Blood Vessel Segmentation from Fundus Images Using Modified U-net Convolutional Neural Network

Afolabi O. Joshua  
Department of Electrical/Engineering Science, University of Johannesburg, Johannesburg, South Africa  
Email: afolabij@uj.ac.za

Fulufhelo V. Nelwamondo and Gugulethu Mabuza-Hocquet  
Council for Scientific and Industrial Research, Pretoria, South Africa  
Email: {fnelwamondo, gmabuza}@csir.co.za

Abstract—The state of blood vessels in the retina is an important factor used to diagnose the presence of several eye diseases and heart conditions. This is the reason why blood vessels segmentation of fundus images has gained wide popularity among researchers. This paper proposes blood vessel segmentation method based on an improved U-net Convolutional Neural Network (CNN) architecture. The proposed method involves very minimal fundus pre-processing and no post-processing of segmented blood vessel, thus, making the method very simple and easy to use. The method is tested on four publicly available databases which are: the DRIVE, CHASE _DB1, STARE and the HRF databases. The proposed method outperforms existing methods in accuracy, specificity and sensitivity on the four databases. A sensitivity of 0.8309, 0.7796, 0.7506, 0.8059 and a specificity of 0.9742, 0.9864, 0.9824 and 0.9826 are achieved on the DRIVE, CHASE_DB1, STARE and HRF databases respectively.

Index Terms—segmentation, blood vessel, convolutional neural network, u-net, retina fundus image

I. INTRODUCTION

Proper analysis of the retinal fundus image is very important for accurate and quick diagnosis of eye diseases such as glaucoma, diabetic retinopathy and macular degeneration. A delayed diagnosis of this ocular diseases is a leading cause of blindness in the world. Retinal blood vessels are of great importance to ophthalmologists when studying fundus images since they hold vital knowledge about the state of the eyes [1]. For proper analysis, the blood vessels must be segmented and properly examined. The manual segmentation of blood vessel is a tedious and time consuming task influenced by emotional instability and fatigue [2]. The tedious segmentation process raises a need for quick, accurate and less tedious automation process that also allows for a large number of fundus images to be segmented for their blood vessel at once. Many of the existing blood vessel segmentation methods [1]-[3] are not adaptive i.e. the methods are context specific and can only be used on fundus images of the same or very similar quality and pathology. This major drawback makes it difficult for the methods to be used on fundus images from another database. Moreover, the methods are also computationally intensive when used on large batches of fundus images.

In recent years, several deep learning architectures have been used for blood vessel segmentation process. This makes the methods adaptable and less context specific. This means the methods can be used on fundus images from different databases without any need for adjustment [4], [5], [6] and [7]. These deep learning methods however, include a lot of pre-processing and post-processing but have achieved high score in metrics like accuracy and area-under-curve. It is also worth stating that the deep learning methods have not been able to achieve high scores in metric like sensitivity.

In this work, we propose a blood vessel segmentation method using a modified U-net architecture. The modified U-net architecture has much less number of parameters than the traditional U-net. The contributions of this research include a blood vessel segmentation method that involve minimal pre-processing and no post-processing. The method yields higher sensitivity, specificity and accuracy score and a very low segmentation time for both online and batch segmentation process.

The rest of this paper is organized as follows: section II discusses the related work, section III discusses the proposed approach of the experiment, section IV presents the results of the experiment and the last section presents the conclusion

