2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Improved Automatic Grading of Diabetic Retinopathy Using Deep Learning and Principal Component Analysis

Eman Mohamed[1], Mai Abd Elmohsen[2] and Tamer Basha[3]

Abstract— **Diabetic retinopathy (DR) is one of the most common chronic diseases around the world. Early screening and diagnosis of DR patients through retinal fundus is always preferred. However, image screening and diagnosis is a highly time-consuming task for clinicians. So, there is a high need for automatic diagnosis. The objective of our study is to develop and validate a new automated deep learning-based approach for diabetic retinopathy multi-class detection and classification. In this study we evaluate the contribution of the DR features in each color channel then we pick the most significant channels and calculate their principal components (PCA) which are then fed to the deep learning model, and the grading decision is decided based on a majority voting scheme applied to the out of the deep learning model. The developed models were trained on a publicly available dataset with around 80K color fundus images and were tested on our local dataset with around 100 images. Our results show a significant improvement in DR multi-class classification with 85% accuracy, 89% sensitivity, and 96% specificity.**

I. INTRODUCTION

Diabetic retinopathy (DR) is one of the most dangerous health complications associated with diabetes [1], causing vision loss and irreversible blindness [2]. DR severity can be divided into five grades or stages: normal, mild, moderate, severe, and proliferative based on the different deformation in the retinal parts of the eye [3], [4]. In clinical routine, ophthalmologists usually determine the stage based on inspecting the damage occurring in retinal blood vessels leading to hemorrhage, microaneurysms, and exudates [5]–[8]. However, extracting these features is usually subjective and time-consuming [9], [10]. Hence, the need for automating this task to improve the efficacy of early detection and treatment [3], [4].

Several machine learning approaches were introduced for DR grading. Adarsh et al. used a vector of retinal features extracted to improve the automated DR detection using SVM achieved an accuracy of 95% on both DIARETDB0 and DIARETDB1 datasets [11]. As most of the generated retinal fundus images have low contrast and non-uniform illumination, so many studies worked to enhance the image quality by applying different preprocessing techniques to improve the retinal features. Contrast local adaptive histogram equalization (CLAHE) is frequently used for vascular architecture enhancement [12]. Another technique for blood vessel enhancement was also used to generate the best contrast [13], [14]. Blind source separation approaches such as principal component analysis (PCA) and independent component analysis (ICA), also used for color pigments, architecture feature extraction, and segmentation [15], [16].

Recently, deep learning algorithms, especially convolutional neural networks (CNN) had a great contribution in automating the DR classification [17], [18]. Gulshan et al. used neural networks to classify retinal fundus images into three classes (normal, moderate, and severe) [2]. Sahlsten et al. Proposed a deep learning model for the classification of both DR and macular edema using Inception-v3 architecture producing a comparable result to the Gulshan group [19]. Furthermore, Pratt et al. proposed a convolutional neural network (CNN) to predict the five classes of DR achieving promising results of 75% accuracy on the public Kaggle data set of 80,000 images [10] [20].

While most of the previous studies used the traditional RGB format of the colored fundus images, other studies have reported that RGB is not the best format for the feature extraction process as it does not provide accurate representation, especially in the vascular architectures and the optic disc [21]. Recently, some studies demonstrated that using the green channel only achieves better performance as a result of enhancing the image contrast [22], [23]. Rajesh et al. proposed a two-stage CNN model, the first one to investigate the presence of DR, and a second model used for DR grading. They used the green component after and CLAHE technique as a preprocessing step for the input data, achieving an average accuracy of 90% [24]. Gadekallu et al. used deep neural networks combined with PCA and firefly techniques to enhance DR grading accuracy, sensitivity, and specificity, but they did not have the same level of improvement in a less dimensional dataset [25].

In this work, we investigate the performance of different color spaces and statistical components on the CNN performance to provide a multi-level classification of retinal fundus images. First, we study the CNN classification performance when fed with the different channels of several color spaces of the images. Then, we pick the color channels that achieve the best classification results, use PCA to extract the highest contributing components of these channels, and use these components as the input to the CNN.

