Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X
IMPACT FACTOR: 7.056

IJCSMC, Vol. 12, Issue. 5, May 2023, pg.54 – 63

Fundus Images Classification of Diabetic Retinopathy using MobileNetV2

Nisreen I. R. Yassin
National Research Centre, Cairo, Egypt
ni.yassin@nrc.sci.eg

DOI: https://doi.org/10.47760/ijcsmc.2023.v12i05.006

Abstract— Diabetic retinopathy (DR) is a chronic eye disease and the main reason of blindness between adults. Manual DR screening methods require skilled readers, effort, and time. Automated DR detection and classification of its severity is important for early and effective disease control. Mobile devices can facilitate regular screening of retinal fundus images using automated lightweight architectures techniques. While, many automated techniques have been proposed in the literature, new models are needed to improve the accuracy and to facilitate the accessibility to such models. In this study, MobileNetV2 is used to classify five classes of DR severity that includes no DR, mild DR, moderate DR, severe DR, and proliferative DR. Colour fundus images are enhanced using contrast-limited adaptive histogram equalization (CLAHE) to defeat bad lighting conditions. MobileNetV2 is used to extract image features. A new classification head based on batch normalization and L2 regularization is applied on the extracted features. The model is tested using augmented data, non-augmented data, fine-tuning, and without fine-tuning. The experiments indicate the importance of augmentation and fine-tuning for improving the evaluation metrics. The experimental results show the superiority of the proposed technique compared to recent techniques in literature with training, validation, and test accuracies of 98.65, 87.94%, and 87.67% respectively.

Keywords— Diabetic retinopathy, Fundus images, Deep learning, MobileNetV2, fine-tuning.

I. INTRODUCTION

Globally, there are about 537 million people suffering from diabetes and the number is estimated to raise to 643 million in 2030 and 783 million in 2045 [1]. Diabetes is a chronic disease occurs when the body becomes resistant to insulin or does not make sufficient insulin, which leads to elevated scales of blood sugar. This long-term condition has major effects on the body systems. One of these effects is the deterioration of the back of the eye (retina) due to diabetic retinopathy (DR) [2]. DR is the most popular complication of diabetes mellitus and estimated as the source of 51% of blindness and 56% of visual impairment cases globally [3]. The retina is the light-sensitive coat of cells at the back of the eye that transforms light into electrical signals. The signals are sent to the brain, which turns them into seen images. A network of tiny blood vessels provides the retina with its needs of blood. High blood sugar level damages these blood vessels cause serious bleeding in the eye, which can lead to blindness. DR occurrence has increased worldwide, so it is one of the most defy issues meeting ophthalmological research. In early stages, DR is not usually noticed and does not have apparent symptoms; however, it can be detected by taking photographs of the eyes during diabetic eye screening. Early detection of DR is important to prevent possible vision loss by effective management and timely treatment of the disease [4, 5].

DR is classified using fundus images into five classes according to its severity which are No DR, Mild, Moderate, Severe and Proliferative DR. DR screening process is the searching process for significant hidden features of DR, which is a time-consuming duty for ophthalmologists. It needs skilled readers, effort, and dense time. For patients, regular DR screening costs high expenses in developed countries. Further, there is always conflict between readers when DR screened manually, which results in more spreading of the disease [6]. Automatic diagnosis systems of DR are able to support ophthalmologists, conserve time and costs, and attain good accuracy through feature extraction and classification of fundus images [7, 8]. An overwhelming number of machine learning (ML) techniques have been proposed for DR detection and classification such as Neural Network (NN) and Support Vector Machine (SVM). These classic ML techniques are inefficient and inflexible in case of large high dimensional data [9]. Recently, deep learning (DL) is vastly used in medical image analysis [10, 11]. It has accomplished impressive results in the field of computer vision. DL is a machine-learning category uses convolutional neural network (CNN) architecture for processing huge amount of data. DL becomes more preferable than other machine learning techniques because of its ability to learn and extract features without any human intervention. Furthermore, DL does not need any kind of segmentation. Deep CNN consists of convolution, pooling, and fully connected layers [12, 13]. DL needs huge amount of data for training and testing process, therefore transfer learning (TL) is an applicable solution in case of small amount of data. Lately, many studies have been proposed to detect and classify DR using Deep learning; however, there is a need for more accurate solutions especially for multi-classification problem of DR. This research aims to improve the performance of automatic classification of DR into five classes. Colour fundus images are first enhanced using contrast-limited adaptive histogram equalization (CLAHE). MobileNetV2 model is applied efficiently on the enhanced fundus images through fine-tuning, and then a new classifier is adopted to classify images into the five classes: No DR, mild DR, moderate DR, severe DR, and proliferative DR. The model is tested using augmented data, non-augmented data, fine-tuning, and without fine-tuning. The rest of the paper is organized as follows; Section 2 presents the related work, Section 3 introduces the proposed methodology; Section 4 discusses the experimental results, and Section 5 provides the conclusions.

