Mansoura University
Faculty of Engineering
Mansoura Engineering Journal

Deep Learning Grading System for Diabetic Retinopathy using Fundus Images

D. Elsawah, *A. Elnakib**, and H. S. Moustafa

KEYWORDS:
Deep learning, Fundus images, Diabetic Retinopathy, Grading

***Abstract*—Diabetic Retinopathy (DR) is one of the main causes of blindness that can be overcome, if it is early detected. This work proposes an automated early detection and grading of DR using fundus images. The proposed detection and grading system investigate different deep learning architectures (i.e., ResNet and AlexNet) that are applied to an augment data to extract deep compact features of the fundus images. The extracted features are input to a pixel-wise Neural Network (NN) classifier or a Support Vector Machine (SVM) classifier for automated DR grading. The performance of the proposed system is evaluated using a publically available fundus Indian Diabetic Retinopathy Image Dataset (IDRiD), collected for ISBI-2018 challenge. The IDRiD dataset consists of 516 retinal images of normal and different DR grades, i.e., mild, moderate, severe, and Proliferative Diabetic Retinopathy (PDR). Our system achieves an overall accuracy of 95.73%, sensitivity of 95.73%, and specificity of 98.51% utilizing an AlexNet-based architecture and a pixel-wise NN classifier. Compared to the previous related work, the proposed system shows promising DR grading performance.**

I. INTRODUCTION

DIABETIC Retinopathy (DR) is one of the main causes of visual blindness. According to the recent report of the Vision Loss Expert Group (VLEG) of the global burden of disease study, DR has increased the number of blind people, during the period from 1990 to 2015, from 0.2 million to 0.4 million [1]. Early recognition of DR is the way to prevent the occurrence of blindness. The evolution of DR in the patient is determined by the presence of anomaly lesions, such as microaneurysms [2], hemorrhages [3], and hard exudates [4],[5]. This paper aims at providing an automated DR grading system that can not only recognize early DR occurrence but can also differentiate between the normal and the standard clinical stages of the disease, i.e., mild, moderate, severe, and PDR.

For the task of automated DR grading and/or related lesions detection, different methodologies have been developed throughout the literature. These methods can be categorized as traditional methods or deep learning methods [6]. Traditional methods attempt to extract handcrafted features from fundus images, and then use a classifier for lesion detection or DR grading. For example, Marin et al. [7] applied a system to detect the retinal exudate lesions based on feature extraction and supervised regression-based classification. Pereira et al. [8] used a multi-agent system model for microaneurys segmentation. On the other hand, deep learning algorithms, based on Convolutional Neural Networks (CNN), were recently applied to retinal images, e.g., for lesion detection [9-13] and

Received: (12 July, 2020) - Revised: (8 October, 2020) - Accepted: (11 October, 2020)

D. Elsawah, is with the Electronics and Communications Engineering Department, Faculty of Engineering, Mansoura University, 35516 Mansoura City, Egypt (e-mail: doaakhalil313@gmail.com)

** **Corresponding author:** A. Elnakib, is with the Electronics and Communications Engineering Department, Faculty of Engineering, Mansoura University, 35516 Mansoura City, Egypt (e-mail: nakib@mans.edu.eg)*

H. S. Moustafa, is with the Electronics and Communications Engineering Department, Faculty of Engineering, Mansoura University, 35516 Mansoura City, Egypt (e-mail: hossam_moustafa@hotmail.com).

for DR grading [14][15], showing promising results. For example, Tan et al. [9] applied an algorithm for lesion detection i.e., microaneurysms, hemorrhages, and exudate, using a CNN model. This model consists of nine layers, three convolutional layers, three max-pooling layers, and three fully connected layers. Lam et al. [10] used a CNN model to generate specific probability maps for localizing pathologic lesions. Quellec et al. [11] used a CNN model to create heatmaps to detect lesion locations. Chudzik et al. [12] applied a patch-based fully CNN model, followed by a decision tree classifier for automated DR detection. Xu et al. [13] utilized a CNN with eight convolution layers, two fully connected layers, and a softmax layer to detect DR from color fundus image. Kori et al. [14] applied an ensemble of DensNet and ResNet models to differentiate between DR stages, resulting in an accuracy of 83.9%. Recently, Elsawah et al. [15] applied a preprocessing step followed by a ResNet model to achieve an accuracy of 88% on IDRiD dataset. While these algorithms have achieved considerable success in DR grading, the task of improving the classification accuracy should be still investigated to expand the clinical benefits.

This study presents an automated DR classification system using CNN modelling and data augmentation. The main contributions of this work are as follows:

- Investigating different deep learning architectures for DR grading (i.e., AlexNet and ResNet)
- Improving the performance of deep learning training using data balancing and augmentation
- Performance evaluation on the challenging ISBI'2018 IDRiD dataset using the standard metrics; i.e., sensitivity, specificity, and accuracy.

Our proposed system shows the ability to differentiate between normal and different DR grades, i.e., mild, moderate, severe, or PDR, with an improved performance over other related works on the same challenging ISBI'2018 IDRiD data. The rest of this paper is as follows: Section 2 illustrates the collected database and the proposed system details. Section 3 demonstrates the results and related discussions. Finally, Section 4 concludes the paper.

