JPPI Vol 9 No 2 (2019) 141-150

Jurnal Penelitian Pos dan Informatika

32a/E/KPT/2017

e-ISSN 2476-9266
p-ISSN: 2088-9402

Doi:10.17933/jppi.2019.090205

COMPARISON OF CONVOLUTIONAL NEURAL NETWORK MODEL IN CLASSIFICATION OF DIABETIC RETINOPATHY

PERBANDINGAN MODEL CONVOLUTIONAL NEURAL NETWORK PADA KLASIFIKASI RETINOPATI DIABETES

[1]Hartanto Ignatius, [2]Ricky Chandra, [3]Nicholas Bohdan, [4]Abdi Dharma

[1234] Universitas Prima Indonesia
Jalan Sekip, Kec. Medan Petisah, Kota Medan, Sumatra Utara 20111, Indonesia
abdidharma@unprimdn.ac.id[4]

Naskah Diterima: 29 August 2019; Direvisi : 26 November 2019; Disetujui : 7 December 2019

Abstract

Untreated diabetes mellitus will cause complications, and one of the diseases caused by it is Diabetic Retinopathy (DR). Machine learning is one of the methods that can be used to classify DR. Convolutional Neural Network (CNN) is a branch of machine learning that can classify images with reasonable accuracy. The Messidor dataset, which has 1,200 images, is often used as a dataset for the DR classification. Before training the model, we carried out several data preprocessing, such as labeling, resizing, cropping, separation of the green channel of images, contrast enhancement, and changing image extensions. In this paper, we proposed three methods of DR classification: Simple CNN, Le-Net, and DRnet model. The accuracy of testing of the several models of test data was 46.7%, 51.1%, and 58.3% Based on the research, we can see that DR classification must use a deep architecture so that the feature of the DR can be recognized. In this DR classification, DRnet achieved better accuracy with an average of 9.4% compared to Simple CNN and Le-Net model.

Keyword: Diabetic Retinopathy, Messidor, Deep Learning, Convolutional Neural Network

INTRODUCTION

There are several diseases that can be caused by complications with diabetes mellitus, one of which is Diabetic Retinopathy (DR). DR is often known as microvascular diseases. This disease attacks blood vessels in the retina of the eye, with marked damage to blood vessels and bleeding in the eye's retina. DR cause decreased vision and blindness to sufferers (Wang & Lo, 2018). The problem regularly occurs with DR is the delay in diagnosis because, in the initial stages, the sufferers do not experience interference with their vision. Observation of patient's retina can detect Diabetic Retinopathy. It takes a long time to detect DR manually. Automatic identification related to DR is one of the methods to detect it (Noronha & Nayak, 2012).

Messidor is a dataset containing 1,200 retinal images. This dataset is often used to evaluate segmentation and assessment of Diabetic Retinopathy and has been distributed since 2008 (Decencière et al., 2014).

Based on an article written by Vogt M., Machine Learning is a field in computer science that studies pattern recognition and computational learning theories on artificial intelligence (AI). Deep Learning is one of the methods of Machine Learning to build machines that have intelligence like humans (Vogt, 2019).

In the article written by Majaj NJ and Pelli DG, it was also explained that Machine Learning could be a tool for automatic classification. Deep learning is better than Artificial Neural Networks in the 1980s (Majaj & Pelli, 2018).

Convolutional Neural Network (CNN) is a method that combines the workings of Artificial Neural Networks and Deep Learning. CNN is a method that develops the workings of the Multilayer Perceptron (MLP), which is useful in two-dimensional data processing (Han & Li, 2015). CNN can classify images with pretty good accuracy (Bora, Chowdhury, Mahanta, Kundu, & Das, 2016). There are various types of models/architectures developed with different accuracy.

