

# Using transfer learning for diabetic retinopathy stage classification

Classifying  
diabetic  
retinopathy  
stage

Enas M.F. El Houby

*Department of Systems and Information, Engineering Division,  
National Research Centre, Giza, Egypt*

## Abstract

**Purpose** – Diabetic retinopathy (DR) is one of the dangerous complications of diabetes. Its grade level must be tracked to manage its progress and to start the appropriate decision for treatment in time. Effective automated methods for the detection of DR and the classification of its severity stage are necessary to reduce the burden on ophthalmologists and diagnostic contradictions among manual readers.

**Design/methodology/approach** – In this research, convolutional neural network (CNN) was used based on colored retinal fundus images for the detection of DR and classification of its stages. CNN can recognize sophisticated features on the retina and provides an automatic diagnosis. The pre-trained VGG-16 CNN model was applied using a transfer learning (TL) approach to utilize the already learned parameters in the detection.

**Findings** – By conducting different experiments set up with different severity groupings, the achieved results are promising. The best-achieved accuracies for 2-class, 3-class, 4-class and 5-class classifications are 86.5, 80.5, 63.5 and 73.7, respectively.

**Originality/value** – In this research, VGG-16 was used to detect and classify DR stages using the TL approach. Different combinations of classes were used in the classification of DR severity stages to illustrate the ability of the model to differentiate between the classes and verify the effect of these changes on the performance of the model.

**Keywords** Deep learning, Diabetic retinopathy, Retinal fundus images, Transfer learning

**Paper type** Research paper

Received 15 July 2021  
Revised 27 August 2021  
Accepted 13 September 2021

## 1. Introduction

Diabetes mellitus is a chronic disease that is caused by the inability of the pancreas to produce a sufficient amount of insulin, which is a hormone that adjusts blood sugar, or it is the inability of the body to use the produced insulin effectively. High blood sugar is a prevalent result of uncontrolled diabetes and eventually affects many systems of the body, such as blood vessels and nerves. Therefore, it is a main cause of blindness, heart attacks, strokes and kidney failures [1]. Diabetic retinopathy (DR) is considered one of the serious complications of diabetes, which is responsible for 2.6% of overall blindness. High levels of blood sugar destroy the blood vessels in the retina. That rises the probability of fluid leakage and bleeding which results in dangerous vision problems that might leads to blindness [2]. To decrease the dangerous effect of DR, early detection, precise diagnosis and appropriate treatment are required [3, 4]. Therefore, an intelligent automated method for early and accurate detection of DR is required to manage the progress of the disease and thus guarantee appropriate treatment.

Classification of DR includes the weighting of many features and finding the position of these features. This is an exhausting time-consuming task for ophthalmologists, and it is

© Enas M.F. El Houby. Published in *Applied Computing and Informatics*. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licences/by/4.0/legalcode>

*Conflict of interest:* The authors have no conflict of interests to declare.



Applied Computing and  
Informatics  
Emerald Publishing Limited  
e-ISSN: 2210-8327  
p-ISSN: 2634-1964  
DOI 10.1108/ACI-07-2021-0191

---

prone to mistakes. Therefore, ophthalmologists can be supported by computer aided diagnosis systems, which can detect abnormalities and classify the severity of different cases. They can decrease the load on ophthalmologists and reduce inconsistencies between manual readers. Great work is achieved on detecting DR automatically using tradition methods such as k-nearest neighbor (K-NN) and support vector machine (SVM) that depend on hand-crafted features extraction and then classifying different cases depending on the selected features [5, 6]. In contrast, features can be learned automatically from the original images through the training phase using deep learning [7].

The advancement in deep learning has motivated researchers to use deep learning in medical images analysis. Convolutional neural network (CNN) is a type of deep learning networks that are specialized in applications of image analysis. Where the layers nearer to the input of the model learn low-level features like lines, the layers in the middle learn convoluted abstract features that integrate the lower level features, and the layers closer to the output interpret the features extracted in the context of the classification [8]. The high-performing CNN models that were recently applied in image classification tasks and achieved high performance can be imported and used for another image classification task using transfer learning (TL) approach.