II. Materials and Methods

The proposed deep learning system, (see Figure 1), is designed based on three processing stages. Initially, data is augmented to improve the training capabilities of the proposed deep learning model. In the second stage, a transfer learning is applied by investigating two architectures, i.e., ResNet50 and AlexNet. Finally, a classification stage is used to grade fundus image.

A. *DR Data Description*

Several datasets have addressed DR grading. The most recent ones include the DDR [16], the Kaggle [17], and the IDRiD [18] datasets. A collection of DR datasets can be found in [19]. Throughout the current study, the proposed system is tested using the IDRiD dataset [18], since it represents an updated recent data and is challenging (a part of the ISBI-2018 segmentation and grading challenge). Fig. 2 shows five samples of the data, representing the five data grades; normal, mild, moderate, severe, and Proliferative Diabetic Retinopathy (PDR). Normal images contains no lesions. Lesions are either dark (small red round spots, i.e., Microaneurysms (MA) or larger irregular red spots, i.e., Haemorrhages (HM)) or bright. Bright lesions include soft exudates (SE) or hard exudates (HE)). SE are cotton wool edema, i.e., are white retina spots caused by the swelling of the nerve fiber, whereas HE are bright-yellow spots due to plasma leakage. Different grade lesion samples are shown to in Fig. 2. Mild grade indicates the existence of MA lesions only [19]. Moderate grade indicates more than just MA lesions but less than severe grade [19]. Severe grade indicates severe lesions with no sign of proliferative DR [19]. PDR indicates one or more pre-retinal/vitreous HM or neovascularization [19]. More details about the difference between the data grades can be found in [19]. More details about the IDRiD data can be found in [18].

B. *Data Augmentation and Balancing*

To improve the performance of the proposed deep learning model, the IDRiD database is augmented. First, the original fundus images are resized from its original size (i.e., 2848×4288) to 256×256 to afford high-speed processing. We follow the same augmentation steps as in [14]; five cropped images are created to the dimension of 224×224 as follows: four images are cropped from each corner of the image and one image is cropped around the center of the image, such as the size of any cropped image is 224x224 (see Figure 3). Further, the five cropped images were flipped horizontally along the vertical axis to create a total of ten cropped images of size 224x224 (see Figure 3).

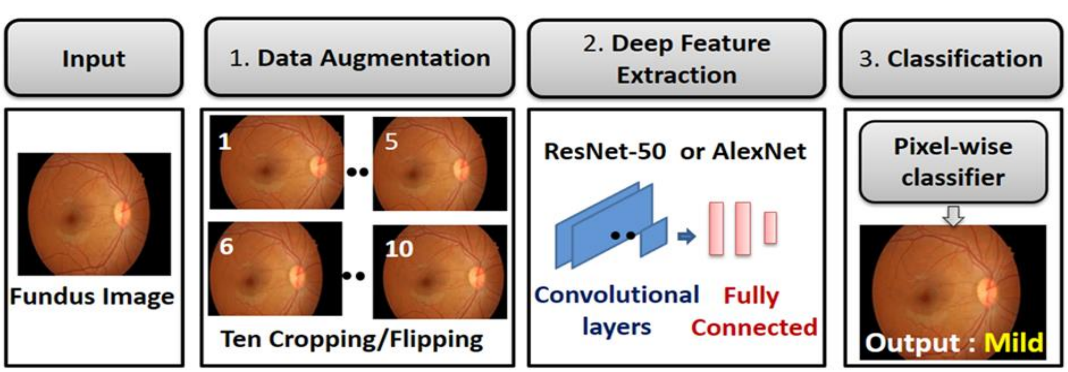
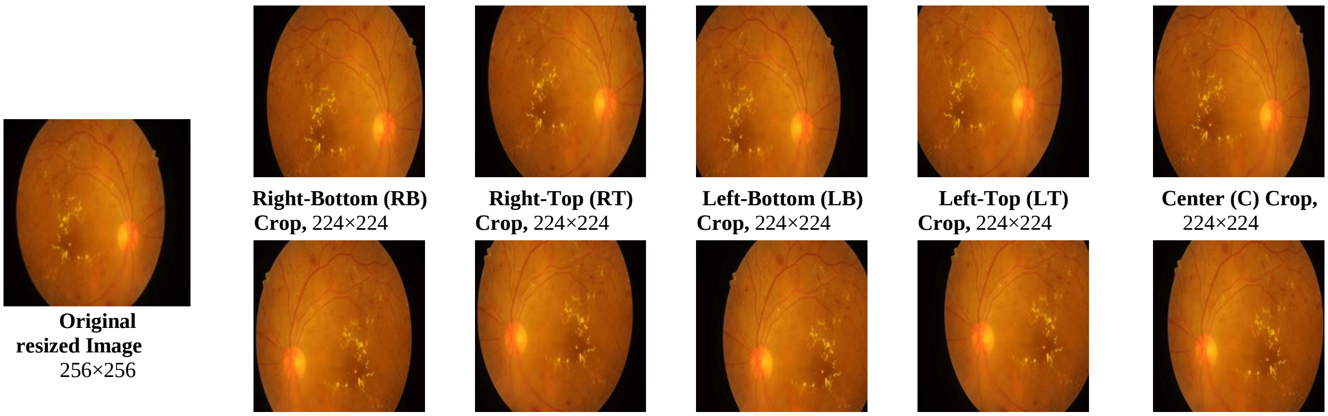


Fig.1. Proposed system for DR grading: Raw input data images are augmented using ten cropping/flipping, then deep features are extracted from augmented images, and finally grading is performed using a pixel-wise classifier.

