

Automatic Fundus Image Segmentation for Diabetic Retinopathy Diagnosis by Multiple Modified U-Nets and SegNets

Swathi Ananda*, Daichi Kitahara[†], Akira Hirabayashi[†], and K. R. Udaya Kumar Reddy*

*Department of Computer Science and Engineering, NMAM Institute of Technology, Nitte, Karkala, Karnataka, India

[†]College of Information Science and Engineering, Ritsumeikan University, Kusatsu, Shiga, Japan

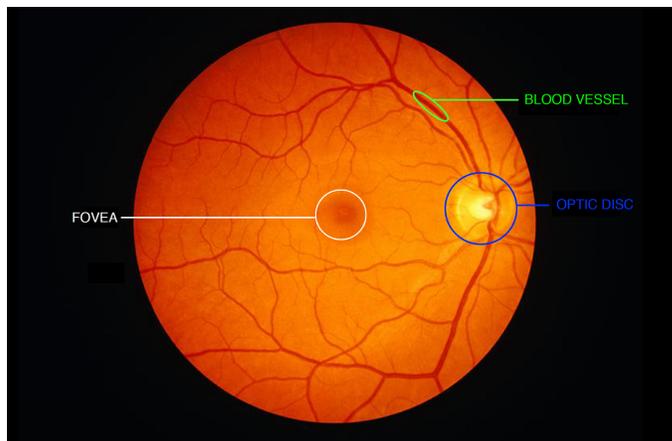
Abstract—Diabetes mellitus leads to damage of the retina by a high blood sugar level. This disease is called diabetic retinopathy (DR), and it is one major cause of blindness among working-aged people. DR affects about 80% of patients who have had diabetes for twenty years or more. The longer a period of diabetes is, the higher the risk of developing DR is. In order to prevent the blindness caused by DR, accurate DR diagnosis from a retinal fundus image is important. Recently, deep learning techniques play a significant role in the field of computer vision. When we apply deep learning to segmentation of abnormal parts in fundus images, two major problems arise. One is that the number of available data is insufficient to train a deep neural network. The other is that the sizes of the abnormal parts are quite different depending on the type of the disease, which leads to low segmentation accuracy of small diseases. These two problems make the fundus image segmentation challenging. In this paper, we propose a segmentation method using multiple deep neural networks. To train the deep neural networks from a small number of data, we use data augmentation as preprocessing and adopt the Dice coefficient with the binary cross entropy as a loss function. Moreover, to improve the segmentation accuracy of small diseases, e.g., microaneurysms, we construct one individual network for each type of the disease. In experiments, the networks are trained from IDRiD dataset and tested for MESSIDOR dataset. We compare and discuss the accuracy of the proposed method with modified U-Nets and SegNets.

I. INTRODUCTION

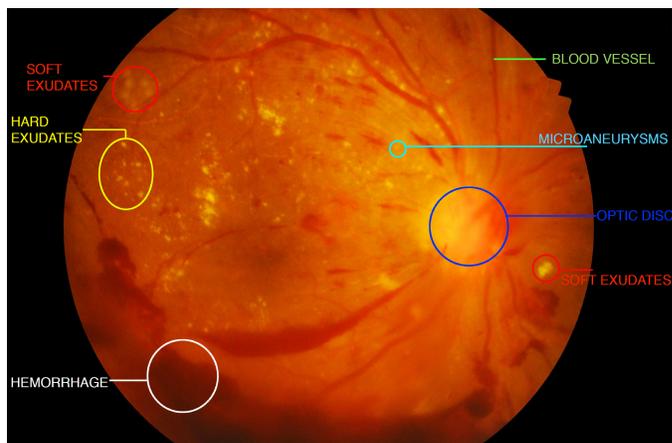
Diabetic retinopathy (DR) is one major cause of blindness. It is a chronic eye disease, and 80% of all people suffering from diabetes will face this problem [1]–[3]. DR causes damage to the retina of eyes, but more than 90% of the blindness can be prevented if DR is detected at an early stage. Diabetic patients must regularly undergo the screening test to check the onset of DR. However, manual examinations by ophthalmologists take much time, and the number of experts is not sufficient to meet the growing demand for screening. As a result, development of an automatic diagnosis system for DR has been desired.

Fundus photography is the most commonly-used screening for DR diagnosis. Fundus images of normal and damaged eyes are shown in Figs. 1(a) and 1(b), respectively. In every fundus image, there always exists one optic disc, which is also called the blind spot because there are no photoreceptors in this part. The optic disc represents the beginning of the optic nerve and is shown as a bright circle in the fundus image. Blood vessels come into the retina from the optic disc and spread in various

This work was done while the first author stayed in Kusatsu, Japan for 6 months as an internship student of Ritsumeikan University Research Program.



(a) Fundus image of a normal eye



(b) Fundus image of a damaged eye

Fig. 1. Typical fundus images of normal and damaged eyes. Microaneurysms, hard & soft exudates, and hemorrhages appear as abnormal parts in (b).

directions as capillaries to supply retinal cells with oxygen and nutrition. If DR develops and progresses, then microaneurysms, hard & soft exudates, and hemorrhages appear as shown in Fig. 1(b). The microaneurysms are the first occurrence of DR. These happen when tiny blood vessels in the retina begin to swell. The hard exudates are lipid leakages from some blood vessels and are the most visible signs of DR at an early stage. Dot or blot hemorrhages also occur from damaged capillaries as less-visible signs at the early stage. At the next stage of DR,

the soft exudates, which are more blurred compared with hard ones, appear as died and dilated ganglion cell axons. Moreover, blood flow is locally stopped, i.e., the ischemia happens, and the hemorrhages become larger. When DR further progresses, abnormal new blood vessels are created. These vessels are thin and fragile, and if they are broken, then vitreous hemorrhages occur and the hard exudates become larger and larger.

