MEKATRONIKA
e-ISSN: 2637-0883
VOL. 2, ISSUE 1, 68 – 72
DOI: https://doi.org/10.15282/mekatronika.v2i1.6741

ORIGINAL ARTICLE

The Diagnosis Of Diabetic Retinopathy By Means Of Transfer Learning And Fine-Tuned Dense Layer Pipeline

Abdulaziz Abdo Salman[1], Ismail Mohd Khairuddin[1]*, Anwar P.P. Abdul Majeed[1] and Mohd Azraai Mohd Razman[1]

[1]Innovative Manufacturing, Mechatronics and Sports Laboratory, Faculty of Manufacturing and Mechatronics Engineering Technology, Universiti Malaysia Pahang (UMP), 26600 Pekan, Pahang Darul Makmur, Malaysia.

ABSTRACT – Diabetes is a global disease that occurs when the body is disabled pancreas to secrete insulin to convert the sugar to power in the blood. As a result, some tiny blood vessels on the part of the body, such as the eyes, are affected by high sugar and cause blocking blood flow in the vessels, which is called diabetic retinopathy. This disease may lead to permanent blindness due to the growth of new vessels in the back of the retina causing it to detach from the eyes. In 2016, 387 million people were diagnosed with Diabetic retinopathy, and the number is growing yearly, and the old detection approach becomes worse. Therefore, the purpose of this paper is to computerize the old method of detecting different classes of DR from 0-4 according to severity by given fundus images. The method is to construct a fine-tuned deep learning model based on transfer learning with dense layers. The used models here are InceptionV3, VGG16, and ResNet50 with a sharpening filter. Subsequently, InceptionV3 has achieved 94% as the highest accuracy among other models.

ARTICLE HISTORY
Received: 13th May 2020
Revised: 2nd June 2020
Accepted: 12th June 2020

KEYWORDS
CNN
Transfer Learning
Fine-Tuning

INTRODUCTION

Diabetes is a frequent condition nowadays as a result of the various dangerous meals that are manufactured. Diabetes is a disease in which the body is unable to generate enough insulin to convert blood sugar into cells for energy generation, resulting in devastating effects [1]. Diabetic Retinopathy (DR) is a complication of diabetes that damages the patient's eyes. It caused damage to the blood vessels on the back of the eye, known as the retina, due to excessive blood sugar levels. It also causes blindness in the patient if it is not diagnosed early enough to prevent blood from going through. Offbeat vessels form on the region in certain circumstances, which is another cause of blindness [2].

Diabetic retinopathy causes serious vision problems and Diabetic Macular Edema (DME). DME is one of these conditions caused by a leaking fluid in vessels, resulting in a swollen Macula and blurred vision. Neovascular Glaucoma is another DR problem that prevents blood from draining from arteries due to aberrant vessels sprouting out of the retina [3]. According to severity, diabetic retinopathy is divided into four categories: mild, moderate, severe nonproliferative, and proliferative. A retina with a swell occurring in blood vessels in a small area is classified as mild nonproliferative.

In contrast, a retina with some blocked blood vessels is classified as moderate nonproliferative. On the other hand, Severe nonproliferative retinal disease progresses to the third stage, with more blocked blood vessels in the retina preventing blood flow. In the final stage, the disease has progressed to where new abnormal vessels may form, causing blood to seep out, potentially resulting in blindness or visual loss [4].

Diabetes was classified as a global illness in 2016, affecting roughly 387 million people worldwide, with estimates predicting that this number will rise to 592 million by 2035. Diabetic retinopathy affects approximately 93 million individuals worldwide as a result of diabetes. Proliferative Diabetic Retinopathy PDR, on the other hand, has been detected in 17 million people. DME and vision-threatening diabetic retinopathy affect 21 and 28 million individuals, respectively [5].

Early identification of DR can prevent disastrous consequences such as blindness. Therefore, to avoid such outcomes, ophthalmologists inspect the eyes for any signs of DR, which is done manually using an antiquated approach. Drops will be inserted into the eyes to broaden the vision and allow for clearer images, and later a dye may be injected into the veins of the eyes. Optical coherence tomography (OCT), on the other hand, is another test for obtaining clear images of the retina and determining its thickness. All of these tests rely on human intelligence and accompanying technology to detect DR correctly, so errors are likely to occur in a significant percentage of cases [6].

DR is a kind of retinopathy caused by diabetes, and it has a high risk of losing vision, while medicine has not been able to find a replacement. Ophthalmologists have been working to develop a quick and accurate method for detecting DR early in order to avert negative repercussions. Therefore, an automated detection system based on Convolution Neural Network (CNN) models will be developed in this study to determine the degree of DR. Furthermore, a transfer learning strategy will aid in training a model using fine-tuned techniques and dense layers for improved performance. This process aids in reducing time and eliminating human errors, which are the primary goals of this paper.