(a) normal **(b) mild** **(c) moderate** **(d) severe** **(e) PDR**

Fig.2. Typical samples of fundus images that represent DR stages: (a) normal stage, (b) mild-NPDR (early) stage, (c) moderate-NPDR stage (d) severe-NPDR stage, and (e) PDR stage. MA, HM, SE, HE stand for Microaneurysms, Haemorrhages, soft exudates, and hard exudates, respectively.



Original resized Image 256×256

Right-Bottom (RB) Crop, 224×224 **Right-Top (RT) Crop,** 224×224 **Left-Bottom (LB) Crop,** 224×224 **Left-Top (LT) Crop,** 224×224 **Center (C) Crop,** 224×224

Flipped RB Crop, 224×224 **Flipped TR Crop,** 224×224 **Flipped LB Crop,** 224×224 **Flipped LT Crop,** 224×224 **Flipped C Crop,** 224×224

Fig.3. Data augmentation process: ten augmented images are created per a raw sample (first column). First row shows four cropped images from each corner of the image, named by the corner, and one cropped image around the center of the image. The second row shows their horizontal flipping around the vertical axis

Since (*i*) the number of image per each grade for the IDRiD is unbalanced, and (*ii*) the size of the augmented data is ten time the size of the raw data, the proposed system attempt to balance the data by randomly selecting an equal number of image for each grade to avoid biasing the training results. This step is referred to in this paper by the term "data balancing".

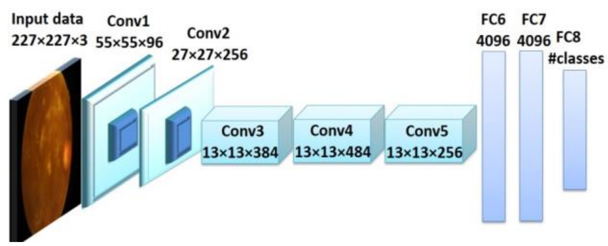


Fig.4. AlexNet architecture developed by Alex Krizhevsky [22], containing five convolutional layers (Conv1 to Conv5) and three fully connected (FC6 to FC8) layers.

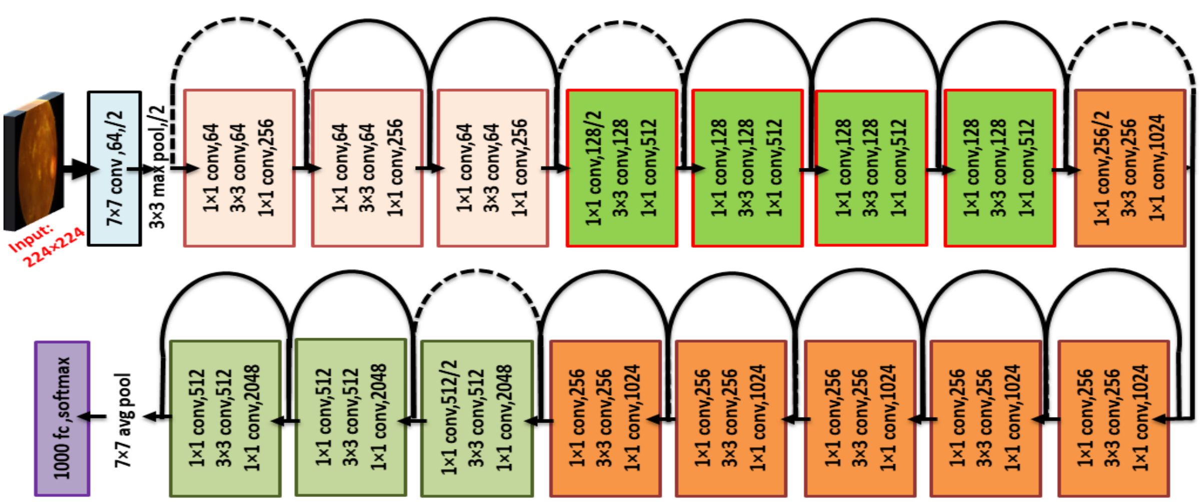


Fig.5. ResNet-50 structure composed of fifty weighted layers, with 23,534,592 trainable parameters. Details can be found in [23].

C. *Deep-learning Architecture*

The proposed system adopts applying transfer learning [20, 21] using the well-known CNN models, i.e., AlexNet (Figure 4, [22]) and ResNet50 (Figure 5, [23]). Transfer learning is based on using pre-trained models that have been trained over a huge training database. The parameters of the convolutional layers of the pre-trained networks are transferred (kept unchanged) to the target model, whereas the last fully connected layers are trained using the new data (i.e., in our case, the 516 fundus images). ResNet is a short name for Residual Network. ResNet50, used in the proposed system, is composed of fifty weighted layers, with 23,534,592 trainable parameters. The basic idea of a ResNet model is to skip blocks of convolutional layers by using shortcut connections (see Figure 5). On the other hand, AlexNet is composed of five convolutional layers and three fully connected layers (see Figure 4). More details about AlexNet and ResNet models can be found in [22] and [23], respectively.

