Shallow Vessel Segmentation Network for Automatic Retinal Vessel Segmentation

1st Tariq M. Khan
Faculty of Sci Eng & Built Env, School of Info Technology
Deakin University, Geelong Waurn Ponds Campus
Victoria, Australia
tariq045@gmail.com

2nd Faizan Abdullah
Department of Electrical and Computer Engineering
COMSATS University Islamabad
Islamabad, Pakistan
faizan.abdullah@outlook.com

3rd Syed S. Naqvi
Department of Electrical and Computer Engineering
COMSATS University Islamabad
Islamabad, Pakistan
saud_naqvi@comsats.edu.pk

4th Muhammad Arsalan
Division of Electronics and Electrical Engineering
Dongguk University
Seoul 100-715, Korea
engineerarsal5@gmail.com

5th Muhamamd Aurangzeb Khan
Computing and Communications
Lancaster University
Lancaster, UK
m.a.khan4@lancaster.ac.uk

***Abstract*—Accurate automatic segmentation of the retinal vessels is crucial for early detection and diagnosis of vision-threatening retinal diseases. This paper presents a lightweight convolutional neural network termed as Shallow Vessel Segmentation Network (SVSN) for vessel segmentation. To achieve semantic segmentation encoder-decoder structures embedded with spatial pyramid pooling modules are used. After checking the input features with pooling through multiple fields of view and rates, it becomes easy for the erstwhile networks to encode multi-scale contextual information. While boundaries for sharper objects are captured by the prevalent networks. Moreover, the need for pre- and post-processing steps are eradicated. Consequently, the detection accuracy is significantly improved with scores of 0.9625 and 0.9645 on DRIVE and STARE datasets respectively.**

I. Introduction

Images obtained from Retinal vasculature structure in the medical field provides important information about the condition of the eye, which can help ophthalmologists to diagnose and predict certain retinal pathologies like, age-related muscular degeneration (AMD), retinal vascular occlusions, hypertension, glaucoma, diabetic retinopathy (DR), chronic systematic hypoxemia[1], [2], [3]. These diseases can be characterized by observing changes in retinal blood vessels such as shape, width, branching pattern, and tortuosity. Early detection, diagnosis, and tracking of disease progression can prevent vision loss in case of AMD and DR, and lead to cost-effective treatment options for other conditions [4], [5], [6]. To achieve this target retinal vascular segmentation is done which is fed to the automated systems.

This task of segmentation is manually done by ophthalmologists which requires a lot of experience and time to segment these eye images properly where there is limited information reliable enough to be extracted. This process configures a complex task, not only due to abrupt variations in the attributes (size, shape, intensity levels) and arrangement (branching, crossing) of the vessels but also to the low quality of retinal images [7], [8], [9], [10]. Whereas the automated systems with modern algorithms have an ability to test large screenings with improved accuracy, less time, and reduced human work. Although the accuracy of these algorithms is trivial due to several challenges faced in automatic vessel segmentation. Automatic vessel segmentation has been widely accepted as a challenging task but is vital for a computer-aided diagnostic system for ophthalmic diseases [11], [12], [13].

Li *et al* [14] presented a supervised method that remolds the segmentation task as a problem of transforming data from the retinal image to the vessel map into cross-modality. Modeling the transformation is proposed by a wide and deep neural network with strong induction capacity, and an effective training strategy is presented. The network can output the label map of all pixels for a given image pat instead of a single label of the center pixel. A comparison of the performance of various methodologies has been presented in [15]. Recently deep neural networks gained a lot of attention in discriminative tasks and outperformed state-of-art methods. Aslani *et al* [16] proposed a hybrid feature vector system which is the combination of robust features of various algorithms. The random forest discriminator is used to efficiently segment the vessels and non-vessel pixels. In [17], a neural

978-1-7281-6926-2/20/$31.00 ©2020 IEEE

network architecture has been presented where a retinal vessel segmentation task is formulated as a multi-label inference task which learned by the joint loss function. A supervised segmentation technique using a deep neural network trained using a large fundus images repository is introduced in [18]. The samples are pre-processed using zero-phase whitening, global contrast normalization, and augmented using geometric transformations and gamma corrections. Most of the retinal vessel segmentation algorithms failed to segment structurally thin and elongated vessels. To solve this problem an efficient conditional random field model is proposed in [19]. Generative Adversarial Network (GAN) has been used for many different image processing tasks. Son *et al* [20] proposed a GAN model to correctly segment the retinal vessels. An extended version of the U-Net architecture is proposed by [21] to segment the retinal vessels efficiently. This extension is the combination of Recurrent Convolutional Neural Networks and Recurrent Residual Convolutional Neural Networks. CNN models have been used for efficient retinal vessels segmentation in [22] and [23]. Hu *et al* [22] proposed a retinal segmentation method which is based on convolutional neural network (CNN) and conditional random field (CRF). The segmentation task is accomplished in two phases. In the first phase, the CNN model is used to produce a probability map. In the second phase, the CRF is applied to segment the retinal vessels. In [23], a novel method is proposed that combines the multi-scale analysis provided by the Stationary Wavelet Transform with a multi-scale CNN model to handle the vessel structure's varying width and direction in the retina. The work of Guo *et al* [24] used ConvNets to segment the retinal blood vessels by deploying global and local features extractors. This method ensured the separation of thin vessels by maintaining spatial consistency.

