Available online at http://scik.org

Commun. Math. Biol. Neurosci. 2021, 2021:42

https://doi.org/10.28919/cmbn/5660

ISSN: 2052-2541

EVALUATION OF CONVOLUTIONAL NEURAL NETWORK VARIANTS FOR DIAGNOSIS OF DIABETIC RETINOPATHY

ALHADI BUSTAMAM[1], DEVVI SARWINDA[1,*], RADIFA H. PARADISA[1], ANDI ARUS VICTOR[2], ANGGUN RAMA YUDANTHA[2], TITIN SISWANTINING[1]

[1]Department of Mathematics, Faculty of Mathematics and Natural Science, Universitas Indonesia, Depok, Indonesia

[2]Department of Ophthalmology, Faculty of Medicine, Universitas Indonesia, Cipto Mangunkusumo National General Hospital, Jakarta, Indonesia

Abstract: Diabetic Retinopathy (DR) is a long-term complication of Diabetes Mellitus (DM) that impairs vision. This stage occurs in visual impairment and blindness if treated late. DR identified through scanning fundus images. A technique on classifying DR in fundus images is the deep learning approach, one of the methods of implementing machine learning. In this study, the Convolutional Neural Networks (CNN) method applied with the ResNet-50 and DenseNet-121 architectures. The data adopted in this analysis was generated from DIARETDB1, an online database containing fundus images. Then, the pre-processing stage is carried out on the fundus image to improve model performance, such as selected the green channel from the images and inverted it, converted the images into grayscale images, and applied Contrast Limited Adaptive Histogram Equalization (CLAHE) for uniform contrast in the images. The outcome of this research indicates that the ResNet-50 model is better than DenseNet-121 in detecting DR. The most reliable results from the ResNet-50 model's case testing are accuracy, precision, and recall of 95%, 98%, and

---

[*]Corresponding author

E-mail address: devvi@sci.ui.ac.id

Received March 8, 2021

96% respectively.

Keywords: diabetic retinopathy; fundus image; deep learning; resnet; densenet.

2010 AMS Subject Classification: 92B20.

1. INTRODUCTION

Diabetes Mellitus (DM) is a chronic disease characterized by an increase in blood glucose levels that exceed average values. It occurs because the body cannot produce insulin, or the effectiveness of using insulin is reduced. Diabetes is a severe threat to global health. On average, there are 463 million adults aged 20-79 years are suffering from diabetes. This number estimated to rise to 578 million adults in 2030 and 700 million adults in 2045. Currently, Indonesia occupies the 7th position in the world with the number of diabetics in 2019, reaching 10.7 million adults and is predicted to increase to 13.7 million adults by 2030 [1].

The excessive-high blood sugar levels will increase the risks of encountering complications, especially macrovascular, microvascular, or both complexities if DM patients do not maintain their blood sugar levels properly. Macrovascular is one of the cardiovascular diseases, such as heart attacks, strokes, and lack of blood flow to the legs. Meanwhile, microvascular is an injury to small blood vessels which includes eye damage (retinopathy), kidney damage (nephropathy), and nerve damage (neuropathy) [2].

One of the most common microvascular complications is Diabetic Retinopathy (DR). A study involving 1,967 DM patients with an average age of 58.4 years in Indonesia reported that there were 29.1% DM sufferers experiencing complications in the eye, where 10.5% experienced complications of Non-Proliferative Diabetic Retinopathy (NPDR), and 5.7% Proliferative Diabetic Retinopathy (PDR) [3]. Qureshi et al. [4] revealed that DR is a disorder identified by the presence of microaneurysms (MA), soft/hard exudates (EXs), and hemorrhages (HMs) on the retina. In general, DR is divided into two stages, specifically Non-Proliferative Diabetic Retinopathy (NPDR) and Proliferative Diabetic Retinopathy (PDR). NPDR is further divided into mild, moderate, and severe stages, while PDR is an advanced DR.

Prior researchers used a branch of science in detecting DR on fundus images. There were particular methods applied, such as the Support Vector Machine (SVM) technique on the STARE and DRIVE dataset [5], the Artificial Neural Network (ANN) method on the MESSIDOR dataset [6], the Complete method of Local Binary Pattern on the STARE dataset [7].

