DEEP LEARNING BASED RETINAL VESSEL SEGMENTATION: A REVIEW

Article · July 2020

DOI: 10.37418/amsj.9.6.62

3 authors:

Navdeep Singh
Punjabi University, Patiala

Danish Bansal
Punjabi University, Patiala

Dheeraj Nagpal
Punjabi University, Patiala

ADV MATH SCI JOURNAL

Advances in Mathematics: Scientific Journal **9** (2020), no.6, 3827–3837
ISSN: 1857-8365 (printed); 1857-8438 (electronic)
https://doi.org/10.37418/amsj.9.6.62 Spec Issiue on ICAML-2020

DEEP LEARNING BASED RETINAL VESSEL SEGMENTATION: A REVIEW

N. SINGH[1], D. BANSAL, AND D. NAGPAL

ABSTRACT. Changes in the blood vascular structure are the first indication that a person is affected by some disease. Various diseases are diagnosed by the structure of retinal blood vessels. Manual extraction is challenging; therefore, the automation of this process is significant. In the review, various supervised blood vessel segmentation techniques along with classifiers and frameworks have been reviewed besides performance measures employed, benefits and limitations. Proposed techniques are measured on the DRIVE, CHASE and STARE datasets. The primary intent of this study is to present a latest reference of information on the state-of-the-art methods on blood vessel segmentation algorithms using deep learning so that suitable methods can be chosen according to the task.

1. INTRODUCTION

Several biological features of retinal such as veins and arteries have their diagnostic significance and can be used in monitoring the disease progression. Ophthalmologic diseases, including glaucoma, diabetic retinopathy, age-related fovea degeneration, hypertension, are the leading causes of blindness. Changes in the structure of retinal blood vessel are the first indication of above mentioned as well as other cardiovascular diseases. Analysis of these modifications can assist the doctor in calling for early action to halt the advancement of the

[1]*corresponding author*

2010 *Mathematics Subject Classification.* 92B20, 68T05.
Key words and phrases. retinal vessel segmentation, deep learning, machine learning, convolutional neural network, classification models.

disease. Automation of this operation not only reduce the time required for manual extraction of the blood vessels but also lessen the price connected with trained clinicians and remove the element of human error associated with manual scoring.

A comprehensive review of various retinal vessel segmentation techniques has been presented in [1], [12], but this review paper aims at most innovative and most recent deep learningbased vessel segmentation methods found in the literature. Vessel segmentation using deep learning has been discussed in section 2, advantages and limitations of the existing techniques has been presented in section 3, various publicly available datasets have been discussed in section 4. Section 5 contains results followed by conclusion in section 6.

2. Vessel Segmentation

Extracting vessels from a retinal image is a very time-consuming process and can only be done by trained professionals under the supervision of the ophthalmologists. Automatic extraction techniques not only reduce the segmentation time but can also be applied by any naïve person. Before using the segmentation technique, to increase the quality of the image, image pre-processing is done which makes it suitable for accurate extraction of the blood-vessels from the retinal fundus image. Multiple features are obtained from patches of the image to create featurevectors, which are used to train the model or the classifier, which ultimately classifies pixels into one of the follwing classes: vessel or non-vessel pixels. The typical vessel segmentation process is shown in Figure 1.

2.1. **Pre-Processing.** First of all, masking is done. Masking means setting boundry pixel values in an image to zero to separate image from background which gives the Reagon of Interest(ROI). Then the images are preprocessed with various techniques such as histogram equalization [14], or Gaussian filtering to lower noise, increase the contrast of vessels [13], and make the intensity of the image more uniform. Multiple other preprocessing techniques like [17], can also be used to enhance the image without losing image features.