To segment vessels at multiple scales, a multiscale lightweight architecture termed as Shallow Vessel Segmentation Network (SVSN) network is proposed. The main idea behind the shallow network is rooted in previous studies that advocate the idea of overthinking and promote shallower networks [25], [26]. Shallow networks have been shown to perform better or at par with deeper neural networks in previous studies [27], [28]. The key ideas behind the proposed work are two-fold: 1) shallower network architecture to retain semantic information, 2) separate multi-scale contextual information extraction for both large and tiny vessels.

The proposed method used a residual shallow network along with spatial pyramid pooling. This work is inspired by DeepLabv3 [29] that applies several parallel atrous convolutions with different rates, and from the PSPNet [30] that performs pooling operation at different grid scales. Finally to recover the boundaries of vessels a decoder module is used. In addition, the network proposed is independent of pre- and post-processing steps. Compared to the above approaches, the proposed architecture is better performing.

II. Proposed Method

Computer vision comprises a major portion of semantic segmentation which is an art of assigning a semantic label to each pixel. Especially Deep convolution networks mitigate the limitations of the existing methods based on fully convolutional networks. This work will comprise two different types of neural networks, obtaining semantic segmentation through encoder-decoder structure, or spatial pyramid pooling module, as shown in Fig. 1. In the proposed method, instead of using 8 residuals, 4 residuals are used in the encoder of each network. One neural network will extract sharp boundaries of the objects while the other will work on pooling features through several resolutions to extract significant contextual information.

To extract multi resolution or multi-scale contextual information, Deep Lab v3 uses Atrous Spatial Pyramid pooling, also to multi grid-scale pooling operations PSPNet is used, as shown in Fig. 2. Although the latest feature map has significant semantic information even then full information of the boundaries of the object remain insignificant due to the convolutions with the striding operations. This problem can be solved by introducing Atrous convolution here which can filter out denser features. However, given the design of state-of-art neural networks and limited GPU memory, it is computationally prohibitive to extract output feature maps that are 8, or even 4 times smaller than the input resolution [29]. Hence if we go for denser features the costs for computations will increase. On contrary encoder-decoder structures proofs to be faster, as there are no dilated images in the encoder path and the decoder gradually extracts the boundaries for sharp objects. Combining the encoder-decoder module and getting the best outcomes of both the approaches it is fruitful to signify the encoder module by feeding contextual multi-scale information.

There are two key approaches that constitute the design of the proposed method. Firstly, the standard DeepLab v3+ method employs a 7×7 receptive field followed by atrous convolutions. It is hypothesized that this step may not be able to capture the tiny vessel information that exists at smaller scales. Therefore a separate encoder-decoder pathway with a receptive field of 3×3 is incorporated in the proposed approach. Also, the repetitive residual processing at the encoder stage destroys important vessel information at smaller scales. To eradicate this issue, residual stages at the encoder end are shredded to retain semantic information related to tiny vessels.

The architecture of the proposed architecture shown in Fig 1. The 3×3 receptive field path extracts contextual information at multiple scales capturing the fine details of the tiny vessels present in the retinal image. On the other hand, the 7×7 path extracts multiscale contextual information of the thick vascular structure. The proposed architecture can, therefore, be seen as an ensemble of feature extractors, where each predictor focuses on a particular modality in the vascular structure. The information of both the tiny and thick vessels is combined at the end of the network. The fusion scheme that worked best in our experiments is the addition of the feature