In this study, the authors propose a DR diagnosis using a deep learning approach, namely the Convolutional Neural Network (CNN) method, to extract features and classify DR on fundus image data. Convolutional Neural Network (CNN) is a Neural Network type with a convolutional concept, where CNN uses a distinctive, well-adapted architecture to classify images [8]. In earlier studies, this approach was utilized to identify pulmonary nodules on Computed Tomography (CT) images [9], Xie et al. detected brain tumors in the Cancer Imaging Archive (TCIA) [10], and others so. CNN also has various architectures such as LeNet, AlexNet, Inception-V3, VGG, ResNet, DenseNet, and others. However, researchers only use the ResNet-50 and DenseNet-121 model architecture, where both architectures are CNN architectures that use the concept of a shortcut connection in their network [11]. Shortcut connections make the training process more convenient and lead to better results than ordinary deep neural networks [12]. In this study, an evaluation of the performance of ResNet-50 and DenseNet-121 in detecting DR on DIARETDB1 and DIARETDB0 data is an online database containing fundus images.

2. Materials and Methods

In evaluating the performance of the CNN method to detect diabetic retinopathy, the researchers conducted several stages of the study, described in Figure 1.

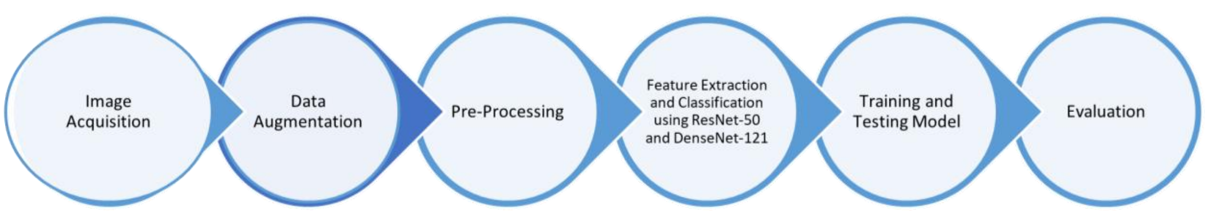
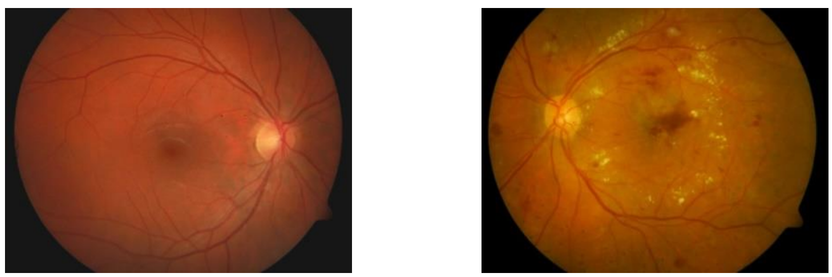


Figure 1. Research flow

2.1. Image Acquisition

The data used in this study is an online database obtained from https://www.it.lut.fi/project/imageret/diaretdb1. DIARETDB1 is a public database as a benchmark for detecting Diabetic Retinopathy (DR). DIARETDB1 database consists of 89 color fundus images with 5 non-DR (normal) images and 84 DR images, with uniform sizes, namely 1500x1152 pixels. The image was captured by a digital fundus camera with a 50-degrees view and adjusted for different imaging settings. The image data obtained in the Kuo-pio university hospital. Images were captured with few 50-degree field-of-view digital fundus cameras with unknown camera settings (flash intensity, shutter speed, aperture, gain) [13]. These data can also be used to assess the overall performance of diagnostic methods.



(a) (b)

Figure 2. Fundus image sample (a) DR, (b) normal

2.2. Data Augmentation

In the DIARETDB1 and DIARETDB0 dataset, both datasets have fewer images in the normal class than the DR class, resulting in an imbalance class. Data augmentation was implemented to expand the number of datasets. The more data that is trained in deep learning, the better the image recognition system will be. With big data, convolutional networks are proven to be very good for medical image analysis tasks [14]. The data augmentation techniques used were rotation, flip, and add light, which was randomly performed.

