SD-Unet: A Structured Dropout U-Net for Retinal Vessel Segmentation

Changlu Guo, Márton Szemenyei
Department of Control Engineering and Information Technology
Budapest University of Technology and Economics
Budapest, Hungary
clguo.ai@gmail.com

Yang Pei, Yugen Yi *, *Member*, *IEEE*
School of Software
Jiangxi Normal University
Nanchang, China
yiyg510@jxnu.edu.cn

Wei Zhou
College of Computer Science
Shenyang Aerospace University
Shenyang, China
zhouweineu@outlook.com

***Abstract*—At present, artificial visual diagnosis of fundus diseases has low manual reading efficiency and strong subjectivity, which easily causes false and missed detections. Automatic segmentation of retinal blood vessels in fundus images is very effective for early diagnosis of diseases such as the hypertension and diabetes. In this paper, we utilize the U-shaped structure to exploit the local features of the retinal vessels and perform retinal vessel segmentation in an end-to-end manner. Inspired by the recently DropBlock, we propose a new method called Structured Dropout U-Net (SD-Unet), which abandons the traditional dropout for convolutional layers, and applies the structured dropout to regularize U-Net. Compared to the state-of-the-art methods, we demonstrate the superior performance of the proposed approach.**

Keywords: retinal segmentation; DropBlock; SD-Unet; U-Net; dropout;

I. INTRODUCTION

The varied structure of retinal blood vessels can be used to diagnose certain pathological diseases such as hypertension or diabetes. Hypertensive Retinopathy (HR) is a retinal disease caused by hypertension [1], which increases blood vessel curvature, or causes the narrowing of blood vessels [2]. Diabetic Retinopathy (DR) is another common retinal disease induced by elevated blood glucose levels, accompanied by the swelling of the retinal blood vessels [3]. Hence, retinal blood vessels distilled from fundus images can help ophthalmologists detect and diagnose early conditions of certain serious diseases.

Many researchers have developed more accurate retinal vessel segmentation algorithms. Existing methods mainly include manual and automated segmentation procedures. Naturally, manual segmentation is not only time-consuming, but also requires a large quantity of specialists. Hence, researchers pay more attention to algorithm-based automatic segmentation to reduce the burden of manual segmentation. In the past two decades, numerous studies have focused on blood vessel segmentation from retinal fundus images, and researchers have proposed numerous fully automated methods [4-6]. Due to the complexity of the retinal vessel structure, there are still various challenges, such as segmentation of the vessel structure, segmentation near the bifurcation and intersection, and segmentation of anomalies.

In recent studies, deep learning was shown to have excellent performance in pathological image analysis tasks. In particular, the retinal blood vessel segmentation has been dominated by a variety of deep learning methods. Liskowski et al. [4] proposed a network consisting of a stack of convolutional layers and three fully connected layers, but required the input of small patches from fundus images to train. Maninis et al. [5] proposed a network named DRIU, which extracted the intermediate feature maps from a pretrained VGG-16 [7] network on ImageNet, then sampled and spliced them using transposed convolution, and applied 1×1 convolutional layer. Fu et al. [6] integrated the convolutional neural network and conditional random field layers to form a new integrated deep network. Compared with traditional Convolutional Neural Networks (CNN), U-Net [8] based on Fully Convolutional Networks (FCN) has attracted considerable attention since it can learn coarse-to-fine representations. Yan et al. [9] simultaneously trained the U-Net model by appending two separate branches with joint losses, one with a segment-level loss and the other with pixel-level losses.

In our study, we propose an FCN-based network called SD-Unet, which greatly reduces the overfitting problem of U-Net and enhances the ability of segmenting blood vessels end-to-end. In addition, we use the whole fundus images to train our networks. Similarly to U-Net, SD-Unet is designed as a U-shape structure, in which the up-sampling operators with a great number of feature channels are symmetrically stacked with the traditional CNN. Therefore, the context information can be captured and propagated to the higher resolution layer. Our specific contributions are as follows:

(1) Inspired by the DropBlock, a structured dropout is added to each convolutional layer to propose the Structured Dropout U-Net (SD-Unet).

(2) SD-Unet is tested on three public retinal fundus image datasets with annotated blood vessel authentic markers: DRIVE [11], STARE [12] and CHASE DB1 [13].