In research conducted by Regina Lourdhu Suganthi S, Hanumanthappa M, and Kavitha S., the CNN method is used with the AlexNet model to classify images. The results of the research are well-classified testing images. This research also shows that deep learning methods are effective in classifying images (Regina Lourdhu Suganthi, Hanumanthappa, & Kavitha, 2018). In research conducted by Al-Jawfi R., they use the LeNet architecture to recognized Arabic script. The LeNet architecture successfully recognizes Arabic script correctly (Al-Jawfi, 2009).

In research conducted by G Alaslani M and A. Elrefaei L, the AlexNet model was used to recognize iris. Data preprocessing is carried out before training the model. The research uses several types of datasets to compare the results. The accuracy results obtained is 98.33% (G Alaslani & A. Elrefaei, 2018).

In research conducted by Sisodia DS, Nair S, and Khobragade P., they made separation of green channel of images, contrast enhancement, cropping, and resizing in the preprocessing data stage (Sisodia, Nair, & Khobragade, 2017).

In research conducted by Adarsh P and Jeyakumari D, the SVM method was used to classify DR according to their severity. This research uses the DIARETB1 and DIARETB0 dataset. The average accuracy of this research was 95.3%. (Adarsh & Jeyakumari, 2013).

In research conducted by Pratt, researchers made their architecture in classifying DR based on their severity. They used the dataset from Kaggle. In the preprocessing stage, they did image resizing. The resizing size was 512 x 512 pixels. The results of this research are 75% accuracy (Pratt, Coenen, Broadbent, HaDiabetes Retinopatiing, & Zheng, 2016).

In research conducted by Xu K, Feng D, and Mi H, researchers made their architecture in classifying DR based on the detection of the Diabetic Retinopathy. In the preprocessing stage, they did image resizing. The size was resized to 512 x 512 pixels. The results of this research are 91.5% accuracy (Xu, Feng, & Mi, 2017).

In research conducted by Arcadu, F et al, researchers predicted the progress of DR using CNN. The Inception V-3 model was used in this study. The results from predictions using CNN are quite high. (Arcadu et al., 2019).

Based on the explanation above, we will compare several CNN models in classifying DR classification.

METHODOLOGY

This is a quantitative research, where research is done by developing mathematical models, theories, or hypotheses. This research will compare three CNN models in classifying DR classification. The classification in this research is binary classification which consists of normal retina and DR retina.

In this research, we use fundus images as objects in the classification of normal retina and DR retina. The DR retina will have one or more features of DR. These feature are Neovascularization, Microaneurysms, Edema, Exudates and Cotton Wool Spots (Duh, Sun, & Stitt, 2017) .

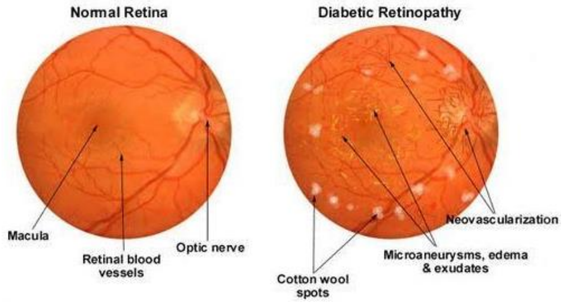

Figure 1. Difference between Normal Retina and Diabetic Retinopathy Retina

The dataset used in this research was obtained from the Messidor Database which can be downloaded on the website http://www.adcis.net/en/third-party/messidor. The Messidor database has 12,00 fundus images, which are divided into 546 normal images that are normal and 654 DR images. Fundus images in the Messidor Database come in various sizes, like 1440 x 960, 2240 x 1488, and 2304 x 1536. The extension of fundus images is *tiff (Decencière et al., 2014).

Table 1. Example of Image on the Messidor Dataset

Fundus Image	Variable
	Normal
	DR

1. Data Preprocessing

Labeling is the first process conducted in this research, followed by separating dataset into two groups: Normal Retina and DR Retina. In the Messidor dataset, Retinopathy grade was given on the fundus image. Retinopathy grade is given a range from 0 to 4. Images that have retinopathy grade 0 is grouped into the normal dataset. Apart from retinopathy grade 0 which is grouped into DR.