2.3. Image Pre-processing

Fundus image data in DIARETDB1 and DIARETDB0 have red (R), green (G), and blue (B) components. The authors only take the green channel and apply an inverted green channel so that the blood vessels in the fundus image appear brighter than the background. In the next step, the image is transformed into grayscale, and the size is reduced to 224 pixels to reduce the number of parameters used during the computation process. The authors also apply image contrast uniformity using Contrast Limited Adaptive Histogram Equalization (CLAHE) [17]. More briefly, the preprocessing stage displayed in Figure 3.

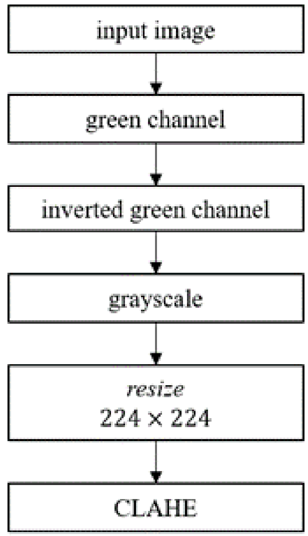


Figure 3. The pre-processing process from input datasets to CLAHE

2.4. Feature Extraction and Classification using Convolutional Neural Network Model

Convolutional Neural Network (CNN) is a deep learning method that was first introduced in 1989 by Yann LeCun, a French scientist. LeCun builds on the Fukushima discovery, which had previously discovered the Neocognitron, a fundamental image recognition neural network. Since then, research in improving CNN's performance has continued to develop, especially in 2015, to date [11]. Patterson and Gibson (2017) explained that the main objective of CNN is to study high-level features of data through convolution operations [15]. This network is suitable for image recognition and image classification. In building the CNN architecture, there are three basic stages required, namely input, feature extraction (convolution layer, ReLU activation function, and pooling layer), and classification (fully connected layer).

In this study, we proposed two architectures model on CNN, ResNet-50, and DenseNet-121. This research used ResNet and DenseNet architecture because both architectures utilize shortcut connections [11]. This shortcut connection improves the training process to obtain better outcomes than other deep neural networks [16].

2.4.1. ResNet-50 Architecture

ResNet is another representative of the deep network. Prior to ResNet, the state-of-the-art DNNs was more extensive. However, deep networks are tough to train because of the notorious vanishing gradient problem as the gradient is back propagated to earlier layers, repeated multiplication may make the gradient infinitely small. By introducing a skip connection (or shortcut connection) to fit the input from the previous layer to the next layer without any modification of the input, ResNet can have a very deep network of up to 152 layers. The architecture of ResNet-50 is shown in Figure 4.

In Figure 4, the meaning of writing "7×7 conv, 64/2" is that the convolution process is carried out using a 64-filter size 7×7 kernel with a stride shift size equal to 2. If "/ 2" is not listed, then the process is carried out with stride is equal to 1. In ResNet-50, Batch Normalization and ReLU are applied consecutively after the convolution process. In this study, the fundus image is encoded by densely joined CNN. The authors resize the fundus image into a 224 x 224 grid. The weights of ResNet are initialized using Adam's standard parameters. Then, we treat the final fully connected layer as a fixed feature extractor.