To apply transfer learning of the CNN models, the last fully connected layer of the pre-trained models (FC8 layer in AlexNet or FC1000 layer in ResNet50) is replaced with a new fully connected layer with four nodes, representing the normal and the DR grades. The parameters of the introduced layers are trained using our IDRiD data. The vectors of activities of the last fully connected layer just before FC8 layer in AlexNet (i.e, the FC7 layer) or just before FC1000 in Res-Net50 (i.e., the flatten layer), represent the feature descriptor of the IDRiD data.

D. *Classification*

Two types of classifiers are investigated in grading (i.e., a pixel-wise NN and SVM [24],[25]). The input of the classifier is the output vector of activity (FC7 in AlexNet or the flatten layer in ResNet50) and the classifier output is the image grade. The pixel-wise NN classifier is a four-neuron output layer that maps the feature space to the grades. The SVM performs multi-linear mapping using a basic binary kernel. The performance is evaluated for either the ResNet50 or AlexNet, each along with either NN classifier or the SVM classifier based on standard performance evaluation metrics, in order to select the best classifier.

E. *Performance Evaluation*

The individual sensitivity, Sen_i, individual specificity, $Spec_i$, and the individual accuracy, AC_i, for each DR grade i, are three standard classification (grading) metrics [26, 27] that are used for evaluating the proposed system performance. For a certain grade i, Let P_i and N_i denote the number of the real positive and negative cases in the data. Therefore, to test a given system, $P_i = TP_i + FN_i$, where P_i is the sum of the number of the true positive cases achieved by the tested algorithm plus the number of the missed positive cases missed by the tested algorithm. Likewise $N_i = TN_i + FP_i$, i.e., N_i is the sum of the number of the true negative evaluated by the tested algorithm plus the number of the mistakenly positive cases pointed out the tested algorithm. Following these notation, the three performance metrics are evaluated as follows:

$$AC_i = \frac{TP_i + TN_i}{P_i + N_i} = \frac{TP_i + TN_i}{TP_i + FN_i + TN_i + FP_i} \quad (1)$$

$$Sen_i = \frac{TP_i}{P_i} = \frac{TP_i}{TP_i + FP_i} \quad (2)$$

$$Spec_i = \frac{TN_i}{N_i} = \frac{TN_i}{TN_i + FP_i} \quad (3)$$

Where i is the grade type, out of ‘normal’ and four possible DR grades, i.e., ‘mild’, ‘moderate’, ‘severe’, or ‘PDR’. Note that, as done in [14], due to the closeness between ‘Severe’ and ‘PDR’ images (see Fig.2), the grade ‘Severe’ and ‘PDR’ are grouped in one class, labelled as ‘severe-PDR’. All experiments in this paper are based on this categorization. The overall accuracy AC, sensitivity Sen, and specificity $Spec$ are evaluated as the mean of the five individual grades’ values.

III. Results and Discussion

The proposed system performance has been evaluated for using either the ResNet50 or AlexNet, and either the NN or the SVM classifiers, using the three-performance metrics in order to select the best DR grading system. This section illustrates, in the best DR grading system. This section illustrates, in details, the experimental setup, results, and related discussions.

TABLE I
COMPARISON RESULTS BETWEEN RELATED METHODS [14],[15] AND OUR PROPOSED DR GRADING SYSTEM FOR DIFFERENT TRAINING DATA: CASE I (NO AUGMENTATION & NO DATA BALANCING), CASE II (AUGMENTATION & NO DATA BALANCING), AND CASE III (AUGMENTATION & DATA BALANCING), USING ONE OF THE TWO DEEP LEARNING MODELS (RESNET50 OR ALEXNET) WITH NN CLASSIFIER OR SVM CLASSIFIER. NA DENOTES NONAPPLICABLE.

Case	Grade	#Images	ResNet50&NN			ResNet50&SVM			AlexNet&SVM			AlexNet&NN		
			AC_i	Sen_i	$Spec_i$	AC_i	Sen_i	$Spec_i$	AC_i	Sen_i	$Spec_i$	AC_i	Sen_i	$Spec_i$
Case I Raw/ Unbalanced	Normal	167	80.51	80.00	80.76	80.28	55.56	92.18	73.61	90.60	65.43	78.71	90.60	81.73
	Mild	25	94.81	0.00	99.32	95.28	0.00	100	95.28	0.00	100	94.81	0.00	99.32
	Moderate	169	68.18	58.82	72.81	60.28	72.03	54.54	68.06	9.32	96.70	71.71	9.32	74.75
	Sever-PDR	155	81.17	56.52	91.67	79.17	52.78	90.67	73.06	75.00	72.22	82.46	75.00	90.74
	Mean	516	81.17	62.34	86.14	78.75	45.09	84.29	77.50	43.73	83.59	81.92	63.71	86.64
Case II Augmented/ Unbalanced	Normal	1670	82.95	59.88	93.98	87.54	70.32	95.78	79.76	93.24	73.31	93.51	90.62	94.36
	Mild	250	98.86	0.00	80.98	91.92	62.86	93.4	94.27	20.57	98.02	97.16	69.33	98.57
	Moderate	1690	69.45	67.06	70.75	78.27	56.72	88.76	75.72	50.30	88.1	90.12	82.64	93.76
	Sever-PDR	1550	81.91	73.55	85.50	84.91	88.48	95.28	84.41	64.70	92.88	94.15	92.26	95.48
	Mean	5160	83.29	63.44	82.80	85.66	69.59	93.31	83.54	57.20	88.08	93.70	87.47	95.54