*CORRESPONDING AUTHOR | Ismail Mohd Khairuddin | ✉ ismalkhai@ump.edu.my

RELATED WORK

Physicians have developed a method of detecting DR by extending their eyesight to take a snapshot of the retina, which can take up to two days because of the complexity of the pictures. Last decade, researchers such as in [7] proposed a method for computerizing the identification of diabetes mellitus (DM) induced by elevated blood glucose levels. Consequently, aberrant vessels form on the back of the eyes, and a level of severity can be recognized accordingly. The author has suggested a CNN-based deep learning approach for detecting DR using colorized pictures of the vessels on the retina. A dataset comprises 166 photos categorized as No DR, Mild NPDR, Moderate NPDR, Severe NPDR, and Proliferative DR from Kaggle online databases. These images were implanted into pre-trained models by transfer learning technique and fine-tuning. AlexNet, VGG16, and Inception-v3 are pre-trained models that are built on CNN feature extractor. Furthermore, an SGD optimizer was utilized to determine the cost's minimal value, and a ReLU activation function was utilized to replace negative pixels with zero. After using 5-fold cross-validation to train the data, Inception-v3 had the highest accuracy of the other models at 63.23%.

Also, [8] has developed a deep learning model by utilizing much more data, where the dataset contains 35,126 retinal images. The author trained the dataset using two pre-trained models, inception-v3 and Xception, with fine-tuning technique. Furthermore, data augmentation aided in avoiding overfitting and improving the robustness of the training models. He examined two activation functions, such as Rectified Linear Unit RELU and Exponential Linear Unit ELU with the SGD and Adaptive Moment Estimation ADAM optimizers. As a result of his research, he managed to obtain an accuracy of 87.12%.

On the other hand, [9] has employed a different approach with the selfsame dataset, where The Gabor filter was used as a feature extractor. This method analyses the texture of photos and converts them to integers as features ranging from 0 to 59. Furthermore, the Gabor approach was combined with several machine learning classifiers to compare and create a flawless automated system, such as C4.5, KNN, random forest, and oneR. The author obtained an accuracy of 86.2% after applying 10-fold cross-validation [10].

METHODOLGY

Dataset

The used dataset here is collected from hosted platform Kaggle. The data involves 5 classes according to severity. It starts from No DR class to Proliferate DR as Figure 1 represents a sample of each class. The dataset consists of 3,662 fundus images with the following portion of each class: 1,805 No DR, 370 Mild, 999 Moderate, 193 Severe and 295 Proliferative DR images [10].

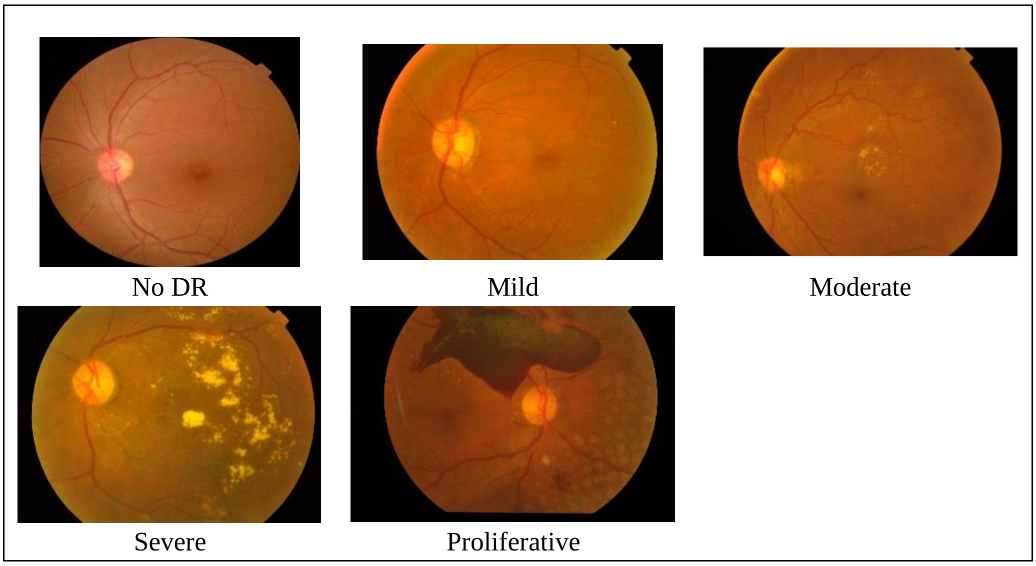

Figure 1: Sample of all DR classes taken from the dataset

Preprocessing Data

The data had been subjected to resizing process and converting each image to a size of 224x224 due to input constraints of pre-trained models. Moreover, data augmentation processes like flipping and rotation are implemented to balance the skewness of different classes. Therefore, the data upsampled to 2,316 fundus images by each class. Furthermore, the data was split into testing and training with a ratio of 80:20, respectively. The augmented data was exposed to a 5x5 sharping filter to increase the clearness of the edges of small objects on the images, as shown clearly in Figure 2.