2.2. **Patch Extraction.** Image patches are the sub-regions of an image that form the basis of various approaches under computer vision and have broad applicability ranging from object detection to texture classification. In patch extraction,

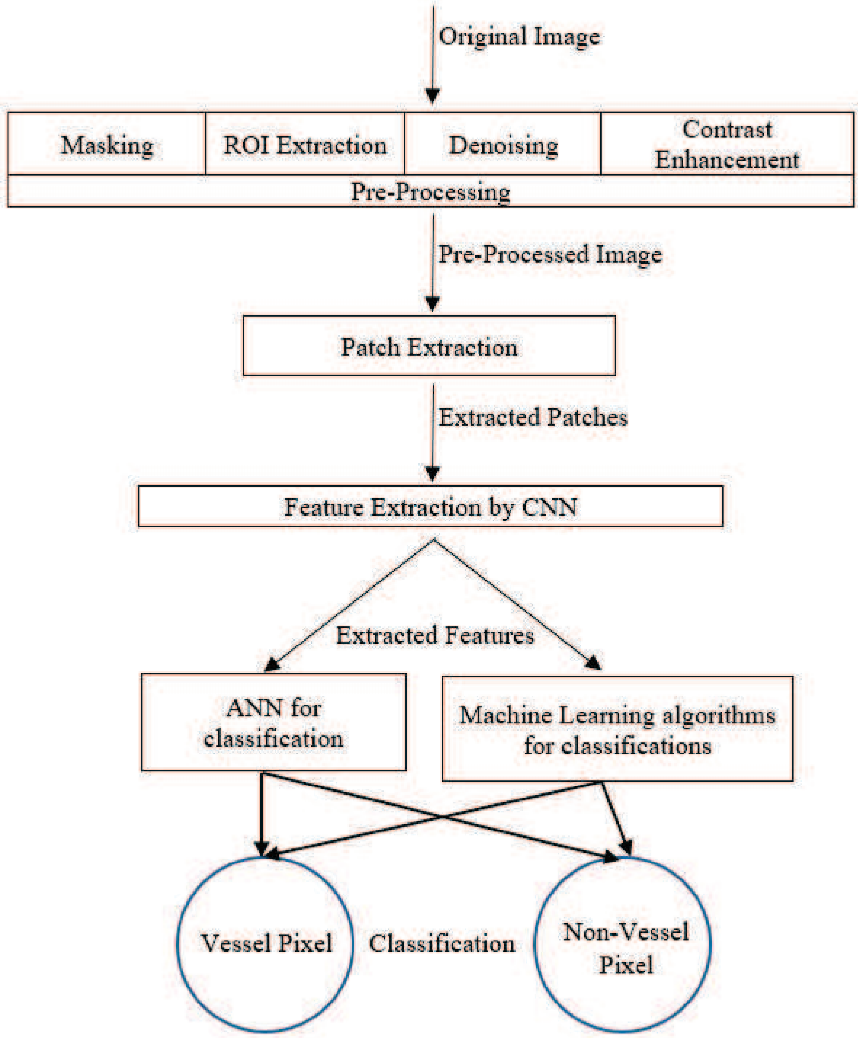

FIGURE 1. Steps of segmentation of blodd-vessels from the original image

a large number of overlapping patches of size $p \times p$ are chosen randomly which are obtained from various images of various datasets. In patch extraction, image is first divided into small patches or blocks, and then a referenced/centered pixel is manipulated based on its neighbouring pixels located in the patch.

2.3. **Feature Extraction.** Each extracted patch passes through various convolution layers, and various feature vectors are obtained from each patch. The neurons of convolutional neural network can detect the end points and edges and each layer in CNN then combine these features to capture higher-order features. Then the features extracted from these layers of CNN used to train the deep learning model.

2.4. **Classification.** After acquiring higher-order features from various convolutions, these feature vectors are passed to different machine learning algorithms such as KNN[1], SVM[2] or RF[3] or are directly given to an artificial neural network by flattening the feature vector. The model trained is able to classify the patches as one of the categories: pixel vessel or non-pixel vessel.

3. TECHNIQUES

In medical imaging, neural networks(NNs) has a magnificent record of applications.

In contrast, the transfer functions used in deep convolutional neural networks (CNN), allows us practical training of neural networks with dozens of layers. There are two ways of using Convolutional Neural Networks (CNN) for segmentation of retinal vessels as shown in Figure 2.

(1) Image features extracted using CNN [6] and classification using fully connected layers, [16].
(2) Image features extracted using CNN and classification using various machine learning algorithms like KNN[1], SVM[2] and RF[3], [2].