After labelling the dataset based on groups, we split the dataset into three parts: Train, Validation, and Test with a ratio of 5:2:3. The number of train, validation and tests dataset are 600, 240, and 360 with a total of 1,200 images.

The process was followed by cropping, resizing, separation of the green channel of images, contrast enhancement, and changing the image extension. The cropping process is useful for cutting out parts of the image that are meaningless, ensuring that the images are consistent with each other, and only displays the relevant parts required for the training. The resizing process is useful to resize the various sized images into one one size (Sisodia et al., 2017).

Figure 2. Cropping and resizing

The Separation of the green channel of images is the process of removing color channels that do not belong to the green channel. At the green channel, all features related to DR can be identified more clearly than in other channels. Increased contrast serves to improve parts of the image. Images that have no similarity in terms of contrast result in some images not being able to show parts clearly. (Sisodia et al., 2017)

Figure 3. Separation of the green channel of images and Contrast Enhancement

In the training process, Keras libraries cannot process *tiff extension data, so that the image extension of the image is changed to *tiff from *jpeg.

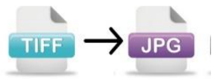

Figure 4. Extension Changing

2. Convolutional Neural Network (CNN)

This research uses three types of architectures/models in the DR classification. The model used is a model which has only a few layers because of limited resources. The training data is carried out for 50 iterations

The first model we use was the simple CNN model, and the model was enough to classify dogs and cats images. The model consists of two Convolutional layers, Average Pooling, Flatenning layers, and two Fully Connected layers. The activation function used is the softmax activation function (M, K, Laskshmi, Madel, & Kurakula, 2018).

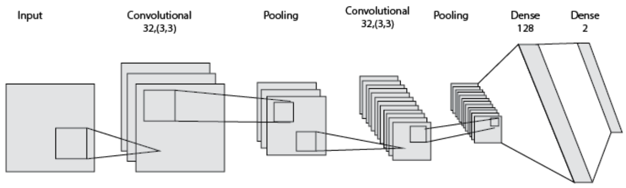

Figure 5. Model 1

The second model we use was Le-Net model. The Le-Net successfully recognizes Arabic script correctly. The model consists of two Convolutional layers, Average Pooling, Flatenning layers, and two Fully Connected layers, and with softmax activation function (Al-Jawfi, 2009).

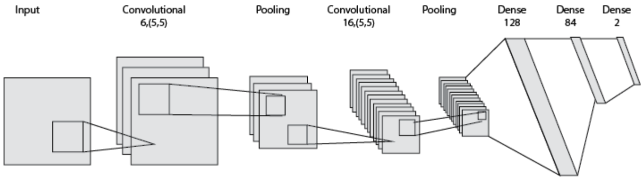

Figure 6. Model 2

The third model that we use was customized model. We want to compare this customized models with those two previous models, and we call this model as DRnet. The model consists of four Convolutional layers, Average Pooling, Flatenning Layers, and two Fully Connected layers with a softmax activation function.

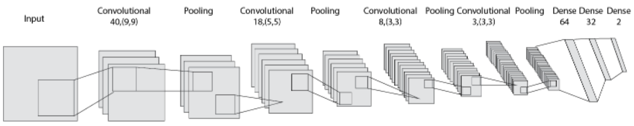

Figure 7. Model 3 (DRnet Model)

3. Confusion Matrix

The Confusion matrix is a method used in calculating accuracy. The confusion matrix is presented in the form of a table stating the amount of correct test data classified and the amount of test data incorrectly classified (Visa, Ramsay, Ralescu, & Van Der Knaap, 2011).