II. RELATED WORK

Many studies have been conducted regarding DR detection and classification. Machine learning techniques (MLTs) have been extensively used to extract features from fundus images either hand crafted features using conventional MLTs, or automatically extracted features using deep learning. Here, we will concentrate on the studies that based on deep learning techniques. Thota and Reddy [14] used the pre-trained network VGG-16 to classify the DR severities based on a public dataset from Kaggle platform. The achieved classification accuracy, sensitivity, and specificity were 74%, 80%, and 65% respectively. Pradhan et al. [15] used three models to classify DR into five severities VGG-16, ResNet50, and Inception V3. The highest accuracy was 78% gained by VGG-16. El Houby [16] classified different groups of severity using VGG-16. The achieved accuracy for 5-class classification was 73.7. Gadekallu et al. [17] classified the extracted features into two classes affected or not affected with DR using deep neural network. The achieved accuracy was 97.3%. Bodapati et al. [18] used 3662 images from Kaggle to classify DR level based on transfer learning. The achieved accuracy was 84.31%. Lam et al. [7] applied GoogLeNet and AlexNet to classify DR stages based on Kaggle and Messidor datasets. Three classification models were tested 2-ary, 3-ary, and 4-ary with achieved accuracies of 74.5%, 68.8% and 57.2% respectively. Mishra et al. [19] proposed DR 5-stages classification using 3662 images based on DenseNet and vgg-16 with achieved accuracies of 0.9611 and 0.7326 respectively. Pratt et al. [20] proposed a CNN approach to diagnose DR from fundus images. The achieved accuracy using 5000 validation images was 75%. Chowdhury and Meem [21] applied Inception v3 classifier to classify DR severities using a dataset collected from Kaggle. They classified DR into 2, 3, and 5 classes with achieved accuracies of 61.3%, 60.3% and 37.7% respectively. Sheikh and Qidwai [22] proposed a DR classification system based on MobileNetV2 architecture. They used a custom dataset to classify DR severities into five classes, where good quality images were selected for training process. The achieved accuracy was 91.68%. Pamadi et al. [23] proposed two CNN models, the first to detect retinopathy through binary classification and the second to classify retinopathy into five classes. They enhanced the accuracy by using transfer-learning technique based on MobileNetV2 model. The achieved accuracy was 97% for binomial classification and 78% for multinomial classification. Gebremariam [24] proposed a model for early detection of DR. The model was applied on two DL architectures MobileNetV2 and EfficientNetB0. According to five stages classification, the accuracy achieved using MobileNetV2 was 60% while 71% using EfficientNetB0. Menaouer et al. [25] proposed a hybrid deep learning method based on three models (CNN, VGG16, and VGG19) to detect and classify DR in fundus images. They achieved an accuracy of 90.60% using APTOS 2019 dataset.

III. MATERIALS AND METHOD

In this section, the dataset used for the training and testing are described along with the transfer-learning model used in this study. The dataset is further described in Section A. The proposed methodology for the classification of DR severities is described in Section B.

A. *Dataset*

Differentiating between healthy retina and a retina affected by DR is difficult and sophisticated problem. The dataset used in this study is obtained from Kaggle [26]. The EyePACS dataset contains 35,126 coloured fundus images with highly diverse levels of illumination used to train and test the proposed model. The dataset is composed of 5-classes representing the level of DR severity (No_DR, mild, moderate, severe, and proliferate_DR). These images are PNG formatted and in size of 224x224. Fundus images are complicated low quality images, they have been taken in different lighting conditions with various kinds of cameras. Fig. 1 shows samples of images from the EyePACS dataset.

IV. PROPOSED METHODOLOGY

In this research, a multiclass categorization method for colour fundus images is proposed. The main objective is to identify retina images infected by DR and the stage of infection efficiently with better accuracy. Classification of DR into five classes automatically is a challenging task as there are inter-class similarities. Two problems encountered with multiclass categorization of DR, the availability of data and high-class imbalance found in publically available datasets. These problems have been solved using different augmentation techniques. Details of the proposed method are explained in the next subsections.

(a) No_DR 22_right.png | (b) mild 15_left.png | (c) moderate 30_right.png | (d) severe 932_left.png | (f) proliferate_DR 6997_right.png

Fig. 1 samples of images from the EyePACS dataset.

1) Data Pre-processing and augmentation: Efficient training of deep learning models requires high quality datasets [27]. Fundus images used in this study contain uninformative black space areas surround the retina region that considered unwanted regions and should be removed. Automatic cropping is applied to all images in the dataset using a bounding box surround the retina region. Fundus images appears dark, it suffers from low contrast and quality, and therefore contrast-limited adaptive histogram equalization (CLAHE) is used to enhance the visual quality of the fundus images. RGB fundus images are converted to HSV format, then only the brightness channel (V) is subjected to CLAHE to preserve colour channels. Images normalization is important step to attain consistency for a set of data. In the proposed study, the range of pixel intensity values for each image in the dataset is mapped from 0, 255 to 0 and 1 by multiplying each pixel by a factor of 1/255.