Automatic fundus image segmentation has been studied and many methods have been developed based on traditional image processing techniques [4]–[8] and machine learning techniques [9]–[21]. In the machine learning approaches, by using training data, the accuracy of segmentation can be improved compared with the traditional image processing approaches. In particular, deep learning techniques [12]–[21] are the most trendy machine learning because they have been achieving tremendous success in the fields of computer vision and medical image processing.

In [12]–[18], deep convolutional neural networks (CNNs) are used. Gargeya and Leng [12] and Lam *et al.* [13] constructed CNNs for detection of the stage of DR, and executed fundus image segmentation roughly as abnormality heat maps by visualization of the learned features. P. Prentašić and S. Lončarić [14], Yu *et al.* [15], Perdomo *et al.* [16], and Khojasteh *et al.* [17] constructed CNNs for exudate segmentation, and achieved 77% sensitivity for DRiDB dataset¹ [22], 88% sensitivity for e-ophtha dataset² [23], 99% sensitivity for e-ophtha dataset, and 99% sensitivity for DIARETDB1 v1 dataset³ [24], respectively. Note that the quality of fundus images and segmentation labels are different depending on the dataset, and we cannot directly compare the above results. Orlando *et al.* [18] constructed a CNN and combined it with a random forest for segmentation of red lesions composed of hemorrhages and microaneurysms, and achieved 48% sensitivity for DIARETDB1 v1 dataset and 36% sensitivity for e-ophtha dataset. From these papers, one finds that segmentation of exudates is relatively easy compared with those of hemorrhages and microaneurysms since the contrast of exudate parts is high while that of red lesions is low.

When we apply deep learning to fundus image segmentation, two major problems arise. One is that the number of available data is insufficient to train deep CNNs. The other is the distribution of data is very imbalanced, i.e., the number of normal pixels is much greater than that of abnormal pixels, and the sizes of the abnormal parts are also different depending on the type of the disease and the stage of DR. These two problems make the fundus image segmentation challenging. Zheng *et al.* [19] constructed a conditional generative adversarial network (CGAN) [25] for resolving the above problems. Based on the idea of Pix2Pix [26], they generated label-preserving virtual fundus images of the minority class by the CGAN as data augmentation. Then, they constructed an ensemble CNN based on U-Net [27] for exudate segmentation and achieved 90% sensitivity for HEI-MED dataset⁴ [4], 90% sensitivity for e-ophtha dataset, 93% sensitivity for DIARETDB1 v1 dataset, and 95%

sensitivity for MESSIDOR dataset⁵ [28]. However, they did not challenge more difficult segmentation of hemorrhages and microaneurysms. Moreover, the data augmentation method by the CGAN cannot be applied, e.g., to IDRiD dataset⁶ [29] since this dataset has no blood vessel labels.

Many researchers have focused on segmentation of only one disease, and there are few papers trying to segmentation of all types of diseases. Tan *et al.* [20] constructed a 10-layer CNN for simultaneous segmentation of exudates, hemorrhages and microaneurysms. For CLEOPATRA dataset⁷ [30], it achieved 87% and 71% sensitivities of exudates and red lesions, respectively. It also achieved 62% and 46% sensitivities of hemorrhages and microaneurysms. Saha *et al.* [21] used SegNet [31] for simultaneous segmentation of hard & soft exudates, hemorrhages, microaneurysms, and the optic disc. For IDRiD dataset, hard exudates and the optic disc can be segmented well while the results of soft exudates, hemorrhages, and microaneurysms are bad. Especially, microaneurysms can hardly be segmented because the number of microaneurysm pixels is very few.

In this paper, for mainly improving the segmentation accuracy of small diseases such as soft exudates, hemorrhages, and microaneurysms, we propose to construct one individual deep CNN for each type of the disease. Moreover, to train the deep CNNs from a small number of data, we use data augmentation as preprocessing and adopt the Dice coefficient with the binary cross entropy as a loss function. The Dice coefficient enables CNNs to learn fundus image features more robustly than the binary cross entropy when the data distribution is imbalanced. In experiments, the deep CNNs are trained from IDRiD dataset and tested for MESSIDOR dataset. We compare and evaluate the performance of the proposed method by using the original U-Nets, modified U-Nets and SegNets as the deep CNNs.

II. PROPOSED FUNDUS IMAGE SEGMENTATION

Simultaneous segmentation of different types of diseases in fundus images by a single CNN is a difficult task as shown in [20], [21]. Hence, in this paper, we construct multiple CNNs, and each CNN segments only a single class, i.e., the optic disc, hard exudates, soft exudates, hemorrhages, or microaneurysms. We train each network by using IDRiD dataset [29]. This dataset is recently released, and includes some fundus images which are difficult to be segmented due to the lower contrast than other datasets and larger abnormal parts caused by the progress of DR. Therefore, it can be expected that if each CNN achieves good segmentation results for IDRiD dataset, it also achieves good results for the other segmentable datasets. Since the data augmentation by a CGAN [19] is not applicable to IDRiD dataset, we increase training data simply by rotation of fundus images. To alleviate the imbalanced data problem, we combine the Dice coefficient with the binary cross entropy as a loss function. We adopt the original U-Net [27], its modified version by us, and SegNet [31] as candidates of CNNs, and compare each result.

¹https://ipg.fer.hr/ipg/resources/image_database

²<http://www.adcis.net/en/third-party/e-ophtha/>

³<http://www.it.lut.fi/project/imageret/diaretdb1/>

⁴<https://github.com/lgiancaUTH/HEI-MED>

⁵<http://www.adcis.net/en/third-party/messidor/>

⁶<https://iee-dataport.org/open-access/indian-diabetic-retinopathy-image-dataset-idrid>

⁷This dataset is not available publicly online.

