IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received May 6, 2020, accepted June 13, 2020, date of publication June 25, 2020, date of current version August 5, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3005044

Automatic Detection and Monitoring of Diabetic Retinopathy Using Efficient Convolutional Neural Networks and Contrast Limited Adaptive Histogram Equalization

ASRA MOMENI POUR[1], HADI SEYEDARABI[1,2], SEYED HASSAN ABBASI JAHROMI[3], AND ALIREZA JAVADZADEH[4]

[1]Department of Biomedical Engineering, Faculty of Advanced Medical Sciences, Tabriz University of Medical Sciences, Tabriz 51666-14766, Iran
[2]Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz 51666-16471, Iran
[3]Medical Engineering and Physics Department, Shiraz University of Medical Sciences, Shiraz 71937-11351, Iran
[4]Nikokari Eye Hospital, Tabriz University of Medical Science, Tabriz 5154645395, Iran

Corresponding author: Hadi Seyedarabi (seyedarabi@tabrizu.ac.ir)

ABSTRACT Diabetic retinopathy is a medical condition of the damaged retina that is caused by diabetes and lack of proper monitoring and treatment, which usually leads to blindness. However, diabetic retinopathy monitoring requires an expert ophthalmologist. Recently, automatic monitoring models with acceptable efficiency are suggested as an alternative for expert ophthalmologists. In this paper, a new diabetic retinopathy monitoring model is proposed by using the Contrast Limited Adaptive Histogram Equalization method to improve the image quality and equalize intensities uniformly as the pre-processing step. Then, EfficientNet-B5 architecture is used for the classification step. The efficiency of this network is in uniformly scaling all dimensions of the network. The final model is trained once on a mixture of two datasets, Messidor-2 and IDRiD, and evaluated on the Messidor dataset. The area under the curve (AUC) is enhanced from 0.936, which is the highest value in all recent works, to 0.945. Also, once again, to further evaluate the performance of the model, it is trained on a mixture of two datasets, Messidor-2 and Messidor, and evaluated on the IDRiD dataset. In this case, the AUC is enhanced from 0.796, which is the highest value in all recent works, to 0.932. In comparison to other studies, our proposed model improves the AUC.

INDEX TERMS Diabetic Retinopathy, contrast limited adaptive histogram equalization, convolutional neural networks, EfficientNet.

I. INTRODUCTION

With an increase in the computing and processing power of computers and the development of image processing techniques in recent years, the idea of using a computer to analyze medical images and automatically diagnose diseases has attracted the attention of many medical and computer specialists. This idea of analyzing retina images and diagnosing ocular and vascular diseases is very efficient and inexpensive.

Retinal vascular structures contain a great deal of medical information. Diabetes, hypertension, cardiovascular diseases, and some others can only be detected through the examination of retinal fundus images. Diabetic Retinopathy (DR) monitoring is one of the acknowledged methods in examining patients. Annual screening of the eye fundus of diabetic patients is one of the essential parts of their treatment. However, the analysis of retinal fundus images demands an ophthalmologist expert.

The requirements for DR monitoring and limitation in resources inspired researchers to develop an automated system to analyze retinal fundus images. Generally, there are five different DR categories: No DR, Mild DR, Moderate DR, Severe DR, and Proliferative DR. An automated detection of DR, classifies each patient into one of these classes through the examination of their retinal fundus images. The early models of such systems were limited to computer-aided systems that provided vessel segmentation of fundus images. Some computer vision DR detection models

The associate editor coordinating the review of this manuscript and approving it for publication was Tao Zhou.

are also proposed, but their accuracy rates are rather low. Convolutional Neural Networks (CNNs) and deep learning introduced a new stage for this problem. They helped researchers to design a reliable end to end DR detection system. Gulshan *et al*. [1], confirmed their CNN model results with a group of US board-certified ophthalmologists to confirm CNN's reliability. One of the challenging aspects of working with image data is the high number of its dimensions. Even a small 100×100 image represents a point in a 10000-dimensional space.