978-1-7281-6926-2/20/$31.00 ©2020 IEEE

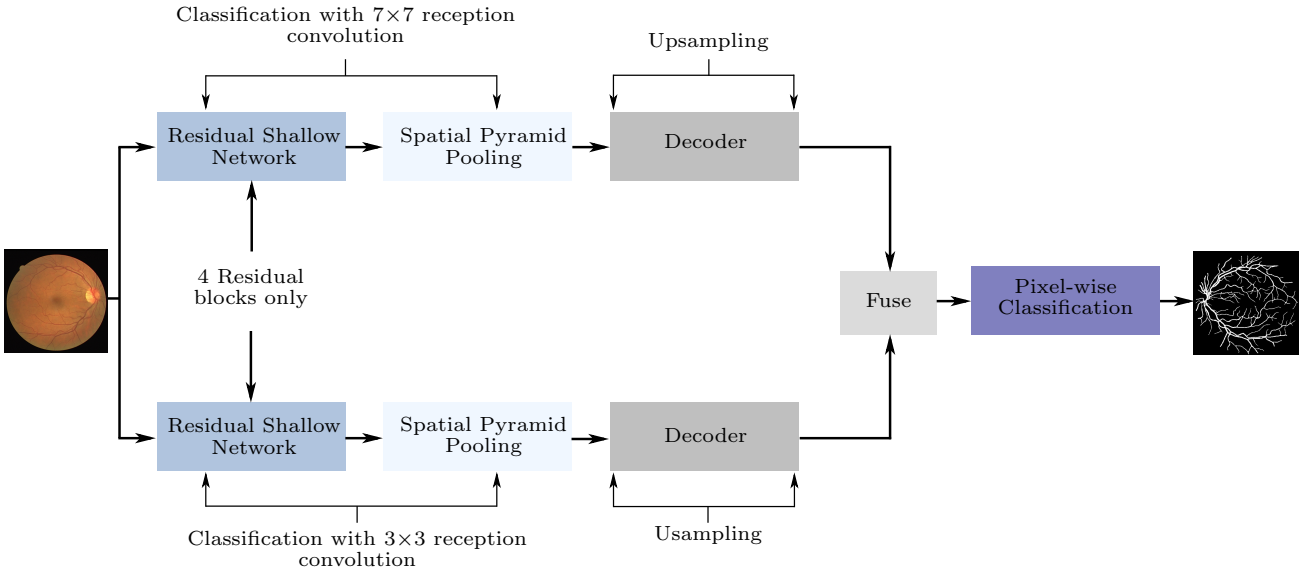

Fig. 1. Block diagram of the proposed architecture.

information originating from both the paths.

The differences between the proposed architecture and DeepLab v3+ are presented in Table I. The major difference is in the way the semantic and contextual information is preserved by the two architectures.

A. Data Augmentation

In a deep neural network, training is dependent on the size of input data. For effective training, a large size data is required. When the size of training data is low, parameters are underdetermined, and the network is poorly trained that badly affects the performance of a model. One way to tackle this problem is by data augmentation, which is used to alleviates this by using existing data more effectively. In this paper, first, all images are resized to 640×640. Then image rotation is used on these resized images to generate synthetic images using original training images. Each image is rotated with a difference of 1 degree from 0 to 360. By doing this we got 360 rotational images for each image and a total of 7200. In order to get rid of artifacts in the rotated images first, we converted binary images into logical images and then used bi-cubic interpolation while rotating these images. After rotation, we used contrast enhancement and generated 1800 more images of different contrasts. Therefore, by augmentation from 20 images to 9000 images are synthetically generated.

B. Training of Networks

In this work, we used the Adam optimizer with an initial learning rate of 0.005, gradient threshold of 5, epsilon of 0.000001, and squared gradient decay factor of 0.95 with global L2 normalization. Our model was trained for 15 epochs with a mini-batch size of 8 images with shuffling after each epoch as our network converges faster. For calculating the loss overall pixels available in the candidate mini-batch according to the vessel and non-vessel class, so we used the cross-entropy loss function. The network tries to learn the dominant class if there is a difference between the numbers of pixels in different classes, which leads to slower convergence of the network and also affects accuracy.

III. Experimental Results

A. Materials

The proposed method is tested on two publicly available datasets as listed below:

1) DRIVE [31], Digital Retinal Images for Vessel Extraction: Retina periphery scans taken from a broad age group diabetics in the Netherlands.
2) STARE [32] Structured Analysis of the Retina: 20 samples from a collection of 400 mid-resolution images taken in USA

In DRIVE, the binary masks are available for each image. The vascular structure is manually marked as vessel or non-vessel. In contrast with DRIVE, FOV mask is not available for the STARE database. The appropriate mask must therefore be generated employing existing procedures [33]. To train and then test the classifier, DRIVE provides with distinct sets of images but STARE database has no such provisions. A common solution to this problem has been to employ an arbitrary combination of images in a situation where training and testing information isn't independently given. This approach appears to be practical, yet it exhibits unrealistically favorable outcomes primarily due to information overlap [34], [35]. This problem is best tackled by a "leave-one-out" routine that skips one sample out of a total of 'n' samples used from the dataset once during the whole 'n-1' iterations [36], [33]. each iteration skips a different image from the set such that all images get left out once and doing so results in remarkable enhancement in execution. This investigation utilized the " leave-one-out "

978-1-7281-6926-2/20/$31.00 ©2020 IEEE

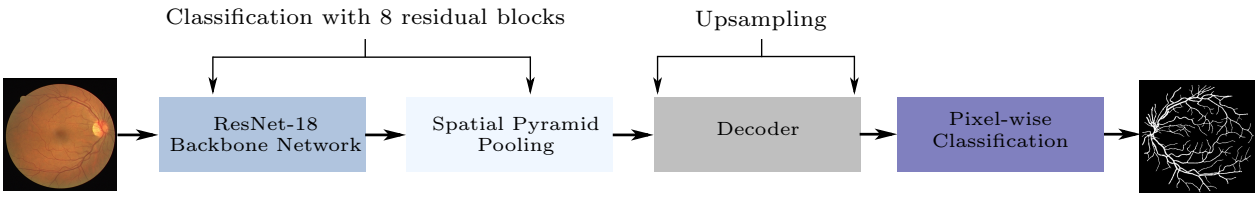

Fig. 2. Block diagram of deep lab v3++ architecture.