TL approach is to utilize a pre-trained model to train a new model. It uses the knowledge obtained during solving one problem and exploits it in solving various but relevant problems. The traits learned by pre-training on the large dataset can be transferred to the new network, where only the classification component is trained on the new smaller dataset, to fine-tune the new data. TL saves considerable time used in developing and training a deep CNN model [9]. There are many high-performing pre-trained models that can be imported and used for image recognition. Most of these models have been developed as part of the annual ImageNet Large Scale Visual Recognition Challenge (ILSVRC). Examples of these models from the published literature are visual geometry group (VGG) [10], inception modules (GoogleNet) [11, 12], residual neural network (ResNet) [13] and neural architecture search network (NasNetLarge) [14] etc. These models were trained using ImageNet data which consists of 1,000,000 images with 1,000 classes, so they have learned to detect generic features and their learned weights are provided and used in similar problems. They achieved state-of-the-art performance and when used to develop other image recognition tasks, they remain effective [15, 16].

In this research, a TL approach using a pre-trained model VGG-16 was utilized to detect DR and classify its stages based on retinal fundus images. The remainder of the paper is organized as follows; related research articles about the detection and diagnosis of DR are reviewed in [section 2](#). The proposed method for detection and classification of DR is introduced in [section 3](#). The experimental results of the proposed model are illustrated in [section 4](#). The discussion and comparison with the literature are presented in [section 5](#). Finally, conclusion and future work are drawn in [section 6](#).

## 2. Related work

Many systems were proposed in the literature for the detection and diagnosis of DR using various machine learning techniques (MLTs). These systems are either based on conventional MLTs, which depend on hand-crafted features extraction, or deep learning where the features are extracted automatically during the training. In the next sections, some of these systems that were found in the literature will be illustrated.

Based on conventional MLTs, random forest (RF) was used to classify fundus images according to DR grades based on 35 extracted features from the detected red lesions and achieved an accuracy of 74.1% on Messidor dataset [17]. Three classifiers which are neural network, RF and SVM were applied on the DIAbetic RETinopathy DataBase fundus images to classify microaneurysms which are early indicators of DR, based on collected patches from

---

images. An AUC of 0.985 and F-measure of 0.926 were achieved using SVM classifier which outperformed the other classifiers [18]. The fuzzy technique was used in different tasks of DR classification such as the preprocessing stage as filtering and histogram equalization and was used as well in the detection of 4 retinal structures. An accuracy of 0.93, specificity of 1 and sensitivity of 0.8679 were achieved using k-NN [19]. Gaussian mixture was used for region segmentation; AlexNet was used for features extraction, linear discriminant analysis and principal component analysis (PCA) for features selection and finally SVM for classification of DR. The best achieved accuracy was 97.93% with sensitivity of 1 and specificity of 0.93 [20].

Based on deep learning, PCA was used to reduce the dimensionality, followed by grey wolf optimization to select the optimal parameters and finally deep neural network was trained on Debrecen dataset from UCI machine learning repository to classify the extracted features into “affected with DR” or not. The achieved accuracy was 97.3% with sensitivity of 91% and specificity of 97% [21]. CNN was trained on the Kaggle fundus images dataset to classify DR and achieved an accuracy of 75%, sensitivity of 30% and specificity of 95% [22]. TL based on pre-trained models was used; GoogLeNet and AlexNet were applied on Kaggle and Messidor-1 datasets. The achieved accuracies were 74.5%, 68.8% and 57.2% for 2-ary, 3-ary and 4-ary classification models, respectively [23]. GoogLeNet Inception v3 classifier was applied on Kaggle dataset. The achieved accuracies were 61.3%, 60.3% and 37.7% for 2-class, 3-class and 5-class, respectively [24]. Synergic deep learning model was applied on Messidor dataset to detect DR and classify its severity. It achieved an accuracy of 99.28, sensitivity of 98 and Specificity of 99 [25]. VGG-16 model was applied on 35,126 images from Kaggle dataset. The achieved 5-class classification accuracy was 74%, the sensitivity was 80%, the specificity was 65% and the AUC was 0.80 [26]. DenseNet and vgg-16 were utilized to classify the fundus images into the 5-stages of DR using 3662 images from Kaggle dataset. The achieved accuracies were 0.9611 and 0.7326 for DenseNet and vgg-16 respectively [27]. vgg-16 was also used 3662 images from Kaggle dataset to classify the severity level of DR with accuracy of 84.31%, F1 score of 84 and an AUC of 97 [28]. AlexNet, VGG-16 and SqueezeNet were applied on 1200 images of MESSIDOR dataset to classify the severity level of DR. The achieved accuracies were 93.46%, 91.82% and 94.49% respectively; specificity of 94.53, 88.54 and 94.54, respectively; and sensitivity of 92.38, 93.47 and 94.47 [29]. Inception V3, vgg-16 and ResNet50 were applied on 35,126 images from Kaggle dataset to classify DR into 5 severity stages; vgg-16 achieved the highest accuracy of 78% [30]. The preprint of this article is available on [31].