Efficient training of deep learning models requires a balanced dataset such that each class in the dataset contains approximately the same number of images. In other words, an imbalanced dataset is a dataset with unequal class distribution [28]. Data imbalance affects the performance of deep learning models through biasing to majority class (negative class) which results in high accuracy but poor performance [29]. The dataset used in this research is highly imbalanced as shown in Fig. 2. Imbalanced ratio (IR) is a way to indicate the degree of data imbalance where it is the ratio between majority classes (negative class) to minority class (positive class) and defined as in eq. (1) [28].

$$IR = no.of\ negative\ class\ instances/\ no.of\ positive\ class\ instances \qquad (1)$$

We can observe that class "0" contains the highest number of instances 25,810 that is much more compared to class "4" which contains 708 instances as depicted in table 1. The IR value is 36.45 that indicates severe imbalance. Two steps were used to solve this problem. The first step is to increase the number of instances in minority classes, which are class "3" and class "4". Data augmentation is a technique used to increase the amount of data and to achieve the required balance between classes. Offline data augmentation is applied through rotating every image in classes "3" and "4" by three angles 90, 180, and 270. In addition, horizontal and vertical flipping are applied to instances in these classes. Observing the data after this step, convergence between the classes 1, 2, 3, and 4 is achieved but class imbalance is still encountered due to large number of images in class "0" compared to the remainder classes. The second step is decreasing the number of instances in the majority class "0" by applying Random Under-sampling method [28]. This method depends on randomly removing of majority class instances until a convergence is achieved. After balancing the data, the total number of instances is decreased from 35,126 to 23,080. Automatic splitting is applied on the balanced data to divide it into training data and test data by 70% to 30%.

Considerable amount of training data is also a requirement for effective training. Therefore, we used online augmentation to increase the number of instances of the training dataset. Geometric and flipping augmentation are applied to training data just before feeding to the CNN model. Rotation within 20 degrees and horizontal flipping are found suitable. Fig. 3 shows image samples after applying data online augmentation. Online augmentation does not include the original data in the training process but includes only the augmented images so the final total number of data increased to 39,848. Table 1 depicts the structure of the dataset and the number of instances before and after balancing and augmentation processes.

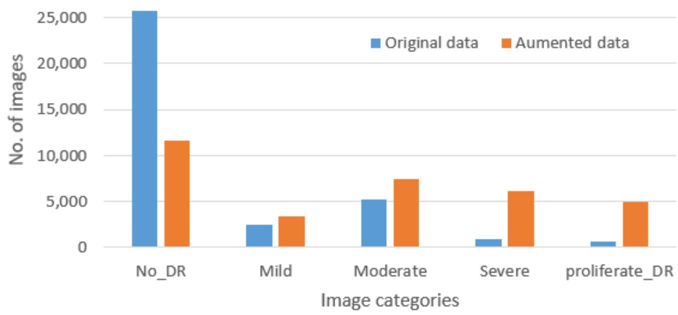


Fig. 2 original data versus balanced data.

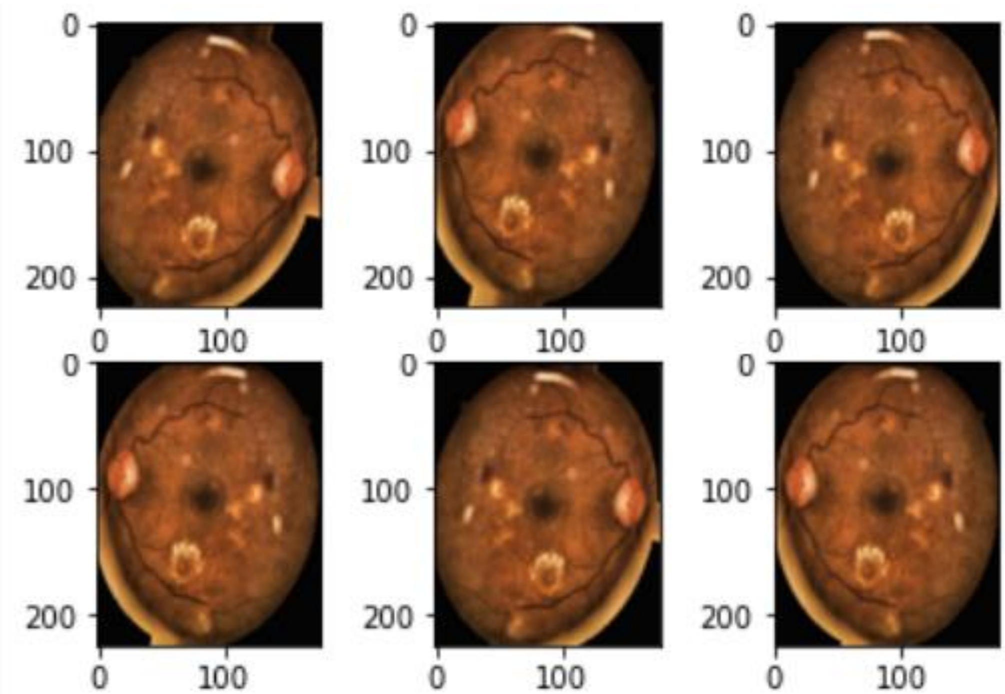


Fig. 3 image samples after applying pre-processing and data online augmentation.