TABLE I
DIFFERENCE BETWEEN DEEPLAB V3+ AND THE PROPOSED ARCHITECTURE.

Sr No.	DeepLab v3+	SVSN (proposed)
1	ResNet-18 as backbone	Train from scratch
2	8 residual blocks	4 residual blocks
3	Contextual information from a 7×7 receptive field	Contextual information from both a 3×3 and 7×7 receptive fields
4	Semantic information loss due to multiple residual blocks	Semantic information preservation through lightweight architecture

technique to train the network with the images loaded from the STARE dataset.

B. Evaluation Criterion

The efficacy of any method of vascular segmentation depends on how well vessels and pixels of the background are properly discriminated. The results of the segmentation are compared with the manually annotated binary ground truth that act as the reference map. This comparison gives true / false and positive / negative core values. A pixel identified as a vessel is labelled positive while recognized as a background pixel puts it in the false category. True means that any pixels are segmented properly as vessels or as non-vessels and vice versa. Thus all four variables play an important role in determining the effectiveness of every technique of vascular classification.

1) True Positive (TP): Vessels classified correctly
2) False Negative (FN): Vessels classified as background
3) True Negative (TN): Non-vessels classified correctly
4) False Positive (FP): Non-vessels classified as vessels

Using the core parameters listed above, specific ratios are evaluated to quantify and contrast the performance of the technique under scrutiny with other state-of-the-art segmentation strategies as follows [37]:

$$Se = \frac{TP}{TP+FN},$$
$$Sp = \frac{TN}{TN+FP},$$
$$Acc = \frac{TP+TN}{TP+FN+TN+FP},$$

C. Comparison with state-of-the-art

To evaluate the proposed method, we conduct experiments on the datasets DRIVE and STARE and compare with the current state-of-the-art methods. For the DRIVE dataset, the proposed framework achieves 0.8295, 0.9755, 0.9625 and 0.9710 for Se, Sp, Acc and AUC respectively. The sensitivity and accuracy of the proposed method on DRIVE data is higher than the existing state-of-the-art. The specificity of [51] is

TABLE II
COMPARISON WITH STATE-OF-THE-ART METHODS ON THE DRIVE DATASET.

Type	Methods	Year	Se	Sp	Acc	AUC
Unsupervised methods	Zhang [37]	2016	0.7743	0.9725	0.9476	
	Aguiree[38]	2018	0.7854	0.9662	0.950	
	Khan [39]	2018	0.730	0.979	0.958	
	Khawaja (CLAHE)[40]	2019	0.8027	0.9733	0.9561	
	Khawaja (GLM) [40]	2019	0.7907	0.9790	0.9603	
	Khawaja[41]	2019	0.8043	0.9730	0.9553	
	Zhou[42]	2020	0.7262	0.9803	0.9475	
Supervised methods	Marin et al. [43]	2011	0.7067	0.9801	0.9452	0.9588
	Fraz et al. [44]	2012	0.7406	0.9807	0.9480	0.9747
	Cheng et al. [35]	2014	0.7252	0.9798	0.9474	0.9648
	Li et al. [45]	2016	0.7569	0.9816	0.9527	0.9738
	Orlando et al. [46] FC	2017	0.7893	0.9792	N.A	0.9507
	Orlando et al. [46] UP	2017	0.7076	0.9870	N.A	0.9474
	Dasgupta and Singh [47]	2017	0.9691	0.9801	0.9533	0.9744
	Yan et al. [48]	2018	0.7653	0.9818	0.9542	0.9752
	Olaf et al. [49] U-Net	2018	0.7537	0.9820	0.9531	0.9755
	Azad et al. [50] BCDU-Net	2019	0.8007	0.9786	0.9560	0.9798
	Alam et al. [51] R U-Net	2019	0.7751	0.9816	0.9556	0.9782
	Alam et al. [51] R2U-Net	2019	0.7792	0.9813	0.9556	0.9784
	Soomro et al. [52] Strided U-Net	2019	0.8020	0.9740	0.9590	0.9480
	Chen et al. [53] Deeplab v3++	2018	0.8220	0.9750	0.9620	0.9680
	Proposed (SVSN)	2020	0.8295	0.9755	0.9625	0.9710