Case III Augmented/ Balanced	Normal	250	94.93	94.67	95.00	95.20	97.33	94.67	96.00	93.33	96.67	97.87	94.67	98.67
	Mild	250	97.60	98.67	97.33	97.60	96.00	98.00	94.67	94.67	94.67	97.67	100	97.33
	Moderate	250	81.97	69.33	99.33	92.00	64.00	99.00	86.93	34.67	100	97.35	93.33	99.34
	Sever-PDR	500	91.73	90.67	88.89	96.00	97.33	95.11	87.73	93.33	82.67	97.33	95.33	98.67
	Mean	1250	91.56	88.80	95.14	95.20	88.67	96.70	91.33	79.50	93.50	97.56	95.73	98.51
Related works	Metric	#Images	AC	Sen	$Spec$									
Ensemble model [14]	Mean	56	83.90	NA	NA									
ResNet [15]	Mean	250	94.00	88.00	96.11									

A. Collected Database

The IDRiD data consists of 516 images, captured by a retinal specialist at an eye clinic, located in Nanded, Maharashtra, India. Images were acquired using a Kowa VX-10 alpha digital fundus camera with a 50-degree field of view (FOV), and all are centralized near to the macula. The original images have a resolution of 2848×4288 pixels and are stored in "jpg" file format. Typical samples are provided in Figure 2. More data details can be found on [18].

B. Experimentation Setting

The system classifier is trained and tested on 70% and 30%, respectively, of the whole dataset. The deep learning models are trained by optimizing the cross-entropy loss with a learning rate of 10^{-4} and a maximum number of epochs of 30.

The performance results of the proposed system are carried over three cases, described in Table I. Case I involves no augmentation and no data balancing, Case II permits augmentation with no data balancing, and finally, Case III permits both augmentation with balancing data. Specific numbers of images for each class label, for all cases, are demonstrated in Table I.

C. Quantitative Results

Table I shows the accuracies, sensitivities, and specificities of the different training data setting of the proposed system. More specifically, the three different cases, Case I (no augmentation & no data balancing), Case II (augmentation & no data balancing), and Case III (augmentation & data balancing), are involved. For each case, the mix between the two utilized models (i.e., ResNe50 and AlexNet) and the two classifiers (NN and SVM) are investigated. Since the training data is relatively of small size, we expect that to avoid the problem of overfitting, the use of the AlexNet model, composed of lower number of layers (five convolutional layers and two fully connected layers), could achieve better results than ResNet50 (composed from 50 layers). In addition, we expect, for the same reason, that using a simple NN could achieve better accuracy than a sophisticated SVM classifier. As expected, Table I demonstrates that using AlexNet with a NN classifier achieves the best performance for all investigated cases, e.g., the mean accuracies AC are 81.92%, 93.70%, and 97.56% for Case I, Case II, and Case III, respectively (see Fig. 6). In addition, it is clear that Case III achieves the best performance among the competing cases, since it takes into account both the augmentation and balancing. Figure 7 exemplifies samples of correctly classified visual results of fundus images from different test subjects. These results highlight the potential of using AlexNet/NN, taking into account both augmentation and balancing, for achieving high performance DR grading of the challenging data.

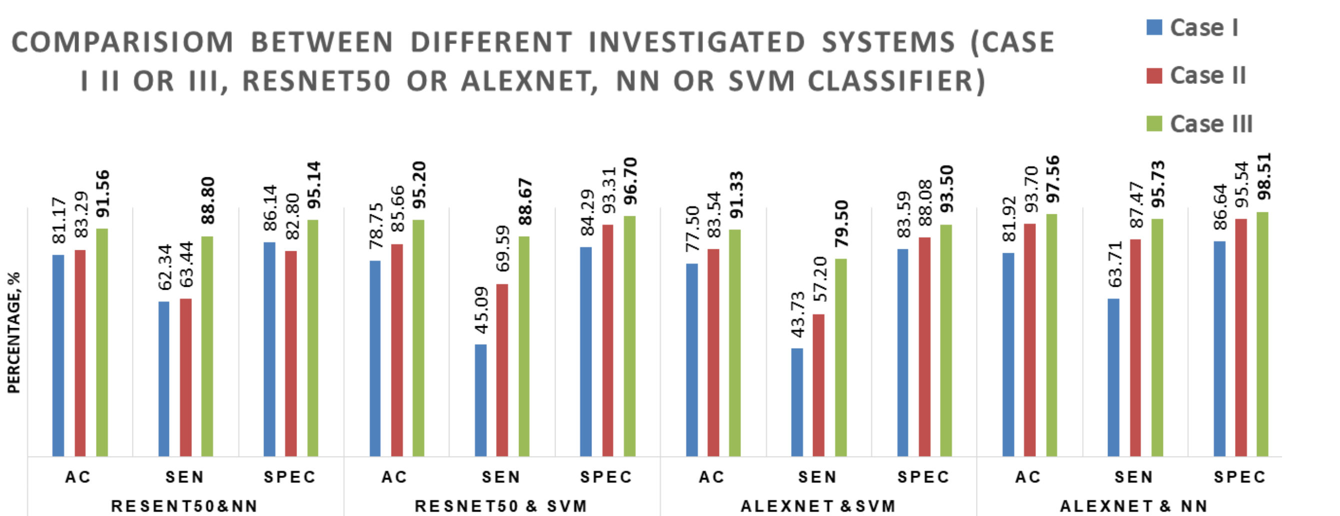


Fig. 6. Comparison results of the different cases of the proposed system, i.e., using different training data: Case I (no augmentation & no data balancing), Case II (augmentation & no data balancing), and Case III (augmentation & data balancing)