### 3. The proposed method

In this section, the used data and the proposed model are described. First, the used dataset for developing the proposed model is presented. Then, the full process, which contains “Pre-processing the data” and “Developing Transfer Learning-based CNN Model” phases, is explained. In “Pre-processing the data” phase, the data is prepared for developing CNN model based on TL approach in the second phase.

#### 3.1 The used dataset

In this research, the proposed model was conducted using the data obtained from the publicly available benchmark dataset which is the Kaggle dataset [32]. The dataset contains colored highly diverse levels of illumination in fundus images. A set of 35,126 retinal images from the Kaggle dataset was used to develop the model.

Kaggle dataset images are in PNG format and they are re-sized to  $224 \times 224$  pixels. Each image is labeled as left or right eye. Each image is categorized according to the level of DR

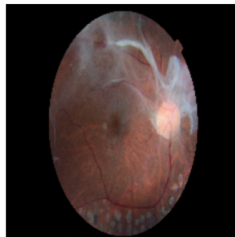
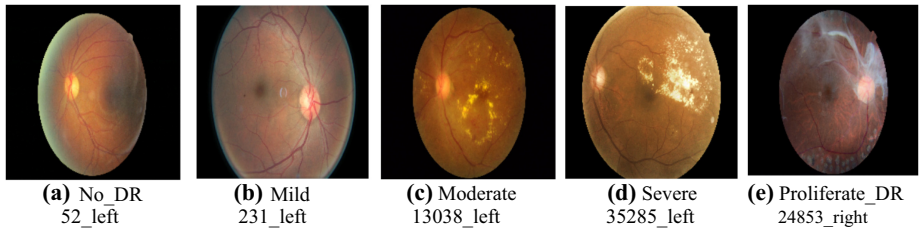
severity into one of 5-class labels (0–4) to represent (normal, mild, moderate, severe, proliferate\_DR) stages. Figure 1 shows different samples from Kaggle dataset representing different stages, where Figure 1(a) is a normal sample and (b)-(e) samples represent different stages of severity.

### 3.2 Pre-processing the data

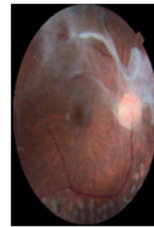
To develop the proposed CNN model based on TL approach, pre-processing steps were applied on the retinal fundus images to prepare the images for the learning phase. The pre-processing steps can be summarized as follow:

- (1) The retinal fundus image region was cropped automatically from each image to remove the background and unwanted region. Figure 2(a) shows a sample of an original image from the Kaggle dataset, while Figure 2(b) shows the same image after removing unwanted region.
- (2) One of the most important challenges in the development of a deep learning model is the unbalanced and limited data size. In this research, the used data does not suffer

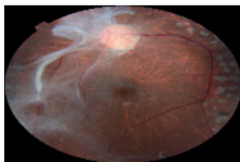
**Figure 1.** Samples from Kaggle dataset representing different stages of DR severity