Many research works have used machine learning algorithms to classify a pixel as vessel pixel or nonvessel pixel, and whereas others have used artificial neural networks(ANN).

Wang et al. in [2] proposed a technique that combines two classifiers; CNN and RF for segmentation of the retinal vessels. For the pre-processing part, initially the green channel was extracted from the original image, and then the image was denoised followed by contrast enhancement using Histogram equalization and Gaussian filtering. Super-pixel-based sample selection, [2],

[1]K-neighbour classifier
[2]Support vector machine
[3]Random-forest

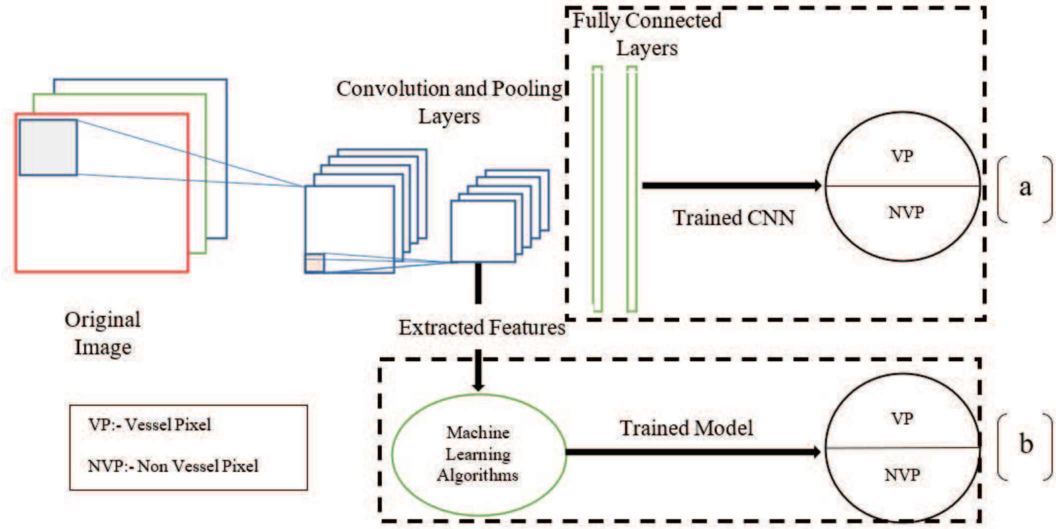

FIGURE 2. Pictorial representation of features extracted using Convolutional Neural Networks being used in (a) ANN (b) ML model

method was used to capture redundancy in the data. It also enforces local consistency while maintaining original image boundaries. The classic LeNet-5 network , [15], with 4×4 convolution matrix and max-pooling by factor 2, was adopted for feature extraction. Extracted features from max-pool and convolution layers were fed to the Random Forest Algorithm. Three predictions were made RF1 – Prediction by feeding features of 2nd max-pooling layer into the random forest, RF2 – Prediction by feeding features of 4th max-pooling layer into the random forest, and RF3 – Prediction by feeding features of 5th Convolutional layer features into the random forest. The winner takes all approach was used that combined all three predictions to get the final prediction. By combining the qualities of feature learning, traditional classifier and ensemble classification, the method learnt better features from raw images. The method was trained over 100 epochs and used default parameters for Random Forest. For the training purpose, performance evaluation was done on DRIVE and STARE data sets. For sampling on raw image, square windows for patch extraction are directly used and for input, original image intensities are used.

P. Liskowski et al. in [3] proposed a technique which is based on a deep neural network trained over a large sample of examples, i.e. 400000 patches.

Zero-phase whitening, [3], was applied to whiten the pixels of the patch and the patches were pre-processed with global contrast normalization to extract better features for learning. Image augmentation with scaling by a factor between 0.7 and 1.2, rotation by an angle from $[-90°, 90°]$, horizontal and vertical flipping and Gamma correction were applied to prevent over-fitting. In the proposed method,initially the learning rate was set to 10^{-3} and then lowered by a factor of 10 every 10000 iterations. The training stopped after 19 epochs. The model was trained using all the training images of the DRIVE data-set and randomly chosen 19 images of the STARE data-set following a leave one out strategy. Remaining images are used for testing. This method is not efficient for thin blood vessels.