A. Candidates of Individual Networks

1) *U-Net*: U-Net [27] was developed as a CNN for biomedical image segmentation. The important change from the standard fully convolutional networks (FCNs) is how to implement residual connections between an encoder and a decoder to preserve the information on boundaries while encoding the input image. The standard FCNs use simple additive operations with hyperparameters which are weights of the residual connections. U-Net uses concatenation operations to store channels of the intermediate images between encoder and decoder layers. The U-Net name signifies that the shape of the above network can be seen as ‘U’ due to the concatenation operations. Important parts and involved basic blocks in U-Net are explained below.

Encoder: The encoder part can be considered to be built from 6 blocks. Each block has two 3×3 convolutional layers with ReLU activation functions. This is followed by one 2×2 max pooling layer with a stride of 2. The pooling layer is not present in the 6th convolutional block because we do not need to reduce the image size any further. After every such blocks, the number of feature channels is doubled.

Decoder: The decoder part is a series of up-sampling operations coupled with convolutions and concatenation operations. The first layer of the decoder receives a highly cropped version, of the input image, separated into different feature channels by the encoder part. The input to each block in the decoder is first passed to one up-sampling layer which doubles the size of the image, and then into one 2×2 convolutional layer. After this, the corresponding outputs from the encoder blocks, which were saved earlier, are concatenated with the results of this 2×2 convolutional layer to preserve the boundary information which is otherwise lost in the pooling operations. Then, two more 3×3 convolutional layers are applied and the results are passed to the next block. Each convolutional layer uses the ReLU activation function.

Classification: For pixel-wise classification, the final layer is one 1×1 convolutional layer with as many channels as the number of segmentation classes. The activation function used here is the sigmoid function since we are only concerned with the binary classification, i.e., whether some disease is present or not present for each pixel. The activation function can be replaced with the softmax function for multiclass segmentation.

2) *Modified U-Net*: Although U-Net was designed to learn features from a relatively small number of data for biomedical image segmentation, it is difficult to obtain good segmentation results for every disease by the original U-Net due to the imbalance in fundus images (see Section III). We reduce the numbers of filters & layers, and modify the encoder as follows, inspired by Pix2Pix [26]. Figure 2 shows our modified U-Net.

- We replace ReLU with Leaky ReLU of a hyperparameter $\alpha = 0.2$, which helps network neurons to go dead (i.e., activation = 0) in the early part of training, and hence the learning speed and the training accuracy are improved.
- Before each pooling operation, we add a layer of batch normalization to smooth the loss function written in the next subsection, which improves the learning speed.

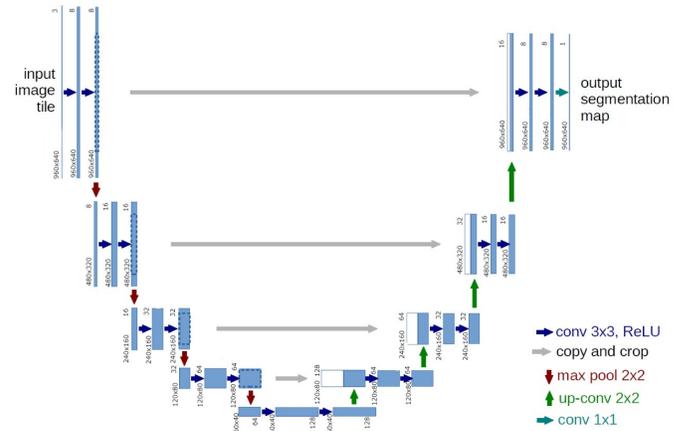


Fig. 2. Modified U-Net architecture.

3) *SegNet*: SegNet [31], in essence, is similar to the above described U-Net architecture. In fact, both models came out in the same year⁸ as further developments of the FCNs. Only one principal difference is the residual connection between the encoder and the corresponding decoder layer. Differently from U-Net, instead of concatenating outputs from the encoder and the decoder layers, SegNet sends only the max pooling indices from the encoder to the decoder. In pooling operations of the encoder, the indices of the maximum activations are selected and saved for later uses in the corresponding decoder layers. Then in the decoder, the same indices are used to place outputs of the previous convolutional layers while the rest of positions are set to 0s. This operation is called the unpooling, and acts a similar role to the up-sampling operation in U-Net. Note that the up-sampling operation in U-Net doubles the image size by 2×2 convolutions while the unpooling operation doubles the image size without any convolution. The other components of SegNet are explained below.

Encoder: Similar to U-Net, the encoder of SegNet is built from 5 basic blocks. Each block has two or three 3×3 convolutional layers followed by the batch normalization layers with ReLU activation functions. Each block is succeeded by one 2×2 max pooling layer with a stride of 2 except the 5th block because we do not need to reduce the image size any further. After every such blocks, the number of feature channels is doubled and the image size is reduced by half.

Decoder: The decoder part is built from 5 basic blocks, which are reversed versions of the blocks in the encoder. At the beginning of each block, the image size is doubled from the output of the previous block by the unpooling operation. After unpooling, each block in the decoder is composed of three or two 3×3 convolutional layers, and each convolutional layer is followed by the batch normalization operation with the ReLU activation function.

Classification: The final layer for pixel-wise classification is the same one as U-Net and the modified U-Net.

⁸SegNet was first published in arXiv:1511.00561, 2015.