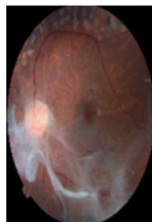
(a) Original image



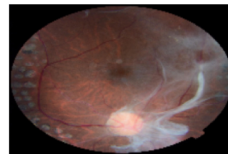
(b) After cropping



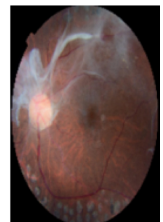
Rotate 90



Rotate 180



Rotate 270



Flipped

(c) Rotation of a cropped image by different orientations

**Figure 2.** Pre-processing steps for one image from the proliferate\_DR images of the Kaggle dataset

from data limitation especially that the adopted approach for learning is TL which relatively overcomes the data limitation problem. However, as it is clear from Table 1 that there is a balancing problem in the used data where the representation of the classes is unbalanced. The images in class 3 and 4 do not have enough representation as the other classes which are an obstacle in the way of detection of the cases belonging to these classes. Therefore, augmentation was applied to the poorly represented classes which are 3 and 4 to solve the data balancing problem. Thus, each training image belonging to these classes was rotated by three angles of 90°, 180° and 270° and then flipped to enlarge the representation of these classes in the dataset. Column 4 in Table 1 shows the number of images in different classes after augmentation. Figure 2(c) shows the augmentation by different orientations of the sample in Figure 2(b) which belongs to proliferate\_DR (class 4) in the Kaggle dataset.

- (3) All images were re-sized to the same size to satisfy the CNN requirement of equally sized images that are provided as input to CNN model.

### 3.3 Developing transfer learning-based CNN model

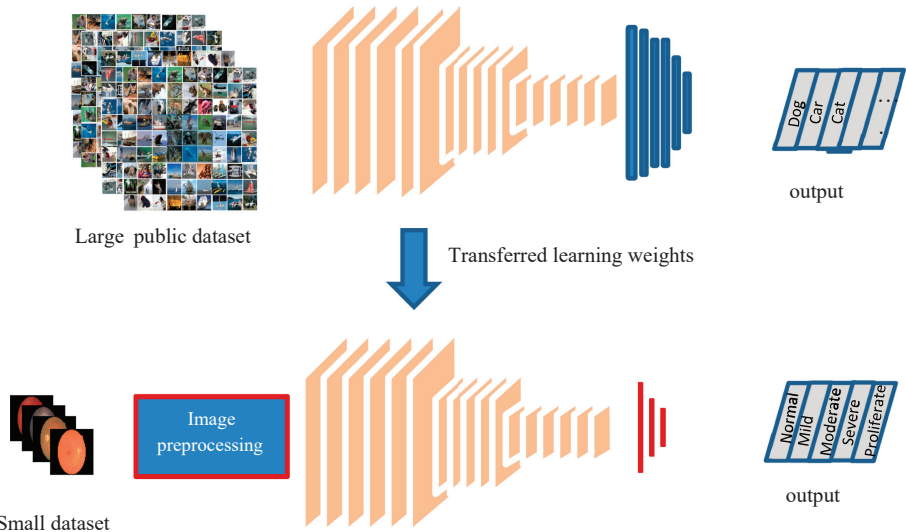
TL is to utilize a pre-trained model to train another model. The pre-trained models such as VGG, ResNet and Inception are trained on ImageNet which is a large dataset. The developers of these models provided their models publicly to enable more research on the use of these representations in computer vision. Where these pre-trained models contain many millions of parameters in their architectures, training them from scratch requires very long computational time and huge number of input images. So, TL is the best solution to many problems where it can exploit pre-trained models to solve other problems such as the one presented in this research. The used TL architecture is shown in Figure 3. As shown in the figure, the pre-trained CNN model was trained using ImageNet which is a large public dataset that contains 1,000,000 images to be classified into 1,000 classes. The retinal fundus dataset was utilized to train the pre-trained network, after applying pre-processing on it.

The top 2 layers of the pre-trained model, which are employed to classify 1000 classes, were removed and replaced by an output layer with SoftMax activation function as a classifier with 5-nodes to supply 5 output classes, which represent the stages of DR. The 5 nodes can be changed to (2–4) nodes according to the different combinations of severity groupings to specify the required output as will be shown in “Experimental Results” section. The residual components of CNN were handled as features extractor for the new dataset, while the pre-trained model weights were kept unchanged. The new network was re-trained with retinal fundus images dataset with learning rate of 0.001 and Adam optimizer, while the number of epochs was (10–20) epochs.