Song Guo et al. in [4] presented a technique which used a deep learning network with short connection (BTS-DSN) for the segmentation of the blood vessels. In this deep neural network low level semantic information is passed to high level using bottom-top-short connection and all the structure information is passed to low level by top-bottom-short connection. This technique reduce the noise in low-level side outputs. For the augmentation part, scaling, flipping and rotation were used. No pre-processing was implied on the DRIVE dataset whereas for STARE dataset green channel was extracted. Two models with the backbone; VGGNet and ResNet were used. Song Guo et al. employed cross-training to show that this model is robust. For the proposed experiments, DRIVE's training dataset was partitioned into 2 sets of 15 images and 5 images for training and validation, respectively. The images of the DRIVE's test dataset were used for testing. While, for the STARE dataset, initially, 20 images were divided into two groups of 10 each. Further, the first 10 images were partitioned into 7 and 3 images for training and validation purpose respectively, and the other 10 were used for testing. For the CHASE initially, 28 images were divided into two groups of 20 and 8. Further, the 20 images were partitioned into 15 and 5 images for training and validation purpose, respectively. Testing was done on the remaining 8 images. By separating the training directories for training and validation purposes, we got an advantage of testing more images. The semantic gap between side output layers is reduced due to usage of BTS, when the backbone is VGG-Net and ResNet-101.

Juntang Zhuang in [5] recommended a ladder net framework that is a chain of recurring U-nets. This ladder-net has numerous pairs of encoder-decoder branches. Ladder-net has more paths for information flow because of skip connections and residual blocks. It is observed as a combination of fully convolutional networks. Convolution, with a stride of 2 was employed. The model in the proposed experiment was trained over 250 epochs, and learning rate was initially set to 10^{-2}, reduced to 10^{-3} on 20$^{\text{th}}$ epoch and finally 10^{-4} on 150$^{\text{th}}$ epoch. Juntang randomly sampled 19000 and 76000 patches of shape 48×48 each from CHASE (from first 20 images out of 28) and DRIVE (training directory) respectively out of which 90% were used for training and 10% for validation respectively. Other 8 images from CHASE and DRIVE's test directory were used for testing purposes. This method has superior performance over the earlier ones.

Americo et al. in [7] proposed a method in which a fully connected convolution neural network and the stationary wavelet transform is combined to give the multiscale analysis. This method copes with varying directions and width of the retinalvessel structure in the retina. Data augmentation was done by performing the rotation of patches at $90°$, $180°$ and $270°$. There are four main stages in this technique. To enrich the input of the FCN, input building was done through the stationary wavelet transforms by adding new channels to input. In the second stage, overlapping patches were extracted. Classification using fully convolution networks with 14-layer architecture containing convolutional layers, max-pooling layers, concatenation layers for skip connections, upsampling layers, cropping layers was done in the third stage. In the last stage, multiple predictions were made. In the experiment, 20 epochs were used to train the network with learning rates 0.05, 0.02, 0.002, 0.0002 on 1st, 10$^{\text{th}}$,14$^{\text{th}}$ and 20$^{\text{th}}$ epoch, respectively and learning rate decay of 10^{-6}. All 20 training images of the DRIVE dataset were used for training the model. For CHASE and STARE, k-fold cross-validation was used, where k was chosen to be 5 and 4 for STARE and CHASE datasets respectively.

Zhengyuna Liu in [8] proposed two methods using fully convolutional neural networks; U-net and Ladder-net. These methods began with the reformation of the original image to the greyscale image. The grayscale image was then standardized with Z-score standardization which centralized the values of pixels

TABLE 1. Datasets

Dataset	Total Images	Image Size	Refrence
DRIVE	40	584×565	[11]
STARE	20	700×605	[10]
CHASE	28	999×960	[9]

and speeded up the training process. For contrast-enhancement, CLAHE[4] and Gamma adjustment were applied. The pre-processed images were then rotated at angles of $90°$, $180°$ and $270°$. Then 160000, 128000, 168000 patches of 48×48 are extracted from DRIVE, STARE and CHASE datasets, out of which 90% were used for training, 10% for validation and data augmentation.
So, in this section, various deep learning techniques have been discussed to segment retinal blood-vessels from the retinal images. Each method presents a unique way to extract the vessels with its own advantages and limitations. These advantages and limitations can be taken into account while designing a novel deep learning framework for the segmentation of the retinal blood-vessels from retinal fundus images.