CNNs are designed at a fixed resource budget to handle such problems and then scaled up for better efficiency and accuracy [2]. Although, there are different methods to perform the process of scaling up CNNs. The common way is to scale up CNNs by their depth [3] or width [4]. Another way is to scale up models by image resolution, which is less common [5]. The scaling of CNNs is applied to obtain higher accuracy. Although scaling two or three dimensions is possible as arbitrary, but it requires tedious manual settings and often reduces accuracy and efficiency [2].

In this paper, an effective pre-processing technique is proposed to increase image contrast which is called Contrast Limited Adaptive Histogram Equalization (CLAHE). This method helps a lot in analyzing the biomedical image. Although, to detect and classify of DR is proposed a lot of processing techniques but among them all, image contrast effect is beneficial in diagnosing DR [6]. After pre-processing the data, the new EfficientNet-B5 architecture is used for the classification step, that the efficiency of the suggested method is in uniform scaling all dimensions of the network.

II. RELATED WORKS

One of the first uses of CNNs for the retinal vessel segmentation is suggested by Melinscak *et al*. [7]. They address the problem as a binary classification and trained with CNN, which classified individual pixels separately and determined that a pixel of the input image is vessel or non-vessel. Maji *et al*. [8] trained 12 different CNNs and used their ensemble as a single model to detect and extract the blood vessels in retinal fundus images, where the average output of all CNNs is used as the output of the model.

A different CNN architecture is proposed by Liskowski and Krawiec [9]. They used a simple CNN architecture with three convolutional layers, one pooling layer, and two fully-connected layers. They extracted more than 200,000 patches randomly from each image to train the model.

Also, Dasgupta and Singh [10] proposed a fully CNN-based approach. Both the input and the output of their proposed CNN architecture have the same dimensions of $1 \times 28 \times 28$.

DR classification is discussed comprehensively in the literature, which is based on Neural Networks [11]. Pratt *et al.* [12] proposed one of the first end-to-end CNN approaches for DR classification. Training a CNN-based DR classification model requires a large number of training images. They trained the CNN model using the Kaggle dataset, which contains more than 50,000 labeled fundus images in five different classes. The pre-processing of their model consists of color normalization and resizing the images.

Gulshan *et al*. [1] proved that the accuracy of a CNN model has a strong relationship with the amount of the data upon which the model is trained. There are a large number of weights that should be acknowledged. To simplify the process of learning, they used a pre-trained version of Google Inception v3 architecture [13] and fine-tuned the network using the medical images. Moreover, they analyzed the performance of the Google Inception v3 network for DR detection and confirmed the results with a group of US board-certified ophthalmologists.

Wang and Yang [14] proposed a CNN model to localize the discriminative regions in the retina image. This was enabled by adding a regression activation map (RAM) after the average pooling layer of the network and reconstructing an area of interest (ROI) by adding up-sampling layers. Wang *et al*. provided an ROI map, which described the areas that the network decision is made based on them. By taking into account it, their model can be used as an assistant for clinicians. Wu *et al*. [15] also proposed a classification of DR and diabetic macular edema.

In diagnosing the DR signs, Hajeb Mohammad Alipour *et al*. [16] employed the CLAHE method to enhance the image, applied curvelet transform to extract features from images, that finally fed the images into a Support Vector Machine (SVM). The CLAHE method performs better than other contrast enhancement methods to augment the vessels in the retinal fundus images with different backgrounds [17].

Since AlexNet [18] won the ImageNet contest in 2012, CNNs have been used for a wide range of deep learning applications, especially for computer vision [3], [19]. From 2012 to the present, researchers have been trying to achieve better accuracy and efficiency of models on various tasks.