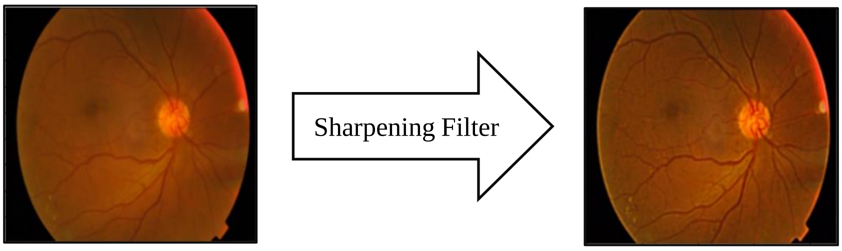

Figure 2: a) Original image, b) Image after applying sharping filter

CNN Pre-trained Models

The training-process structure is divided into a) pre-trained models, b) the neural network methodology is a powerful tool for simulating neurons in human brains. CNN is a Neural Network (NN) that excels in picture analysis because it can distinguish between them by passing them through convolution layers. Convolution and pooling are the two operational aspects of this type of NN. Layers with filters such as the kernel filter are used in the convolution process to extract particular characteristics from pictures. Pooling operation is constructing additional layers following convolution layers to minimize feature selection [11].

Keras is a Python framework that includes a number of pre-trained models, such as ResNet50, which achieves high accuracy when using the ImageNet database [10]. This model was produced by solving the gradient problem, and it is made up of a 50-layer stack of blocks. VGG16, on the other hand, is made up of three entirely interconnected layers, the first two of which have 4096 channels each while the third level has 1000 channels. These channels serve the same purpose as RBG, but they reflect different aspects of the input image. VGG16 employed Linear Response Normalization LRN with the ReLU activation function in all hidden layers, which improved accuracy but increased training time [12]. Moreover, convolutions, average pooling, max pooling, concatenates, dropouts, and fully linked layers are among the symmetric and asymmetric building elements that make up InceptionV3. Batch normalization is applied to activation inputs throughout the model. Softmax is used to calculate the loss [13].

These models are followed with fully dense connected layers, as illustrated clearly in Figure 3. The FCL had constructed with self-normalized NN by Selu activation function and L2 regularization in all dense layers. All the out parameters from pre-trained models undergoing global average pooling to flatten the dimension with batch normalization. Sequences of dense layers were connected to each other, starting with two layers with 512 neurons flowing it with batch normalization. Then, a dense layer with 256 neurons, flowing with dropout dense 128 layers with batch normalization. Lastly, two dense layers with 64 neurons with dropout with softmax classifier. In addition, all models are trained with 1650 GTX GPU graphic card with 4 GB manufactured by AMD company, and the models are build by Tensoflow Keras.