Table 2. Confusion Matrix

Correct Classification	Classified as	
	Predicted +	Predicted -
Actual +	True Positives (TP)	False Negatives (FN)
Actual -	False Positives (FP)	True Negatives (TN)

The value resulting from Confusion Matrix is Accuracy and Misclassification:

a. Accuracy, the percentage of the amount of data that is classified correctly.

$$Accuracy = \frac{(TP + TN)}{Total\ Data}$$

b. Misclassification (Error) Rate, percentage of the amount of data classified incorrectly.

$$Missclassification\ Rate = \frac{(FP + FN)}{Total\ Data}$$

RESULTS AND DISCUSSION

Results

A. Data Preprocessing

Data preprocessing in this research consists of data labeling, cropping, resizing, separation of the green channel of images, contrast enhancement, and changing the extensions of images.

1. Labeling

   In this stage, we split the dataset into three parts, Train, Validation, and Test set. The amount of data in the Train, Validation, and Test set is 600, 240, 360.

Table 3. Labeling

No	Group	Number of Data
1.	Train Normal	273
2.	Train Diabetic Retinopathy	327

No	Group	Number of Data
3.	Validation Normal	109
4.	Validation Diabetic Retinopathy	131
5.	Test Normal	165
6.	Test Diabetic Retinopathy	195

2. Cropping

   In this stage, we cut the image into squares, and the cut part is in the middle right on the retina object.

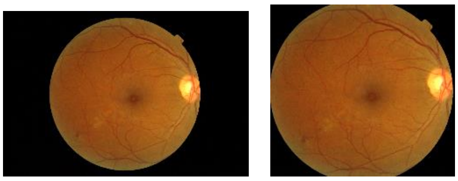

Figure 8. Original Image and Cropping

3. Resizing

   In this stage, the size of the image are changed to 512 x 512 pixels.

Figure 9. Resizing

4. Separation of the green channel of images

   In this stage, the image turns green due to the removal of the red channel and the green channel. This process leaves the green channel in the image.

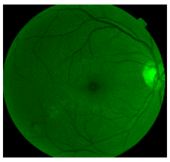

Figure 10. Separation of the green channel of images

5. Contrast Enhancement

   In this stage, increasing the contrast of the image is done so that the image looks brighter than usual. Parts of the image become more clearly visible than before this process.

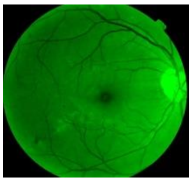

Figure 11. Contrast Enhancement

B. Training

The number of parameters that we do training is 61,326, 66,666, and 513,346. The parameters trained can be describe.

Table 4. Number of Simple CNN Model's Parameter Trained

No	Layer	Output	Parameter
1	Convolutional	62,62,32	896
2	Max Pooling	31,31,32	0
3	Convolutional	29,29,32	9248
4	Max Pooling	14,14,32	0
5	Flattening	6272	0
6	Dense	128	502944
7	Dense	2	258

Table 5. Number of Le-Net Model's Parameter Trained

No	Layer	Output	Parameter
1	Convolutional	28,28,6	456
2	Max Pooling	14,14,6	0
3	Convolutional	10,10,16	2416
4	Max Pooling	5,5,16	0
5	Flattening	400	0
6	Dense	120	48120
7	Dense	84	10164
8	Dense	2	170

Table 6. Number of DRnet Model's Parameter

Trained

No	Layer	Output	Parameter
1	Convolutional	120,120,40	9760
2	Convolutional	116,116,18	18018
3	Max Pooling	58,58,18	0
4	Convolutional	56,56,8	1304
5	Max Pooling	28,28,8	0
6	Convolutional	26,26,3	219
7	Max Pooling	13,13,3	0
8	Flattening	507	0
9	Dense	64	35219
10	Dense	32	2080
11	Dense	2	66

The training time for each model is 4,386s, 4,101s, and 7,121s. The highest accuracy when training in models is 99.91%, 75.53%, and 87.73%,. The graph below shows the accuracy of the training models.