There are different ways to scale CNNs. ResNet [3] can be scaled up or down by setting layers or depth of network, while WideResNet [4] and MobileNets [20] can be scaled by adjusting width of network. Adjusting the depth and width of the network are both crucial to express the potency of CNNs [21]–[24].

Tan and Le [2] proposed a family of models called EfficientNets that achieve better accuracy and efficiency than popular CNNs. Generally, they have examined eight models from this family. These models were scaled up using a principled way based on the resources available.

The main objective of this study is to present a model for DR monitoring in fundus retinal images. It can be used to monitor DR without the need for a specialist, and also significantly reduce costs. On the other hand, the goal is to provide a higher accuracy model compared to other methods.

III. MATERIALS AND METHODS

DR is an important technique, which requires experienced ophthalmologists. They analyze the retinal fundus images and look for subtle lesions and symptoms. Recently, lots of researchers designed various models to replace or assist ophthalmologists based on computer vision algorithms.

In the following, we propose a new DR model that utilizes a useful pre-processing technique called CLAHE, and new EfficientNet architecture to achieve an accurate model. The final model is trained once on a mixture of two Messidor-2 and IDRiD datasets and evaluated on the Messidor dataset. Also, to further evaluate the performance of the model, it is once again trained on a mixture of two Messidor-2 and Messidor datasets and evaluated on the IDRiD dataset.

A. CLAHE METHOD

There are some common ways to improve image contrast. The Histogram Equalization (HE) enhances image contrast. This method can keep the brightness of the background, but it does not work well when we work with color images [25]. The Brightness Preserving by Histogram Equalization (BBHE) method is another image enhancement method. This approach is useful when brightness conservation is needed. The problem with this method is that it needs more time for calculation [26]. In the Brightness Preserving Dynamic Fuzzy Histogram Equalization (BPDHE) method, the mean intensity of the output and input images are approximately equal. The problem with this method is that transparency of the image contrast cannot be achieved [27]. Adaptive Histogram Equalization (AHE) divides the image into tiles and then applies HE in each of these tiles. Although doing this way augments the contrast in near-constant areas of the image, it may cause augmented noise in these areas. Also, with this method, when there are areas in the image that are darker or brighter than most parts of the image, the contrast is low [28]. The CLAHE is one of the ways of image quality increasing that improves the contrast and the visible surface of the dark image. Also, it reduces the noise amplification of the tiles by limiting the contrast. Although the CLAHE method does not completely eliminate artifacts, it is better from AHE. Block size (BS) and clip limit (CL) are the two main parameters of CLAHE, which control the image quality and are determined by the users. CL is the amount at which the histogram is clipped, and it depends on two factors of the normalization of the histogram and the size of the neighborhood area. Generally, before computing the Cumulative Distribution Function (CDF), the CLAHE restricts the augmentation by clipping the histogram at a predetermined amount. If the user defines improper parameters for CLAHE, the results may not be even as good as compared to HE [29], [30].

Therefore, we use CLAHE to improve image contrast and to equalize intensities uniformly. Using this method, the image is divided into continuous tiles without overlap. Then the histogram above a threshold is clipped, and clipped pixels distribute to each grey level. Finally, HE is applied to each of the tiles, and the mapping is interpolated between the neighboring tiles. The resulting mapping to each pixel is interpolated from the intensity mappings of eight neighboring tiles. Therefore, by doing this method, the vessels are displayed better in the image, and the network can better learn.

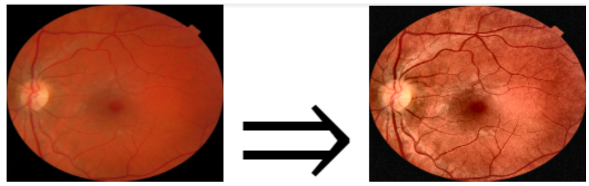

FIGURE 1. Depicts the increase in image contrast with the CLAHE method.

Figure 1 depicts the increase in image contrast with the CLAHE method.