(3) We demonstrate that the SD-Unet performs better than U-Net on three datasets. Moreover, the SD-Unet

* Corresponding author

This work is supported by the China Scholarship Council, China, the Stipendium Hungaricum Scholarship, Hungary and the National Natural Science Foundation of China under Grants 61602221 and 61672150, the Scientific Research Fund Project of Liaoning Provincial Department of Education under Grant JYT19040, Liaoning Provincial Department of Science and Technology Natural Fund Guidance Program under Grant 2019-ZD-0234.

2471-7819/19/$31.00 ©2019 IEEE
DOI 10.1109/BIBE.2019.00085

achieves the highest accuracy on DRIVE and CHISE DB1 compared to some state-of-the-art methods.

II. METHODOLOGY

Our goal is to establish a deep learning model to segment the retinal blood vessels. Inspired by U-Net [8] and DropBlock [14], we propose a Structured Dropout U-net (SD-Unet), which is designed by combining the DropBlock and U-Net architectures. In this section, we detail the proposed method. Figure 1 shows the architecture of SD-Unet. SD-Unet consists of three down-sampling blocks and three up-sampling blocks, which are connected by skip connections. Each down-sampling block consists of two consecutive 3×3 convolutional layers and a max pooling layer. For each convolutional layer, there is a DropBlock layer and a ReLU layer to follow it. The down-sampling block is almost identical to the up-sampling block, except that it uses the transposed convolutional layer instead of the pooling layer.

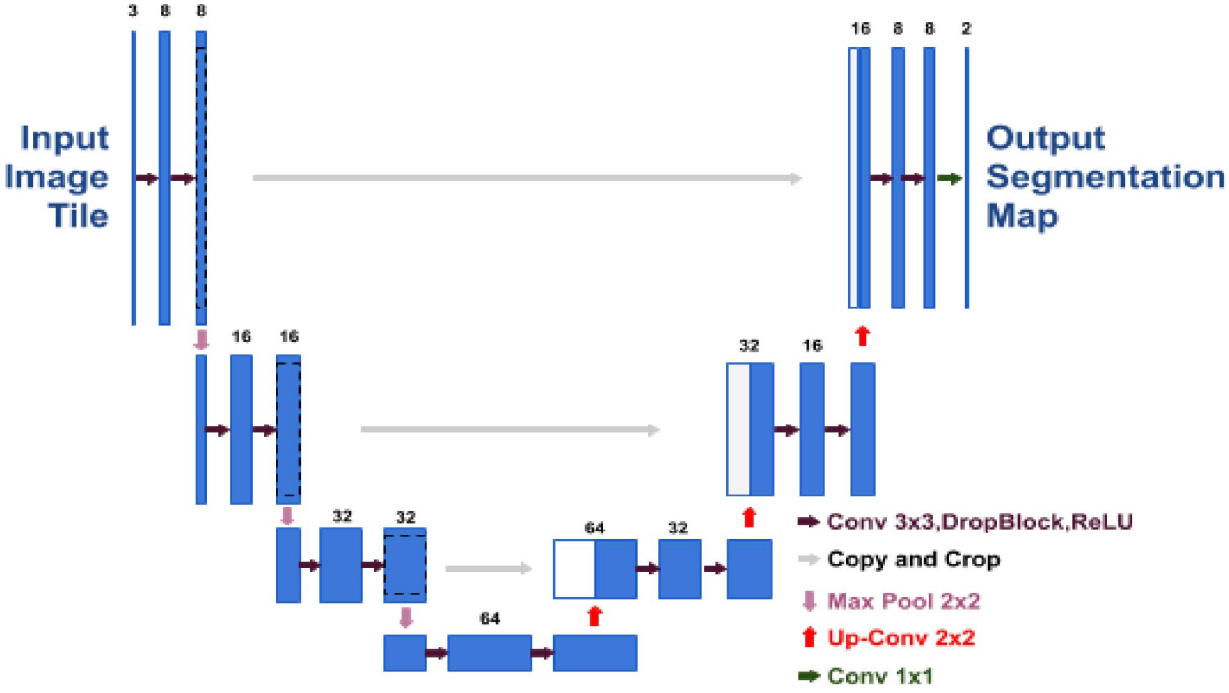

Figure 1. The architecture of SD-Unet.