II. RELATED WORK

Sunil et al. [4], [5] proposed the use of a pre-trained model to segment blood vessels from fundus images. The DEEPLAB-COCO-LARGEFOV [8] model was pre-trained on the Microsoft COCO data-set [9]. The pre-trained model was afterwards trained with 800 image patches extracted from 66 fundus images. The outputs of
the model which were in patches were then recombined to form the desired segmented blood vessels. Sunil et al. recorded an accuracy of 0.9394 and area under curve of 0.894 when they tested their model on 23 fundus images. Sonro et al. [6] proposed a method that involves the pre-processing of fundus images before blood vessel segmentation can be done. The pre-processing steps include removal of non-uniform illumination, conversion of fundus images to single gray scale images and image rescaling. The pre-processed images were fed into a Convolutional Neural Network (CNN) model. The output of the model was further processed using a double threshold method. Sonro et al. achieved a sensitivity of 0.75 and an accuracy of 0.947 on the DRIVE and STARE databases. Oliveira et al. [7] proposed a method that combined stationary wavelet transformations with a fully convolutional neural network for blood vessel segmentation process. Stationary wavelet transform processes were done on the fundus images before patches were extracted from the fundus images. The patches were fed into the proposed fully convolutional neural network. The method was evaluated on three databases and good results were obtained (as seen in Table II). Xiancheng et al. [10] proposed a pipeline for blood vessel segmentation using the U-net architecture. In order to do this, the fundus images were first pre-processed and the U-net trained with patches of the pre-processed images so as to increase the accuracy of the architecture. A total of 190000 patches were extracted from the DRIVE database. Xiancheng et al. achieved area under curve score of 0.9790 and 0.9805 on the DRIVE and STARE databases respectively. Liskowski et al. [11] proposed a supervised segmentation algorithm using deep learning framework to segment blood vessels from fundus images. In order to do this, fundus images were pre-processed with zero-phase component analysis and contrast normalization. The proposed algorithm was trained on over 5 million image patches and tested on the DRIVE, CHASE and STARE databases. Liskowski et al. reported that their model has great performances (0.979) in the accuracy metric. Qiaoliang et al. [12] proposed a cross modality learning approach in which the retinal image was transformed to a vessel map. Qiaoliang et al. used deep learning architecture with strong induction ability trained with patches from the fundus images. A high specificity, sensitivity and accuracy score was reported.

It should be noted that many of the methods discussed above used patches from the fundus images to train the chosen deep learning model and the output of the model recombined to form the desired segmented blood vessel. It should also be noted that these approaches come with a lot of pre-processing and post-processing to obtain the desired segmentation.

III. PROPOSED EXPERIMENTAL APPROACH

The approach proposed for the blood vessel segmentation task is based on deep learning techniques. Deep learning with Convolutional Neural Network (CNN) has excelled in segmentation tasks because of its robust framework and better accuracy when trained with large dataset [8], [9].

In our proposed method, fundus images from all the datasets are scaled down so as to reduce computation cost and time. The re-scaling has no negative impact on the training process, it rather increases the training speed. The fundus images are scaled down to 512 x 512 for the DRIVE [13] and STARE [14], [15] databases and are scaled down to 960x960 in the CHASE_DB1 [16]-[22] and HRF [23] database. The contrast of the fundus images is further enhanced using the histograms calculated over several tile regions of the image. Scikit equalize adapt_hist is used for this process. Improving the contrast makes the fundus images to have uniform contrast and hence, better training of model. The improved contrast fundus images are fed into the CNN model.

The CNN architecture proposed is a modified and improved version of the U-Net [24]. Keras framework with tensor flow backend is used. When compared to the original U-net, the proposed architecture as shown in Fig. 1 has more convolutional layers. Though the proposed CNN architecture has more layers, the filter size (3x3) is kept the same in all layers except the output layer which has a filter size of 1x1. The proposed architecture has much less number of parameters than the earliest U-Net. Our experiment revealed that networks with large parameters over-fit quickly on the training data and therefore generalizes poorly for segmentation tasks. Batch-Normalization by Ioffe and Szeged [25] is used on each layer so as to bring the mean activation close to zero in all layers. Leaky version of the Rectified linear unit (ReLU) is used as the activation function (f(x) = 0.018 * x for x < 0, f(x) = x for x >= 0) [26]. Leaky ReLU is used because it does not saturate and it makes the network to converge faster. The loss function used is shown in (1).

\[
C(X,Y) = -logf(X,Y) \tag{1}
\]

\[
f(X,Y) = \frac{2\sum_{i,j} w_{i,j} x_{i,j} y_{i,j}}{\sum_{i,j} w_{i,j} \sum_{i,j} y_{i,j}} \tag{2}
\]
where the probability that the pixels predicted belongs to the foreground is \( X = (x_{ij}) \) and the given output is \( Y = (y_{ij}) \), \( h, w \) are the height and width respectively.