TABLE 1
DATASET STRUCTURE BEFORE AND AFTER BALANCING AND AUGMENTATION

Classes no.	Classes labels	No of images	No. of balanced images	Training data	Test data	Training data after online augmentation
0	No_DR	25,810	7440	5820	1620	11640
1	Mild	2,443	2443	1710	733	3420
2	Moderate	5,292	5292	3704	1588	7408
3	Severe	873	4365	3056	1309	6112

4	proliferate_DR	708	3540	2478	1062	4956
Total		35,126	23,080	16768	6312	33536
					39848	

2) *Feature extraction and classification of colour fundus images:* Convolution neural network extracts features from images automatically depending on the training process. Deep learning models consist of many convolutional layers. Initial layers extract edges and contours; however, deeper layers extract detailed features. The extracted features are used by fully connected layers to classify the input [30]. In the proposed study, MobileNetV2 is a pre-trained deep learning model used for feature extraction and classification of fundus images. MobileNetV2 is a convolutional neural network designed for mobile devices and embedded systems. It is a lightweight model based on pointwise and depthwise convolutions layers. The main feature of mobileNetV2 is the significant reduction in the number of its parameters and operations that reduces the size of the network and optimizes the speed [31, 32].

Deep learning models consist of consecutive layers in which various features are learned at various layers. The general features are learned in initial layers then become fine-grained features in deeper layers. Based on the extracted features, the classification process takes place to get the result using the last layer of the model that is the fully connected layer. In TL, the pre-trained network is used as feature extractor without its final layer, which is replaced by a shallow classifier. Pre-trained models are trained from scratch using a huge and renowned dataset, which requires extended computational time. During this training, millions of weights are learned and saved. Transfer learning to support small datasets to gain superior performance [33] uses the learned knowledge from this training. Therefore, when applying TL for a new task, the learned weights are used without updating during the training of new small dataset. In transfer learning with fine-tuning, the model conserves the weights of a number of specified layers while updates the weights of remaining layers. MobileNetV2 have been trained on the ImageNet dataset that contains one million of images and has one thousand categories. The concept of TL is illustrated in fig. 4.

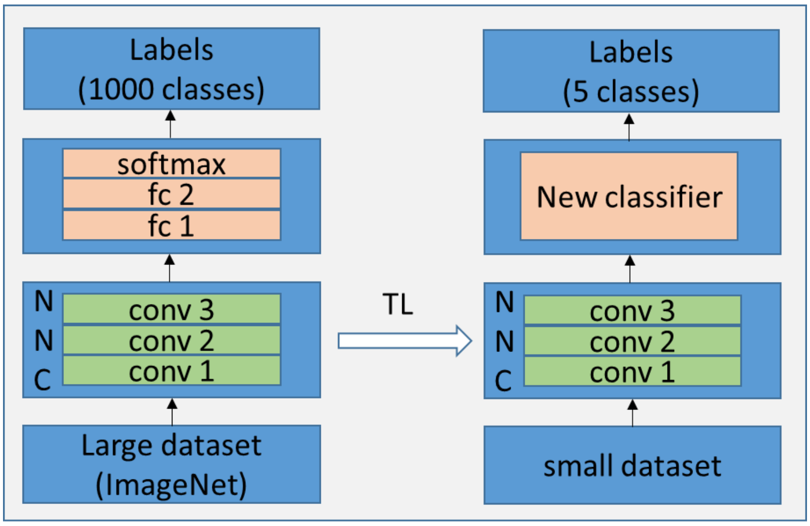


Fig. 4 Concept of Transfer learning.

MobileNetV2 is retrained on small dataset of fundus images to classify DR stages. To classify the data, batch normalization and L2 regularization techniques are used in the new classifier. The extracted features are first normalized using a batch normalization layer. The used classifier consists of a fully connected layer of size 256 with L2 regularization, a dropout layer of 0.5, and a fully connected layer with softmax activation function of five outputs. The proposed model is shown in Fig. 5. The batch normalization layer is used to speed up the training at higher learning rates. Mini batches of data are normalized between the CNN layers by forcing them to have a zero mean and a standard deviation of one. Therefore, each data point simulates the standard normal distribution. Batch normalization can be expressed mathematically as follows [34]:

$$z^N = \frac{z - m_z}{s_z} \quad (2)$$

Where z^N is the normalized data point, z is the output data point of the neuron, m_z is the mean of output data points of the neurons, and s_z is the standard deviation of output data points of the neurons.

