Frangi-Net: A Neural Network Approach to Vessel Segmentation

Weilin Fu[1], Katharina Breininger[1], Roman Schaffert[1], Nishant Ravikumar[1], Tobias Würfl[1], Jim Fujimoto[3], Eric Moult[3], Andreas Maier[1,2]

[1]Pattern Recognition Lab, Department of Computer Science, Friedrich-Alexander-University Erlangen-Nuremberg, Germany
[2]Erlangen Graduate School in Advanced Optical Technologies (SAOT)
[3]Department Electrical Engineering and Computer Science and Research Laboratory of Electronics Massachusetts Institute of Technology
weilin.fu@fau.de

Abstract. In this paper, we reformulate the conventional 2-D Frangi vesselness measure into a pre-weighted neural network ("Frangi-Net"), and illustrate that the Frangi-Net is equivalent to the original Frangi filter. Furthermore, we show that, as a neural network, Frangi-Net is trainable. We evaluate the proposed method on a set of 45 high resolution fundus images. After fine-tuning, we observe both qualitative and quantitative improvements in the segmentation quality compared to the original Frangi measure, with an increase up to 17% in F1 score.

1 Introduction

Fundus imaging can help to diagnose and monitor a number of diseases, such as diabetic retinopathy, glaucoma, age-related macular degeneration [1]. Visual analysis by a trained ophtamologist can be extremely time-consuming and hinder broad clinical application. To support this workflow, automatic segmentation [2] of the retinal vessel tree has been studies for decades, among which a vessel enhancement filter by Frangi et al. [3] is the most popular and forms the basis to various other strategies [4]. However, this task is particularly challenging due to the complex structure of retinal vessels (e.g. branching, crossing) and low image quality (e.g. noise, artifacts, low resolution). In recent years, deep learning techniques have been exploited in the field of retinal vessel segmentation [5,6]. These approaches are data-driven, and tend to share a similar structure, where a classifier follows a feature extractor. For instance, Maji et al. [5] use an auto-encoder as a feature extractor, and random forests as a classifier; Tetteh et al. [6] apply a convolutional neural network (CNN) inception model for feature extraction and another CNN model for classification. These novel, data-driven approaches perform comparably to conventional methods, but do not yield significant improvements.

In this paper, we also investigate deep learning techniques on retinal vessel segmentation. To avoid an explicit separation into "feature extractor" and "classification", we propose an alternative approach, which formulates the Frangi filter

as a pre-weighted network ("Frangi-Net"). The intuitive reasoning behind such an approach is that, by representing the Frangi filter as a pre-weighted neural network, subsequent training of the latter should improve segmentation quality. We aim to utilize prior knowledge of tube segmentation as a basis, and further improve it using a data-driven approach.

2 Materials and Methods

2.1 Frangi Vessel Enhancement Filter

The Frangi vesselness filter [3] is based on the eigen-value analysis of the Hessian matrix in multiple Gaussian scales. Coarse scale structures are typically obtained by smoothing the image with a Gaussian filter g_σ where σ is the standard deviation. The Hessian matrix is a square matrix of second-order partial derivatives of the smoothed image. Therefore, it can be alternatively calculated by convolving the original image patch directly with the 3 2-D kernels, namely $\frac{\partial^2 g_\sigma}{\partial x^2}, \frac{\partial^2 g_\sigma}{\partial y^2}, \frac{\partial^2 g_\sigma}{\partial x \partial y}$, which are the second-order partial derivatives of the Gaussian filter g_σ.

$$G_\sigma = \begin{bmatrix} \frac{\partial^2 g_\sigma}{\partial x^2} & \frac{\partial^2 g_\sigma}{\partial x \partial y} \\ \frac{\partial^2 g_\sigma}{\partial x \partial y} & \frac{\partial^2 g_\sigma}{\partial y^2} \end{bmatrix} \tag{1}$$

The Hessian matrix at each pixel of image f is calculated as,

$$H_\sigma = G_\sigma * f = \begin{bmatrix} H_{xx} & H_{xy} \\ H_{xy} & H_{yy} \end{bmatrix} \tag{2}$$