In this research, TL-based model was developed to classify the retinal fundus images dataset into the different stages of DR severity levels. The three most common pre-trained models which are VGG, ResNet and Inception were used to classify the Kaggle dataset into its five severity levels. Where VGG achieved the best result in this task, it was used as the

No. of classes	DR classification	No. of images	No. of images after augmentation
0	No_DR	25,810	25,810
1	Mild	2,443	2,443
2	Moderate	5,292	5,292
3	Severe	873	4,365
4	proliferate_DR	708	3,540
<i>Total</i>		<i>35,126</i>	<i>41,450</i>

**Table 1.**  
The distribution of the  
different classes in  
Kaggle database



**Figure 3.**  
Transfer learning  
architecture

adopted model in this research and so it was further investigated by performing several experiments where they classify the dataset into different combinations of classes.

VGG-16 is composed of 16 depth layers. The input to VGG-16 is a  $(224 \times 224)$  size image. The network contains a set of convolutional filters with  $(3 \times 3)$  size. A stride of 1 pixel is used for all convolution filters, the padding is 1 pixel for  $(3 \times 3)$  convolutional filters. The rectification (ReLU) activation function is used for all hidden layers. Five of convolutional layers are followed by  $(2 \times 2)$  max-pooling layers with a stride of 2. Finally, 2 fully connected (FC) layers with 4096 channels each are applied, followed by the output of 1000 channels (one for each class) with soft-max activation function.

#### 4. Experimental results

This section demonstrates the analysis and the experimental results of the proposed model. To validate the efficiency of the proposed model and to compare the results with others, benchmark dataset was used for implementation. Keras Python deep learning library on top of TensorFlow framework was used for implementing the model based on VGG with 16 layers (VGG-16) on a machine with an Intel® Core™ i7 CPU@ 3.6 GHz with 32 GB RAM and a Titan X Pascal Graphics Processing Unit (GPU). Extensive experiments were conducted to get the best setting that achieves the best results.

The dataset was randomly split into training and test sets, where the training dataset represents 70% of the whole data and the remaining 30% was used to test the model. The classification model was implemented according to the proposed architecture previously described using the training dataset and tested using test data. As mentioned before, the proposed model was applied using the three most common pre-trained models which are ResNet, Inception and VGG to classify the retinal fundus images of the Kaggle dataset into the five severity levels. The achieved accuracies were 66.24%, 63.41% and 73.7% for ResNet50, Inception and VGG-16, respectively. Table 2 shows the achieved accuracies of the used pre-trained models and the input shape required for each model. Since VGG-16 achieved the best result, it was used as the adopted model and hence used for more experiments where it classifies the dataset into different combinations of classes, as will be shown below.

First, to test the capability of the proposed model to detect abnormality in general, experiment #1 was conducted. It is a binary classification task, where it classifies the cases into normal and abnormal that includes the other 4-classes which are {mild, moderate, severe, proliferate\_DR}. So, these 4-classes were merged into 1-class which is abnormal class. The achieved accuracy of this experiment was 75.5% in detecting abnormality.

As mentioned before, each image in the Kaggle dataset is categorized into one of the 5-classes (0–4) according to the level of severity to represent (normal, mild, moderate, severe, proliferate\_DR) stages. To test the capability of the model to classify the cases into the different 5 severity levels, experiment #2 was conducted. The achieved accuracy was 73.7%.

According to the consulted ophthalmologists, in Kaggle database the mild cases did not form an obvious class as some of its cases could be classified as normal while others were more likely to belong to moderate. So, it was suspected that the model might not be able to distinguish them from normal and moderate classes. Likewise, the severe and proliferative were not easily distinguishable. Therefore, in experiment #3, “normal” and “mild” cases; and “severe” and “proliferative” cases were merged. And, in experiment #4, “mild” and “moderate” cases; and “severe” and “proliferative” cases were merged. The achieved accuracies were 80.5% and 76.4%, respectively, which proves that the differentiating traits for the mild class are not evident and the classification of classes of the dataset is not accurate. The increment in the accuracy in experiment #3 compared to experiment #4, clarifies that the mild class is closer to normal class, which may be translated that the number of cases in the mild class that inclines to the normal class are more than the ones that inclines to the moderate class.

In another approach, there was the intention to determine the severity level among the abnormal cases only, neglecting the normal cases. In experiment #5, the model classified the 4 abnormal classes. The achieved accuracy was 63.5%. The reduction in the accuracy compared to experiment #2 is due to the absence of the “normal” class which is proved to be clearly distinct. Due to the mentioned classes convergence, merging between classes was applied excluding the “normal” cases again. So, in experiment #6 mild was merged with moderate, while severe was merged with proliferate\_DR. The achieved accuracy improved to 85.78%.