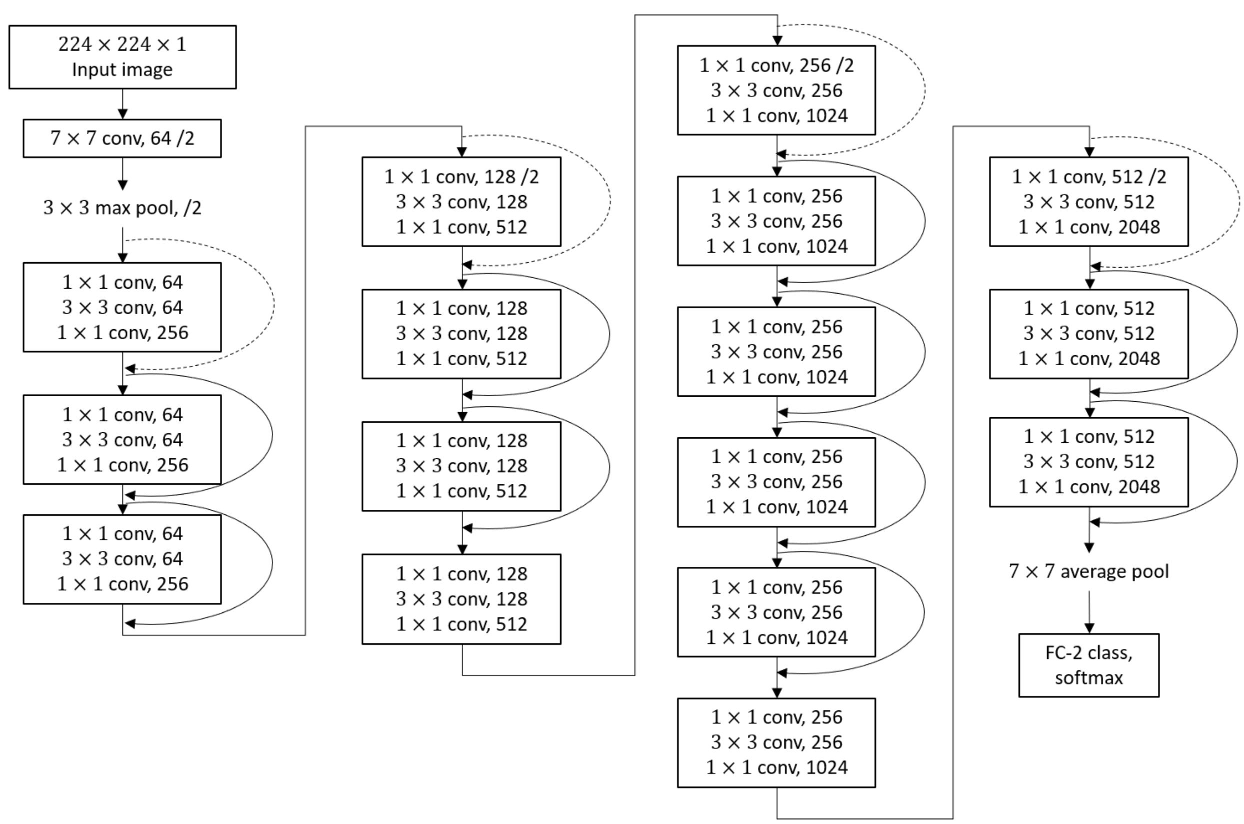


Figure 4. Architecture of ResNet-50

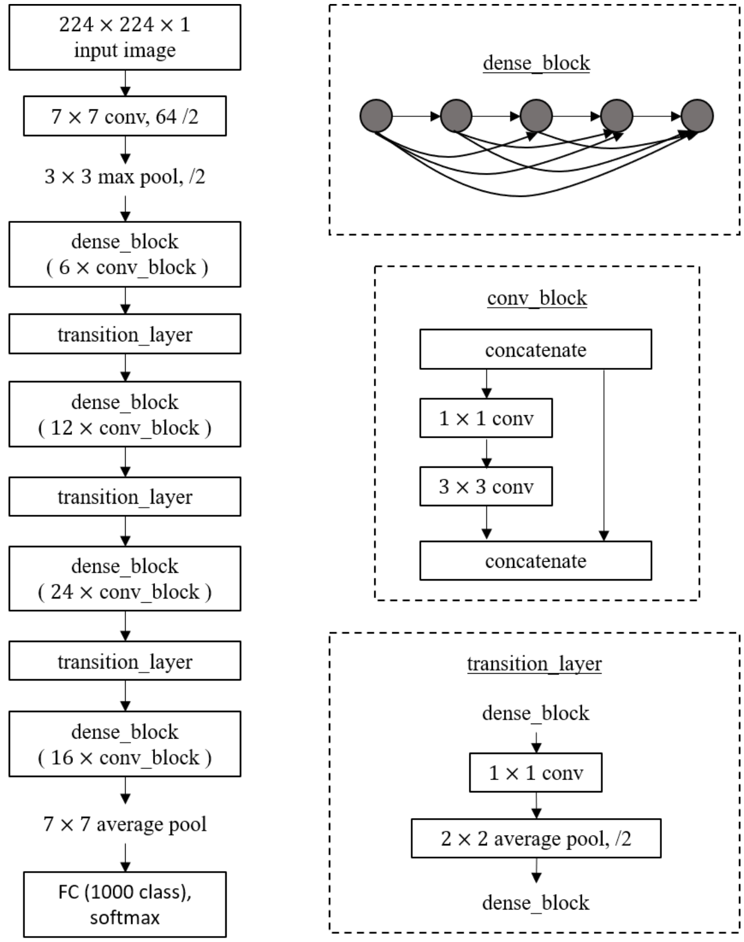


Figure 5. Architecture of DenseNet-121

2.4.2. DenseNet-121 Architecture

In ResNet, gradients can proceed rapidly through the identity function from later layers to the earlier layers. To enhance the information flow between layers, DenseNet utilizes direct connections from any layer to all of its subsequent layers. Consequently, all layer receives a concatenation of the feature maps of all preceding layers.

Huang, Liu, Van Der Maaten, and Weinberger (2017) proposed a new architecture for CNN namely DenseNet [16]. The main idea of this architecture is to connect each layer to every other layer. The l^{th} layer receives the feature map of all previous layers, $x_0, \ldots, x_{l-1}$. The l^{th} layer output is defined as:

$$x_l = H_l([x_0, x_1, \ldots, x_{l-1}])$$

where $[x_0, x_1, \ldots, x_{l-1}]$ is a concatenation operation (feature map combination) at layers

$0, \dots, l-1$, and H_l is defined as a composite function of operations such as batch normalization, ReLU, pooling, or convolution. Concatenation operation is also a shortcut connection on DenseNet. Meanwhile, in this study, the normalization was carried out using BN and activating ReLU. Then, a dense block is applied by combining several stack convolutions layers. From the combined results of the convolution layer on the dense block, the next step is the transition layer which is used to adjust the dimensions between dense blocks. For more details, an illustration of the DenseNet-121 architecture can be seen in Figure 5.

2.5. Experimental Design