A. U-NET

In the field of medical image analysis, CNNs perform better on the benchmark datasets than traditional methods. This is because the convolution operations take advantage of the structural information in medical images well [25, 26]. U-Net (as well as some other U-shaped architectures) has been shown to achieve robust and accurate performance in cardiac MR, brain tumors [15] and abdominal CT [16, 17] image segmentation tasks. The emergence of U-Net has brought great prospects for deep learning in the field of medical image analysis. For this reason, numerous deep learning methods based on U-Net were proposed.

U-Nets have shown superb performance and can be run efficiently by using GPUs, so they are often used for image segmentation tasks. The later advantage is primarily related to extract image features from different image scales. U-Nets implement the fusion of features at different scales, which improves the accuracy of the model. The coarse feature map captures the context information and highlights the classification and position of the foreground object. In order to link the coarse-level and fine-level dense predictions, the feature maps distilled from different scales are merged by a skip connection.

B. DropBlock for Image Analysis

To avoid the problem of overfitting in computer vision, various regularization methods are employed in deep neural networks, such as dropout [19]. Dropout is a simple method to prevent neural networks from overfitting. The main feature of dropout is that it randomly discards features during training. This strategy is suitable for fully connected layers, while it is less suitable for convolutional layers because of the spatial correlations among the active cells. However, the semantic information about the input can still be sent to the following layer, which may also cause overfitting. For this reason, a structured dropout is needed to regularize convolutional networks. Recently, Golnaz Ghiasi et al. [14] proposed DropBlock to regularize convolutional architectures. In DropBlock, a contiguous region of a feature map, (i.e.: features in a block), is discarded together. Since DropBlock drops features in contiguous area, the network necessarily counts on evidence from the other areas that fits the data (see Figure 2).

Similar to dropout, DropBlock is a simple regularization method. The main difference is that DropBlock removes contiguous areas from a layer's feature map instead of discarding separate random units. The DropBlock layer has two essential parameters: $block\_size$ and γ. The parameter $block\_size$ is the size of the block to be deleted. When $block\_size$=1, DropBlock is identical to dropout [18], and when $block\_size$ covers the complete feature map, DropBlock is identical to Spatial Dropout [19]. The parameter γ controls the number of activation units to be deleted which can be calculated as:

$$\gamma = \frac{(1-keep\_prob)\times(w\times h)}{block\_size^2\times(w-block\_size+1)\times(h-block\_size+1)} \quad (1)$$

where $keep\_prob$ is the probability of keeping the neuron in a classical dropout state. The size of the valid seed area is $(w-block\_size+1)\times(h-block\_size+1)$ where w and h are the width and height of the feature map, respectively. DropBlock has an important nuance, that is, there may be some overlap in the removed blocks, and the previous equation is just an approximation.

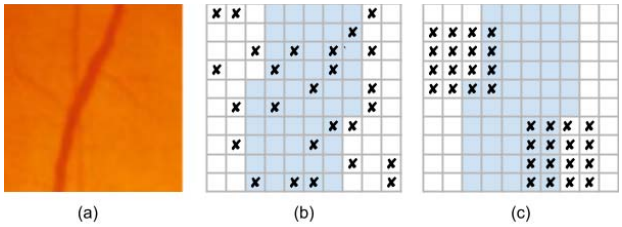

Figure 2. (a) A region of the original image given to the convolutional neural network. The blue areas in (b) and (c) have activation units that include information in (a). Randomly deleting activations is not sufficient for discarding semantic information because neighboring activations have nearly relevant information. Alternately, discarding successive areas may drop some semantic information (e.g., non-vessel regions) and thus force the rest units to discover features for blood vessel and background classification of the input image.

C. DropBlock in U-Net Model

Because of the success of dropout, many CNN architectures adopt the method. In most cases, dropout is primarily employed in the fully connected layer of CNNs. Dropout can randomly discard features, which makes it effective for the fully connected layer. However, the dropout is not very effective for FCN models. Therefore, FCN architectures, such as U-Net require an effective way to remove certain semantic information to prevent network from overfitting.

In order to solve this problem, we propose structured dropout to regularize U-Net. In our proposed architecture, a new convolutional unit is constructed, in which each convolutional layer is followed by a DropBlock layer and a ReLU layer, as shown in Figure 3. Dropping contiguous regions using DropBlock may remove certain semantic information and thus force the rest units to discover features for categorizing the input image [14].

