Fig. 8 shows that Gabor fails to distinguish it with a real vessel and therefore it causes false detection after Otsu is applied. The CNN does not have this problem because it is not based on thresholding the response of vessel-like structures. The improvement from CNN-RGB to CNN- CLAHE is mainly focused on the detection of thin and vulnerable vessels. CLAHE proves to be successful in enhancing vessels without causing serious noise. With vessels enhanced, the performance improvement of CNN is natural (Figure 17).

A notable improvement from Gabor to CNN- RGB is that CNN gets rid of false detection around the FOV border. For the DRIVE dataset, sometimes there is a thin line along the border inside the FOV. An example can be viewed in Fig. 16. The thin line has high intensity and resembles a vessel structure. Hence, it is likely to have large response when a gaussian filter is applied. Fig. 8 shows that Gabor fails to distinguish it with a real vessel and therefore it causes false detection after Otsu is applied. The CNN does not have this problem because it is not based on thresholding the response of vessel-like structures. The improvement from CNN-RGB to CNN-CLAHE is mainly focused on the detection of thin and vulnerable vessels. CLAHE proves to be successful in enhancing vessels without causing serious noise. With vessels enhanced, the performance improvement of CNN is natural.

Conclusion

The area of vessel segmentation in computer vision is infamously challenging. Advancing the field by achieving human-level performance and beyond in this task does not just aid the field of ophthalmology and the innumerable patients whose prognosis is significantly improved as a result, but the field of computer vision as a whole given the applicability of top-performing approaches to other problems needing direct attention.

Those behind the collection and dissemination of valuable biomedical images and expert-level manual segmentations for open-sourced competitions such as DRIVE and STARE should be given enormous credit for their effort and generosity. As incremental progression is made in the field by the groups developing ways to achieve gradually increasing DC scores, many invaluable contributions to the field at large are bound to result.

References

1. Staal JJ, Abramoff MD, Niemeijer M, Viergever MA, van Ginneken B. Ridge based vessel segmentation in color images of the retina. *IEEE Transactions on Medical Imaging*. 2004; 23: 501- 509.
2. Hoover A, Kouznetsova V, Goldbaum M. Locating Blood Vessels in Retinal Images by Piece-wise Threhsold Probing of a Matched Filter Response. *IEEE Transactions on Medical Imaging. 2000;* 19: 203-210.

3. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*. 2012: 1-9.
4. Madabhushi A, Lee G. Image analysis and machine learning in digital pathology: Challenges and opportunities. *Med Image Anal*. 2016; 33: 170-175.
5. Haskins G, Kruecker J, Kruger U, Xu S, Pinto PA, et al. Learning deep similarity metric for 3D MR-TRUS registration. *arXiv preprint arXiv. 2018; 1806.04548.*
6. Jin KH, Unser M, Yi KM. Self-Supervised Deep Active Accelerated MRI. *arXiv preprint arXiv. 2019; 1901.04547.*
7. Yao R, Ochoa M, Yan PK, Intes X. Net-FLICS: fast quantitative wide-field fluorescence lifetime imaging with compressed sensing–a deep learning approach. *Light: Science & Applications* 8.1. 2019: 26.
8. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. *International Conference on Medical image computing and computer-assisted intervention*. Springer, Cham, 2015.
9. Gao X, Cai Y, Qiu C, Cui Y. Retinal blood vessel segmentation based on the Gaussian matched filter and U- net. *10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI).* IEEE. 2017.
10. Alom Z, Hasan M, Yakopcic C, Taha TM, Asari VK. Recurrent residual convolutional neural network based on u-net (r2u-net) for medical image segmentation. *arXiv preprint arXi*. 2018: 1802.06955.
11. Jin Q, Meng Z, Pham TD, Chen Q, Wei L, et al. DUNet: A deformable network for retinal vessel segmentation. *arXiv preprint arXiv. 2018:* 1811.01206.
12. Dai J, Qi H, Xiong Y, Li Y, Zhang G, et al. Deformable convolutional networks. *Proceedings of the IEEE international conference on computer vision*. 2017.
13. Kassim YM, Maude RJ, Palaniappan K. Sensitivity of Cross-Trained Deep CNNs for Retinal Vessel Extraction. *40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2018.
14. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv*. 2014: 1409.1556.
15. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
16. F. Chollet. Keras. Ed: 2015.
17. Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, et al. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *arXiv preprint arXiv*. 2016: 1603.04467.
18. Nugroho HA, Lestari T, Aras RA, Ardiyanto I. Segmentation of Retinal Blood Vessels Using Gabor Wavelet and Morphological Reconstruction. *3rd International Conference on Science in Information Technology*. 2017.
19. Chang CC, Lin CC, Pai PY, Chen YC. A Novel Retinal Blood Vessel Segmentation Method Based on Line Operator and Edge Detector. *Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*. 2009.

Correspondence: Jason T. Smith, Department of Biomedical Engineering, Rensselaer Polytechnic Institute, Troy, NY, 12180, USA, E-mail: smithj28@rpi.edu

Rec: May 05, 2020; Acc: May 23, 2020; Pub: May 26, 2020

Int Clin Med Therp. 2020;2(1):11
DOI: 10.36879/ICMT.20.000113