Metrics used to measure the performance of the model are sensitivity, specificity, accuracy, area under the receiver operating characteristic curve (AUC), harmonic mean of precision and the Matthew’s correlation coefficient.

Sensitivity (3) is a measure of the proportion of the true positives that are rightly detected while specificity (4) is a measure of the proportion of the true negatives correctly detected as negative. Accuracy (5) is a measure of how close a result is to its actual true value. Receiver operating curve is a plot of sensitivity against specificity while harmonic mean of precision and sensitivity (6) is the harmonic mean of sensitivity and specificity (F1). Matthew’s correlation coefficient (MCC) (7) is a measure of the quality of how well a binary classification is carried out.

\[
\text{sensitivity} = \frac{TP}{TP + FN} \tag{3} \\
\text{specificity} = \frac{TN}{TN + FP} \tag{4} \\
\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{5} \\
\begin{align*}
\text{f1} = \frac{2TP}{2TP + FP + FN} \\
\text{mcc} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \tag{7}
\end{align*}
\]

where: TP = True Positive, TN = True Negative, FP = False Positive, FN = False Negative, f1 = harmonic mean of precision and sensitivity and mcc = Matthew’s correlation coefficient.

The optimization of the model is done using the Stochastic Gradient Descent (SGD) with a momentum of 0.90 with nesterov. SGD is used because it gives a better result when compared to other optimizers.

### IV. EXPERIMENT RESULTS

The modified U-net model is trained for 100 epochs on the DRIVE database and 50 epochs on the three remaining databases i.e. CHASE_DB1, STARE and HRF. DRIVE database has 40 fundus images; 20 for training and 20 for testing. Each fundus image has a size of 768x584 pixels. The CHASE_DB1 has 28 fundus images with each image having a size of 960x999 pixels. In our experiment, the 20 fundus images are split into four folds of training and testing. Hence, the model is trained and tested on all the 20 images. The STARE database has 20 fundus images, each of size 605x700. The HRF database has 45 fundus, each of size 3504x2336. We used the same set up used for CHASE_DB1 fundus images on both HRF and STARE only that the fundus images in both are split into 5 folds (not 4 folds as in CHASE_DB1).

The model is trained using Kaggle’s 2 CPU cores, 14 GB RAM, 1 NVIDIA Tesla K80 GPU. A batch-size of 4 is used on both the DRIVE and STARE fundus images and a batch-size of 2 is used on both the CHASE_DB1 and HRF fundus images. An image size of 512x512 is used on both STARE and DRIVE while an image size of 960x960 is used on both CHASE_DB1 and HRF fundus images.

We present our result in Table I, and compare our results with other state-of-the-art methods in Table II-V using the sensitivity, specificity and accuracy metrics. Prediction time for our model is given by Kaggle’s 2 CPU cores, 14 GB RAM, 1 NVIDIA Tesla K80 GPU.

### TABLE I. RESULT OF PROPOSED BLOOD VESSEL SEGMENTATION METHOD ON DIFFERENT DATABASES

<table>
<thead>
<tr>
<th>Database</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRIVE</td>
<td>0.8309</td>
<td>0.7796</td>
<td>0.7506</td>
</tr>
<tr>
<td>CHASE_DB1</td>
<td>0.9742</td>
<td>0.9864</td>
<td>0.9824</td>
</tr>
<tr>
<td>STARE</td>
<td>0.9615</td>
<td>0.9722</td>
<td>0.9658</td>
</tr>
<tr>
<td>HRF</td>
<td>0.9026</td>
<td>0.8830</td>
<td>0.8668</td>
</tr>
</tbody>
</table>