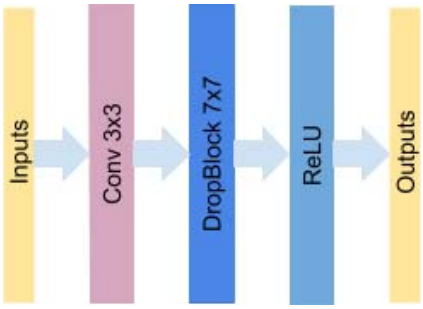

Figure 3. Convolution unit after adding DropBlock

III. EXPERIMENTS AND RESULTS

In this section, we introduce the implementation of SD-Unet and systematically compare SD-Unet and U-Net on different datasets using various evaluation metrics. We first present the results achieved on test set using different metrics, and illustrate the segmentation results on a few images from the test set. Moreover, we compare SD-Unet with state-of-the-art methods, most of which are deep neural networks as well as a small number of traditional segmentation methods. All experiments are performed using the NVIDIA TITAN Xp GPU and the Tensorflow and Keras frameworks.

A. Datasets and Preparation

We evaluate the performance on three publicly available datasets: DRIVE [11], STARE [12] and CHASE [13]. DRIVE was obtained from the Dutch diabetic retinopathy (DR) screening program, which includes 20 training images and 20 test images with a resolution of 565×548. STARE (Structural Analysis of the Retina) contains 20 fundus images with a resolution of 700×605. We manually divide the STARE dataset into 10 training images and 10 test images. CHASE dataset contains 28 images of 14 patients and each patient has 2 images with a resolution of 999×960. The manual annotations of retinal blood vessels created by experts can be used as the ground truth (Figures 4-6).

U-Net's VALID convolution reduces the resolution of the Feature Map, but we want to be able to retain the appropriate resolution to the last layer Feature Map. For DRIVE, we resize it to 700×700 by padding it with zero in four margins. For STARE and CHASE, we resize them to 790×790 and 1100×1100, respectively. In order to make the results more reasonable, we crop the segmentation results to the original size.

To augment the data, we perform the following four operations on the original images of the three datasets: 1) random arbitrary angle (0~360 degrees) rotation; 2) color jitter; 3) adding Gaussian noise; 4) horizontal and vertical flips. Therefore, three new images are produced for each and 12 images are generated from an individual image.

B. Evaluation Metrics

We use *PPV, TNR, TPR, ACC* and *AUC* to evaluate the performance of our model on retinal vessel segmentation. *PPV* represents the Positive Predictive Value and shows the ratio of true positive samples to all predicted positive samples. *TNR* represents the True Negative Rate and

indicates the negative proportion of the correct identification. *TPR* represents the True Positive Rate and demonstrates the positive proportion of the correct identification. *ACC* represents the Accuracy which measures the ratio of the correctly categorized pixels to all the pixels in the dataset. *AUC* represents the Area Under Curve, which is defined as the area enclosed by the Receiver Operating Characteristic (ROC) curve and the coordinate axis. These metrics are calculated as follows:

$$PPV = \frac{TP}{TP + FP} \tag{2}$$

$$TNR = \frac{TN}{TN + FP} \tag{3}$$

$$TPR = \frac{TP}{TP + FN} \tag{4}$$

$$ACC = \frac{TP + TN}{TP + FP + TN + FN} \tag{5}$$

where *TP* means the count of true positive samples; *TN* is the count of true negative samples; while *FN* and *FP* represent the number of false negative and false positive samples respectively. In addition, we also evaluated SD-Unet and U-Net performance using F-measure (*F1*) and Jaccard Similarity (*JS*). Here *SR* represents the segmentation result and *GT* represents the ground truth.

$$F1 = 2 \times \frac{PPV \times TPR}{PPV + TPR} \tag{6}$$

$$JS = \frac{|SR \cap GT|}{|SR \cup GT|} \tag{7}$$

C. Comparisons with U-Net

We compared U-Net and SD-Unet using the DRIVE, STARE and CHASE datasets. We evaluate two versions of U-Net: U-Net means that dropout is not used, that is, we keep each neuron output with probability 1. U-Net* is the case of using dropout, we keep each neuron output with probability 0.75. As mentioned earlier, we divide the data into training, validation, and test sets. We use the training set to train the three models from scratch and initialize the weights with random values. Training is performed with a min-batch size of 4 and the number of epochs is 100. Since the number of training data for the three data sets is different, we set the training iterations of DRIVE, STARE and CHASE to 64, 48 and 40 respectively. We set the feature map size of the first convolutional layer to 8 and we use binary cross-entropy as the loss function. An Adam optimizer is used, and the learning rate is set to 0.001. For SD-Unet, we set the size of the block to be deleted from all feature maps to 7, regardless of the resolution of the feature map; for the best performance for each dataset, we keep each neuron output with probability 0.92, 0.93 and 0.93 for the DRIVE, STARE, and CHASE datasets respectively.