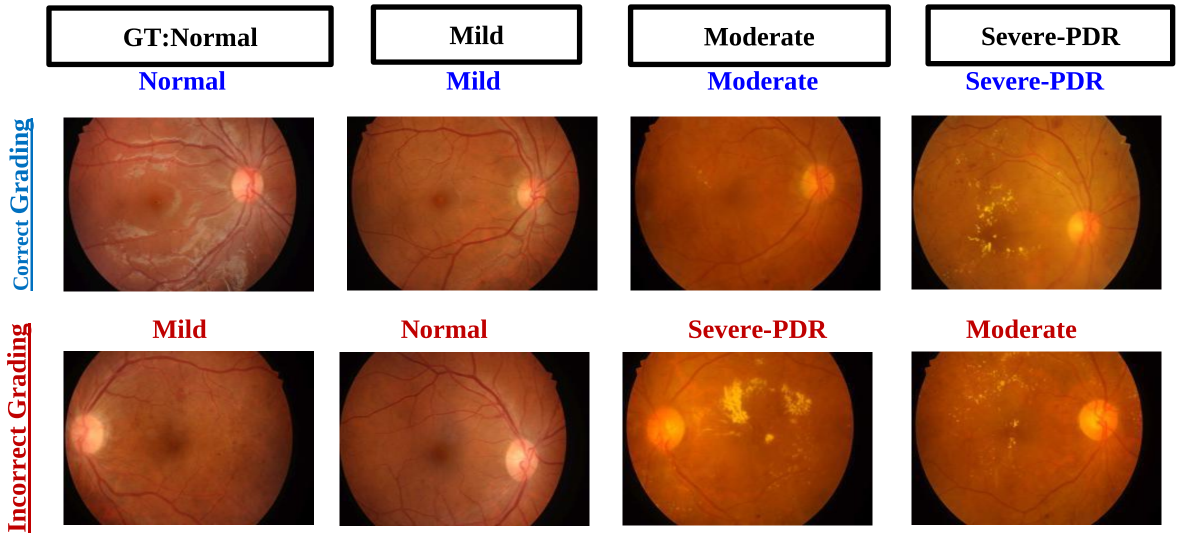


Fig.7. Classification examples of retinal images of different DR stages using our proposed system (Alexnet/NN) with Case III setting (data augmentation & data balancing). First row shows correctly classified images, and second row shows incorrectly classified images. GT denotes the ground truth grading; each column has a different GT grade arranged as: "normal", then "mild", "moderate", and finally "severe-PDR".

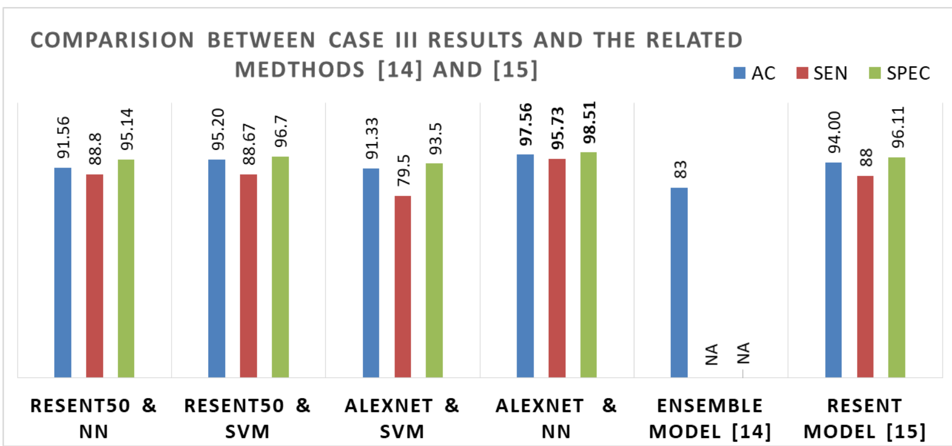


Figure 8. Comparison results between the proposed system Case III results (using ResNet50 or AlexNet with a NN classifier or SVM classifier) and related methods [14] and [15].

In order to justify quantitatively the need for the data augmentation and balancing, the performance results of the proposed system for the best deep-learning-model/classifier (i.e., AlexNet/NN) have been compared over the three different cases, described in Table I. As expected, data augmentation helps to significantly increase the size of the training data ten times, and therefore significantly improves the overall AC from 81.92% to 93.70%. In addition, data balancing further improves the overall AC from 93.70% to 97.56%. These results highlight the strength of the proposed data augmentation and balancing in improving the quality of the training process (See Figure 6). To further investigate the advantages of the proposed DR grading system, it has been compared with the related state-of-the-art methods, on the same database. In Table I and Fig. 7, the proposed grading system achieves a superior performance over the related methods [14], [15]. These results highlights its advantages.