B. Loss Function

Since the segmentation problem is spanning six classes, i.e., the optical disc, hard exudates, soft exudates, hemorrhages, microaneurysms, and non-disease pixels, Saha *et al.* [21] used the softmax function in the final layer of SegNet. Let $Z = (z_{i,j,c}) \in (0, 1)^{I \times J \times C}$ be the final output of SegNet, where I and J respectively are the height and the width of the input image, and C is the number of classes, i.e., $C = 6$. The softmax function converts the previous output $X = (x_{i,j,c}) \in \mathbb{R}^{I \times J \times C}$ into Z by

$$z_{i,j,c} = \frac{e^{x_{i,j,c}}}{\sum_{c'=1}^C e^{x_{i,j,c'}}} \quad \text{for all } i, j, \text{ and } c. \quad (1)$$

From Equation (1), $z_{i,j} = (z_{i,j,1}, z_{i,j,2}, \dots, z_{i,j,C})$ satisfies $\sum_{c=1}^C z_{i,j,c} = 1$ for all i and j , and $z_{i,j,c} \in (0, 1)$ represents the probability of the c th class at the (i, j) pixel. True labels $t_{i,j} = (t_{i,j,1}, t_{i,j,2}, \dots, t_{i,j,C}) \in \{0, 1\}^C$ are given as one-hot vectors of size C , and the multiclass binary cross entropy

$$\text{BCE} = \frac{-1}{IJ} \sum_{i,j,c} t_{i,j,c} \log(z_{i,j,c}) \quad (2)$$

was used as a loss function. If we conduct such multiclass segmentation by a single CNN, with the use of IDRiD dataset [29] as training data, the results are not satisfactory and are subpar especially for microaneurysms as shown in [21].

Hence, in this paper, we construct multiple CNNs, and each CNN segments only a single class, i.e., each CNN determines whether some disease is present or not present for each pixel. This strategy is inspired by the loss function used for training Mask R-CNN [32] on 20 classes of COCO dataset [33]. In the proposed method, we swap the softmax function for the sigmoid function. Let $Z = (z_{i,j}) \in (0, 1)^{I \times J}$ and $X = (x_{i,j}) \in \mathbb{R}^{I \times J}$ be the final output and the previous output, respectively. The sigmoid function converts X into Z by

$$z_{i,j} = \frac{1}{1 + e^{-x_{i,j}}} = \frac{e^{x_{i,j}}}{e^{x_{i,j}} + 1} \quad \text{for all } i \text{ and } j. \quad (3)$$

From Equation (3), $z_{i,j} \in (0, 1)$ holds, which represents the probability of some disease at the (i, j) pixel. The binary cross entropy in (2) is replaced with the following single class one

$$\text{BCE} = \frac{-1}{IJ} \sum_{i,j} [t_{i,j} \log(z_{i,j}) + (1 - t_{i,j}) \log(1 - z_{i,j})]. \quad (4)$$

Although individual CNNs, using the single class binary cross entropy in (4) as a loss function, can significantly improve the performance, there still exist artifacts and spots which should be corrected. This is due to the imbalanced data problem, i.e., the number of non-disease pixels is much greater than that of disease pixels, especially for microaneurysms, and hence training would be finished before learning features of the diseases.

To alleviate the above problem, we use the Dice coefficient

$$\text{DC} = \frac{2\langle T, Z \rangle}{\|T\|_{\text{F}}^2 + \|Z\|_{\text{F}}^2} = \frac{2 \sum_{i,j} t_{i,j} z_{i,j}}{\sum_{i,j} t_{i,j}^2 + \sum_{i,j} z_{i,j}^2} \quad (5)$$

in conjunction with the binary cross entropy as a loss function, where $\langle \cdot, \cdot \rangle$ is the inner product, and $\|\cdot\|_{\text{F}}$ is the Frobenius norm. Since the Dice coefficient can penalize each pixel which

has been incorrectly labeled, each CNN learns features more robustly than the binary cross entropy when the data distribution is imbalanced. Note that the Dice coefficient equals 0 if $Z = T$, and is positive otherwise. Therefore we have to make $(1 - \text{DC})$ as small as possible. By using Equations (4) and (5), finally the proposed loss function to be minimized is given by

$$\text{Loss} = \text{BCE} + (1 - \text{DC}). \quad (6)$$

III. NUMERICAL EXPERIMENTS

A. Experimental Settings

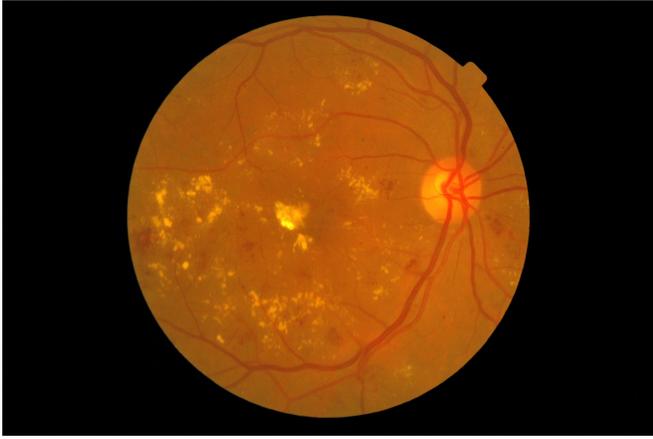
We used all 81 fundus images in IDRiD dataset [29] and a few fundus images in MESSIDOR dataset [28] to evaluate the performance of the proposed method. Each individual CNN for segmentation of the optical disc, hard exudates, soft exudates, hemorrhages, or microaneurysms, was trained from 54 images in IDRiD dataset, and tested for the rest 27 images in IDRiD dataset and the images in MESSIDOR dataset. The image size in IDRiD dataset is 4288×2848 , and it is too large to fit the images into the limited GPU memory as one minibatch. We resized the original images between 1920×1280 and 400×400 and checked each training result. From these results, 960×640 was selected as the best image size with regard to the balance between the quality of the resized images, training speed, and segmentation accuracy. To obtain better results from 54 images, we used data augmentation techniques to create more images. The augmentation was executed in Keras by using the options `rotation_range = 30`, `zoom_range = 0.05`, `width_shift_range = 0.2`, `height_shift_range = 0.2`, `horizontal_flip = True`, `vertical_flip = False`, `fill_mode = 'constant'`, and `cval = 0`.