We evaluate our models utilizing test data from three datasets. *PPV*, *TNR*, *TPR*, *JS*, *F1-score*, *ACC* and *AUC* were compared, and the experimental results are shown in Tables I-III. As the results clearly demonstrate, SD-Unet accomplishes the best performance according to most of the indicators. Notably, SD-Unet reaches the highest global accuracy among the three models. The *ACC* values for U-Net, U-Net*, and SD-Unet are 0.9656, 0.9655, 0.9674 on DRIVE, 0.9705, 0.9674, 0.9725 on STARE and 0.9735, 0.9675, 0.9738 on CHASE, respectively.

Retinal blood vessels segmentation results are shown in Figures 4-6. As the results show, SD-Unet produces more pronounced blood vessel segmentation. The proposed SD-Unet performs better on non-vascular areas, and it can also detect weak blood vessels that may be lost in U-Net and U-Net*, so the produced segmentation retains more details. It is worth mentioning that U-Net* apparently loses some weak blood vessel areas. By using the DropBlock, SD-Unet overcomes this shortcoming. The experimental results lead to the conclusion that the introduction of the structured dropout form of the SD-Unet architecture has superior performance for retinal vessel segmentation.

TABLE I. TESTED RESULTS ON **DRIVE**

Models	DRIVE						
	PPV	*TNR*	*TPR*	*JS*	*F1*	*ACC*	*AUC*
U-Net	0.8511	0.9872	0.7554	0.9656	0.7979	0.9656	0.9811
U-Net*	**0.8845**	**0.9911**	0.7003	0.9655	0.7782	0.9655	0.9735
SD-Unet	0.8335	0.9848	**0.7891**	**0.9674**	**0.8042**	**0.9674**	**0.9836**

TABLE II. TESTED RESULTS ON **STARE**

Models	STARE						
	PPV	*TNR*	*TPR*	*JS*	*F1*	*ACC*	*AUC*
U-Net	**0.8628**	**0.9907**	0.7221	0.9745	0.7843	0.9705	0.9830
U-Net*	0.8302	0.9871	0.7144	0.9718	0.7605	0.9674	0.9714
SD-Unet	0.8605	0.9899	**0.7548**	**0.9763**	**0.8020**	**0.9725**	**0.9850**

TABLE III. TESTED RESULTS ON **CHASE**

Models	CHASE						
	PPV	*TNR*	*TPR*	*JS*	*F1*	*ACC*	*AUC*
U-Net	0.7941	0.9867	0.7955	0.9735	0.8043	0.9735	0.9869
U-Net*	**0.8536**	0.9866	**0.7786**	0.9675	**0.8123**	0.9675	0.9843
SD-Unet	0.8486	**0.9900**	0.7559	**0.9738**	0.7978	**0.9738**	**0.9872**

D. Comparisons with Exiting Methods

Finally, we compare our approach with several state-of-the-art methods. In Tables IV-VI, we summarize the release year and the performance on the DRIVE, STARE, and CHASE datasets. The results demonstrate that the SD-Unet architecture performs best on DRIVE and CHASE datasets. It achieves the highest ACC of 0.9674 and 0.9738 and the highest AUC of 0.9836 and 0.9872, which indicates that the performance of SD-Unet is better than previous works of retinal vessel segmentation. The above results clearly show that SD-Unet is an effective method for retinal vessel segmentation.