D. Limitations

Although the current study achieves a superior performance over the related methods, the second row in Fig. 7 shows that the proposed system is not able to find the correct grading for specific images. As shown in Fig. 7, although the system output grade is incorrect, it points to an adjacent grade (e.g., mild is classified to be either normal or moderate), which may be visually close. These results show the possibility of fine-tuning the proposed system to fix these errors in the future, by seeking the integration of new relevant discriminant features.

IV. CONCLUSION AND FUTURE WORK

In this work, we present an automated DR grading system based on deep learning, data augmentation, and data balancing. First, the data is augmented and balanced in order to improve the training capabilities of the proposed deep learning model. Second, a transfer learning is applied by investigating two architectures, i.e., ResNet50 and AlexNet, to extract retinal image features. Finally, a classifier is used to grade fundus images into one of four grades: normal, mild, moderate, or severe-PRD. This system has been tested over the challenging

ISBI'2018 IDRiD database, achieving superior performance over related methods, evidenced by the obtained higher classification accuracy, sensitivity and specificity. Later on, further features will be investigated to improve the accuracy. In addition, patch-based analysis will be investigated to improve the results. Furthermore, other databases will be targeted in order to quantify the robustness of the proposed DR grading system.

References

[1] M. S. Nowak and A. Grzybowski. (2018), "Review of the epidemiology of diabetic retinopathy", Modern Reina from Ophthalmology Times. [Online]. Available: http://www.modernretina.com/diabetic-retinopathy/review-epidemiology-diabetic-retinopathy/.

[2] P. Costa, A. Galdran, A. Smailagic, and A. Campilho, "A weakly-supervised framework for interpretable diabetic retinopathy detection on retinal images," IEEE Access, vol. 6, pp. 18747-18758, 2018.

[3] G. Ometto, P. Assheton, F. Calivá, P. Chudzik, B. Al-Diri, A. Hunter, and T. Bek, "Spatial distribution of early red lesions is a risk factor for development of vision-threatening diabetic retinopathy, " Diabetologia, vol. 60, pp. 2361-2367, 2017

[4] Z. Wang, Y. Yin, J. Shi, W. Fang, H. Li, and X. Wang, "Zoom-in-net: Deep mining lesions for diabetic retinopathy detection," in International Conference on Medical Image Computing and Computer-Assisted Intervention, Cham, Springer, September 2017, pp. 267-275.

[5] C. Pratheeba, and N.N. Singh, "A Novel Approach for Detection of Hard Exudates Using Random Forest Classifier," Journal of medical systems, vol. 43, pp.180, 2019.

[6] X. Li, X. Hu, L. Yu, L. Zhu, C. W. Fu, and P. A. Heng, " CANet: Cross-disease Attention Network for Joint Diabetic Retinopathy and Diabetic Macular Edema Grading, " IEEE Transactions on Medical Imaging, pp. 1-1 , November 2019.

[7] D. Marin, M. E. Gegundez-Arias, B. Ponte, F. Alvarez, J. Garrido, C. Ortega, M. J. Vasallo, and J. M. Bravo, "An exudate detection method for diagnosis risk of diabetic macular edema in retinal images using feature-based and supervised classification, " Medical and Biological Engineering and Computing, vol.56, pp. 1379-1390, August 2018.

[8] C. Pereira, D. Veiga, J. Mahdjoub, Z. Guessoum, L. Gonçalves, M. Ferreira, and J. Monteiro, "Using a multi-agent system approach for microaneurysm detection in fundus images," Artificial Intelligence in Medicine, vol. 60, pp. 179-188, March 2014.

[9] J. H. Tan, H. Fujita, S. Sivaprasad, S. V. Bhandary, A. K. Rao, K. C. Chua, and U. R. Acharya, "Automated segmentation of exudates, haemorrhages, microaneurysms using single convolutional neural network," Information Sciences, vol. 420, pp. 66-76, December 2017.

[10] C. Yu. C. Lam, L. Huang, and D. Rubin, "Retinal lesion detection with deep learning using image patches," Investigative Ophthalmology and Visual Science, vol. 59, pp. 590-596, January 2018.

[11] G. Quellec, K. Charrière, Y. Boudi, B. Cochener and, M. Lamard, "Deep image mining for diabetic retinopathy screening," Medical Image Analysis, vol. 39, pp. 178-193, July 2017.

[12] P. Chudzik, S. Majumdar, F. Calivá, B. Al-Diri and A. Hunter, "Microaneurysm detection using fully convolutional neural networks," Computer Methods and Programs in Biomedicine, vol. 158, pp.185-192, May 2018.

[13] K. Xu, D. Feng, and H. Mi, "Deep convolutional neural network-based early automated detection of diabetic retinopathy using fundus image," Molecules, vol. 22, pp. 2054, 2017.

[14] A. Kori, S. S. Chennamsetty and V. Alex, "Ensemble of convolutional neural networks for automatic grading of diabetic retinopathy and macular edema," arXiv preprint arXiv, pp.1809.0422, 2018.

[15] D. K. Elswah, A. Elnakib, H. Moustafa, "Automated Diabetic Retinopathy Grading using ResNET," in: Proceeding of 37th National Radio Science Conference (NRSC 2020), September 8-10, 2020.