The two eigenvalues of the Hessian matrix are denoted by λ_1 and λ_2 ($|\lambda_2| \geq |\lambda_1|$), which are calculated as

$$\lambda_{1,2} = \frac{(H_{xx} + H_{yy}) \pm \sqrt{(H_{xx} - H_{yy})^2 + 4H_{xy}^2}}{2} \tag{3}$$

A high vesselness response is obtained if λ_1 and λ_2 satisfy the following conditions: $\|\lambda_1\| \approx 0$, and $\|\lambda_2\| \gg \|\lambda_1\|$. A mathematical description of the vesselness response is presented in [3],

$$V_0(\sigma) = \begin{cases} 0, & \text{for dark tubes if } \lambda_2 < 0, \\ \exp(-\frac{R_B^2}{2\beta^2})(1 - \exp(-\frac{S^2}{2c^2})), & \text{otherwise,} \end{cases} \tag{4}$$

where $S = \sqrt{\lambda_1^2 + \lambda_2^2}$ is the second order structureness, $R_B = \frac{\|\lambda_1\|}{\|\lambda_2\|}$ is the blobness measure, and V_0 stands for the vesselness value. β, c are image-dependent parameters for blobness and structureness terms, and are set to 0.5, 1, respectively.

When using a Gaussian kernel with a standard deviation σ, vessels whose diameters equal to $2\sqrt{2}\sigma$ have the highest vesselness response. The diameter of

retinal vessels in this work range from 6 to 35 pixels. Hence we choose a series of σ as $3, 6, 12$ pixels accordingly. Instead of convolving a patch with three different kernels (i. e., $G_3, G_6,$ and G_{12}), we create a three-level resolution hierarchy [4], and convolve each level with G_3 only. The resolution hierarchy consists of the original patch and two downsampled versions, using factors 2 and 4, respectively.

2.2 Pre-weighted Frangi-Net

Inspired by [7], we implement the multi-scale Frangi filter as a neural network called Frangi-Net on the basis of the previous section. The architecture of Frangi-Net is described in Fig. 1. The proposed method is used to analyze image patches, with an output size set to 128×128 pixels. Input patches for the three sub-nets are cropped as in Fig. 1 (a).

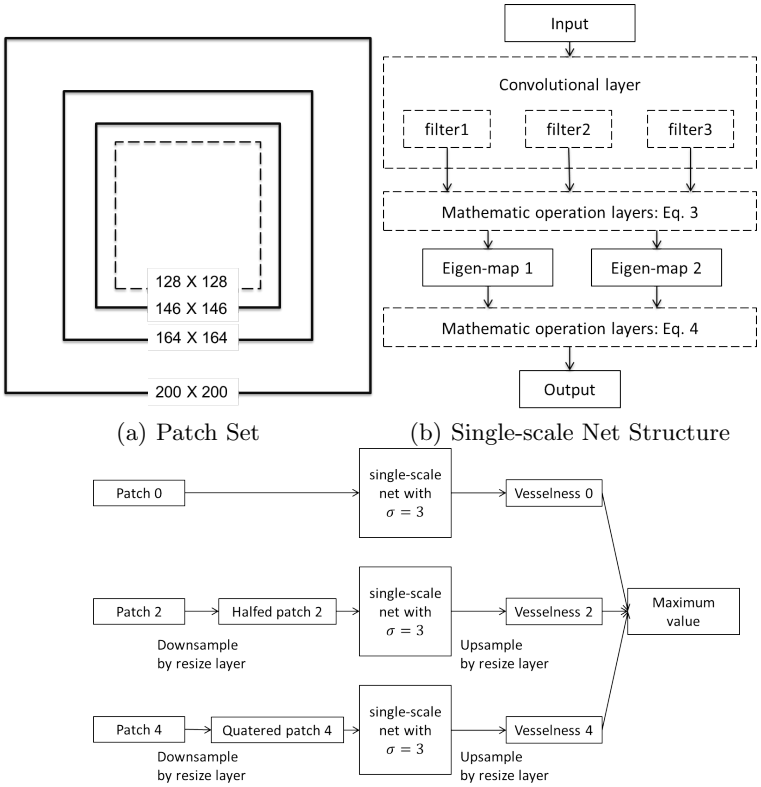
For a single-scale sub-net, we first use a resize layer to downsample images to the corresponding size. Secondly, we use a 2-D convolution layer with three filters to get the Hessian matrix. These filters are initialized as second partial derivatives of the two-dimensional Gaussian function with $\sigma = 3$. After that, a combination of mathematic operation layers are applied, based on Eqs. 3 and 4 to calculate eigen-values and vesselness responses of this scale.