TABLE IV. COMPARISONS WITH EXISTING METHODS ON **DRIVE** DATASET

Method	Year	*PPV*	*TNR*	*TPR*	*ACC*	*AUC*
Li et al. [20]	2015	-	0.9816	0.7569	0.9527	0.9738
Azzopardi et al.[21]	2015	-	0.9704	0.7655	0.9442	0.9614
Liskowski et al. [4]	2016	-	0.9807	0.7811	0.9535	0.9790
Fu et al. [7]	2016	-	-	0.7603	0.9523	-
Chen et al. [22]	2017	-	0.9735	0.7426	0.9453	0.9516
Roychowdhury et al. [23]	2017	–	0.9830	0.7250	0.9520	0.9620
Alom et al. [24]	2018	-	0.9813	0.7792	0.9556	0.9784
SD-Unet	**2019**	**0.8335**	**0.9848**	**0.7891**	**0.9674**	**0.9836**

TABLE V. COMPARISONS WITH EXISTING METHODS ON **STARE** DATASET

Method	Year	*PPV*	*TNR*	*TPR*	*ACC*	*AUC*
Li et al. [20]	2015	-	0.9844	0.7726	0.9628	0.9879
Azzopardi et al.[21]	2015	-	0.9701	0.7716	0.9497	0.9563
Fu et al. [7]	2016	-	-	0.7412	0.9585	-
Liskowski et al. [4]	2016	-	0.9862	0.8554	0.9729	0.9928
Chen et al. [22]	2017	-	0.9696	0.7295	0.9449	0.9557
Roychowdhury et al. [23]	2017	-	0.9730	0.7720	0.9510	0.9690
Alom et al. [24]	2018	-	0.9862	0.8298	0.9712	0.9914
SD-Unet	**2019**	**0.8605**	**0.9899**	**0.7548**	**0.9725**	**0.9850**

TABLE VI. COMPARISONS WITH EXISTING METHODS ON **CHASE** DATASET

Method	Year	*PPV*	*TNR*	*TPR*	*ACC*	*AUC*
Li et al. [20]	2015	-	0.9793	0.7507	0.9581	0.9716
Azzopardi et al.[21]	2015	-	0.9587	0.7585	0.9387	0.9487
Fu et al. [7]	2016	-	-	0.7130	0.9489	-
Roychowdhury et al. [23]	2017	-	0.9824	0.7201	0.9530	0.9532
Alom et al. [24]	2018	-	0.9820	0.7759	0.9634	0.9715
SD-Unet	**2019**	**0.8486**	**0.9900**	**0.7559**	**0.9738**	**0.9872**

IV. CONCLUSION

Deep learning is a method used to characterize and to learn from data, and it is broadly used in the field of medical image analysis. In this study, we proposed a fully convolutional neural network based on the U-Net architecture called SD-Unet, which is applied to retinal vessel segmentation (pixel-wise classification). Our work integrates a structured form dropout called DropBlock into standard U-Net to improve the performance. DropBlock is a simple regularization method for convolutional neural networks. As a U-shaped architecture, SD-Unet aims to capture the context through the down-sampling and combine the coarse-level feature mapping with the fine-level mapping to achieve accurate positioning of the up-sampling. Moreover, with DropBlock, SD-Unet can discard some features of contiguous regions during training, thus it can alleviate the overfitting problem.

It is the first time that SD-Unet has been used for retinal vessel segmentation. To verify the effectiveness of the proposed approach, three publicly available datasets (DRIVE, STARE and CHASE DB1) were used to train and test in our experiments. Meanwhile, comparisons of SD-Unet and the existing state-of-the-art methods were presented. The results demonstrate that the performance of SD-Unet surpasses the state-of-the-art methods in retinal vessel segmentation.

In our future work, we plan to use more retinal datasets to verify the performance of SD-Unet. Moreover, we also plan to extend SD-Unet to other aspects of medical image segmentation, such as nuclear segmentation, lung segmentation, and more.