The choice of an optimizer is crucial because if the choice is poor, then overfitting, slow training speed, or low performance happens. We used ADAM [34] as the optimizer with a learning rate of 0.0001 and a minibatch size of 4. Since Leaky ReLU ($\alpha = 0.2$) is used in the modified U-Net, careful consideration has to be taken for the initialization of network parameters. We used He's initialization [32] already implemented in Keras.

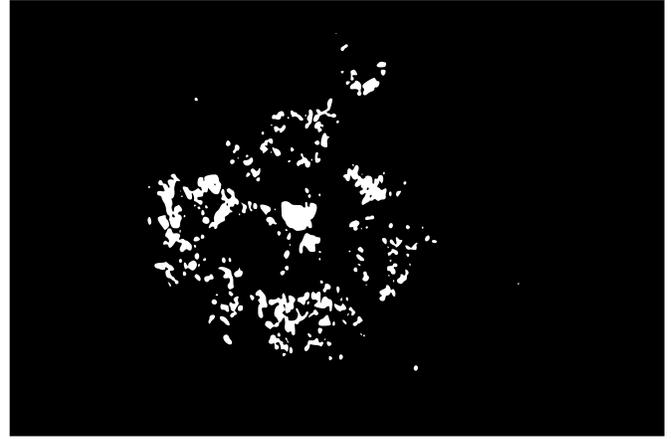
B. Results and Discussion

We compared the performance of the proposed method by using the original U-Nets, the modified U-Nets and SegNets as the individual CNNs. The results were evaluated using the Dice coefficient in (5) and the loss function in (6) for training and test data. Tables I, II, and III show the values of the Dice coefficient and the loss function of each class by the original U-Net, our modified U-Net, and SegNet, respectively. From Table I, we can observe that the original U-Net achieved good results for the optic disc and hard exudates as shown in Fig. 3. On the other hand, for hemorrhages, the Dice coefficient of test data is lower than 0.5, and soft exudates and microaneurysms are not detected at all. Moreover, the original U-Net required 4 hours for training while the modified one required 1 hour.

Figures 4, 5, 6, 7, and 8 show the segmentation results, by the modified U-Net and SegNet, of the optic disc, hard exudates, soft exudates, hemorrhages, and microaneurysms, respectively. From these figures and Tables III, we can observe that SegNet achieved better results than the original U-Net, and



(a) Original fundus image



(b) Segmentation result by the original U-Net

Fig. 3. Segmentation result of hard exudates by the original U-Net.

TABLE I

VALUES OF THE DICE COEFFICIENT AND THE LOSS FUNCTION OF EACH CLASS FOR TRAINING AND TEST DATA BY THE ORIGINAL U-NET

Type of DR Disease	DC for Training Data	Loss for Training Data	DC for Test Data	Loss for Test Data
Optic Disc	0.9337	0.0765	0.9204	0.0911
Hard Exudates	0.7556	0.2662	0.6837	0.3640
Soft Exudates	0	1.0579	0	1.1480
Hemorrhages	0.7244	0.3007	0.4968	0.5657
Microaneurysms	0	1.0252	0	1.0234

TABLE II

VALUES OF THE DICE COEFFICIENT AND THE LOSS FUNCTION OF EACH CLASS FOR TRAINING AND TEST DATA BY THE MODIFIED U-NET

Type of DR Disease	DC for Training Data	Loss for Training Data	DC for Test Data	Loss for Test Data
Optic Disc	0.9904	0.0109	0.9565	0.0549
Hard Exudates	0.8522	0.1582	0.7082	0.3336
Soft Exudates	0.9401	0.0615	0.7125	0.3267
Hemorrhages	0.8653	0.1461	0.7793	0.2240
Microaneurysms	0.7950	0.2093	0.6920	0.4430

TABLE III

VALUES OF THE DICE COEFFICIENT AND THE LOSS FUNCTION OF EACH CLASS FOR TRAINING AND TEST DATA BY SEGNET

Type of DR Disease	DC for Training Data	Loss for Training Data	DC for Test Data	Loss for Test Data
Optic Disc	0.9880	0.0207	0.9632	0.0439
Hard Exudates	0.5560	0.4321	0.5250	0.4320
Soft Exudates	0.8852	0.1191	0.6198	0.4829
Hemorrhages	0.6988	0.3336	0.5281	0.5260
Microaneurysms	0.7520	0.2333	0.6523	0.3561

the Dice coefficient of test data is larger than 0.5 for all the classes. Furthermore, from the figures and Tables II, we can observe that the modified U-Net achieved the best results, and the Dice coefficient is larger than about 0.7 for all the classes. Especially, the modified U-Net could detect microaneurysms, which is very small but important to specify early stage DR.

IV. CONCLUSION

Early stage DR detection is important to prevent the blindness, but regular screening tests are becoming heavy burdens for ophthalmologists with increasing number of diabetic patients. To reduce the burdens, in this paper, we addressed the fundus image segmentation problem by deep learning for DR diagnosis. In this problem, the most challenging task is to construct a deep CNN from a limited number of imbalanced data. Since it is difficult to simultaneously segment different types of diseases by a only one CNN, we proposed to construct one individual CNN for each type of disease. Moreover, to alleviate the imbalanced data problem, we proposed to use data augmentation and adopt the Dice coefficient in conjunction with the binary cross entropy as a loss function. Numerical experiments showed that the modified U-Net achieved the good results even for soft exudates, hemorrhages, and microaneurysm.