TABLE III
COMPARISON WITH STATE-OF-THE-ART METHODS ON THE STARE DATASET

Type	Methods	Year	Se	Sp	Acc	AUC
Unsupervised methods	Zhang [37]	2016	0.7791	0.9758	0.9554	
	Aguiree[38]	2018	0.7116	0.9454	0.9231	
	Khan [39]	2018	0.790	0.965	0.951	
	Khawaja (CLAHE)[40]	2019	0.7980	0.9732	0.9561	
	Khawaja (GLM) [40]	2019	0.7860	0.9725	0.9583	
	Khawaja [41]	2019	0.8011	0.9694	0.9545	
Supervised methods	Marin et al. [43]	2011	0.6944	0.9819	0.9526	0.9769
	Fraz et al. [44]	2012	0.7548	0.9763	0.9534	0.9768
	Li at al. [45]	2016	0.7726	0.9844	0.9628	0.9879
	Orlando et al. [46] FC	2017	0.7680	0.9738	N.A	N.A
	Orlando et al. [46] UP	2017	0.7692	0.9675	N.A	N.A
	Yan et al. [48]	2018	0.7581	0.9846	0.9612	0.9801
	Olaf et al. [49] U-Net	2018	0.8270	0.9842	0.9690	0.9898
	Soomro et al. [52] Strided U-Net	2019	0.8010	0.9690	0.9610	0.9450
	Chen et al. [53] Deeplab v3++	2018	0.8320	0.9760	0.9650	0.9735
	Proposed (SVSN)	2020	0.8382	0.9749	0.9645	0.9740

978-1-7281-6926-2/20/$31.00 ©2020 IEEE

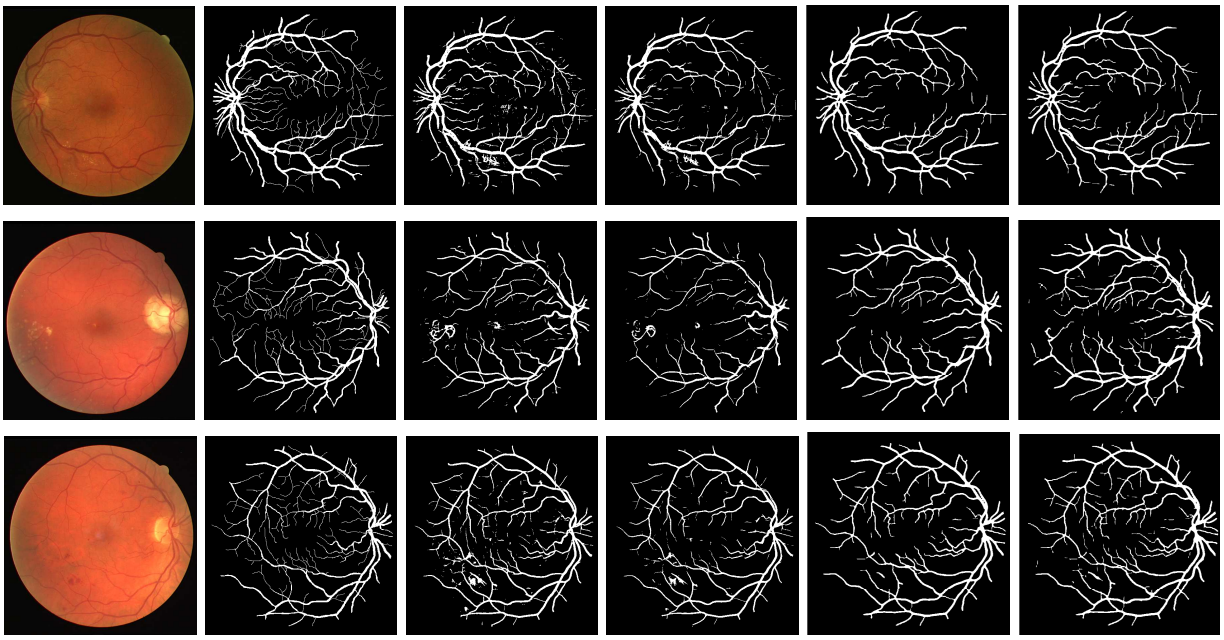

Fig. 3. Analysis of Segmented Output of three sample images (test image 3, 8, 14) of DRIVE dataset. Second column shows the ground truth images. Column 3 and 4 shows the output of [46] UP and [46] FC , respectively. Column 5 shows the output of Chen et al. [53] Deeplab v3++. The visual output of the proposed method is shown in the last column.

highest among all methods. If we compare proposed method with baseline [53] Deeplab v3++ method then we can see that proposed method performed better. The sensitivity and AUC of the proposed is much higher than [53] Deeplab v3++ while specificity and accuracy are comparable. It can be observed that the balanced accuracy of the proposed method is higher than all existing methods on DRIVE dataset.

For the STARE dataset, the FOV masks are generated by applying a thresholding method. According to the results in Table III, the proposed framework achieves 0.8382, 0.9749, 0.9645 and 0.9740 for Se, Sp, Acc and AUC respectively. The Sensitivity of the proposed method is highest than all existing state-of-the-art methods. The accuracy of Olaf et al. [49] is highest and [53] Deeplab v3++ is second highest and the propose methods accuracy is at third slightly lower than [53] Deeplab v3++. It can be observed that the balanced accuracy of the proposed method is higher than all existing methods on the STARE dataset.

The visual results on the DRIVE database in Fig. 3 show that the proposed method captures tiny vessels that are missed by the approach of Chen et al. [53] Deeplab v3++ and [46]. Also, the pathological noise and the part of the optic disc erroneously included by the output of Orlando et al. [46] in test images 3, 8, and 14 are successfully suppressed by the proposed method as well as Chen et al. [53] Deeplab v3++. In conclusion, the visual results on the DRIVE dataset in Fig. 3 clearly demonstrate that the output of the proposed captures more tiny vessels than Chen et al. [53] Deeplab v3++, in terms of accuracy both propose and Chen et al. [53] Deeplab v3++ performed well under the noisy condition where the approach of Orlando et al. struggles to suppress pathological noise and the optic disc boundary.