The raw vesselness scores from Frangi-Net typically range from 0 to roughly 0.4. To create a binary vessel mask, a threshold t is used, which is manually set to 10^{-3} in this work. This means that most background pixels are squashed inbetween $[0, t)$, and the rest few vessel pixels spread along $(t, 0.4)$. To obtain a probability map, we subtract the data with the threshold t, and asymmetrically scale it, so that raw scores are more evenly redistributed. Positive values are multiplied with scale s_{pos} and negative values with s_{neg}, where s_{pos}, s_{neg} are set to 2000 and $20,000$ in this work. In this way, raw scores are redistributed to $[-20, 80)$ and after a sigmoid layer, raw vesselness values of $0, t$, and 0.4 are converted to $0, 0.5$, and 1, respectively.

2.3 Trainable Frangi-Net

With the proposed network structure, we observe two sets of trainable parameters in the network: first, the covolution kernels responsible for the computation of the Hessian matrix; second, the parameters β, c, that control the influence of structureness and blobness features. For each scale, this results in three trainable convolution kernels and two additional parameters that can be adapted during training. In Frangi-Net, single-scale nets are structured in parallel and updated independently, which means in total we have nine kernels and six parameters to update during training.

A high resolution fundus (HRF) image database [8] is used in this work. The dataset includes 15 images of each healthy, diabetic retinopathy, and glaucomatous subjects. The HRF images are high resolution 3504×2336 pixel RGB fundus photographs. Only the green channels are used where vascular structures manifest a good contrast [9]. The intensity values are normalized to [0, 1]. We train Frangi-Net on each cartegory independently, randomly dividing from the



(a) Patch Set (b) Single-scale Net Structure

(c) Frangi-Net Overall Structure

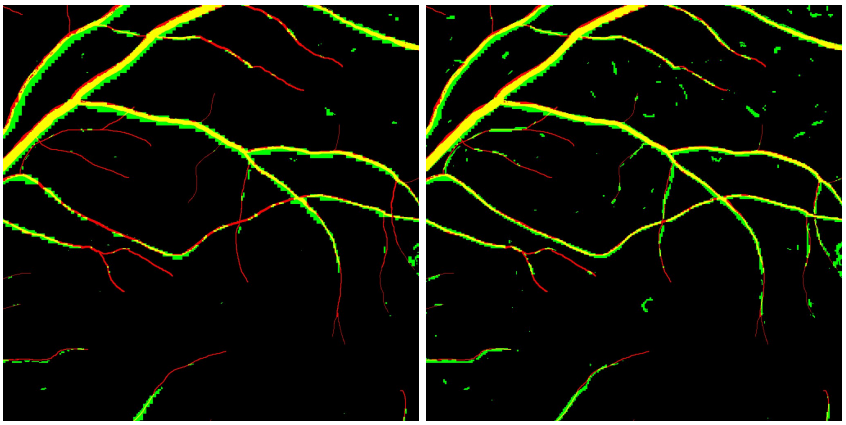
Fig. 1. The Frangi-Net architecture: In (a), dash line represents output patch size 128×128; solid lines represent the patch set in the resolution hierarchy. (b) describes the dataflow through the framework of a single-scale net, where eigen-map 1 and 2 are maps of λ_1, λ_2 respectively; (c) describes the overall structure of Frangi-Net, where patch 0, patch 2, patch 4 correspond to size 146×146, 164×164, 200×200 in (a).

15 images 10 for training, 2 for validation and 3 for testing. We also train the network with all 45 images, randomly taking 30 for training, 6 for validation and 9 for testing. As segmentation of retinal vessels from fundus images is an unbalanced classification problem, where the background pixels outnumber vessel pixels by approximately 10 to 1, dice coefficient loss [10] instead of the common cross entropy loss is used. The optimizer we utilize here is gradient descent with learning rate 10^{-6} and momentum 0.5. 1000 steps are used for training on each category independently, while 3000 steps are used for training on the whole database. Batch size is chosen as 250. We use python and tensorflow framework for the whole implementation.