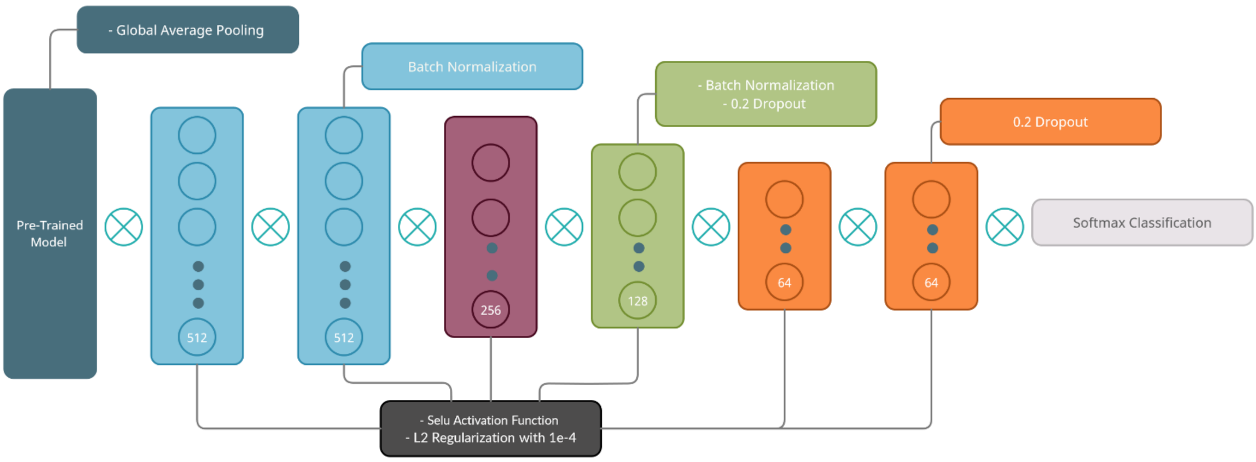

Figure 3: Full Connected Dense Layers

RESULT AND DISCUSSION

The first stage of building the model is to pick the optimal dataset to obtain the best results, as shown in Table 1. The upsample dataset was found to be the best chunk of data for building the models, with an accuracy of 89.53% using the normal transfer learning TF methodology of VGG16.

Table 1: Accuracy of all kind of data

Data	No.	Accuracy
Normal	3,662	81.17%
Downsample	965	67.69%
Upsample	11,580	89.53%

The models were trained using a variety of strategies, starting with Sub-Fine-Tuned SFT layers and a learning rate of $1e^{-5}$. This impacted the model's accuracy since it had to overcome the local minimum to achieve the global minimum. When 12 layers were unfrozen, VGG16 outperformed other models, achieving the best accuracy of 91.59 %. Another aspect was that all models were subjected to fully fine-tuned FT layers, which resulted in Inceptionv3 having the most excellent performance metrics.

L2 regularisation, on the other hand, was used to alleviate the model's overfitting problem, and all models' performance matrices improved to approximately 94 %. However, as a result of using L2, the cross-entropy had grown by approximately 2%. All completely fine-tuned FT models with L2 regularisation models produced high results, as shown in Table 2, in all tests. As a result, InceptionV3 had a 94.08 %, 93.99 % recall, and 94.02% precision.

Table 2: Different stage Results

Model	Method	Performance Metrics			
		Cross entropy	Accuracy	Recall	Precision
VGG16	TF+SFT	0.4915	89.53%	89.22%	89.73%
VGG16	TF+ ($\alpha = 1e^{-5}$)+SFT	0.2872	91.59%	91.29%	92.00%
Inceptionv3	TF+ ($\alpha = 1e^{-5}$)+FT	0.2576	93.83%	93.78%	93.90%
Inceptionv3	TF+ ($\alpha = 1e^{-5}$)+FT+L2	0.4132	94.08%	93.99%	94.20%
RestNet50	TF+ ($\alpha = 1e^{-5}$)+FT+L2	0.5017	92.65%	92.44%	92.76%

InceptionV3's best result is the best result that could be produced so far. Another useful way to look at the output of the inceptionV3 model is to utilize the confusion matrix, as shown in Fig 4.1. The number of corrected and missing images predicted by the model is represented using this visualization style. No DR region has a lighter area since the model correctly predicted 484 pictures out of 486 and missed two photos in the Mild class. In contrast, the model fared poorly in the prediction of Proliferate DR, where it correctly identified 400 photos out of 451 and was primarily confused with 22 photos in the Moderate class. The remaining courses are in a respectable range, with an average percentage of the inaccuracy of roughly 6%.