IV. Conclusion

This paper presented a new approach to retinal vessel segmentation by adapting ideas from the DeepLab v3+ method. One of the key goals was to capture information at multi-scales to improve the sensitivity of the method to both large and tiny vessels. This was successfully achieved by incorporating a 3×3 reception pathway in addition to the 7×7 paths. The second goal was to retain semantic information regarding tiny vessels that otherwise gets lost in the repetitive convolution process. This was successfully achieved by shredding the relativize residual processing in the standard encoder of DeepLab v3 +. The results of the proposed method demonstrate its efficacy over the standard DeepLab v3+ method and more than 20 state-of-the-art retinal vessel segmentation approaches.

References

[1] Sindri Traustason, Annette Schophuus Jensen, Henrik Sven Arvidsson, Inger Christine Munch, Lars Søndergaard, and Michael Larsen, "Retinal oxygen saturation in patients with systemic hypoxemia," *Investigative ophthalmology & visual science*, vol. 52, no. 8, pp. 5064–5067, 2011.

[2] Mohammad AU Khan, Tariq M Khan, Kahtan I Aziz, Sayed S Ahmad, Nighat Mir, and Emhemed Elbakush, "The use of fourier phase symmetry for thin vessel detection in retinal fundus images," in *2019 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*. IEEE, 2019, pp. 1–6.

[3] Mohammad AU Khan, Tariq M Khan, DG Bailey, and Toufique A Soomro, "A generalized multi-scale line-detection method to boost retinal vessel segmentation sensitivity," *Pattern Analysis and Applications*, vol. 22, no. 3, pp. 1177–1196, 2019.

[4] Mohammad AU Khan, Tariq M Khan, Syed S Naqvi, and M Aurangzeb Khan, "Ggm classifier with multi-scale line detectors for retinal vessel segmentation," *Signal, Image and Video Processing*, vol. 13, no. 8, pp. 1667–1675, 2019.

978-1-7281-6926-2/20/$31.00 ©2020 IEEE

[5] Mohammad AU Khan, Tariq M Khan, Toufique Ahmed Soomro, Nighat Mir, and Junbin Gao, "Boosting sensitivity of a retinal vessel segmentation algorithm," *Pattern Analysis and Applications*, vol. 22, no. 2, pp. 583–599, 2019.

[6] Mohammad A.U. Khan Tariq M. Khan Manoranjan Paul Nighat Mir Toufique Soomro, Junbin Gao, "Role of image contrast enhancement technique for ophthalmologist as a diagnostic tool for the diabetic retinopathy," in *IEEE International Conference on Digital Image Computing: Techniques and Applications (DICTA), At Gold Coast, Autralia*, 2016, pp. 1–8.

[7] Muhammad Moazam Fraz, Paolo Remagnino, Andreas Hoppe, Bunyarit Uyyanonvara, Alicja R Rudnicka, Christopher G Owen, and Sarah A Barman, "Blood vessel segmentation methodologies in retinal images–a survey," *Computer methods and programs in biomedicine*, vol. 108, no. 1, pp. 407–433, 2012.

[8] Mehwish Mehmood, Tariq M Khan, Mohammad AU Khan, Syed S Naqvi, and Wadee Alhalabi, "Vessel intensity profile uniformity improvement for retinal vessel segmentation," *Procedia Computer Science*, vol. 163, pp. 370–380, 2019.

[9] Toufique Ahmed Soomro, Junbin Gao, Tariq Khan, Ahmad Fadzil M Hani, Mohammad AU Khan, and Manoranjan Paul, "Computerised approaches for the detection of diabetic retinopathy using retinal fundus images: a survey," *Pattern Analysis and Applications*, vol. 20, no. 4, pp. 927–961, 2017.

[10] T. A. Soomro, T. Mahmood Khan, M. A. U. Khan, J. Gao, M. Paul, and L. Zheng, "Impact of ica-based image enhancement technique on retinal blood vessels segmentation," *IEEE Access*, vol. 6, pp. 3524–3538, 2018.

[11] Cemil Kirbas and Francis Quek, "A review of vessel extraction techniques and algorithms," *ACM Computing Surveys (CSUR)*, vol. 36, no. 2, pp. 81–121, 2004.

[12] Toufique Ahmed Soomro, Mohammad AU Khan, Junbin Gao, Tariq M Khan, and Manoranjan Paul, "Contrast normalization steps for increased sensitivity of a retinal image segmentation method," *Signal, Image and Video Processing*, vol. 11, no. 8, pp. 1509–1517, 2017.

[13] Toufique A Soomro, Mohammad AU Khan, Junbin Gao, Tariq M Khan, Manoranjan Paul, and Nighat Mir, "Automatic retinal vessel extraction algorithm," in *2016 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*. IEEE, 2016, pp. 1–8.