In this study, several scenarios were carried out, including in the distribution of datasets, both training data, validation data, and testing data. Validation data consists of 15% from training datasets are chosen randomly. There are three divisions of the dataset which can be seen in Table 1.

Table 1. Dataset Classification

Case	Training Data	Testing Data
Case 1	70%	30%
Case 2	80%	20%
Case 3	90%	10%

Table 2. Confusion Matrix

Actual Class	Prediction Class	
	Positive	Negative
Positive	TP	FN
Negative	FP	TN

Meanwhile, to evaluate the performance of the proposed method, researchers used three measurements, namely accuracy, precision, and recall with the Confusion Matrix in Table 2 and the following formulas:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{1}$$

$$Precison = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

Accuracy describes how accurate the model can classify data correctly. Precision (positive predictive value) is the ratio of true positive predictions to the overall positive predicted results. Recall (true positive rate) is the ratio of true positive predictions compared to the overall true positive data.

3. Results and Discussion

The results of this study are divided into several results according to the flow of research. The results of the first experiment were a pre-processing process carried out on the two datasets, such as DIARETDB0 and DIARETDB1 Datasets. The results of preprocessing are shown in Figure 6.

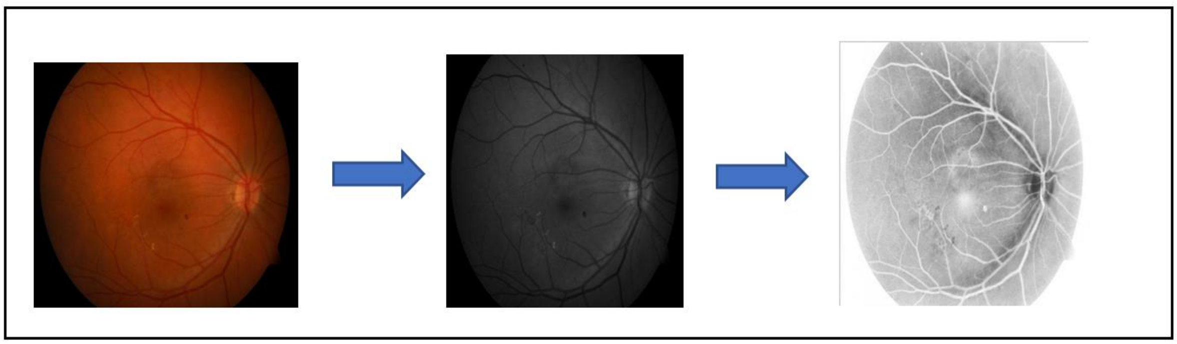


Figure 6. The pre-processing process from input datasets to CLAHE

In building the ResNet-50 and DenseNet-121 implementation systems, Google Colaboratory is used which is connected to Python 3 and the Graphical Processing Unit (GPU). GPU performance is faster than CPU due to its parallel processing architecture, which allows it to perform multiple calculations across data streams simultaneously [18, 19].