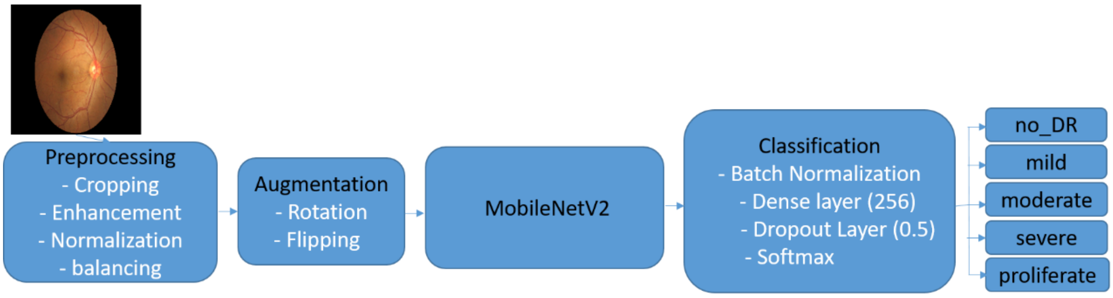


Fig. 5 proposed model for fundus images classification.

V. EXPERIMENTAL RESULTS

Several experiments were performed to evaluate the proposed model. Experiments of this study were applied on Intel®Core™i7 CPU at 2.60 GHz computer using TensorFlow platform for machine learning and Matlab®2017 software for some image processing operations. Fundus image dataset collected from Kaggle was used in the simulation process. MobileNetV2 is a pre-trained network used to classify fundus images into five classes of DR severity.

The proposed model is based on using the pre-trained network as feature extractor that extracts the features of the images according to learned weights gained from training the network using large datasets. Some image pre-processing operations were performed to improve the final classification process. The dark region surrounding the region of interest in the fundus images was cropped. Image enhancement is important for increasing contrast and reducing darkness found in the used dataset. RGB fundus images were converted to HSV colour format. CLAHE was applied on the V channel while the colour channels H and S were saved. Enhanced RGB images were obtained by inverse conversion of the enhanced HSV. The used dataset was unbalanced where one class has many more samples than the other classes; therefore, data augmentation was used to solve this problem. After augmentation, the dataset was divided into 2 groups; 70% for training data and 30% for test data. During the training process, the training data was automatically augmented to get different images and to increase the number of the data samples, which increase the model generalization and robustness. The hyper-parameters used in the training of MobileNetV2 are: initial learning rate of 0.001 decreasing by half every 15 epochs, Adam optimizer, batch size of 32, epochs of 100 and patience of 10. The evaluation of the proposed model was performed using the saved model that gained the best results during the training process. The used evaluation metrics are accuracy, precision, recall, and f1-score. In addition, the confusion matrix is used to validate the proposed model. The used metrics are presented as follows:

$$Recall = T_P/(T_P + F_N)\ \% \tag{3}$$

$$Precision = T_P/(T_P + F_P)\ \% \tag{4}$$

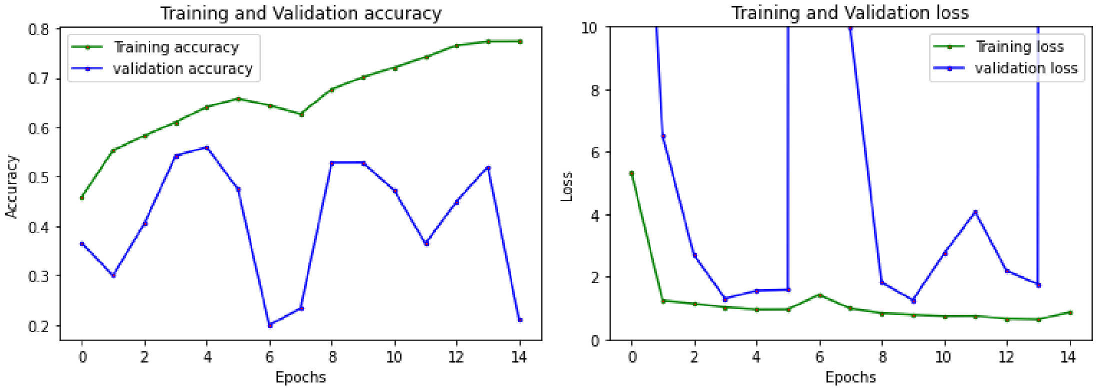
$$Accuracy = (T_P + T_N)/(T_P + F_P + T_N + F_N)\ \% \tag{5}$$

$$F1 - score = 2T_P/(2T_P + F_P + F_N) \tag{6}$$

Where, T_P is true-positive value, F_P is false-positive value, T_N is true negative value and F_N is false-negative value.