B. CNN MODEL

Recently, CNNs are applied to various problems, which have obtained excellent results. They significantly improved the image processing results in terms of accuracy and generalization. However, designing the proper CNN architecture for such problems is a vital, yet complex task.

There are three scalable dimensions for CNNs: depth, width, and resolution. Depth is the number of layers in a network. Width is the number of channels in a convolutional layer, whereas the resolution is the same as the input image resolution. Generally, network all dimensions scaling improves accuracy compared to network single or two dimensions scaling [2].

There are lots of CNN models with state-of-the-art results in computer vision problems. To the best of our knowledge, one of the most successful CNN models, which is proposed recently is EfficientNets [2].

The advantages of EfficientNets compared to the most popular CNNs are in terms of reducing the number of parameters and FLOPS, furthermore, increasing accuracy, and speed [2].

First, we experiment with the efficient-B0 network and then scale all three dimensions of our network by an effective compound factor. Also, our scaling is uniform, because arbitrarily scaling needs tedious manual settings, and sometimes results in less accuracy and efficiency [2].

Tan and Le [2] proposed an efficient scaling way by a compound factor to scaling dimensions of depth, width, and resolution as uniform. It is formulated as Eq. (1):

$$\text{depth}: d = \alpha^{\phi}, \quad \text{width}: w = \beta^{\phi}, \ \text{resolution}: r = \gamma^{\phi} \quad (1)$$

where ϕ is a coefficient determined by the user that controls how many more resources are available. Parameters of α, β, and γ determine how to allocate these resources to network d, w, and r, respectively. They constrained $(\alpha . \beta^2 . \gamma^2) \approx 2$.

Generally, there are four scaling parameters(α, β, γ, and ϕ) to search. Tan and Le [2] Constructed EfficientNets in two stages:

STAGE 1: They fixed $\phi = 1$, and then searched the optimal values for α, β, and γ scaling.

TABLE 1. The efficient-B0 baseline network architecture.

Operator	Resolution	# Channels	# Layers
Conv3 × 3	224 × 224	32	1
MBConv1, k3 × 3	112 × 112	16	1
MBConv6, k3 × 3	112 × 112	24	2
MBConv6, k5 × 5	56 × 56	40	2
MBConv6, k3 × 3	28 × 28	80	3
MBConv6, k5 × 5	14 × 14	112	3
MBConv6, k5 × 5	14 × 14	192	4
MBConv6, k3 × 3	7 × 7	320	1
Conv1x1 & Pooling & FC	7 × 7	1280	1

TABLE 2. The value of the resolution in efficient networks.

Model	resolution
EfficientNet-B0 [2]	224 × 224
EfficientNet-B1 [2]	240 × 240
EfficientNet-B2 [2]	260 × 260
EfficientNet-B3 [2]	300 × 300
EfficientNet-B4 [2]	380 × 380
EfficientNet-B5 [2]	456 × 456
EfficientNet-B6 [2]	528 × 528
EfficientNet-B7 [2]	600 × 600

STAGE 2: According to the optimal values obtained in the previous step, they fixed α, β, and γ as constants and produced EfficientNet-B1 to B7 with increasing ϕ from 2 to 7.

The architecture of the efficient-B0 is shown in Table 1.

The MBConv block is a reverse spectral block with a Squeeze and Excite block and is used in MobileNetV2. Unlike inverted Residuals Original ResNet blocks, blocks in MBConv are compounded with a layer that first expands channels, then squeezes them [19], [31], [32].

The difference between the efficient networks is in terms of depth, width, and resolution. The value of the resolution in these networks is shown in Table 2.

We observe in doing experiments with EfficientNets-B0 to B4 that useful image information is lost because these networks greatly reduce the image resolution. Therefore, we use the efficient-B5 network architecture and consider the values of the parameters as follows:

$\alpha = 1.2, \beta = 1.1, \gamma = 1.15, w = 1.6, d = 2.2$, and $r = 456$.