3.1. Evaluation of ResNet-50 and DenseNet-121 using DIARETDB1

In this study, ResNet-50 architectural model was implemented in two stages, such as the model training and model testing, where each of them was carried out five times running experiments. Table 3 shows that the ResNet-50 training evaluation model recognized the training data image with an average accuracy of 97% in case 2 and successfully realized the data validation image with an average accuracy of 92.8% in case 3. However, the best running time for Model training occurs in case 1 because it uses less data than the other cases.

Table 3. Evaluation of Training Model in ResNet-50 for DIARETDB1

Case	**Trial**	**Accuracy**		**Running Time Training (second)**
		Training	**Validation**	
Case 1	1	98%	95%	637,31
	2	100%	95%	341,66
	3	93%	90%	312,37
	4	100%	91%	309,78
	5	88%	81%	305,36
	Average	**95,8%**	**90,4%**	**381,296**
Case 2	1	100%	93%	723,08
	2	91%	86%	367,27
	3	100%	93%	382,86
	4	99%	95%	359,54
	5	95%	90%	189,05
	Average	**97%**	**91,4%**	**404,36**
Case 3	1	100%	97%	764,99
	2	100%	99%	805,56
	3	90%	93%	436,2
	4	91%	90%	428,6
	5	90%	85%	420,11
	Average	**94,2%**	**92,8%**	**571,092**

Table 4. Evaluation of Testing Model in ResNet-50 for DIARETDB1

Case	Trial	Accuracy	Precision	Recall	Running Time Testing (second)
Case 1	1	94%	92%	96%	3.56
	2	90%	98%	83%	1.4
	3	86%	82%	92%	1.31
	4	92%	98%	86%	1.56
	5	79%	78%	83%	1.26
	Average	**88.2%**	**89.6%**	**88%**	**1.818**
Case 2	1	95%	96%	94%	2.98
	2	87%	81%	96%	0.64
	3	95%	94%	97%	0.65
	4	90%	97%	83%	1.15
	5	93%	92%	93%	0.73
	Average	92%	92%	92.6%	1.23
Case 3	1	95%	98%	92%	2.32
	2	98%	98%	98%	1.56
	3	85%	81%	92%	0.81
	4	85%	91%	78%	0.8
	5	98%	100%	96%	0.38
	Average	92.2%	93.6%	91.2%	1.174

In the testing model stage, a confusion matrix is used to evaluate the performance of the ResNet-50 in classifying new data that has never been involved in the training model, namely data testing. As in the previous discussion, the testing model is divided into three cases. Based on Table 4, the results of the testing evaluation result of the ResNet-50 model can be known as the best accuracy and precision, namely when using data separation in case 3 with an average accuracy of

92.2% and an average precision of 93.6. Meanwhile, the best average recall is obtained from the separation of case 2 data of 92.6%. Of the three cases, the fastest testing time was case 3 which took 1.174 seconds.

Table 5. Evaluation of Training Model in DenseNet-121 for DIARETDB1

Case	Trial	Accuracy		Running Time Training (second)
		Training	Validation	
Case 1	1	79%	78%	669,3
	2	95%	90%	350,95
	3	57%	63%	338,97
	4	63%	64%	329,73
	5	94%	90%	325,67
	Average	**77,6%**	**77%**	**402,924**
Case 2	1	82%	75%	758,59
	2	73%	74%	384,86
	3	85%	80%	425,5
	4	84%	77%	388,2
	5	90%	85%	208,36
	Average	**82,8%**	**78,2%**	**433,102**
Case 3	1	92%	88%	800,06
	2	91%	90%	837,2
	3	63%	58%	454,72
	4	80%	80%	426,27
	5	75%	69%	440,52
	Average	**80,2%**	**77%**	**591,754**

In Table 5, the evaluation of the training model DenseNet-121 obtained the best results in case 2 where the model succeeded in recognizing the training data image with an average accuracy of 82.8% and succeeded in recognizing the data validation image with an average accuracy of 78.2%. In addition, the best running time for model training occurs in case 1 which takes 402.924 seconds.