4. DATASETS

The proposed methods have been evaluated on three well established public databases: Digital Retinal Images for Vessel Extraction (DRIVE), Child Heart and Health Study in England (CHASE DB1) and Structured Analysis of the Retina (STARE). Table 1 gives us the required information about datasets.

5. EVALUATION RESULTS

As discussed earlier, performance evaluation of various techniques has been done on three datasets; DRIVE, STARE and CHASE. Table 2 shows the results based on various performance measurements on these datasets. The measurements used are Acc(Accuracy),Se(Sensitivity), Sp(Specificity), AUC(Area Under Curve) and F1 Score.

[4]Contrast Limited Adaptive Histogram Equalization
[5]Convolutional Neural Network

TABLE 2. Comparative Analysis of various deep learning techniques

Method	Classifier/ Frame-work	Datasets used	Acc	Se	Sp	AUC	F1 Score
Wang et al., [2]	Random Forest	DRIVE	0.9767	0.8173	0.9733	0.9475	N.A.
		STARE	0.9813	0.8104	0.9791	0.9751	N.A.
Liskowski et al., [3]	CNN[5]	DRIVE	0.9507	0.8460	0.9673	0.9787	N.A.
		STARE	0.9667	0.9289	0.9710	0.9930	N.A.
		CHASE	0.9610	0.7544	0.9846	0.9801	N.A.
Song Guo et al., [4]	BTS-DSN	DRIVE	0.9561	0.7891	0.9804	0.9806	0.8249
		STARE	0.9674	0.8212	0.9843	0.9859	0.8421
		CHASE	0.9627	0.7888	0.9801	0.9840	0.7983
Juntang Zhuang, [5]	Ladder Net	DRIVE	0.9561	0.7856	0.9810	0.9793	0.8202
		CHASE	0.9656	0.7978	0.9818	0.9839	0.8031
Americo et al., [7]	CNN[5]	DRIVE	0.9639	0.8405	0.9814	N.A.	N.A.
		STARE	0.9365	0.6329	0.9924	N.A.	N.A.
		CHASE	0.9600	0.7731	0.9813	N.A.	N.A.
Zhengyuan Liu, [8]	U-Net	DRIVE	0.9559	0.7728	0.9826	0.9794	0.8169
		STARE	0.9638	0.7139	0.9867	0.9846	0.8219
		CHASE	0.9600	0.7240	0.986	0.9784	0.7815
	Ladder Net	DRIVE	0.9566	0.7728	0.9813	0.9805	0.8219
		STARE	0.9529	0.7513	0.9764	0.966	0.7694
		CHASE	0.9656	0.7978	0.9812	0.9839	0.8031

6. CONCLUSION

A comprehensive analysis of various deep learning-based retinal blood vessel segmentation techniques has been provided in this paper. All the frameworks that have been used along with the number of patches, patch sizes, number of epochs for which the models have been trained are discussed in detail along with the advantages and limitations of each and every technique so as to provide the reader a ready reference to various novel ideas. The results of various techniques have been summarized, and different datasets and metrics have also been discussed in detail. Talking about the future scope of the retinal vessel segmentation, performance is the key factor for any technology. This research is focused mainly on enhancing the performance of retinal vessel segmentation process i.e. increase the accuracy of our segmentation process and decrease the computational load and power needed to compute the required results.

REFERENCES

[1] N. SINGH, L. KAUR, K. SINGH: *Segmentation of retinal blood vessels based on feature-oriented dictionary learning and sparse coding using ensemble classification approach*, J. Med. Imag., **6**(4) (2019), 044006.

[2] S. WANG, Y. YIN, G. CAO, B. WEI, Y. ZHENG, G. YANG: *Hierarchical retinal blood vessel segmentation based on feature and ensemble learning*, Neurocomputing, **149** (2015), 708–717.