By consulting ophthalmologists, we found that the 4 stages of abnormality can be mainly categorized into proliferative diabetic retinopathy (PDR) and non-proliferative diabetic retinopathy (NPDR) according to severity level. Therefore experiment#7 was conducted by classifying cases into proliferate\_DR and Non-Proliferative {mild, moderate, severe}. The achieved accuracy was 86.5%. Table 3 shows the results of different experiments of the model built using TL based on VGG-16. The used metrics for evaluation are as follow:

$$\text{Recall}(\%) = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{1}$$

$$\text{Precision}(\%) = \frac{\text{Tp}}{\text{TP} + \text{FP}} \tag{2}$$

Model	Input shape	Accuracy%
ResNet50	224 × 224	66.24
Inception (GoogLeNet)	299 × 299	63.41
VGG-16	224 × 224	73.7

**Table 2.**  
The accuracies of different models in the classification of Kaggle dataset

ACI

$$\text{Accuracy (\%)} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \quad (3)$$

$$\text{F1} = \frac{2 \text{TP}}{2 \text{TP} + \text{FP} + \text{FN}} \quad (4)$$

where TP is true-positive value, FP is false-positive value, TN is true negative value and FN is false-negative value.

## 5. Discussion

Detecting DR and the classification of its severity stages are one of the biggest challenges for ophthalmologists. The contribution of this work is to develop a model that helps in detecting DR and classifying its different stages. Since the available number of cases in different DR datasets are relatively limited, so, TL is the suitable approach for the proposed work to employ pre-trained models to build the model that can classify the DR cases using the available data.

To evaluate the proposed work, it was compared with previous works that used TL applied on the same dataset, which is Kaggle dataset, for a fair comparison as illustrated in Table 4. Chowdhury *et al.* [24] used GoogLeNet to classify DR into 2, 3 and 5 classes and the achieved accuracies are 61.3, 60.3 and 37.7 respectively. Lam *et al.* [23] applied AlexNet and GoogLeNet TL approach, but they stated that GoogLeNet achieved better accuracies than AlexNet, which were 74.5, 68.8 and 57.2 for 2, 3 and 4 classes respectively. Although the two researches [23, 24] used the same model which is GoogLeNet and the same dataset, the results of the two researches are different. That may be resulted from different preprocessing steps and changes in the setting of TL-network. Pratt *et al.* [22] built CNN using Kaggle dataset to

Exp	Categorization	No of classes	Accuracy	Recall	Precision	F1 score
1	(Normal; Mild + Moderate + Severe + Proliferation)	2	75.5	71.7	72.97	70.48
2	(Normal; mild; moderate; severe; proliferation)	5	73.7	67.82	66.85	64.28
3	(Normal + Mild; moderate; Severe + Proliferation)	3	80.5	76.14	74.5	73.86
4	(Normal; Mild + Moderate; Severe + Proliferation)	3	76.4	69.6	70.23	67.94
5	(Mild; moderate; severe; proliferation)	4	63.5	63.5	64.3	62.5
6	(Mild + Moderate; Severe + Proliferation)	2	85.78	84.9	87.14	85.8
7	(Mild + Moderate + Severe; proliferation)	2	86.5	88.05	92.18	90.067

**Table 3.** The results of different experiments built using VGG-16

Reviewed studies	Used technique	Accuracy			
		2-Class	3-Class	4-Class	5-Class
Chowdhury <i>et al.</i> [24]	TL Google's Inception v3	61.3%	60.3%		37.7%
Lam <i>et al.</i> [23]	TL Google's Inception	74.5%	68.8%	57.2%	
Pratt <i>et al.</i> [22]	CNN				75%
Thota and Reddy [26]	VGG-16				75%
Pradhan <i>et al.</i> [30]	VGG-16				78%
<i>Proposed model</i>	<i>TL VGG-16</i>	<i>75.5%</i>	<i>80.5%</i>	<i>63.5%</i>	<i>73.7%</i>