Table 6. Evaluation of Testing Model in DenseNet-121 for DIARETDB1

Case	**Trial**	**Accuracy**	**Precision**	**Recall**	**Running Time Testing (second)**
Case 1	1	79%	77%	83%	5,01
	2	93%	95%	90%	1,8
	3	57%	55%	74%	1,82
	4	63%	60%	77%	1,47
	5	90%	96%	84%	2,24
	Average	**76,4%**	**76,6%**	**81,6%**	**2,468**
Case 2	1	82%	76%	93%	4,53
	2	72%	64%	100%	0,6
	3	86%	100%	72%	0,6
	4	82%	76%	94%	0,66
	5	89%	98%	80%	1,19
	Average	**82,2%**	**82,8%**	**87,8%**	**1,516**
Case 3	1	90%	95%	84%	3,84
	2	93%	89%	98%	1,28
	3	60%	56%	100%	0,86
	4	72%	100%	44%	0,36
	5	74%	71%	80%	0,35
	Average	**77,8%**	**82,2%**	**81,2%**	**1,338**

Evaluation of the testing performance of the DenseNet-121 model in classifying testing data was using a confusion matrix. As in the previous discussion, the testing model is divided into three cases. Based on Table 6, the best average accuracy, precision, and recall occurred in case 2 with their respective values, namely 82.2%, 82.8%, and 87.8%. Of the three cases, the fastest training time was in case 3 which took 1.338 seconds.

4. CONCLUSION

The best model training and testing results were obtained according to the implementation of the CNN method with the ResNet-50 and DenseNet-121 architectures in classifying Diabetic Retinopathy (DR) on the fundus image data. In this study, the implementation of ResNet-50 and DenseNet-121 was carried out five times on executing program trials. The results of the performance evaluation of the models are calculated from the average value of the experiments.

Evaluation of training models based on data separation cases, accuracy training ResNet-50 and DenseNet-121, obtained the best results in case 2 with ResNet-50 at 97% and DenseNet-121 at 82.8%. On accuracy validation, the best result was obtained in case 3 with ResNet-50 of 92.8%. The best model running time training occurs in case 1 with ResNet-50 for 381.296 seconds or 6 minutes 21.296 seconds. Overall, the evaluation of the training model ResNet-50 is better than DenseNet-121. The best model training results for ResNet-50 have been previously described. In contrast, the best results from DenseNet-121, namely the best accuracy training and validation, occur in case 2, respectively, at 82.8% and 78.2%, and the best running time occurs in case. 1 for 402,294 seconds or 6 minutes 42,294 seconds.

Based on the testing model evaluation, the best results on the ResNet-50 model are the accuracy of 92.2% (case 3), precision of 93.6% (case 3), recall of 92.6% (case 2), and running time testing takes 1.174 seconds (case 3). Meanwhile, the best results on the DenseNet-121 model are accuracy, precision, and recall of 82.2%, 82.8 %%, and 87.8% for case 2, respectively, and running time testing takes 1.338 seconds for case 3. Overall, the evaluation of training and testing performance of ResNet-50 is better than DenseNet-121 on identifying DR using the DIARETDB1 dataset with 1000 augmented data.

In addition to classification tasks, analyzing fundus images can be done for segmentation tasks. It can assist the doctor in recognizing and detecting DR lesions on images. The segmentation technique can be done using a clustering approach [19, 20]. Recently, many studies have developed this technique based on deep learning. It is one of the prospective research areas in medical image segmentation.

ACKNOWLEDGMENT

This research was supported by the PDUPT 2020 research grant with contract number NKB-325/UN2.RST/HKP.05.00/2020. The authors appreciate colleagues from the Directorate General of Higher Education (BRIN/DIKTI) and the Directorate of Research and Community Engagement Universitas Indonesia who contributed insights and expertise to advance this research in innumerable ways.

CONFLICT OF INTERESTS

The author(s) declare that there is no conflict of interests.

REFERENCES

[1] International Diabetes Federation. IDF Diabetes Atlas, 9th edn. https://www.diabetesatlas.org. (2020).