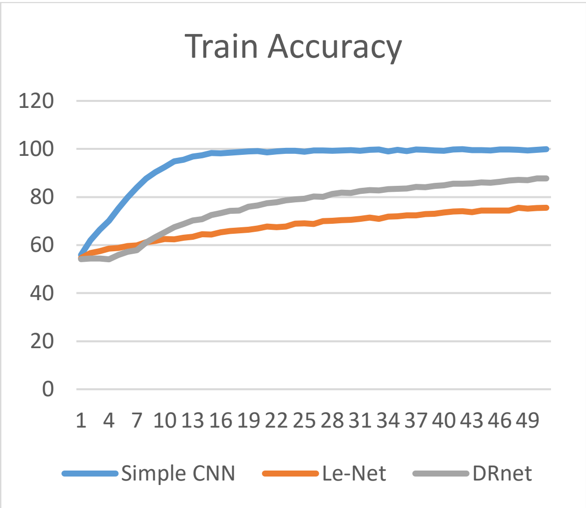

Figure 12. Training Accuracy Graph

The highest accuracy on validation is 64.58%, 71.255, 70.1%. The graph below shows the accuracy of the models at the time of data validation.

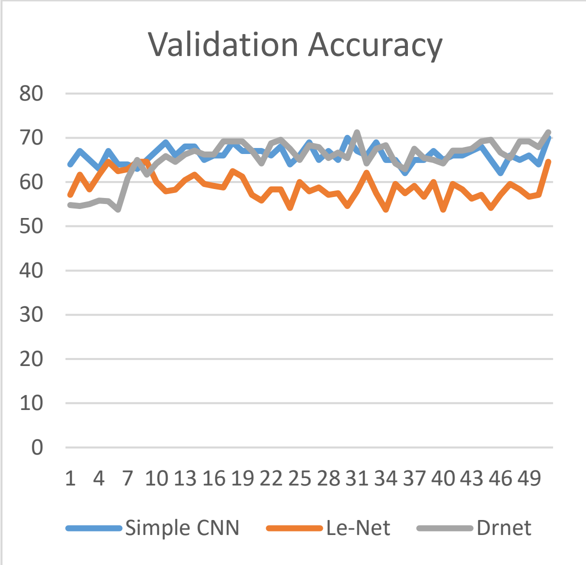

Figure 13. Validation Accuracy Graph

C. Testing

The test model of dataset presented in the Confusion Matrix as follows:

Simple CNN Model

Table 7. Confusion Matrix Simple CNN

Correct Classification	Classified as	
	Predicted +	Predicted -
Actual +	126 (TP)	107 (FN)
Actual -	69 (FP)	58 (TN)

1. Le-Net Model

Table 8. Confusion Matrix Le-Net Model

Correct Classification	Classified as	
	Predicted +	Predicted -
Actual +	77 (TP)	74 (FN)
Actual -	118 (FP)	91 (TN)

2. DRnet Model

Table 9. Confussion Matrix DRnet Model

Correct Classification	Classified as	
	Predicted +	Predicted -
Actual +	139 (TP)	93 (FN)
Actual -	56 (FP)	71(TN)

From Confusion Matrix, the Simple CNN Model acquired an accuracy of 51.1% and the misclassification of 49.9%. The Le-Net obtained an accuracy of 46.7% and the misclassification of 53.3%. Whereas DRnet obtained an accuracy of 58.3% and the misclassification of 41.2%.

The following table summarized the comparison of the models:

Table 10. Comparison of CNN Models

No	Comparison Object	Simple CNN	Le-Net	DRnet
1.	Number of parameters trained	513.346	61.326	66.666
2.	Total Training Time	4.386 s	4.101 s	7.121 s
3.	Average Total Training Time	87.72 s	82 s	118.68 s
4.	The highest value of accuracy in training	99,91%.	75.53%	87.73%
5.	The highest value of accuracy in validation	70.1%	64.58%	71,255%
6.	Percentage of Classification Accuracy	51.1%	46.7%	58.3%
7.	Percentage of Misclassification Classification	49.9%	53.3%	41.2%

Discussion

We created DRnet model architecture based on simple CNN by changing the architecture in several ways, such as by increasing or decreasing the number of layers, the convolution filter value, and the amount of dense.