### TABLE II. BLOOD VESSEL SEGMENTATION COMPARISON ON DRIVE DATABASE

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oliveira et al. [9]</td>
<td>0.8039</td>
<td>0.9804</td>
<td>0.9576</td>
</tr>
<tr>
<td>Liskowski et al. [11]</td>
<td>0.7520</td>
<td>0.9806</td>
<td>0.9515</td>
</tr>
<tr>
<td>Qiaoqiang et al. [12]</td>
<td>0.7569</td>
<td>0.9816</td>
<td>0.9527</td>
</tr>
<tr>
<td>Roychowdhury et al. [18]</td>
<td>0.7249</td>
<td>0.9830</td>
<td>0.9519</td>
</tr>
<tr>
<td>Fraz et al. [20]</td>
<td>0.7406</td>
<td>0.9807</td>
<td>0.9480</td>
</tr>
<tr>
<td>Soares et al. [1]</td>
<td>0.7332</td>
<td>0.9782</td>
<td>0.9466</td>
</tr>
<tr>
<td>Jiang et al. [27]</td>
<td>0.7540</td>
<td>0.9825</td>
<td>0.9624</td>
</tr>
<tr>
<td>Proposed method</td>
<td>0.8309</td>
<td>0.9742</td>
<td>0.9615</td>
</tr>
</tbody>
</table>

### TABLE III. BLOOD VESSEL SEGMENTATION COMPARISON ON CHASE_DB1 DATABASE

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oliveira et al. [9]</td>
<td>0.7779</td>
<td>0.9864</td>
<td>0.9653</td>
</tr>
<tr>
<td>Qiaoqiang et al. [12]</td>
<td>0.7507</td>
<td>0.9793</td>
<td>0.9581</td>
</tr>
<tr>
<td>Roychowdhury et al. [18]</td>
<td>0.7201</td>
<td>0.9824</td>
<td>0.9532</td>
</tr>
<tr>
<td>Fraz et al. [20]</td>
<td>0.7224</td>
<td>0.9711</td>
<td>0.9469</td>
</tr>
<tr>
<td>Zang et al. [16]</td>
<td>0.7644</td>
<td>0.9716</td>
<td>0.9502</td>
</tr>
<tr>
<td>Jiang et al. [27]</td>
<td>0.8640</td>
<td>0.9745</td>
<td>0.9668</td>
</tr>
<tr>
<td>Proposed method</td>
<td>0.7796</td>
<td>0.9864</td>
<td>0.9722</td>
</tr>
</tbody>
</table>

### TABLE IV. BLOOD VESSEL SEGMENTATION COMPARISON ON STARE DATABASE

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oliveira et al. [9]</td>
<td>0.8315</td>
<td>0.9858</td>
<td>0.9694</td>
</tr>
<tr>
<td>Liskowski et al. [11]</td>
<td>0.8145</td>
<td>0.9866</td>
<td>0.9696</td>
</tr>
<tr>
<td>Qiaoqiang et al. [12]</td>
<td>0.7726</td>
<td>0.9844</td>
<td>0.9628</td>
</tr>
<tr>
<td>Roychowdhury et al. [18]</td>
<td>0.7719</td>
<td>0.9726</td>
<td>0.9515</td>
</tr>
<tr>
<td>Fraz et al. [20]</td>
<td>0.7548</td>
<td>0.9763</td>
<td>0.9534</td>
</tr>
<tr>
<td>Soares et al. [1]</td>
<td>0.7207</td>
<td>0.9747</td>
<td>0.9480</td>
</tr>
<tr>
<td>Jiang et al. [27]</td>
<td>0.8352</td>
<td>0.9846</td>
<td>0.9734</td>
</tr>
<tr>
<td>Proposed method</td>
<td>0.7506</td>
<td>0.9824</td>
<td>0.9658</td>
</tr>
</tbody>
</table>

©2020 Journal of Image and Graphics
help alleviate the effect of these damaged fundus images on the database. Some pre-processing will need the contrast of the fundus image to be improved and good quality. The proposed model architecture only needs the contrast of the fundus image to be improved and no other pre-processing is needed. Also, the model is trained with the full fundus images not patches making post-processing not needed. These advantages make the model to be computationally efficient and undemanding while having low prediction time for batch or online segmentation tasks. Future work can be carried out to improve the model’s score in the area under curve metric (AUC).