[2] E.F. Brutsaert, Complications of diabetes mellitus - msd manual professional edition. https://www.msdmanuals.com/professional/endocrine-and-metabolic-disorders/diabetes-mellitus-and-disorders-of-carbohydrate-metabolism/complications-of-diabetes-mellitus. (2020)

[3] A.R. Cholil, D. Lindarto, T.G.D. Pemayun, W. Wisnu, P. Kumala, H.H.S. Puteri, DiabCare Asia 2012: diabetes management, control, and complications in patients with type 2 diabetes in Indonesia, Med. J. Indones. 28 (2019), 47–56.

[4] I. Qureshi, J. Ma, Q. Abbas, Recent Development on Detection Methods for the Diagnosis of Diabetic Retinopathy, Symmetry. 11 (2019), 749.

[5] D.B. Mule, S.S. Chowhan, D.R. Somwanshi, Detection and Classfication of Non-proliferative Diabetic Retinopathy Using Retinal Images, in: K.C. Santosh, R.S. Hegadi (Eds.), Recent Trends in Image Processing and Pattern Recognition, Springer Singapore, Singapore, 2019: pp. 312–320.

[6] S. Chakraborty, G.C. Jana, D. Kumari, A. Swetapadma, An improved method using supervised learning technique for diabetic retinopathy detection, Int. J. Inform. Technol. 12 (2020), 473–477.

[7] D. Sarwinda, A. Bustamam, A. Wibisono, A complete modelling of Local Binary Pattern for detection of diabetic retinopathy, in: 2017 1st International Conference on Informatics and Computational Sciences (ICICoS), IEEE, Semarang, 2017: pp. 7–10.

[8] M.A. Nielsen, Neural networks and deep learning. Determination Press, San Francisco, 2015.

[9] F. Özyurt, E. Sert, E. Avci, E. Dogantekin, Brain tumor detection based on Convolutional Neural Network with neutrosophic expert maximum fuzzy sure entropy, Measurement. 147 (2019), 106830.

[10] H. Xie, D. Yang, N. Sun, Z. Chen, Y. Zhang, Automated pulmonary nodule detection in CT images using deep convolutional neural networks, Pattern Recognit. 85 (2019), 109–119.

[11] A. Khan, A. Sohail, U. Zahoora, A.S. Qureshi, A survey of the recent architectures of deep convolutional neural networks, Artif. Intell. Rev. 53 (2020), 5455–5516.

[12] K. He, X. Zhang, S. Ren, J. Sun; Deep Residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770-778.

[13] T. Kauppi, V. Kalesnykiene, J.K. Kamarainen, et al. DIARETDB1-standard diabetic retinopathy database calibration level 1, 2007.

[14] C. Shorten, T.M. Khoshgoftaar, A survey on Image Data Augmentation for Deep Learning, J. Big Data. 6 (2019), 60.

[15] J. Patterson, A. Gibson, Deep learning: A practitioner's approach, O'Reilly Media, Inc, Sebastopol, 2017.

[16] G. Huang, Z. Liu, L. van der Maaten, K.Q. Weinberger, Densely connected convolutional networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 4700-4708

[17] K. Zuiderveld, Contrast Limited Adaptive Histogram Equalization, Chapter VIII.5, Graphics Gems IV. P.S. Heckbert (Eds.), Cambridge, MA, Academic Press, 1994, pp. 474-485.

[18] B. Alhadi, G. Ardaneswari, H. Tasman, D. Lestari, Performance evaluation of fast smith-waterman algorithm for sequence database searches using CUDA GPU-based parallel computing, J. Next Generat. Inform. Technol. 5 (2014), 38-46.

[19] A. Bustamam, M. S. Sehgal, N. A. Hamilton, S. Wong, M. A. Ragan, and K. Burrage, An efficient parallel

implementation of markov clustering algorithm for large-scale protein-protein interaction networks that uses MPI, in 5th IMT-GT International Conference on Mathematics, Statistics and Their Applications, 2009.

[20] T.S. Permata, A. Bustamam, Clustering protein-protein interaction network of TP53 tumor suppressor protein using Markov clustering algorithm, in: 2015 International Conference on Advanced Computer Science and Information Systems (ICACSIS), 2015: IEEE, pp. 221-226.