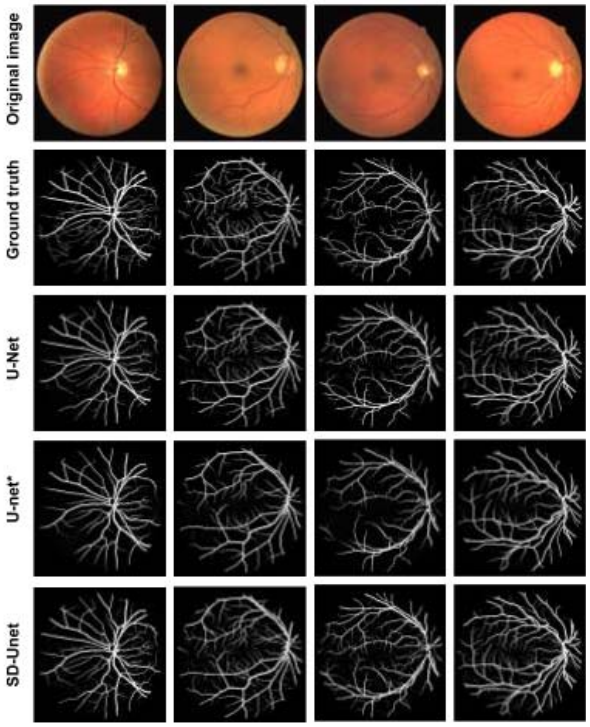

Figure 4. Segmentation results on DRIVE

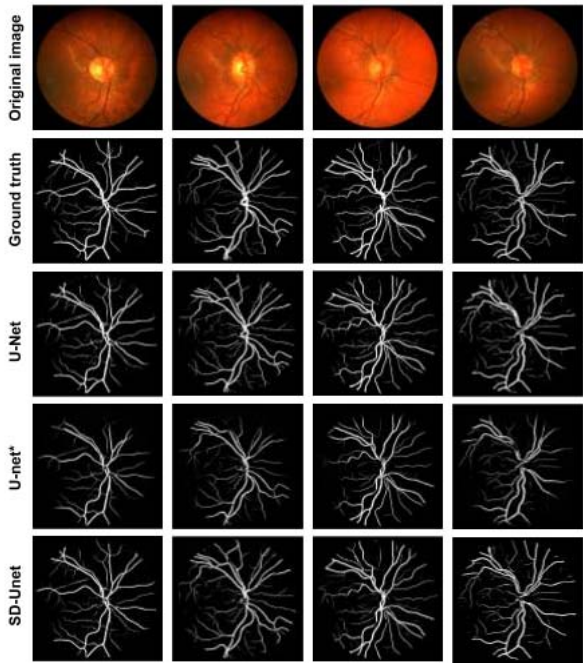

Figure 5. Segmentation results on CHASE

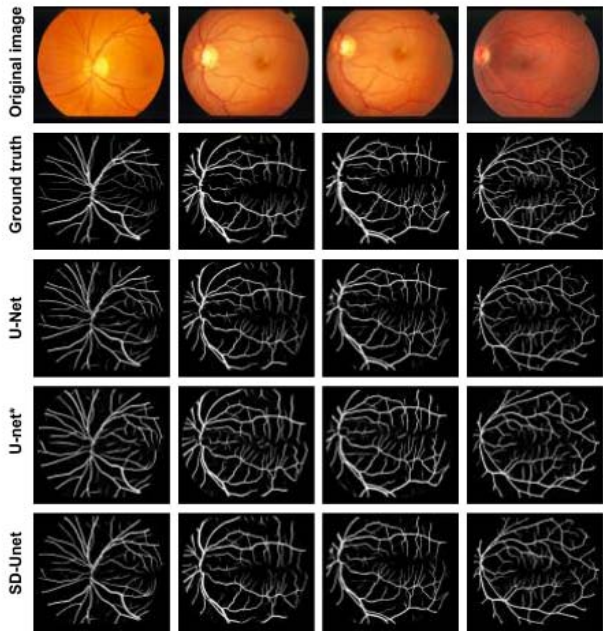

Figure 6. Segmentation results on STARE

REFERENCES

[1] S. Irshad and M. U. Akram, Classification of retinal vessels into arteries and veins for detection of hypertensive retinopathy, in Biomedical Engineering Conference (CIBEC), 2014 Cairo International. IEEE, pp. 133–136, 2014.

[2] C. Y.-l. Cheung, Y. Zheng, W. Hsu, M. L. Lee, Q. P. Lau, P. Mitchell, J. J. Wang, R. Klein, and T. Y. Wong, Retinal vascular tortuosity, blood pressure, and cardiovascular risk factors. Ophthalmology, vol. 118, no. 5, pp. 812–818, 2011.

[3] I. S. Jacobs and C. P. Bean, Fine particles, thin films and exchange anisotropy, in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, pp. 271-350, 1963.

[4] P. Liskowski and K. Krawiec. Segmenting retinal blood vessels with deep neural networks. IEEE transactions on medical imaging, 35(11):2369-2380, 2016.

[5] K. K. Maninis, J. Pont-Tuset, P. Arbelez, and L. V. Gool. Deep Retinal Image Understanding. In Medical Image Computing and Computer-Assisted Intervention (MICCAI), pp. 140-148, 2016.