The values of these parameters are according to the default values suggested by Tan and Le [2].

Also, the selected network is pre-trained on the ImageNet dataset [33]. Using the pre-trained network saves time and helps to achieve better accuracy with the limited available dataset.

IV. DISCUSSION AND RESULT

The main aim of this study is to present a model for DR monitoring, and to reduce human errors in retinal fundus images by improving automated diagnostic methods, and preparing a real-time and highly generalized system. These systems do not require specialist expertise, and significantly reduce the cost of DR monitoring. We can learn more about such diseases by using more specialized knowledge and applying different methods of image pre-processing and using deep artificial NN as the closest human example to the human brain and the human NN to increase the accuracy in the diagnosis of different diseases.

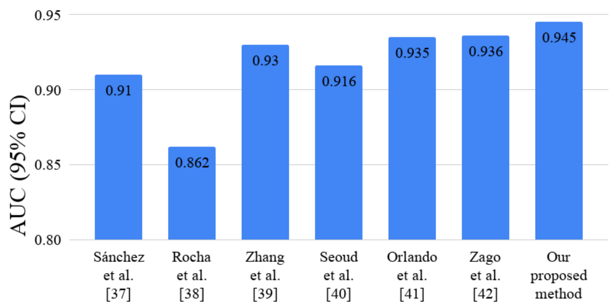

FIGURE 2. Comparison of DR results using the Messidor dataset.

First, the model is trained on a mixture of two Messidor-2 and IDRiD datasets and evaluated on the Messidor dataset. The Messidor-2 and Messidor datasets include DR labels with various levels for automatic DR detection. The Messidor-2 dataset includes 1,748 retinal color images of 874 subjects. Subjects were imaged with a color video 3CCD camera on a Topcon fundus camera. This dataset has two levels: referable subjects and non-referable subjects [34]. Also, the Messidor test set includes 1200 images that classified into segregate groups from 0 (No DR) to 3 (Severe DR) [35]. In our experiment, the test dataset is divided into two groups, {0, 1} versus {2, 3}, that the first group is of type non-referable, and the second group is of type referable. Our model is able to achieve 92% sensitivity and 94.5% AUC in test data.

Finally, our proposed method was compared with the state-of-the-art models on the Messidor DR dataset including Porwal *et al.* [42], that the results of their model were the highest value in all recent works. They designed a lesion localization model using a deep network patch-based approach and used two convolutional neural network models for selecting the training patches. Their model on the Standard Diabetic Retinopathy Database, Calibration Level 1 (DIARETDB1) database trained and their AUC value on the Messidor DR dataset reached 0.936. Therefore, our AUC from 0.936, is enhanced to 0.945.

The comparison gained results with the state-of-the-art models on the Messidor dataset in Figure 2 show that our AUC value is the highest compared to recent works.

Again, our model was trained on a mixture of two Messidor-2 and Messidor datasets and evaluated on the IDRiD dataset. The IDRiD test set includes 516 images that classified into segregate groups from 0 (No DR) to 4 (Severe DR) [36]. In the DR need for referral experiment, we divided the test dataset into two groups, {0, 1} versus {2, 3, 4}. In this case, the AUC from 0.796, which was the highest value in all recent works, and performed by Zago and Porwal [42], is enhanced to 0.932. Also, our model obtained the sensitivity of 93%. As a result, in comparison to this study, our proposed model improves the AUC.

The gained results are significant and confirm the efficiency of the suggested method. Our results are achieved without the data augmentation.

Generally, we used the efficient network architecture that is a lightweight CNN. Useful information in our image with EfficientNet-B0 to B4 is lost because these networks reduce the size of the image much. Therefore, we employ EfficientNet-B5. This network allows the scaling process to be accomplished more systematically by using a compound factor to scale up the network. The goal of designing such a network is to expand as much as possible and provide more effective training.