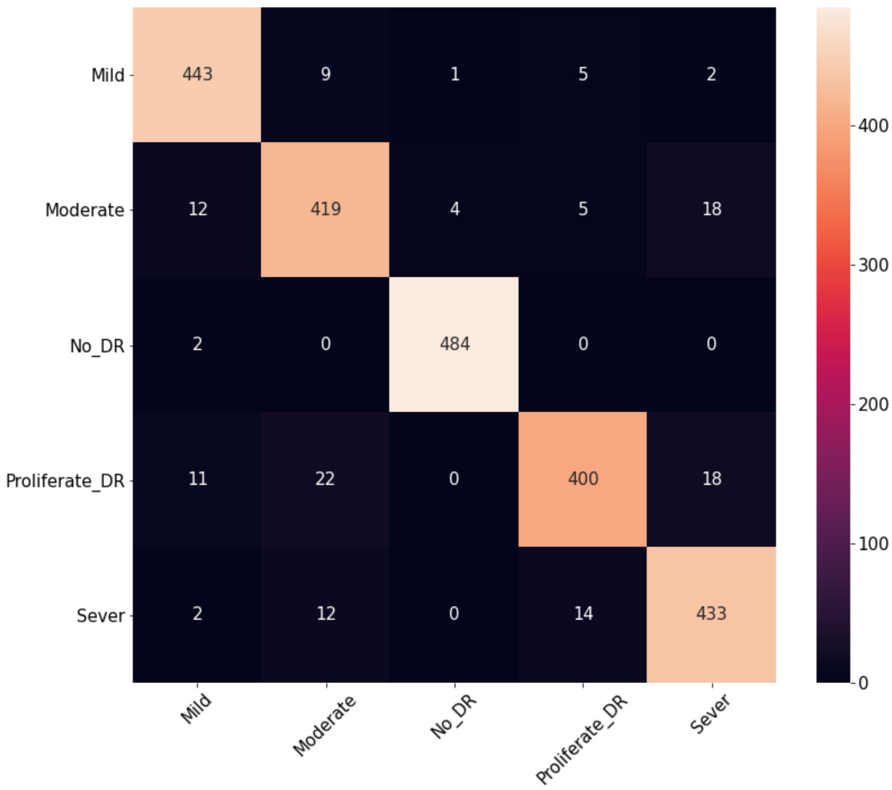

Figure 4: Confusion matrix of InceptionV3

CONCLUSION

To summarize, DR is a sensitive condition that causes visual loss in individuals. The old method of diagnosing took a long time. Thus deep learning is a contemporary method to computerize the process. CNN's deep learning model was proposed, and it consists of dense layers that played a key part in stabilizing the system, while transfer learning is a potent strategy to speed up the process. The fine-tuning technique is a part of transfer learning, and it boosts the model to obtain optimal results. Therefore, When the entire model was fine-tuned, InceptionV3 produced excellent results, with an accuracy of 94%. It is, however, trapped due to mislabeled images and a resemblance between classes. As a result, it is recommended that the images should be relabeled and using other complicated techniques, such as the ensemble approach, and that can be boosted by using a different classifier, like Support Vector Machine (SVM) or Random Forests.

REFERENCES

[1] P. Z. Zimmet, D. J. Magliano, W. H. Herman, and J. E. Shaw, "Diabetes: A 21st century challenge," *Lancet Diabetes Endocrinol.*, vol. 2, no. 1, pp. 56–64, 2014.

[2] M. Chan, "Global Report on Diabetes," *Isbn*, vol. 978, no. April, pp. 6–86, 2016.

[3] M. Whitehead, S. Wickremasinghe, A. Osborne, P. Van Wijngaarden, and K. R. Martin, "Diabetic retinopathy: a complex pathophysiology requiring novel therapeutic strategies," *Expert Opin. Biol. Ther.*, vol. 18, no. 12, pp. 1257–1270, 2018.

[4] A. W. Stitt *et al.*, "The progress in understanding and treatment of diabetic retinopathy," *Prog. Retin. Eye Res.*, vol. 51, pp. 156–186, 2016.

[5] R. Gargeya and T. Leng, "Automated Identification of Diabetic Retinopathy Using Deep Learning," *Ophthalmology*, vol. 124, no. 7, pp. 962–969, 2017.

[6] N. Asiri, M. Hussain, F. Al Adel, and N. Alzaidi, "Deep learning based computer-aided diagnosis systems for diabetic retinopathy: A survey," *Artif. Intell. Med.*, vol. 99, no. December 2018, p. 101701, 2019.

[7] X. Wang, Y. Lu, Y. Wang, and W. B. Chen, "Diabetic retinopathy stage classification using convolutional neural networks," *Proc. - 2018 IEEE 19th Int. Conf. Inf. Reuse Integr. Data Sci. IRI 2018*, pp. 465–471, 2018.

[8] S. Mohammadian, A. Karsaz, and Y. M. Roshan, "Comparative Study of Fine-Tuning of Pre-Trained Convolutional Neural Networks for Diabetic Retinopathy Screening," *2017 24th Iran. Conf. Biomed. Eng. 2017 2nd Int. Iran. Conf. Biomed. Eng. ICBME 2017*, no. December, pp. 1–6, 2018.

[9] V. T, S. M, A. Kumaravel, and K. B, "Gabor filter and machine learning based diabetic retinopathy analysis and detection," *Microprocess. Microsyst.*, no. October, p. 103353, 2020.

[10] B. Li and Y. He, "An Improved ResNet Based on the Adjustable Shortcut Connections," *IEEE Access*, vol. 6, pp. 18967–18974, 2018.

[11] S. Albawi, T. A. M. Mohammed, and S. Alzawi, "Layers of a Convolutional Neural Network," *Ieee*, 2017.

[12] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 2015, pp. 1–14.

[13] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-December, pp. 2818–2826, 2016.