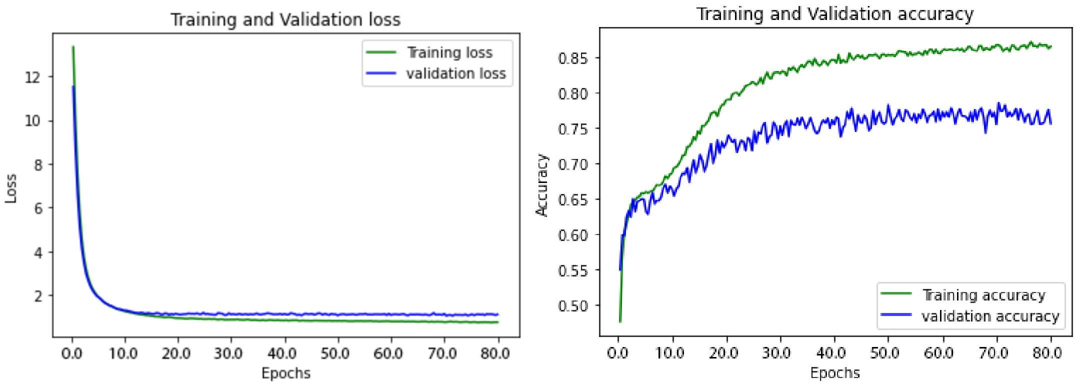
Three experiments have been conducted, using MobileNetV2 transfer learning model without fine-tuning, using MobileNetV2 with fine-tuning and data without augmentation, and MobileNetV2 with fine-tuning and augmented data. Using fine-tuning, the original learned weights of the first eighty layers are used without updating to learn abstract features. Then, the remaining layers are trained from scratch to better learning of DR features. The same hyper-parameters were used in the three experiments. First, we have tested the model without fine-tuning using the augmented data. The training process stopped at epoch 13 with very depressed train and validation accuracies of 42.45% and 39.88% respectively. The stopping occurred because of the patience condition of 10. The obtained results have been improved to 87.63% and 77.93% after removing the patience condition. Second, fine-tuning was added to the same model, tested using the same hyper-parameters for augmented data, and not augmented data. The obtained results show the efficiency of the proposed model using the augmented data, where the training and validation accuracies achieved are 98.65% and 87.94% respectively. While, the corresponding accuracies using the data without augmentation are 77.34% and 55.95%. It can be seen that, the achieved results according to augmented data are much better than that achieved using data without augmentation. In addition, Fine-tuning has extremely improved the results than using only transfer learning. Fig. 6 and fig. 7 show the loss and accuracy curves of MobileNetV2 using non-augmented data and augmented data. Fig. 8 shows the loss and accuracy curves of MobileNetV2 with fine-tuning using augmented data.

As we see, data augmentation is an essential step for achieving considerable validation accuracy. Using the augmented data, MobileNetV2 models with fine-tuning and without fine-tuning were trained and tested using the saved best weights. Table 2 shows the achieved train, validation, and test accuracy of MobileNetV2 for the classification of five classes of DR severities. As seen, test accuracies of MobileNetV2 with fine-tuning and without fine-tuning are 87.67% and 76.42% respectively. Comparing the two models, it is shown that MobileNetV2 with fine-tuning is performing better than the model without fine-tuning. Precision, recall, and F1-score for each class using MobileNetV2 with fine-tuning are shown in table 3. In addition, confusion matrix is shown in fig. 9.



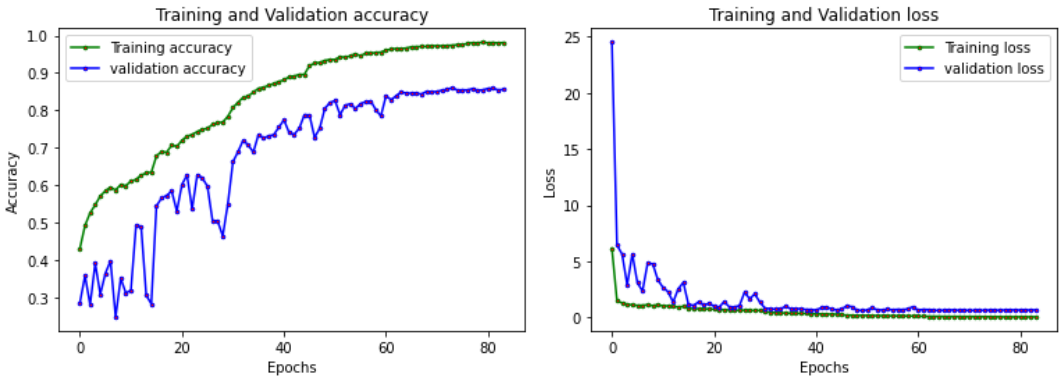
(a) (b)

Fig. 6 Training and validation plots of MobileNetV2 using non-augmented data: (a) accuracy plot, (b) loss plot.



(a) (b)

Fig. 7 Training and validation plots of MobileNetV2 without fine-tuning using augmented data: (a) loss plot, (b) accuracy plot.



(a) (b)

Fig. 8 Training and validation plots of MobileNetV2 with fine-tuning using augmented data: (a) accuracy plot, (b) loss plot.