3 Results

A quantitative comparison of the trained Frangi-Nets with the original Frangi filter is summarized in Table 1. Each quality metric is calculated as an average of that of the testing datasets. After fine-tuning with four different training datasets, the F1 score, accuracy, precision and recall of the segmentation results on the testing datasets all get improved. Segmentation of heathy fundus images performs best, achieving a highest post-training F1 score as 0.712. Comparing the diseased cases to the healthy cases, the F1 score has a more significant improvement after training, up to 17%.

A comparison of the segmentation results before and after training using the healthy datasets is shown in Fig.2. Fig. 2(a) displays that Frangi-Net can segment main vessels well before training. However, it only partially segments the middle-sized vessels and fails to recognize most of the tiny ones. In contrast, after training, Frangi-Net is able to detect the middle-sized vessels well, and it also detects most of the tiny ones (Fig. 2(b)).



(a) Before (b) After

Fig. 2. Vessel segmentation results before and after training. The red, green and yellow colors represent the manual labels, the segmentation results and the overlaps between them, respectively.

4 Discussion

In this work, we proposed to combine the prior knowlegde about retinal vessel that is encoded in the Frangi-Filter with the data-driven capabilities of neural networks. We constructed a net that is equivalent to the multi-scale Frangi filter. We identified the trainable parts of Frangi-Net as convolutional kernels and parameters when computing vesselness. We redistributed the raw output vesselness

Table 1. Segmentation quality before and after training

Dataset	F1 score		accuracy		precision		recall	
	before	after	before	after	before	after	before	after
Healthy	0.669	**0.712**	0.836	**0.843**	0.606	**0.675**	0.746	**0.753**
Diabetic retinopathy	0.495	**0.532**	0.819	**0.822**	0.524	**0.588**	0.468	**0.486**
Glaucomatous	0.495	**0.584**	0.838	**0.841**	0.477	**0.562**	0.608	**0.618**
Whole dataset	0.612	**0.672**	0.847	**0.855**	0.603	**0.675**	0.623	**0.670**

values into a probability map and trained the net by optimizing dice coefficient loss. In experiments with high resolution fundus images of healthy and diseased patients, the trained Frangi-Net performs better than the original formulation both quantitatively and qualitatively.

Future work will investigate on different network architectures for Frangi-Net, or the combination of Frangi-Net with other networks. We will also evaluate Frangi-Net with other datasets.

References

1. Maji D, Santara A, Mitra P, Sheet D. Ensemble of deep convolutional neural networks for learning to detect retinal vessels in fundus images. arXiv. 2016;.
2. Kirbas C, Quek F. A review of vessel extraction techniques and algorithms. ACM Computing Surveys (CSUR). 2004;36(2):81–121.
3. Frangi AF, Niessen WJ, Vincken KL, Viergever MA; Springer. Multiscale vessel enhancement filtering. MICCAI. 1998; p. 130–137.
4. Budai A, Bock R, Maier A, Hornegger J, Michelson G. Robust vessel segmentation in fundus images. Int J Biomed Imaging. 2013;.
5. Maji D, Santara A, Ghosh S, Sheet D, Mitra P; IEEE. Deep neural network and random forest hybrid architecture for learning to detect retinal vessels in fundus images. EMBC. 2015; p. 3029–3032.
6. Tetteh G, Rempfler M, Zimmer C, Menze BH; Springer. Deep-FExt: Deep Feature Extraction for Vessel Segmentation and Centerline Prediction. International Workshop on Machine Learning in Medical Imaging. 2017; p. 344–352.
7. Würfl T, Ghesu FC, Christlein V, Maier A; Springer. Deep learning computed tomography. Med Image Comput Comput Assist Interv. 2016; p. 432–440.
8. Budai A, Odstrcilik J. High Resolution Fundus Image Database; 2013.
9. Yin B, Li H, Sheng B, Hou X, Chen Y, Wu W, et al. Vessel extraction from non-fluorescein fundus images using orientation-aware detector. Med Image Anal. 2015;26(1):232–242.
10. Taha AA, Hanbury A. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. BMC Med Imaging. 2015;15(1):29.