REFERENCES

- [1] D. S. Fong, L. Aiello, T. S. W. Gardner, G. L. King, G. Blankenship, J. D. Cavallerano, F. L. Ferris, and R. Klein, "Retinopathy in diabetes," *Diabetes Care*, vol. 27, no. 1, pp. s84–s87, 2004.
- [2] J. W. Y. Yau, S. L. Rogers, R. Kawasaki, E. L. Lamoureux, J. W. Kowalski, T. Bek, S. J. Chen, J. M. Dekker, A. Fletcher, J. Grauslund, S. Haffner, R. F. Hamman, M. K. Ikram, T. Kayama, B. E. K. Klein, R. Klein, S. Krishnaiah, K. Mayurasakorn, J. P. O'hare, T. J. Orchard, M. Porta, M. Rema, M. S. Roy, T. Sharma, J. Shaw, H. Taylor, J. M. Tielsch, R. Varma, J. J. Wang, N. Wang, S. West, L. Xu, M. Yasuda, X. Zhang, P. Mitchell, and T. Y. Wong, "Global prevalence and major risk factors of diabetic retinopathy," *Diabetes Care*, vol. 35, no. 3, pp. 556–564, 2012.
- [3] J. K. H. Goh, C. Y. Cheung, S. S. Sim, P. C. Tan, G. S. W. Tan, and T. Y. Wong, "Retinal imaging techniques for diabetic retinopathy screening," *Journal of Diabetes Science and Technology*, vol. 10, no. 2, pp. 282–294, 2016.
- [4] L. Giancardo, F. Meriaudeau, T. P. Karnowski, Y. Li, S. Garg, K. W. Tobin Jr., and E. Chaum, "Exudate-based diabetic macular edema detection in fundus images using publicly available datasets," *Medical Image Analysis*, vol. 16, no. 1, pp. 216–226, 2012.
- [5] A. S. E. Sisy, N. M. Salem, and A. F. Seddik, "Automatic detection of exudates from digital color fundus images," *International Journal of Computer Applications*, vol. 122, no. 7, pp. 18–22, 2015.
- [6] M. Partovi, S. H. Rasta, and A. Javadzadeh, "Automatic detection of retinal exudates in fundus images of diabetic retinopathy patients," *Journal of Analytical Research in Clinical Medicine*, vol. 4, no. 2, pp. 104–109, 2016.

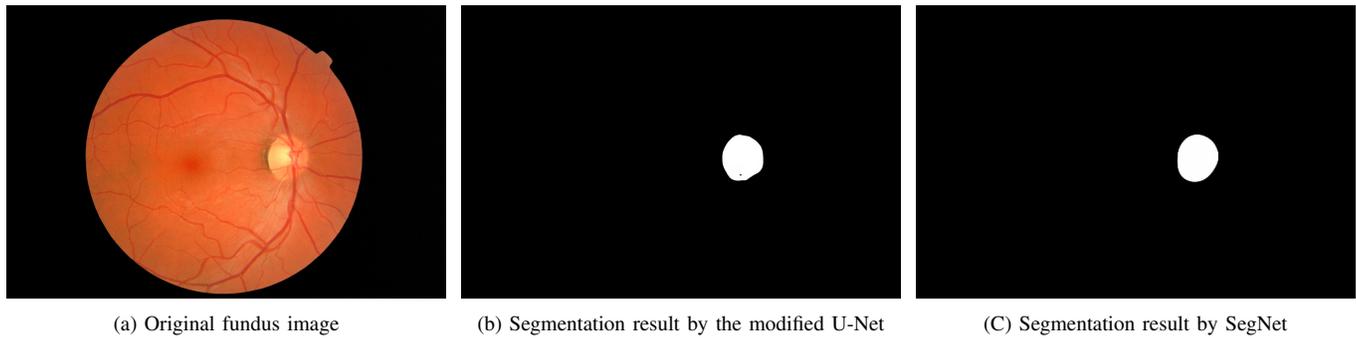


Fig. 4. Segmentation results of the optic disc by the modified U-Net and SegNet.

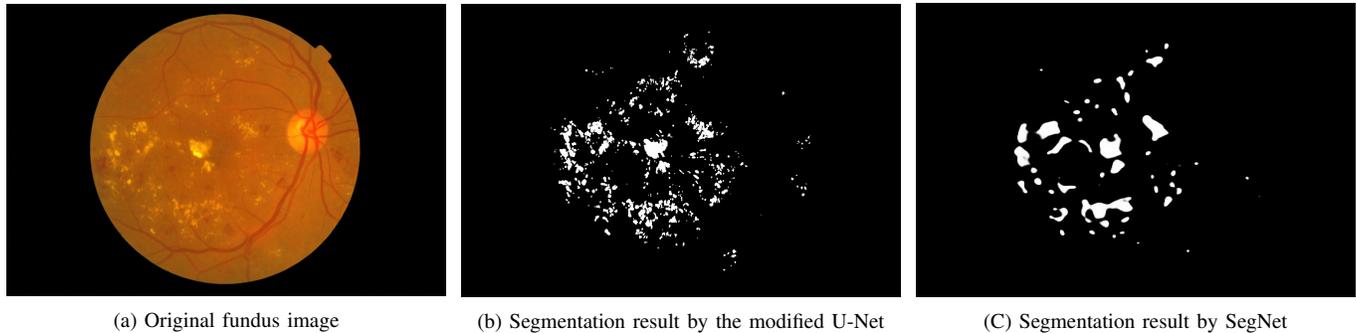


Fig. 5. Segmentation results of hard exudates by the modified U-Net and SegNet.

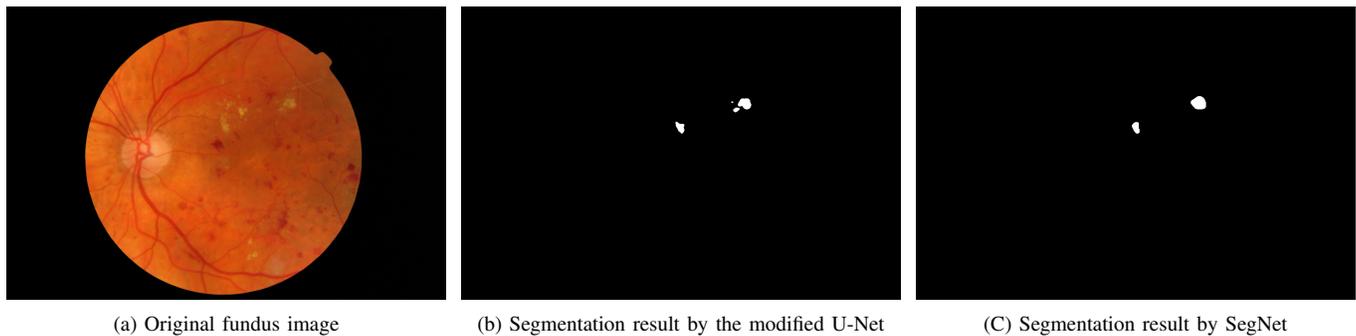


Fig. 6. Segmentation results of soft exudates by the modified U-Net and SegNet.