Our model was trained using an NVIDIA 1080ti GPU, which has 11GB of RAM and uses a core i7 CPU with 32GB RAM. This implementation runs using the TensorFlow backend. We employ Keras v2.3.1,[1] which are popular neural API networks.

V. CONCLUSION

In this paper, a useful technique called CLAHE is presented for the amplification of the vessels in retinal fundus images as the pre-processing step. Therefore, by increasing the contrast are improved the important information inside the images. Also, the new EfficientNet-B5 architecture is employed for the classification step. The efficiency of this network is in uniform scaling all dimensions of the network.

REFERENCES

[1] V. Gulshan, L. Peng, M. Coram, M. C. Stumpe, D. Wu, A. Narayanaswamy, S. Venugopalan, K. Widner, T. Madams, J. Madams, and R. Kim, "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," *J. Amer. Med. Assoc.*, vol. 316, no. 22, pp. 2402–2410, 2016.

[2] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," 2019, *arXiv:1905.11946*. [Online]. Available: http://arxiv.org/abs/1905.11946

[3] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[4] S. Zagoruyko and N. Komodakis, "Wide residual networks," 2016, *arXiv:1605.07146*. [Online]. Available: http://arxiv.org/abs/1605.07146

[5] Y. Huang, Y. Cheng, A. Bapna, O. Firat, D. Chen, M. Chen, H. Lee, J. Ngiam, Q. V. Le, and Y. Wu, "Gpipe: Efficient training of giant neural networks using pipeline parallelism," in *Proc. Adv. Neural Inf. Process. Syst.*, 2019, pp. 103–112.

[6] S. K. Shome and S. R. K. Vadali, "Enhancement of diabetic retinopathy imagery using contrast limited adaptive histogram equalization," *Int. J. Comput. Sci. Inf. Technol.*, vol. 2, no. 6, pp. 2694–2699, 2011.

[7] M. Melinscak, P. Prentasic, and S. Loncaric, "Retinal vessel segmentation using deep neural networks," in *Proc. 10th Int. Conf. Comput. Vis. Theory Appl. (VISAPP)*, 2015, PP. 1–6.

[8] D. Maji, A. Santara, P. Mitra, and D. Sheet, "Ensemble of deep convolutional neural networks for learning to detect retinal vessels in fundus images," 2016, *arXiv:1603.04833*. [Online]. Available: http://arxiv.org/abs/1603.04833

[9] P. Liskowski and K. Krawiec, "Segmenting retinal blood vessels with deep neural networks," *IEEE Trans. Med. Imag.*, vol. 35, no. 11, pp. 2369–2380, Nov. 2016.

[10] A. Dasgupta and S. Singh, "A fully convolutional neural network based structured prediction approach towards the retinal vessel segmentation," in *Proc. IEEE 14th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2017, pp. 248–251.

[11] J. Nayak, P. S. Bhat, R. Acharya U, C. M. Lim, and M. Kagathi, "Automated identification of diabetic retinopathy stages using digital fundus images," *J. Med. Syst.*, vol. 32, no. 2, pp. 107–115, Apr. 2008.

[1] https://keras.io/

[12] H. Pratt, F. Coenen, D. M. Broadbent, S. P. Harding, and Y. Zheng, "Convolutional neural networks for diabetic retinopathy," *Procedia Comput. Sci.*, vol. 90, pp. 200–205, Jan. 2016.

[13] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2818–2826.

[14] Z. Wang and J. Yang, "Diabetic retinopathy detection via deep convolutional networks for discriminative localization and visual explanation," 2017, *arXiv:1703.10757*. [Online]. Available: http://arxiv.org/abs/1703.10757

[15] L. Wu, "Classification of diabetic retinopathy and diabetic macular edema," *World J. Diabetes*, vol. 4, no. 6, p. 290, 2013.

[16] S. H. M. Alipour, H. Rabbani, and M. R. Akhlaghi, "Diabetic retinopathy grading by digital curvelet transform," *Comput. Math. Methods Med.*, vol. 2012, Jan. 2012.