Eman Mohamed[1]. Faculty of Engineering, Systems and Biomedical Engineering, Cairo University, Giza, Egypt, emanibrahim@eng1.cu.edu.eg.
Mai Abd Elmohsen[2], Faculty of Medicine, Ophthalmology Department, Cairo University, Giza, Egypt, Mai_nasser_ophth@yahoo.com.
Tamer Basha[3], Faculty of Engineering, Systems and Biomedical Engineering, Cairo University, Giza, Egypt, tamer.basha@eng1.cu.edu.eg

II. METHODOLOGY

A. Dataset

We used a public DR dataset from the Kaggle competition [20] in the training and validation processes. It contains the five levels of DR. Additional dataset obtained from our university hospital was used for independent testing of the proposed technique. The images were obtained through a research protocol that was approved by the Research Ethics Committee, Faculty of Medicine, Cairo University, which approved retrospective analysis of pre-existing datasets utilized for testing the proposed technique. The dataset contains around one hundred color images divided into the same five classes in the Kaggle dataset. Table 1 shows the number of images in each dataset along with their distribution across the five DR classes.

B. Data preparation and Preprocessing

The investigation of the proposed method efficiency in DR classification includes further steps. First, the color channels of RGB (Red, Green, and Blue), HSV (Hue, Saturation and Value), LAB (Lightness, Red/Green and Blue/Yellow values), YIQ (Luma, and Chrominance) were extracted. Second, image CLAHE was performed on all extract images to improve the contrast and equalize the ranges of pixels' values [26]. Finally, the images of each color channel were fed to the customized neural network proposed in [10]. It consisted of ten convolutional layers starting with a convolutional layer with a rectified linear unit (RELU) activation function and 0.001 learning rate, then each convolutional layer followed by a max-pooling layer with kernel size 3×3 and strides of 2×2. The next two layers were flattened (dense layers) with a dropout of 0.5 to overcome the overfitting problem in our data set. Then, another dense layer with a softmax activation function is used to perform the classification process. The output from the final layer was five classes. During the training process, we used the L2 regularizer to prevent overfitting. The employed loss function was the mean square error (MSE) and the gradient descent optimizer was used to optimize the training process. Fig1 shows example images for the different color channels where different features appear in different strengths in the different channels.

For each DR class, color channels are sorted according to their obtained classification accuracies. PCA was then applied on two different combinations of the color channels: first, the natural color channels systems (i.e., Red and Green and Blue, Hue and Saturation and Value, Lightness, Red/Green, and Blue/Yellow values and Y, I, and Q). Second, the highest three components in terms of accuracy were obtained for each DR class. The output PCA components in each combination were then fed to the same neural network architecture. Finally, the classification decision of each channel in each group was then fused using a simple majority vote [27] to obtain the final classification class of each image.

The proposed approach was implemented using TensorFlow-Keras Framework on an Intel Xeon 2.9. GHz with NVidia GeForce RTX 2070 GPU and 64 GB of RAM. Our algorithm was evaluated in terms of accuracy, sensitivity, and specificity to compare with other studies' results in DR classification.

TABLE I. Description of the public and locally acquired datasets.

	Public Dataset	Local Dataset
Total	80,000	88
Class 1	70684	13
Class 2	2443	11
Class 3	5292	45
Class 4	873	13
Class 5	708	6

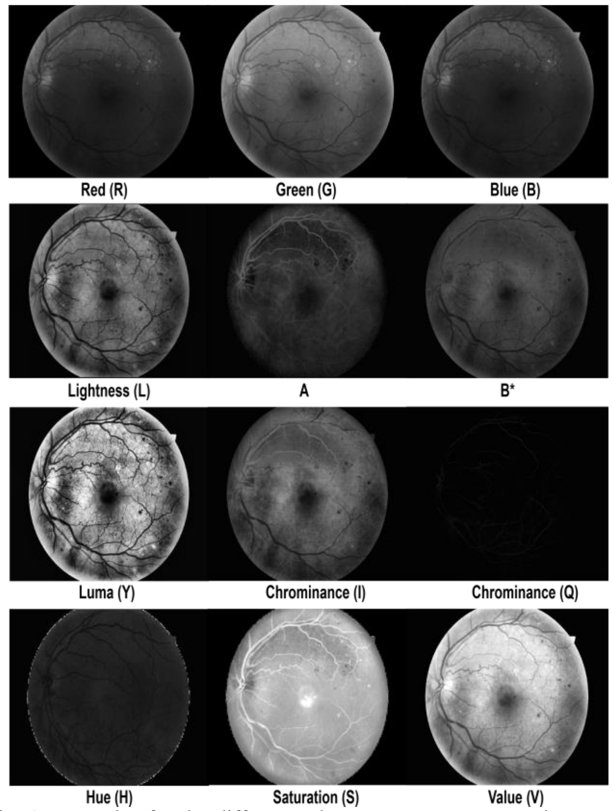

Fig. 1. Examples for the different color spaces components where some features are prominent in some channels more than other channels