**Table 4.** Comparison between proposed work and related works



---

classify the cases into the 5 classes and the achieved accuracy was 75%. Thota and Reddy [26] used VGG-16 to classify Kaggle dataset into the 5 classes and the achieved accuracy was 75%. Pradhan *et al.* [30] also used the VGG-16 to classify Kaggle dataset into 5 severity stages with an accuracy of 78%. As it is shown, the proposed model results outperform the two works that used the same dataset and TL approach but with GoogLeNet. The third, fourth and fifth works are better than the proposed model in classifying the cases into 5 classes which is the only applied classification in these researches. The third and fourth achieved an accuracy of 75%, although they used different models which are CNN and TL using VGG-16, respectively, but the proposed model achieved an accuracy of 73.7%.

It is worth noting that, by applying the proposed model without augmentation the results of different experiments were better, but it was suffering from overfitting. That was clear from that the results of different experiments that remained with the same accuracy through all epochs, even by using different models which are VGG-16, ResNet50 and Inception. As an example, they all achieved 75% to classify DR into 5 severity stages (experiment # 2) and 91.93% for PDR and NPDR (experiment # 7).

## 6. Conclusion and future work

Recently, the number of diabetes patients has increased dramatically and consequently the number of DR patients has increased. To help in the detection of DR and classification of its grade stages, deep learning was used, in this research. The TL approach, which utilized the VGG-16 pre-trained CNN model, was applied on Kaggle retinal fundus images dataset. The pre-trained VGG-16 model was used for feature extraction, then the top 2-layers were replaced by SoftMax activation function with the new output layer which changed to 2–5 classes according to the experiment.

According to the results of different experiments, it was concluded that the borderline between different classes is not sharp specially between mild and normal, and also between severe and proliferative. And even the ophthalmologists suffer from the difficult distinguishing between these classes. So, more work in the future is needed to find more accurate techniques or models to extract the subtle features that can distinguish between different classes. The proposed architecture can also be applied on other datasets to investigate the behavior of the model with the similar severity groupings. This is a possible topic for future work. Application of the model in real life can be beneficial in diagnosis and prevention of complications of DR.

## References

1. <https://www.who.int/news-room/fact-sheets/detail/diabetes>.
2. Yau J.W., *et al.* Global prevalence and major risk factors of diabetic retinopathy. *Diabetes Care*. 2012; 35(3): 556-64.
3. Flaxman S.R., *et al.* Global causes of blindness and distance vision impairment 1990–2020: a systematic review and meta-analysis. *The Lancet Glob Health*. 2017; 5(12): e1221-34.
4. Bourne R.R., *et al.* Causes of vision loss worldwide, 1990–2010: a systematic analysis. *The Lancet Glob Health*. 2013; 1(6): e339-49.
5. Philip S., *et al.* The efficacy of automated “disease/no disease” grading for diabetic retinopathy in a systematic screening programme. *Br J Ophthalmol*. 2007; 91(11): 1512-7.
6. Mookiah M.R.K., *et al.* Computer-aided diagnosis of diabetic retinopathy: a review. *Comput Biol Med*. 2013; 43(12): 2136-55.
7. Litjens G., *et al.* A survey on deep learning in medical image analysis. *Med Image Anal*. 2017; 42: 60-88.