TABLE 2
ACCURACY OF THE MOBILENETV2

MobileNetV2	Evaluation metrics		
	Train accuracy	**Validation accuracy**	**Test Accuracy**
Without fine-tuning	87.63%	77.93%	76.42%
With fine-tuning	98.65%	87.94%	87.67%

TABLE 3
EVALUATION METRICS PER CLASS

	Precision	**Recall**	**F1-score**
No DR	0.86	0.74	0.80
Mild	0.80	0.83	0.81
Moderate	0.80	0.90	0.84
Sever	1.00	0.93	0.97
Proliferate	0.93	0.98	0.95

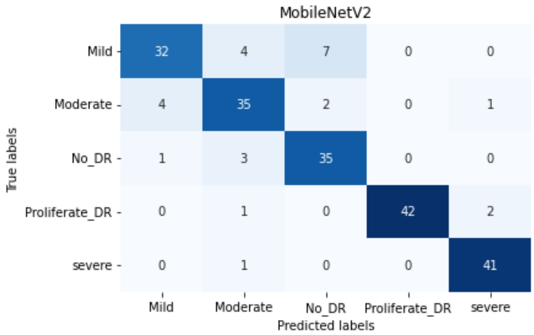


Figure 9 Confusion matrix of MobileNetV2 model with fine-tuning.

The proposed model was compared to models in literature that built on deep learning and used the same dataset. The references [14], [15], and [16] classified the 5 severities of DR using VGG-16 where the achieved accuracies are 74% 78%, and 73.7% respectively. The proposed model outperforms these models where the achieved test accuracy is 88% using mobileNetV2. According to MobileNetV2 models, it is shown that the proposed model achieved higher accuracy than the references [23] and [24] discussed in section 2. We have tested the model using different hyper-parameters to improve the performance. Some studies in the literature such as [22] has gained an accuracy value of 91.68% which is higher than the proposed model. In [22], the authors used custom data that have been selected in good quality to contribute positively in the training process. In the proposed study, we did not follow any selection criteria, which may affected the results and indicates the effect of images quality in the training process. Table 4 summarizes the models that based on MobileNetV2 with their corresponding accuracy values.

TABLE 3
COMPARING THE PROPOSED MODEL TO MODELS IN LITERATURE.

Reference	**Classes**	**Model**	**Accuracy**
[24]	5	MobileNetV2	60%
[23]	5	MobileNetV2	78%
Proposed	5	MobileNetV2	88%

VI. CONCLUSIONS

Modern technological evolution has flatten the road for deep learning to be applied extensively in the medical field for detection and classification of different diseases. In this research, an automated system for detection and classification of DR based on fundus images was proposed. Fundus images were enhanced and normalized to increase quality and uniformity of the images. To avoid class biasing, data augmentation and under-sampling are used. During the training process, on- line data augmentation is applied to increase the number of training instances. MobileNetV2 was applied on the processed images as a feature extractor. A new classifier based on batch normalization and regularization is used to differentiate between 5 DR severities. The model was tested with, without data augmentation, with, and without fine-tuning. The obtained results show the superiority of the proposed model compared to other models. The initial rating of DR using the proposed model might be helpful for ophthalmologists. In the future, the performance of proposed model can be enhanced using different optimization techniques and different hyper-parameters.

REFERENCES

[1]. I. D. A. Reports, "https://diabetesatlas.org"./

[2]. P. Saeedi, I. Petersohn, P. Salpea, B. Malanda, S. Karuranga, N. Unwin, *et al.*, "Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas," *Diabetes research and clinical practice*, vol. 157, p. 107843, 2019.

[3]. J. Chua, C. X. Y. Lim, T. Y. Wong, and C. Sabanayagam, "Diabetic retinopathy in the Asia-Pacific," *The Asia-Pacific Journal of Ophthalmology*, vol. 7, pp. 3-16, 2018.

[4]. G. Parati, G. Stergiou, E. O'Brien, R. Asmar, L. Beilin, G. Bilo, *et al.*, "European Society of Hypertension practice guidelines for ambulatory blood pressure monitoring," *Journal of hypertension*, vol. 32, pp. 1359-1366, 2014.

[5]. S. R. Flaxman, R. R. Bourne, S. Resnikoff, P. Ackland, T. Braithwaite, M. V. Cicinelli, *et al.*, "Global causes of blindness and distance vision impairment 1990–2020: a systematic review and meta-analysis," *The Lancet Global Health*, vol. 5, pp. e1221-e1234, 2017.

[6]. J. Gao, C. Leung, and C. Miao, "Diabetic retinopathy classification using an efficient convolutional neural network," in *2019 IEEE International Conference on Agents (ICA)*, 2019, pp. 80-85.

[7]. C. Lam, D. Yi, M. Guo, and T. Lindsey, "Automated detection of diabetic retinopathy using deep learning," *AMIA summits on translational science proceedings*, vol. 2018, p. 147, 2018.

[8]. Y. Sun and D. Zhang, "Diagnosis and analysis of diabetic retinopathy based on electronic health records," *Ieee Access*, vol. 7, pp. 86115-86120, 2019.

[9]. D. Das, S. K. Biswas, and S. Bandyopadhyay, "A critical review on diagnosis of diabetic retinopathy using machine learning and deep learning," *Multimedia Tools and Applications*, pp. 1-43, 2022.

[10]. E. I. ÜNLÜ and A. ÇINAR, "Segmentation of Benign and Malign lesions on skin images using U-Net," in *2021 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT)*, 2021, pp. 165-169.