- [7] Q. Liu, B. Zou, J. Chen, W. Ke, K. Yue, Z. Chen, and G. Zhao, "A location-to-segmentation strategy for automatic exudate segmentation in color retinal fundus images," *Computerized Medical Imaging and Graphics*, vol. 55, pp. 78–86, 2017.
- [8] M. Canche, O. Dalmau, and M. G. Gadañón, "Automatic detection of hard exudates in retinal images with diabetic retinopathy," in *Proceedings of Mexican International Conference on Artificial Intelligence (MICAI)*, 2017, pp. 53–59.
- [9] A. Sopharak, M. N. Dailey, B. Uyyanonvara, S. Barman, T. Williamson, K. T. Nwe, and Y. A. Moe, "Machine learning approach to automatic exudate detection in retinal images from diabetic patients," *Journal of Modern Optics*, vol. 57, no. 2, pp. 124–135, Jan. 2010.
- [10] E. V. Carrera, A. González, and R. Carrera, "Automated detection of diabetic retinopathy using SVM," in *Proceedings of IEEE International Conference on Electronics, Electrical Engineering and Computing (INTERCON)*, 2017, 4pages.
- [11] S. Long, X. Huang, Z. Chen, S. Pardhan, and D. Zheng, "Automatic detection of hard exudates in color retinal images using dynamic threshold and SVM classification: Algorithm development and evaluation," *BioMed Research International*, vol. 2019, 13 pages, 2019.
- [12] R. Gargeya and T. Leng, "Automated identification of diabetic retinopathy using deep learning," *Ophthalmology*, vol. 124, no. 7, pp. 962–969, 2017.
- [13] C. Lam, D. Yi, M. Guo, and T. Lindsey, "Automated detection of diabetic retinopathy using deep learning," in *Proceedings of AMIA Joints Summits on Translational Science*, 2018, pp. 147–155.
- [14] P. Prentašić and S. Lončarić, "Detection of exudates in fundus photographs using convolutional neural networks," in *Proceedings of the 9th International Symposium on Image and Signal Processing and Analysis (ISPA)*, 2015, pp. 188–192.
- [15] S. Yu, D. Xiao, and Y. Kanagasigam, "Exudate detection for diabetic retinopathy with convolutional neural networks," in *Proceedings of the 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2017, pp. 1744–1747.
- [16] O. Perdomo, J. Arevalo, and F. A. González, "Convolutional network to detect exudates in eye fundus images of diabetic subjects," in *Proceedings of the 12th International Symposium on Medical Information Processing and Analysis*, 2017, 6 pages.
- [17] P. Khojasteh, L. A. P. Júnior, T. Carvalho, E. Rezende, B. Aliahmad, J. P. Papa, and D. K. Kumar, "Exudate detection in fundus images

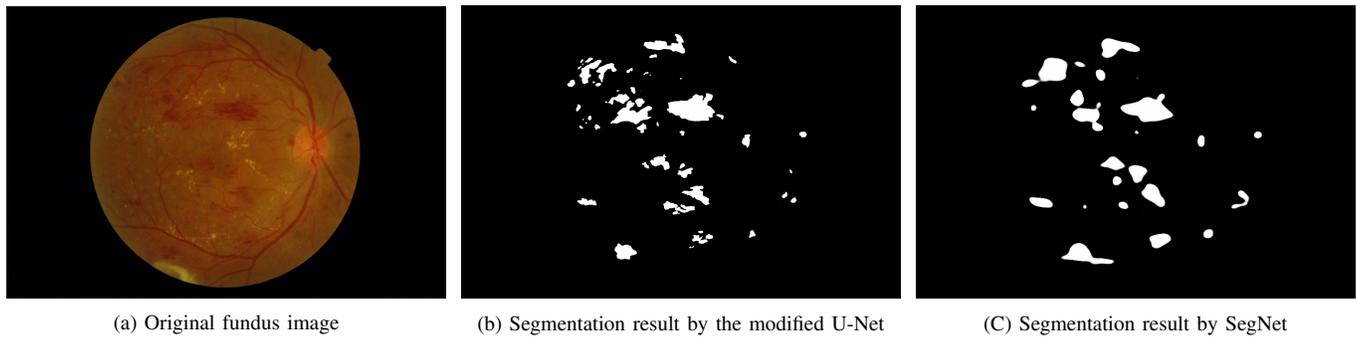


Fig. 7. Segmentation results of hemorrhages by the modified U-Net and SegNet.