III. RESULTS

Fig 2 shows the classification accuracy for each DR class when using the different channels of the images separately. The channels combinations that achieved the best accuracy in each class were L, Y, and G for class 1, L, Q, and Y for class 2, L, Y, and G for class 3, Q, an L, and Y for class 4, L, I, and G for class 5. Fig 2.f shows the average accuracy over all classes using the different color channels, where the L, Y, and Q channels achieved the highest average accuracy over all classes. The majority voting on the top 3 channels in each class obtained an accuracy of 89% for class 1, 95.7% for class 2, 90% for class 3, 100% for class 4, 99% for class 5, and 91.9% for the average overall classes.

TABLE II. Comparing Classification performance across the different DR classes when using the Green channel (used in [23], [24]), the Y channel (used in [30]) and the majority voting on the Y, S, and L channels proposed in this work. Accuracy (Acc), Sensitivity (Sens) and Specificity (Spec).

	Green Channel (%)			Y Channel (%)			YSL Voting (%)		
	Sens	Spec	Acc	Sens	Spec	Acc	Sens	Spec	Acc
Class 1	72	84.2	87.5	74.6	85.9	87.5	74.4	92.4	87.7
Class 2	85.4	96.5	92.8	90.9	90.3	94.3	86.5	97.4	95.7
Class 3	85.4	88.7	90.3	81.1	90	90.9	80	96.9	90
Class 4	98.3	99.2	99.2	99.8	99.8	99.3	96.2	99.6	100
Class 5	99.7	99.9	99.6	100	99	99.5	97.4	99.9	99.9
Total	88.1	93.7	84.7	89.3	93	85.9	86.9	97.3	91.9

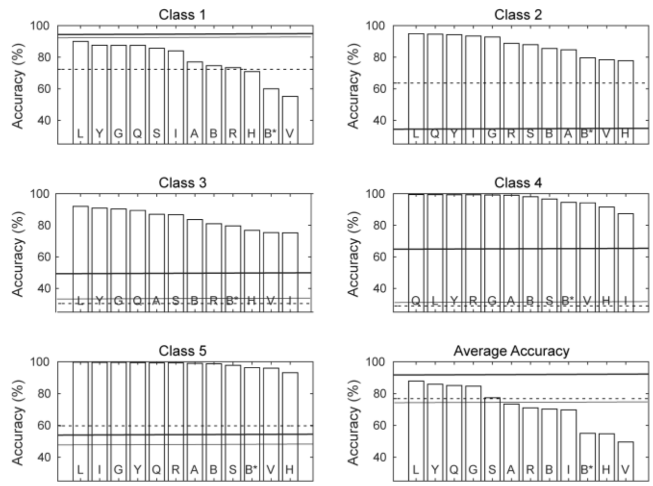

Fig. 2. Classification Accuracy for DR classes using different color components (regular line for Pratt model [10], bold one for Raju [28], and dashed one for Dutta [29]).

Table 2 shows the detailed accuracy, sensitivity, and specificity across all DR classes when using the Green channel (used in [23], [24]), the Y channel (used in [30]), and the majority voting on the Y, S, and L channels proposed in this work. Fig 3 shows exemplary images for PCA components for the typical RGB channels as well as for the selected channel combinations YIQ and HSV. Table 3 shows the detailed accuracy, sensitivity, and specificity across all DR classes when using the PCA components of the YSL and HSV channels. The PCA components of the YIQ appear to obtain the highest accuracy for class 1 while the PCA components of the HSV achieve the highest performance for the rest of the DR classes