[16] T. Li, Y. Gao, K. Wang, S. Guo, H. Liu, and H. Kang. Diagnostic assessment of deep learning algorithms for diabetic retinopathy screening. Inf Sci 2019; 501:511–22.

[17] Kaggle dataset [Online]. Available, https://kaggle.com/c/diabetic-retinopat hy-detection.

[18] Diabetic Retinopathy: Segmentation and Grading Challenge, Indian Diabetic Retinopathy Image Dataset (IDRiD), workshop at IEEE International Symposium on Biomedical Imaging (ISBI-2018) (2018), ISBI challenge website. [Online]. Available: https://idrid.grand-challenge.org/.

[19] W. L. Alyoubi, W. M. Shalash, and M. F. Abulkhair, Diabetic retinopathy detection through deep learning techniques: A review. Informatics in Medicine Unlocked, p.100377, 2020.

[20] K. Weiss, T.M. Khoshgoftaar, and D. Wang, "A survey of transfer learning, " Journal of Big data, vol.3, pp. 9, 2016

[21] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?" Advances in Neural Information Processing Systems, pp. 3320–3328, 2014.

[22] A.Krizhevsky, I. Sutskever, and G.E. Hinton, "Imagenet classification with deep convolutional neural networks," Advances in neural information processing systems, pp. 1097-1105, 2012.

[23] E. Rezende, G. Ruppert, T. Carvalho, F. Ramos and P. De Geus, "Malicious software classification using transfer learning of resnet-50 deep neural network, " in the 16th IEEE International Conference on Machine Learning and Applications, ICMLA, IEEE, December 2017, pp. 1011-1014.

[24] U. Shrivastava, "Multi-class Diagnosis of Diabetic Retinopathy using Deep Learning," Doctoral dissertation, Dhirubhai Ambani Institute of Information and Communication Technology, 2018.

[25] E. V. Carrera, A. González, and R. Carrera, "Automated detection of diabetic retinopathy using SVM, " in 2017 IEEE XXIV International Conference on Electronics, Electrical Engineering and Computing (INTERCON) , IEEE, August 2017,pp. 1-4.

[26] H. H Vo, and A. Verma, "New deep neural nets for fine-grained diabetic retinopathy recognition on hybrid color space, " in 2016 IEEE International Symposium on Multimedia (ISM), IEEE, December 2016, pp. 209-215.

[27] M. Chen, Y. Hao, K. Hwang, L.Wang and L.Wang, "Disease prediction by machine learning over big data from healthcare communities, " IEEE Access, vol. 5, pp. 8869-8879, 2017.

Doaa Elsawah is a graduate student in Electronics and Communications Engineering Department, Faulty of Engineering, Mansoura University since 2017. She received her BSc from the Electronics and Communications Engineering Department, Faulty of Engineering, Al-Azhar University in 2015. Her research interests include deep learning and medical image analysis. She has gained a two-year hand-on experience on medical data analysis during her research studies.

Ahmed A. Elnakib is an associate professor in ECE department, Mansoura University. D. Elnakib have authored or co-authored more than 30 journal articles, 7 book chapters, and 35 peer-reviewed conference papers. D. Elnakib is a regular reviewer for top international medical signal analysis journals that include Medical Image Analysis, IEEE Transactions on Medical Imaging, and Neurocomputing. In 2013, he has awarded the John Houchens Prize for the best outstanding dissertation. In 2020, he has awarded both the Mansoura University and the country encouragement awards in Engineering

Hossam El-Din Moustafa is an associate Professor at the Department of Electronics and Communications Engineering, and the founder and executive manager of Biomedical Engineering Program (BME) at the Faculty of Engineering, Mansoura University. The main research points include biomedical image and signal processing and deep learning applications.

Title Arabic:

نظام التعليم العميق لتصنيف اعتلال الشبكية السكري باستخدام صور قاع العين

Arabic Abstract:

يعد اعتلال الشبكية السكري أحد الأسباب الرئيسية للعمى التي يمكن التغلب عليها إذا تم الكشف عنها مبكرًا. يقترح هذا العمل الكشف الآلي المبكر وتصنيف درجات اعتلال الشبكية السكري (DR) باستخدام صور قاع العين. يفحص نظام الكشف والتصنيف المقترح بنيات تعلم عميق مختلفة (مثل ResNet و AlexNet) التي يتم تطبيقها على بيانات محسنة لإستخراج الميزات المدمجة العميقة لصور قاع العين. يتم إدخال الميزات المستخرجة إما إلى مصنف الشبكة العصبية (NN) أو المصنف دعم آلة ناقل (SVM) لتصنيف DR الآلي. يتم تقييم أداء النظام المقترح باستخدام مجموعة من صور اعتلال الشبكية السكري الهندية (IDRiD) ، التي تم جمعها لتحدي ISBI-2018. تتكون مجموعة بيانات IDRiD من 516 صورة لشبكية العين من درجات DR الطبيعية والمختلفة، هذه الدرجات هي اعتلال الشبكية السكري الخفيف والمعتدل والشديد والتكاثري. يحقق نظامنا إجمالي دقة 95.73٪ وحساسية 95.73٪ وانتقائية98.51% باستخدام بنية AlexNet ومصنف NN. مقارنة باللأعمال السابقة ذي الصلة، يظهر النظام المقترح أداء تصنيف واعد لدرجات اعتلال الشبكية السكري (DR).