[14] Qiaoliang Li, Bowei Feng, LinPei Xie, Ping Liang, Huisheng Zhang, and Tianfu Wang, "A cross-modality learning approach for vessel segmentation in retinal images," *IEEE transactions on medical imaging*, vol. 35, no. 1, pp. 109–118, 2015.

[15] Buket D Barkana, Inci Saricicek, and Burak Yildirim, "Performance analysis of descriptive statistical features in retinal vessel segmentation via fuzzy logic, ann, svm, and classifier fusion," *Knowledge-Based Systems*, vol. 118, pp. 165–176, 2017.

[16] Shahab Aslani and Haldun Sarnel, "A new supervised retinal vessel segmentation method based on robust hybrid features," *Biomedical Signal Processing and Control*, vol. 30, pp. 1–12, 2016.

[17] Avijit Dasgupta and Sonam Singh, "A fully convolutional neural network based structured prediction approach towards the retinal vessel segmentation," in *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*. IEEE, 2017, pp. 248–251.

[18] Paweł Liskowski and Krzysztof Krawiec, "Segmenting retinal blood vessels with deep neural networks," *IEEE transactions on medical imaging*, vol. 35, no. 11, pp. 2369–2380, 2016.

[19] José Ignacio Orlando, Elena Prokofyeva, and Matthew B Blaschko, "A discriminatively trained fully connected conditional random field model for blood vessel segmentation in fundus images," *IEEE transactions on Biomedical Engineering*, vol. 64, no. 1, pp. 16–27, 2016.

[20] Jaemin Son, Sang Jun Park, and Kyu-Hwan Jung, "Retinal vessel segmentation in fundoscopic imaages with generative adversarial networks," *arXiv preprint arXiv:1706.09318*, 2017.

[21] Md Zahangir Alom, Mahmudul Hasan, Chris Yakopcic, Tarek M Taha, and Vijayan K Asari, "Recurrent residual convolutional neural network based on u-net (r2u-net) for medical image segmentation," *arXiv preprint arXiv:1802.06955*, 2018.

[22] Kai Hu, Zhenzhen Zhang, Xiaorui Niu, Yuan Zhang, Chunhong Cao, Fen Xiao, and Xieping Gao, "Retinal vessel segmentation of color fundus images using multiscale convolutional neural network with an improved cross-entropy loss function," *Neurocomputing*, vol. 309, pp. 179–191, 2018.

[23] Américo Oliveira, Sérgio Pereira, and Carlos A Silva, "Retinal vessel segmentation based on fully convolutional neural networks," *Expert Systems with Applications*, vol. 112, pp. 229–242, 2018.

[24] Jinnan Guo, Shiwei Ren, Yueting Shi, and Haoyu Wang, "Automatic retinal blood vessel segmentation based on multi-level convolutional neural network," in *2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*. IEEE, 2018, pp. 1–5.

[25] Jimmy Ba and Rich Caruana, "Do deep nets really need to be deep?," in *Advances in Neural Information Processing Systems 27*, Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger, Eds., pp. 2654–2662. Curran Associates, Inc., 2014.

[26] Yigitcan Kaya, Sanghyun Hong, and Tudor Dumitras, "Shallow-Deep Networks: Understanding and mitigating network overthinking," in *Proceedings of the 2019 International Conference on Machine Learning (ICML)*, Long Beach, CA, Jun 2019.

[27] Z. Shi, L. Zhang, Y. Liu, X. Cao, Y. Ye, M. Cheng, and G. Zheng, "Crowd counting with deep negative correlation learning," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 5382–5390.

[28] J. Pan, E. Sayrol, X. Giro-I-Nieto, K. McGuinness, and N. E. O'Connor, "Shallow and deep convolutional networks for saliency prediction," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 598–606.

[29] Liang-Chieh Chen, George Papandreou, Florian Schroff, and Hartwig Adam, "Rethinking atrous convolution for semantic image segmentation," 2017.

[30] Hengshuang Zhao, Jianping Shi, Xiaojuan Qi, Xiaogang Wang, and Jiaya Jia, "Pyramid scene parsing network," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jul 2017.

[31] Meindert Niemeijer, Joes Staal, Bram van Ginneken, Marco Loog, and Michael D Abramoff, "Comparative study of retinal vessel segmentation methods on a new publicly available database," in *Medical imaging 2004: image processing*. International Society for Optics and Photonics, 2004, vol. 5370, pp. 648–657.

[32] A. D. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *IEEE Transactions on Medical Imaging*, vol. 19, no. 3, pp. 203–210, 2000.

[33] J. V. B. Soares, J. J. G. Leandro, R. M. Cesar, H. F. Jelinek, and M. J. Cree, "Retinal vessel segmentation using the 2-d gabor wavelet and supervised classification," *IEEE Transactions on Medical Imaging*, vol. 25, no. 9, pp. 1214–1222, 2006.