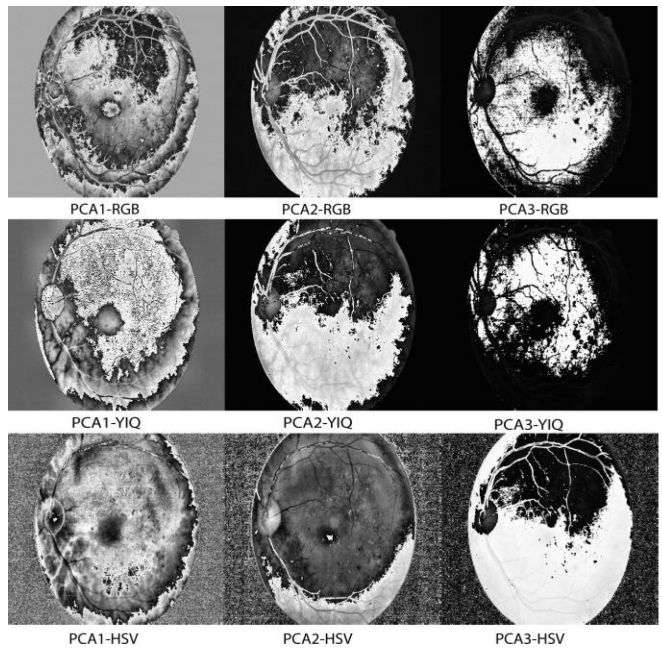

Fig. 3. Example images for the PCA components of the RGB, YIQ, and HSV color channels. Each component illustrates the image features in a different way giving different weights for the features that distinguish the associated DR class

TABLE III. Classification Results for Different DR Classes when using the PCA components of the YSL and HSV channels.

	PCA of YSL (%)			PCA of HSV (%)		
	Sens	Spec	Acc	Sens	Spec	Acc
Class 1	78.9	57	90.5	73.4	90.3	82.1
Class 2	52.1	96.4	60.6	85	95.3	93.6
Class 3	46.4	94	49	78	93.4	87.7
Class 4	64.7	99.3	93.9	98.9	99.1	99.9
Class 5	64	99.7	93.6	99.4	99.3	100
Total	61.3	89.3	72.1	86.9	95.5	90.2

IV. DISCUSSION

In this work, we first evaluated the contribution of each color channel and each principal component in the classification process. Then, the best channels and components, which likely contain more of the indicative features of the different DR classes are picked to be fed to CNN. Our results show a significant improvement in the classification and grading of the DR images. To the best of our knowledge, the proposed approach obtained higher performance metrics than previous studies [10], [17], [29].

The results showed that the combination between the Y channel from YIQ space, saturation component from HSV space and Lightness component from LAB space provides the best combination for class 1 classification. However, the PCA components of the YIQ space are more representative of the rest of the DR classes (class 2-class 5). This comes in agreement with the simple nature of class 1 where the DR is still at an early stage and thus, is likely to be picked directly from the different color channels of the image. On the other hand, the class 2-class 5 classes have developed more complications in the associated features and thus can be easily confused in the different color channels while being easier to be decoupled in the PCA components.

V. Conclusion

A novel multi-level classification approach that combines classical CNN with PCA analysis for DR detection and grading was proposed. Feeding the CNN model with the color channel and/or image principal component that best describes the DR classes helps to achieve better classification for the different DR classes. While DR class 1 is best represented by the YSL channels combination, the rest of the DR classes (class 2-class 5) are best represented in the different PCA components of the YIQ color domain.

Acknowledgment

This work was supported by the Information Technology Industry Development Agency (ITIDA), Cairo, Egypt, under the ITAC program, CFP #146.

References

[1] J. E. Shaw, R. A. Sicree, and P. Z. Zimmet, "and Clinical Practice Global estimates of the prevalence of diabetes for 2010 and 2030," vol. 87, pp. 4–14, 2010.

[2] V. Gulshan et al., "Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs," vol. 94043, pp. 1–9, 2016.

[3] Y. Yang, T. Li, W. Li, H. Wu, and W. Fan, "Lesion Detection and Grading of Diabetic Retinopathy via Two-Stages Deep Convolutional Neural Networks," vol. 2, pp. 533–540, 2017.

[4] M. D. Abr et al., "Improved Automated Detection of Diabetic Retinopa- thy on a Publicly Available Dataset Through Integration of Deep Learning."

[5] Ministry of Health Malaysia Diabetic Retinopathy Screening Team (2012) Diabetes mellitus and complications—module 3-2012. Ministry of Health Malaysia, Putrajaya.

[6] S. Suhaila, R. Vasile, J. Shuttleworth, and C. Jayne, "Automatic screen- ing and classification of diabetic retinopathy and maculopathy using fuzzy image processing," in Brain Informatics, 2016.

[7] D. Welfer, "AUTOMATIC DETECTION OF MICROANEURYSMS AND HEMORRHAGES IN COLOR EYE FUNDUS," vol. 5, no. 5, 2013.

[8] A. M. Informatics, "Diagnosis System for Diabetic Retinopathy to Prevent Vision," vol. 33, no. 3, pp. 1–11, 2013.

[9] M. N. Ozieh and K. G. Bishu, "Trends in Healthcare Expenditure in United States Adults With Diabetes: 2002 – 2011," no. June, pp. 1–8, 2015.