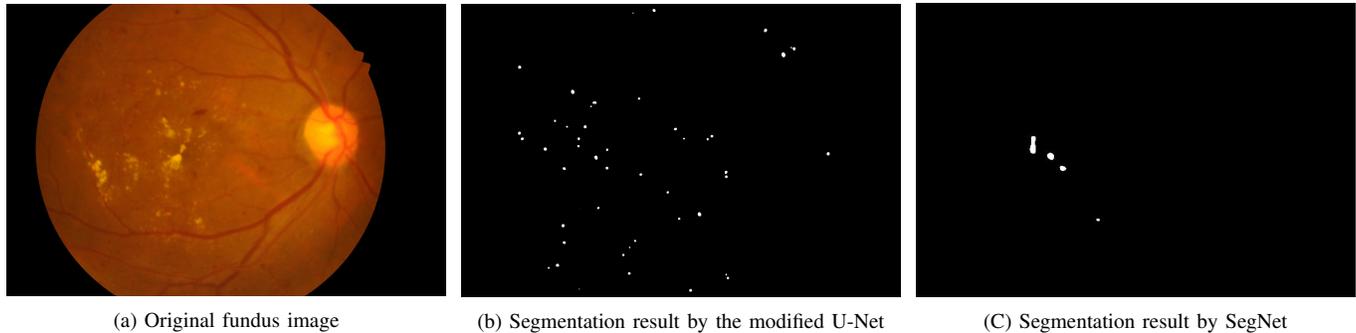


Fig. 8. Segmentation results of microaneurysms by the modified U-Net and SegNet.

- using deeply-learnable features,” *Computers in Biology and Medicine*, vol. 104, pp. 62–69, 2019.
- [18] J. I. Orlando, E. Prokofyeva, M. del Fresno, and M. B. Blaschko, “An ensemble deep learning based approach for red lesion detection in fundus images,” *Computer Methods and Programs in Biomedicine*, vol. 153, pp. 115–127, 2018.
- [19] R. Zheng, L. Liu, S. Zhang, C. Zheng, F. Bunyak, R. Xu, B. Li, and M. Sun, “Detection of exudates in fundus photographs with imbalanced learning using conditional generative adversarial network,” *Biomedical Optics Express*, vol. 8, no. 10, pp. 4863–4878, 2018.
- [20] J. H. Tan, H. Fujita, S. Sivaprasad, S. V. Bhandary, A. K. Rao, K. C. Chua, and U. R. Acharya, “Automated segmentation of exudates, hemorrhages, microaneurysms using single convolutional neural network,” *Information Sciences*, vol. 420, pp. 66–76, 2017.
- [21] O. Saha, R. Sathish, and D. Sheet, “Fully convolutional neural network for semantic segmentation of anatomical structure and pathologies in color fundus images associated with diabetic retinopathy,” *arXiv: 1902.03122*, 4 pages, 2019.
- [22] P. Prentašić, S. Lončarić, Z. Vatavuk, G. Benčić, M. Subašić, T. Petković, L. Dujmović, M. Malenica-Ravlić, N. Budimlija, and R. Tadić, “Diabetic retinopathy image database (DRiDB): A new database for diabetic retinopathy screening programs research,” in *Proceedings of International Symposium on Image and Signal Processing and Analysis (ISPA)*, 2013, pp. 711–716.
- [23] E. Decencière, G. Cazuguel, X. Zhang, G. Thibault, J. C. Klein, F. Meyer, B. Marcotegui, G. Quellec, M. Lamard, R. Danno, D. Elie, P. Massin, Z. Viktor, A. Erginay, B. Lay, and A. Chabouis, “TeleOphta: Machine learning and image processing methods for teleophthalmology,” *IRBM*, vol. 34, no. 2, pp. 196–203, 2013.
- [24] T. Kauppi, V. Kalesnykiene, J. K. Kamarainen, L. Lensu, I. Sorri, A. Raninen, R. Voutilainen, H. Uusitalo, H. Kälviäinen, and J. Pietilä, “DIARETDB1 diabetic retinopathy database and evaluation protocol,” in *Proceedings of Annual Conference on Medical Image Understanding and Analysis (MIUA)*, 2007, pp. 61–65.
- [25] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial nets,” in *Proceedings of Annual Conference on Neural Information Processing Systems (NIPS)*, 2014, pp. 2672–2680.
- [26] P. Isola, J. Y. Zhu, T. Zhou, and A. A. Efros, “Image-to-image translation with conditional adversarial networks,” in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 1125–1134.
- [27] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional networks for biomedical image segmentation,” in *Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2015, pp. 234–241.
- [28] E. Decencière, X. Zhang, G. Cazuguel, B. Lay, B. Cochener, C. Trone, P. Gain, J. R. Ordóñez-Varela, P. Massin, A. Erginay, B. Charton, and J. C. Klein, “Feedback on a publicly distributed database: The Messidor database,” *Image Analysis & Stereology*, vol. 33, no. 3, pp. 231–234, 2014.
- [29] P. Porwal, S. Pachade, R. Kamble, M. Kokare, G. Deshmukh, V. Sahasrabudde, and F. Meriaudeau, “Indian diabetic retinopathy image dataset (IDRiD): A database for diabetic retinopathy screening research,” *Data*, vol. 3, no. 3, 8 pages, 2018.
- [30] S. Sivaprasad, G. Arden, A. T. Prevost, R. Crosby-Nwaobi, H. Holmes, J. Kelly, C. Murphy, G. Rubin, J. Vasconcelos, and P. Hykin, “A multicentre phase III randomised controlled single-masked clinical trial evaluating the clinical efficacy and safety of light-masks at preventing dark-adaptation in the treatment of early diabetic macular oedema (CLEOPATRA): Study protocol for a randomised controlled trial,” *Trials*, vol. 15, no. 458, 10 pages, 2014.
- [31] V. Badrinarayanan, A. Kendall, and R. Cipolla, “SegNet: A deep convolutional encoder-decoder architecture for image segmentation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017.
- [32] K. He, G. Gkioxari, P. Dollár, and R. Girshick, “Mask R-CNN,” in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 2961–2969.
- [33] T. Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, “Microsoft COCO: Common objects in context,” in *Proceedings of European Conference on Computer Vision (ECCV)*, 2014, pp. 740–755.
- [34] D. P. Kingma and J. L. Ba, “Adam: A method for stochastic optimization,” in *Proceedings of International Conference on Learning Representations (ICLR)*, 2015, 15 pages.