[11]. F. Ullah, S. U. Ansari, M. Hanif, M. A. Ayari, M. E. H. Chowdhury, A. A. Khandakar, *et al.*, "Brain MR image enhancement for tumor segmentation using 3D U-Net," *Sensors*, vol. 21, p. 7528, 2021.

[12]. P. Kharazmi, J. Zheng, H. Lui, Z. Jane Wang, and T. K. Lee, "A computer-aided decision support system for detection and localization of cutaneous vasculature in dermoscopy images via deep feature learning," *Journal of medical systems*, vol. 42, pp. 1-11, 2018.

[13]. W. L. Alyoubi, W. M. Shalash, and M. F. Abulkhair, "Diabetic retinopathy detection through deep learning techniques: A review," *Informatics in Medicine Unlocked*, vol. 20, p. 100377, 2020.

[14]. N. B. Thota and D. U. Reddy, "Improving the accuracy of diabetic retinopathy severity classification with transfer learning," in *2020 IEEE 63rd International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2020, pp. 1003-1006.

[15]. A. Pradhan, B. K. Dey, R. K. Nath, and B. Sarma, "Transfer learning based classification of diabetic retinopathy stages," in *2020 International Conference on Computational Performance Evaluation (ComPE)*, 2020, pp. 813-819.

[16]. E. M. El Houby, "Using transfer learning for diabetic retinopathy stage classification," *Applied Computing and Informatics*, 2021.

[17]. T. R. Gadekallu, N. Khare, S. Bhattacharya, S. Singh, P. K. R. Maddikunta, and G. Srivastava, "Deep neural networks to predict diabetic retinopathy," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1-14, 2020.

[18]. J. D. Bodapati, N. S. Shaik, and V. Naralasetti, "Deep convolution feature aggregation: an application to diabetic retinopathy severity level prediction," *Signal, Image and Video Processing*, vol. 15, pp. 923-930, 2021.

[19]. S. Mishra, S. Hanchate, and Z. Saquib, "Diabetic retinopathy detection using deep learning," in *2020 International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE)*, 2020, pp. 515-520.

[20]. H. Pratt, F. Coenen, D. M. Broadbent, S. P. Harding, and Y. Zheng, "Convolutional neural networks for diabetic retinopathy," *Procedia computer science*, vol. 90, pp. 200-205, 2016.

[21]. M. M. H. Chowdhury and N. T. A. Meem, "A machine learning approach to detect diabetic retinopathy using convolutional neural network," in *Proceedings of International Joint Conference on Computational Intelligence*, 2020, pp. 255-264.

[22]. S. Sheikh and U. Qidwai, "Using MobileNetV2 to classify the severity of diabetic retinopathy," *Int. J. Simul. Syst. Sci. Technol*, pp. 1-6, 2020.

[23]. A. M. Pamadi, A. Ravishankar, P. A. Nithya, G. Jahnavi, and S. Kathavate, "Diabetic Retinopathy Detection using MobileNetV2 Architecture," in *2022 International Conference on Smart Technologies and Systems for Next Generation Computing (ICSTSN)*, 2022, pp. 1-5.

[24]. Y. Gebremariam, "Detection Of Diabetic Retinopathy Using Deep Convolutional Neural Network On Mobile Devices," ed, 2022.

[25]. B. Menaouer, Z. Dermane, N. El Houda Kebir, and N. Matta, "Diabetic Retinopathy Classification Using Hybrid Deep Learning Approach," *SN Computer Science,* vol. 3, p. 357, 2022.

[26]. https://www.kaggle.com./

[27]. Z. u. Rehman, M. A. Khan, F. Ahmed, R. Damaševičius, S. R. Naqvi, W. Nisar*, et al.*, "Recognizing apple leaf diseases using a novel parallel real- time processing framework based on MASK RCNN and transfer learning: An application for smart agriculture," *IET Image Processing,* vol. 15, pp. 2157-2168, 2021.

[28]. A. Fernández, S. García, M. Galar, R. C. Prati, B. Krawczyk, and F. Herrera, *Learning from imbalanced data sets* vol. 10: Springer, 2018.

[29]. A. Kulkarni, D. Chong, and F. A. Batarseh" ,Foundations of data imbalance and solutions for a data democracy," in *data democracy*, ed: Elsevier, 2020, pp. 83-106.

[30]. J. Wu, "Introduction to convolutional neural networks," *National Key Lab for Novel Software Technology. Nanjing University. China* ,vol. 5, p. 495, 2017.

[31]. M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 4510-4520.

[32]. A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand*, et al.*, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *arXiv preprint arXiv:1704.04861*, 2017.

[33]. M. Z. U. Rehman, F. Ahmed, M. A. Khan, U. Tariq ,S. S. Jamal, J. Ahmad*, et al.*, "Classification of citrus plant diseases using deep transfer learning," *Computers, Materials & Continua,* vol. 70, pp. 1401-1417, 2022.

[34]. S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *International conference on machine learning*, 2015, pp. 448-456.