[10] H. Pratt, F. Coenen, D. M. Broadbent, S. P. Harding, and Y. Zheng, "Convolutional Neural Networks for Diabetic Retinopathy," Procedia - Procedia Comput. Sci., vol. 90, no. July, pp. 200–205, 2016.

[11] P. Adarsh and D. Jeyakumari, "Multiclass SVM-Based Automated Diagnosis of Diabetic Retinopathy," 2013, pp. 206–210.

[12] T. Shimahara, T. Okatani and K. Deguchi," Contrast enhancement of fundus images using regional histograms for medical diagnosis," pp. 650-653, 2004.

[13] D. S. S. Raja and S. Vasuki, "Automatic Detection of Blood Vessels in Retinal Images for Diabetic Retinopathy Diagnosis," vol. 2015, pp. 8–12, 2015.

[14] A. F. M. Hani and H. A. Nugroho, "Retinal vasculature enhancement using independent component analysis," vol. 2009, no. November, pp. 543–549, 2009.

[15] A. A. Mudassar and S. Butt, "Application of Principal Component Analysis in Automatic Localization of Optic Disc and Fovea in Retinal Images," vol. 2013, 2013.

[16] M. H. A. Fadzil, H. A. Nugroho, P. A. Venkatachalam, H. Nugroho, and L. I. Izhar, "Determination of Retinal Pigments froc1m Fundus Images using Independent Component Analysis," vol. 21, pp. 555–558, 2008.

[17] S. Qummar et al., "A Deep Learning Ensemble Approach for Diabetic Retinopathy Detection," in IEEE Access, vol. 7, pp. 150530-150539, 2019.

[18] S. Suhaila, R. Vasile, J. Shuttleworth, and C. Jayne, "Automatic screening and classification of diabetic retinopathy and maculopathy using fuzzy image processing," in Brain Informatics, 2016.

[19] Sahlsten, J., Jaskari, J., Kivinen, J. et al. Deep Learning Fundus Image Analysis for Diabetic Retinopathy and Macular Edema Grading. Sci Rep 9, 10750 (2019).

[20] https://www.kaggle.com/c/diabetic-retinopathy-detection/data.

[21] J. Diaz-Cely, C. Arce-Lopera, J. C. Mena and L. Quintero, "The Effect of Color Channel Representations on the Transferability of Convolu- tional Neural Networks," in Science and information Conference, Cham, 2020.

[22] Y. Jiang, H. Zhang, and N. Tan, "SS symmetry Automatic Retinal Blood Vessel Segmentation Based on Fully Convolutional Neural Net- works," 2019.

[23] Gultepe, E., Makrehchi, M. Improving clustering performance using independent component analysis and unsupervised feature learning. Hum. Cent. Comput. Inf. Sci. 8, 25 (2018).

[24] Rajesh N.V, Durga Devi G.Y., A. Mari Kirthima and Ambika G N, Diabetic Retinopathy Detection using Deep Learning, International Journal of Advanced Research in Engineering and Technology, 11(7), pp. 208-216, 2020.

[25] Gadekallu TR, Khare N, Bhattacharya S, Singh S, Maddikunta PKR, Ra I-H, Alazab M. Early Detection of Diabetic Retinopathy Using PCA-Firefly Based Deep Learning Model.Electronics. 2020; 9(2):274

[26] A. W. Setiawan, T. R. Mengko, O. S. Santoso, and A. B. Suksmono,"Color Retinal Image Enhancement using CLAHE," no. March 2015, pp. 1–4, 2013.

[27] A. Sinha, H. Chen, D. G. Danu, T. Kirubarajan, and M. Farooq, "Neurocomputing Estimation and decision fusion: A survey," vol. 71, pp. 2650–2656, 2008.

[28] M. Raju, V. Pagidimarri, R. Barreto, et al.Development of a deep learning algorithm for automatic diagnosis of diabetic retinopathy Stud Health Technol Inform, 245 (2017), pp. 559-563.

[29] Dutta S, Manideep BC, Basha SM, Caytiles RD, NCSN. Iyen- gar.Classification of diabetic retinopathy images by using deep learning models. Int J Grid Distrib Comput. 2018;11:89–106.

[30] Shu-I Pao, Hong-Zin Lin, Ke-Hung Chien, Ming-Cheng Tai, Jiann-Torng Chen, Gen-Min Lin, "Detection of Diabetic Retinopathy Using Bichannel Convolutional Neural Network", Journal of Ophthalmology, vol. 2020, Article ID 9139713, 7 pages, 2020