- 
8. Kharazmi P., *et al.* A computer-aided decision support system for detection and localization of cutaneous vasculature in dermoscopy images via deep feature learning. *J Med Syst.* 2018; 42(2): 33.
  9. Oquab M., *et al.* Learning and transferring mid-level image representations using convolutional neural networks. *Proceedings of the IEEE conference on computer vision and pattern recognition*; 2014.
  10. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. University of Oxford and Google DeepMind. arXiv preprint arXiv:1409.1556; 2014.
  11. Szegedy C, *et al.* Rethinking the inception architecture for computer vision. *Proceedings of the IEEE conference on computer vision and pattern recognition*; 2016.
  12. Szegedy C, *et al.* Going deeper with convolutions. *Proceedings of the IEEE conference on computer vision and pattern recognition*; 2015.
  13. He K, *et al.* Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*; 2016.
  14. Zoph B, *et al.* Learning transferable architectures for scalable image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*; 2018.
  15. Russakovsky O., *et al.* Imagenet large scale visual recognition challenge. *Int J Comput Vis.* 2015; 115(3): 211-52.
  16. Krizhevsky A., Sutskever I., Hinton G.E. Imagenet classification with deep convolutional neural networks. *Adv Neural Inf Process Syst.* 2012; 25: 1097-105.
  17. Seoud L, C. J., Cheriet F. Automatic grading of diabetic retinopathy on a public database. *Proceedings of the ophthalmic medical image analysis second international workshop.* Munich; 9 October 2015.
  18. Cao W, Shan J., Czarnek N., Li L. Microaneurysm detection in fundus images using small image patches and machine learning methods. *Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 13–16 November 2017, Kansas City, MO, USA. p. 325-31.
  19. Rahim S.S., *et al.* Automatic screening and classification of diabetic retinopathy and maculopathy using fuzzy image processing. *Brain Inform.* 2016; 3(4): 249-267.
  20. Mansour R.F. Deep-learning-based automatic computer-aided diagnosis system for diabetic retinopathy. *Bio Eng Lett.* 2018; 8(1): 41-57.
  21. Gadekallu T.R., *et al.* Deep neural networks to predict diabetic retinopathy. *J. Ambient Intell. Humaniz. Comput.* 2020.
  22. Pratt H, *et al.* Convolutional neural networks for diabetic retinopathy. *Procedia Comp Sci.* 2016; 90: 200-5.
  23. Lam C, *et al.* Automated detection of diabetic retinopathy using deep learning. *AMIA summits on translational science proceedings.* 2018; 2018: 147.
  24. Chowdhury M.M.H., Meem N.T.A. A machine learning approach to detect diabetic retinopathy using convolutional neural network. *Proceedings of International Joint Conference on Computational Intelligence.* Springer; 2020.
  25. Kathiresan S., *et al.* Automated detection and classification of fundus diabetic retinopathy images using synergic deep learning model. *Pattern Recogn. Lett.* May 2020; 133: 210-216.
  26. Thota N.B., Reddy D.U. Improving the accuracy of diabetic retinopathy severity classification with transfer learning. *2020 IEEE 63rd international midwest symposium on circuits and systems (MWSCAS).* IEEE; 2020.
  27. Mishra S., Hanchate S., Saquib Z. Diabetic retinopathy detection using deep learning. *2020 International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE).* IEEE; 2020.

28. Bodapati J.D., Shaik N.S., Naralasetti V. Deep convolution feature aggregation: an application to diabetic retinopathy severity level prediction. *Signal, Image and Video Proc.* 2021: 1-8.
29. Rehman M.U., *et al.* Classification of diabetic retinopathy images based on customised CNN architecture. 2019 Amity International Conference on Artificial Intelligence (AICAI). IEEE; 2019.
30. Pradhan A., *et al.* Transfer learning based classification of diabetic retinopathy stages. 2020 International Conference on Computational Performance Evaluation (ComPE). IEEE; 2020.
31. El Houby M.F. Using transfer learning for diabetic retinopathy stages classification. PREPRINT (Version 1) available at Research Square; 30 Apr, 2021.
32. <https://www.kaggle.com>.

#### About the author



Enas M.F. El Houby has received Electrical Communication Engineering B.Sc. degree from Mansura University; M.Sc. in Computer science from Cairo University where the thesis was an output of a collaboration project with Michigan state University in USA, and Ph.D. degrees in Computer science from Cairo University. She is currently an Associate Professor in the Department of Systems and Information, Engineering Research Division, National Research Centre. Her research interests are focused on Artificial intelligence, knowledge discovery and data mining, machine learning techniques, bioinformatics, semantics and information retrieval. Also, she is interested in computer-aided diagnosis (CAD) systems. Enas M.F. El Houby can be contacted at: [enas\\_mfahmy@yahoo.com](mailto:enas_mfahmy@yahoo.com); [em.fahmy@nrc.sci.eg](mailto:em.fahmy@nrc.sci.eg)

---

For instructions on how to order reprints of this article, please visit our website:

[www.emeraldgrouppublishing.com/licensing/reprints.htm](http://www.emeraldgrouppublishing.com/licensing/reprints.htm)

Or contact us for further details: [permissions@emeraldinsight.com](mailto:permissions@emeraldinsight.com)