[34] E. Ricci and R. Perfetti, "Retinal blood vessel segmentation using line operators and support vector classification," *IEEE Transactions on Medical Imaging*, vol. 26, no. 10, pp. 1357–1365, 2007.

[35] E. Cheng, L. Du, Y. Wu, Y. J. Zhu, V. Megalooikonomou, and H. Ling, "Discriminative vessel segmentation in retinal images by fusing context-aware hybrid features," *Machine Vision and Applications*, vol. 25, no. 7, pp. 1779–1792, Oct. 2014.

[36] 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, 2004.

[37] Jiong Zhang, Behdad Dashtbozorg, Erik Bekkers, Josien PW Pluim, Remco Duits, and Bart M ter Haar Romeny, "Robust retinal vessel segmentation via locally adaptive derivative frames in orientation scores," *IEEE transactions on medical imaging*, vol. 35, no. 12, pp. 2631–2644, 2016.

[38] Hugo Aguirre-Ramos, Juan Gabriel Avina-Cervantes, Ivan Cruz-Aceves, José Ruiz-Pinales, and Sergio Ledesma, "Blood vessel segmentation in retinal fundus images using gabor filters, fractional derivatives, and expectation maximization," *Applied Mathematics and Computation*, vol. 339, pp. 568–587, 2018.

[39] Khan Bahadar Khan, Amir A Khaliq, Abdul Jalil, and Muhammad Shahid, "A robust technique based on vlm and frangi filter for retinal vessel extraction and denoising," *PloS one*, vol. 13, no. 2, pp. e0192203, 2018.

[40] A. Khawaja, T. M. Khan, K. Naveed, S. S. Naqvi, N. U. Rehman, and S. Junaid Nawaz, "An improved retinal vessel segmentation framework using frangi filter coupled with the probabilistic patch based denoiser," *IEEE Access*, vol. 7, pp. 164344–164361, 2019.

978-1-7281-6926-2/20/$31.00 ©2020 IEEE

[41] Ahsan Khawaja, Tariq M. Khan, Mohammad A. U. Khan, and Syed Junaid Nawaz, "A multi-scale directional line detector for retinal vessel segmentation," *Sensors*, vol. 19, no. 22, 2019.

[42] Chao Zhou, Xiaogang Zhang, and Hua Chen, "A new robust method for blood vessel segmentation in retinal fundus images based on weighted line detector and hidden markov model," *Computer Methods and Programs in Biomedicine*, vol. 187, pp. 105231, 2020.

[43] D. Marin, A. Aquino, M. E. Gegundez-Arias, and J. M. Bravo, "A new supervised method for blood vessel segmentation in retinal images by using gray-level and moment invariants-based features," *IEEE Transactions on Medical Imaging*, vol. 30, no. 1, pp. 146–158, 2011.

[44] Fraz M. M., Remagnino P., Hoppe A., Uyyanonvara B., Rudnicka A. R., Owen C. G., and Barman S. A., "An ensemble classification-based approach applied to retinal blood vessel segmentation," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 9, pp. 2538–2548, 2012.

[45] Q. Li, B. Feng, L. Xie, P. Liang, H. Zhang, and T. Wang, "A cross-modality learning approach for vessel segmentation in retinal images," *IEEE Transactions on Medical Imaging*, vol. 35, no. 1, pp. 109–118, 2016.

[46] J. I. Orlando, E. Prokofyeva, and M. B. Blaschko, "A discriminatively trained fully connected conditional random field model for blood vessel segmentation in fundus images," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 1, pp. 16–27, 2016.

[47] A. Dasgupta and S. Singh, "A fully convolutional neural network based structured prediction approach towards the retinal vessel segmentation," in *IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, 2017, pp. 248–251.

[48] Z. Yan, X. Yang, and K. T. Cheng, "Joint segment-level and pixel-wise losses for deep learning based retinal vessel segmentation," *IEEE Transactions on Biomedical Engineering*, pp. 1–1, 2018.

[49] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi, Eds., Cham, 2015, pp. 234–241, Springer International Publishing.

[50] Reza Azad, Maryam Asadi-Aghbolaghi, Mahmood Fathy, and Sergio Escalera, "Bi-directional convlstm u-net with densley connected convolutions," *CoRR*, vol. abs/1909.00166, 2019.

[51] Md Zahangir Alom, Chris Yakopcic, Mahmudul Hasan, Tarek M. Taha, and Vijayan K. Asari, "Recurrent residual U-Net for medical image segmentation," *Journal of Medical Imaging*, vol. 6, no. 1, pp. 1 – 16, 2019.

[52] T. A. Soomro, A. J. Afifi, A. Ali Shah, S. Soomro, G. A. Baloch, L. Zheng, M. Yin, and J. Gao, "Impact of image enhancement technique on cnn model for retinal blood vessels segmentation," *IEEE Access*, vol. 7, pp. 158183–158197, 2019.

[53] Liang-Chieh Chen, Yukun Zhu, George Papandreou, Florian Schroff, and Hartwig Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," *Lecture Notes in Computer Science*, p. 833–851, 2018.

978-1-7281-6926-2/20/$31.00 ©2020 IEEE