[17] W. K. Pratt, *Digital Image Processing*. New York, NY, USA: Wiley, 1978.

[18] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. neural Inf. Process. Syst.*, 2012, pp. 1097–1105.

[19] M. Tan, B. Chen, R. Pang, V. Vasudevan, M. Sandler, A. Howard, and Q. V. Le, "MnasNet: Platform-aware neural architecture search for mobile," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 2820–2828.

[20] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient convolutional neural networks for mobile vision applications," 2017, *arXiv:1704.04861*. [Online]. Available: http://arxiv.org/abs/1704.04861

[21] M. Raghu, B. Poole, J. Kleinberg, S. Ganguli, and J. S. Dickstein, "On the expressive power of deep neural networks," in *Proc. 34th Int. Conf. Mach. Learn. (JMLR)*, vol. 70, 2017, pp. 2847–2854.

[22] H. Lin and S. Jegelka, "Resnet with one-neuron hidden layers is a universal approximator," in *Proc. Adv. Neural Inf. Process. Syst.*, 2018, pp. 6169–6178.

[23] O. Sharir and A. Shashua, "On the expressive power of overlapping architectures of deep learning," 2017, *arXiv:1703.02065*. [Online]. Available: http://arxiv.org/abs/1703.02065

[24] Z. Lu, H. Pu, F. Wang, Z. Hu, and L. Wang, "The expressive power of neural networks: A view from the width," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 6231–6239.

[25] K.-Q. Huang, Q. Wang, and Z.-Y. Wu, "Natural color image enhancement and evaluation algorithm based on human visual system," *Comput. Vis. Image Understand.*, vol. 103, no. 1, pp. 52–63, Jul. 2006.

[26] Y.-T. Kim, "Contrast enhancement using brightness preserving bi-histogram equalization," *IEEE Trans. Consum. Electron.*, vol. 43, no. 1, pp. 1–8, Feb. 1997.

[27] D. Sheet, H. Garud, A. Suveer, M. Mahadevappa, and J. Chatterjee, "Brightness preserving dynamic fuzzy histogram equalization," *IEEE Trans. Consum. Electron.*, vol. 56, no. 4, pp. 2475–2480, Nov. 2010.

[28] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. Ter Haar Romeny, J. B. Zimmerman, and K. Zuiderveld, "Adaptive histogram equalization and its variations," *Comput. Vis., Graph., image Process.*, vol. 39, no. 3, pp. 355–368, 1987.

[29] K. Zuiderveld, "Contrast limited adaptive histogram equalization," in *Graphics Gems IV*. New York, NY, USA: Academic, 1994, pp. 474–485.

[30] B. S. Min, D. K. Lim, S. J. Kim, and J. H. Lee, "A novel method of determining parameters of CLAHE based on image entropy," *Int. J. Softw. Eng. Appl.*, vol. 7, no. 5, pp. 113–120, Sep. 2013.

[31] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 4510–4520.

[32] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7132–7141.

[33] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 248–255.

[34] *Messidor-2 Data is Available for Researchers in the Public Domain At*. Accessed: Jun. 18, 2019. [Online]. Available: https://medicine.uiowa.edu/eye/abramoff

[35] E. Decencière, X. Zhang, G. Cazuguel, B. Lay, B. Cochener, C. Trone, P. Gain, R. Ordonez, P. Massin, A. Erginay, and B. Charton, "Feedback on a publicly distributed image database: The messidor database," *Image Anal. Stereology*, vol. 33, no. 3, pp. 231–234, 2014.