V. CONCLUSION.

The proposed model achieved state-of-the-art score in the sensitivity, specificity and accuracy metric. The high score is best seen in the HRF dataset. This is because fundus images in the HRF database have high resolution and good quality. The proposed model architecture only needs the contrast of the fundus image to be improved and no other pre-processing is needed. Also, the model is trained with the full fundus images not patches making post-processing not needed. These advantages make the model to be computationally efficient and undemanding while having low prediction time for batch or online segmentation tasks. Future work can be carried out to improve the model’s score in the area under curve metric (AUC).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHORS CONTRIBUTION

Afolabi O. Joshua conducted the research; Gugulethu Mabuza-Hocquet proof read the paper, arranged contents and provided paper presentation guidance; Fulufhelo V. Nelwamondo provided the research guidance and method of implementation. All authors had approved the final version.

REFERENCES


Copyright © 2020 by the authors. This is an open access article distributed under the Creative Commons Attribution License (CC BY-NC-ND 4.0), which permits use, distribution and reproduction in any medium, provided that the article is properly cited, the use is non-commercial and no modifications or adaptations are made.

Afolabi Oluwatosi Joshua has a Master’s Degree in Electrical and Electronic Engineering from the University of Johannesburg, South Africa. He is currently a doctoral student at the same University. His current research focuses on using deep learning techniques to detect glaucoma from fundus images. He has published and presented his research work both locally and internationally.

Gugulethu Mabuza-Hocquet has a PhD in Electrical and Electronic Engineering from the University of Johannesburg, South Africa. She is currently employed as a senior biometrics researcher at the Council for Scientific and Industrial Research (CSIR). Her current research focuses on studying the science behind the human eye and the iris by using computer vision and machine learning techniques to extract features for the automated authentication of children and adults’ identity, as well as the possible detection of certain diseases. Gugulethu is a supervisor for students enrolled for Masters and PhD degrees and a mentor for permanent employees within the organization. She has published and presented her research work both locally and internationally.

Fulufhelo Nelwamondo is an electrical engineer by training, and holds a Bachelor of Science and a PhD in Electrical Engineering, in the area of Computational Intelligence, both from the University of the Witwatersrand, in South Africa. He is a registered Professional Engineer, the Executive Manager for the CSIR Cluster: Next Generation Enterprises and Institutions, and was the Executive Director for the CSIR Modelling and Digital Science Unit. He is a Senior Member of the Institute of Electrical and Electronic Engineers (IEEE), a senior member of the Association of Computing Machinery (ACM), a member of the South African Institute of Electrical Engineers, Visiting Professor of Electrical Engineering, at the Institute of Intelligent Systems, University of Johannesburg. He was a post-doctoral fellow at the Graduate School of Arts and Sciences, of Harvard University. He is the youngest South African ever to receive the Harvard-South Africa fellowship and has been awarded many national and international research accedades, from organizations such as the IEEE, South African Institute of Electrical Engineers, amongst others.

In 2017, Prof. Nelwamondo was awarded the Order of Mapungubwe in Silver, highest civilian honour bestowed by the President of the Republic of South Africa. In 2016, he was recognised by the Operation Research Society of South Africa, for outstanding contributions to the science and profession of Operation Research. Prof Fulufhelo Nelwamondo has research and practical experience in software engineering, computational intelligence and optimisation in various applications. He has interests in emerging and exciting areas of software and technology applications including data science, modelling of complex systems, machine learning, optimisation and mechanism design. He has presented his work in various countries across the world. Prof Nelwamondo has successfully supervised over 20 Masters and 5 PhD students in Electrical Engineering and Computer Science. Prof Nelwamondo has served on a number of Boards of entities, and holds a number of Directorships.

©2020 Journal of Image and Graphics 25