[6] H. Fu, Y. Xu, D. W. K. Wong, and J. Liu, Retinal vessel segmentation via deep learning network and fully-connected conditional random fields, in IEEE International Symposium on Biomedical Imaging, pp. 698-701, 2016.

[7] K. Simonyan and A. Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition. CoRR, abs/1409.1556, 2014.

[8] O. Ronneberger, P. Fischer, and T. Brox, U-Net: Convolutional networks for biomedical image segmentation, in International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 234-241, 2015.

[9] Z. Yan, X. Yang, and K. Cheng. Joint Segment-Level and Pixel-Wise Losses for Deep Learning Based Retinal Vessel Segmentation. IEEE Transactions on Biomedical Engineering, 65(9):1912-1923, 2018.

[10] J. Dai, H. Qi, Y. Xiong, Y. Li, G. Zhang, H. Hu, and Y. Wei, Deformable Convolutional Networks, in International Conference on Computer Vision, pp. 764-773, 2017.

[11] 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.

[12] A. 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, 19(3):203-210, 2000.

[13] M. M. Fraz, P. Remagnino, A. Hoppe, et al. An Ensemble Classification-Based Approach Applied to Retinal Blood Vessel Segmentation. IEEE Transactions on Biomedical Engineering, 59(9):2538-2548, 2012.

[14] G. Ghiasi, T.-Y. Lin, and Q. V. Le. DropBlock: A regularization method for convolutional networks. In Neural Information Processing Systems, 2018.

[15] Kamnitsas, K., Bai, W., Ferrante, E., et al. Ensembles of multiple models and architectures for robust brain tumour segmentation. In: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. pp. 450-462. Cham ,2018.

[16] Roth, H.R., Lu, L., Ferrante, E., et al. Spatial aggregation of holistically-nested convolutional neural networks for automated pancreas localization and segmentation. Medical Image Analysis 45, 94-107, 2018.

[17] Roth, H.R., Oda, H., Hayashi, Y., et al. Hierarchical 3D fully convolutional networks for multi-organ segmentation. arXiv preprint arXiv:1704.06382, 2017.

[18] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, et al. Dropout: A simple way to prevent neural networks from overfitting. The Journal of Machine Learning Research, 15(1):1929-1958, 2014.

[19] Jonathan Tompson, Ross Goroshin, Arjun Jain, Yann LeCun, and Christoph Bregler. Efficient object localization using convolutional networks. In Proceedings IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015.

[20] Q. Li, B. Feng, L. P., et al. A Cross-Modality Learning Approach for Vessel Segmentation in Retinal Images, IEEE Transactions on Medical Imaging, vol. 35, no. 1, pp. 109-118, 2015.

[21] G. Azzopardi, N. Strisciuglio, M. Vento, and N. Petkov, Trainable COSFIRE filters for vessel delineation with application to retinal images, Medical Image Analysis, vol. 19, no. 1, pp. 46-57, 2015.

[22] Y. Chen, A Labeling-Free Approach to Supervising Deep Neural Networks for Retinal Blood Vessel Segmentation. arXiv preprint arXiv:1704.07502, 2017.

[23] S. Roychowdhury, D. D. Koozekanani, and K. K. Parhi, Blood Vessel Segmentation of Fundus Images by Major Vessel Extraction and Subimage Classification, IEEE Journal of Biomedical & Health Informatics, vol. 19, no. 3, pp. 1118-1128, 2017.

[24] M. Z. Alom, M. Hasan, C. Yakopcic, T. M. Taha, and V. K. Asari, Recurrent Residual Convolutional Neural Network based on U-Net (R2U-Net) for Medical Image Segmentation, arXiv preprint arXiv:1802.06955, 2018.

[25] M. Khened, V. A. Kollerathu, G. Krishnamurthi. Fully convolutional multi-scale residual DenseNets for cardiac segmentation and automated cardiac diagnosis using ensemble of classifiers. Medical image analysis, 2019, 51: 21-45.

[26] F. Liao, M. Liang, Z. Li, et al. Evaluate the Malignancy of Pulmonary Nodules Using the 3-D Deep Leaky Noisy-or Network IEEE transactions on neural networks and learning systems, DOI: 10.1109/TNNLS.2019.2892409, 2019.