[36] P. Porwal, S. Pachade, R. Kamble, M. Kokare, G. Deshmukh, V. Sahasrabuddhe, and F. Meriaudeau, ''Indian diabetic retinopathy image dataset (IDRiD): A database for diabetic retinopathy screening research,'' *Data*, vol. 3, no. 3, p. 25, Jul. 2018.
[37] C. I. Sánchez, M. Niemeijer, A. V. Dumitrescu, M. S. Suttorp-Schulten, M. D. Abràmoff, and B. van Ginneken, ''Evaluation of a computer-aided diagnosis system for diabetic retinopathy screening on public data,'' *Investigative Ophthalmol. Vis. Sci.*, vol. 52, no. 7, pp. 4866–4871, Jun. 2011.
[38] A. Rocha, T. Carvalho, H. F. Jelinek, S. Goldenstein, and J. Wainer, ''Points of interest and visual dictionaries for automatic retinal lesion detection,'' *IEEE Trans. Biomed. Eng.*, vol. 59, no. 8, pp. 2244–2253, Aug. 2012.
[39] X. Zhang, G. Thibault, E. Decencière, B. Marcotegui, B. Laÿ, R. Danno, G. Cazuguel, G. Quellec, M. Lamard, P. Massin, A. Chabouis, Z. Victor, and A. Erginay, ''Exudate detection in color retinal images for mass screening of diabetic retinopathy,'' *Med. Image Anal.*, vol. 18, no. 7, pp. 1026–1043, Oct. 2014.
[40] L. Seoud, T. Hurtut, J. Chelbi, F. Cheriet, and J. M. P. Langlois, ''Red lesion detection using dynamic shape features for diabetic retinopathy screening,'' *IEEE Trans. Med. Imag.*, vol. 35, no. 4, pp. 1116–1126, Apr. 2016.
[41] 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,'' *Comput. Methods Programs Biomed.*, vol. 153, pp. 115–127, Jan. 2018.
[42] G. T. Zago, R. V. Andreão, B. Dorizzi, and E. O. T. Salles, ''Diabetic retinopathy detection using red lesion localization and convolutional neural networks,'' *Comput. Biol. Med.*, vol. 116, Jan. 2020, Art. no. 103537.

ASRA MOMENI POUR is currently pursuing the degree in biomedical engineering with the Department of Biomedical Engineering, Faculty of Advanced Medical Sciences, Tabriz University of Medical Sciences, Tabriz, Iran. Her research interests include image processing, data science, machine learning, spectral and kernel clustering, and pattern recognition.

HADI SEYEDARABI received the B.S. degree from the University of Tabriz, Tabriz, Iran, in 1993, the M.S. degree from the K. N. Toosi University of Technology, Tehran, Iran, in 1996, and the Ph.D. degree from the University of Tabriz, Iran, in 2006, all in electrical engineering. He is currently a Professor at the Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz. His research interests include image processing, computer vision, video coding, human–computer interaction, facial expression recognition, and deep learning applications in image processing and computer vision.

SEYED HASSAN ABBASI JAHROMI received the B.Sc. degree in mechanical engineering from the Engineering School, Shiraz University, Shiraz, Iran, in May 1976, the M.Sc. degree in biomedical engineering from the Engineering School, Northwestern University, Evanston, IL, USA, in March 1979, and the M.Sc. degree in structural engineering and the D.Sc. degree in biological engineering from the Engineering School, Washington University, St. Louis, MO, USA, in March 1995. His Quality Factor Analysis (QFA), a new clinic ultra-sensitive acoustic technique, for predication of bone strength and fracture risk with application to osteoporosis.

ALIREZA JAVADZADEH received the M.D. degree from the Tabriz University of Medical Sciences, in 1987. He completed the Ophthalmology Residency and Vitreoretinal Fellowship at Nikookari Eye Hospital, Tabriz University of Medical Sciences, and Hazrat Rasoul Akram Hospital, Iran University of Medical Sciences, in 1994 and 1998, respectively. He is currently a Professor of ophthalmology. Since 1994, he has been a Faculty Member of the Tabriz University of Medical Sciences, where he was the Head of Ophthalmology, for five years.