[3] P. LISKOWSKI, K. KRAWIEC: *Segmenting retinal blood vessels with deep neural networks*, IEEE Trans. Med. Imaging, **35**(11) (2016), 2369–2380.

[4] S. GUO, K. WANG, H. KANG, Y. ZHANG, Y. GAO, T. LI: *BTS-DSN: Deeply supervised neural network with short connections for retinal vessel segmentation*, International Journal of Medical Informatics, **126**(11) (2019), 105–113.

[5] J. ZHUANG: *LadderNet: Multi-path networks based on U-Net for medical image segmentation*, arXiv:1810.07810, (2018).

[6] Q. JIN, Z. MENG, T. D. PHAM, Q. CHEN, L. WEI, R. SU: *DUNet: A deformable network for retinal vessel segmentation*, Knowledge-Based Systems, **178**(2019), 149–162.

[7] A. OLIVEIRA, S. PEREIRA, C. A. SILVA: *Retinal vessel segmentation based on Fully Convolutional Neural Networks*, Applications, **112**(1) (2018), 149–162.

[8] Z. LIU: *Retinal Vessel Segmentation based on Fully Convolutional Networks*, arXiv abs/1911.09915v1, (2019).

[9] C. G. OWEN, A. R. RUDNICKA, R. MULLEN, S. A. BARMAN, D. MONEKOSSO, P.H. WHINCUP, J. NG, C. PETERSON: *Measuring Retinal Vessel Tortuosity in 10-Year-Old Children: Validation of the Computer-Assisted Image Analysis of the Retina (CAIAR) Program*, Investigative Ophthalmology & Visual Science, **50**(5) (2009), 2004-2010.

[10] A. D. HOOVER, V. KOUZNETSOVA, M. GOLDBAUM: *Locating Blood Vessels in Retinal Images by Piece-wise Threshold Probing of a Matched Filter Response*, IEEE Transactions on Medical Imaging, **19** (2000), 203–210.

[11] J. STAAL, M. D. ABRÀMOFF, M. NEIMEIJER, M. A. VIERGEVER, B. VAN GINNEKEN: *Ridge-Based Vessel Segmentation in Color Images of the Retina*, IEEE Transactions on Medical Imaging **23**(4) (2004), 501–509.

[12] N. SINGH, L. KAUR: *A survey on blood vessel segmentation methods in retinal images*, International Conference on Electronic Design, Computer Networks & Automated Verification (EDCAV), Shillong, (2015), 23–28.

[13] N. SINGH, L. KAUR, K. SINGH: *Histogram equalization techniques for enhancement of low radiance retinal images for early detection of diabetic retinopathy*, Engineering Science and Technology, an International Journal, **22**(3) (2019), 736–745.

[14] S. M. PIZER, E. P. AMBURN, J. D. AUSTIN, R. CROMARTIE, A. GESELOWITZ, T. GREER, B. H. ROMENY, J. B. ZIMMERMAN, K. ZUDERVELD: *Adaptive histogram equalization and its variations*, Computer Vision, Graphics, and Image Processing, **39**(3) (1987), 355–368.

[15] Y. LECUN, B. BOSER, J. S. DENKER, D. HENDERSON, R. E. HOWARD, W. HUBBARD, L. D. JACKEL: *Backpropagation Applied to Handwritten Zip Code Recognition*, Neural Computation **1**(4) (1989), 541–551.

[16] E. KAYA, I. SARITAS, I. A. OZKAN: *Supervised Segmentation of Retinal Vessel Structures Using ANN*, arXiv 2001.05549 (**eess.IV**) (2020).

[17] K. KUMAR, S. SURAJ, D. SAMAL: *Automated retinal vessel segmentation based on morphological preprocessing and 2D-Gabor wavelets*, Advanced Computing and Intelligent Engineering, **1082**(2020), 411–423.

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
PUNJABI UNIVERSITY
Email address: navdeepsony@gmail.com

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
PUNJABI UNIVERSITY
Email address: danishbansal808@gmail.com

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
PUNJABI UNIVERSITY
Email address: dheerajnagpal21081999@yahoo.com