CNN architecture needs to be deep enough to classify DR images. DR features such as microaneurysms are tiny enough and do not spread evenly in the image. In DR grades 1 and 2, microaneurysms show up very small in the picture. CNN model that is not deep enough is not capable of detecting these features.

The simple CNN and Le-net are models that can classify images precisely if the features of the problem are visible enough in the image. Because the microaneurysms feature is not spread evenly on the image, and this makes the models output a wrong result.

In this research, we developed a model based on the simple CNN model, which we call DRnet. This model has 4 Convolutional Layers, two layers more than those of Simple CNN and Le-Net. The results of this model have higher accuracy than the previous model, 7.2% against SimpleNet, and 11.6% than Le-Net, with an average of 9.4% increased accuracy.

Compared to the research conducted by other researchers, accuracy results in this research with DRnet is fair enough, with only 11 layers compared to the other research. In Arcadu research, they developd an Inception V-3 model with 42 layers. In Xu-Fu research, their model had eight layers, whereas, in Pratt research, their model had ten layers. It is proven that the DR classification requires a model with a layer that is deep enough to obtain good results.

Based on table 10, parameters trained by Simple CNN, Le-net, and DRnet are 513,346, 61,325, and 66,666. DRnet has a smaller trained parameter than Simple CNN, which is 15,320, but the accuracy of DRnet is higher than SimpleNet. Furthermore, although DRnet has more trained parameters than Le-net, which is 5,341, by adding more parameters, the accuracy resulted from DRnet is higher than Le-Net. Based on this, we can conclude that the number of parameters trained in the model does not guarantee better accuracy, so we cannot rely on the number of parameters as a benchmark to get good results.

CONCLUSIONS

The conclusion in this research, DR classification must use a deep architecture so that the feature of the DR can be recognized.

It is necessary to readjust the proven model according to the specific case, for example, in this DR classification, DRnet achieves better accuracy compare to Simple CNN and Le-Net model.

Another conclusion from this study is that the number of training parameters does not determine qualifications. It is essential to adjust the amount of training to get the right accuracy and the DRnet model with 66,666 parameters have a higher percentage accuracy compare to the other models.

ACKNOWLEDGMENT

The authors would like to express our heartfelt appreciation to our lecturers in Universitas Prima Indonesia, who have provided guidance, encouragement, and criticism that are useful for this research and also to Universitas Prima Indonesia for its support in conducting this research. Moreover, we want to thank our parents for their support and encouragement during our research.

REFERENCES

Adarsh, P., & Jeyakumari, D. (2013). Multiclass SVM-based automated diagnosis of diabetic retinopathy. *International Conference on Communication and Signal Processing, ICCSP 2013 - Proceedings*, 206–210. https://doi.org/10.1109/iccsp.2013.6577044

Al-Jawfi, R. (2009). Handwriting Arabic character recognition lenet using neural network. *International Arab Journal of Information Technology*, *6*(3), 304–309.

Arcadu, F., Benmansour, F., Maunz, A., Willis, J., Haskova, Z., & Prunotto, M. (2019). Deep learning algorithm predicts diabetic retinopathy progression in individual patients. *Npj Digital Medicine*, *2*(1). https://doi.org/10.1038/s41746-019-0172-3

Bora, K., Chowdhury, M., Mahanta, L. B., Kundu, M. K., & Das, A. K. (2016). Image Classification Using Convolutional Neural Networks. *ACM International Conference Proceeding Series*, (June). https://doi.org/10.1145/3009977.3010068

Decencière, E., Zhang, X., Cazuguel, G., Laÿ, B., Cochener, B., Trone, C., … Klein, J. C. (2014). Feedback on a publicly distributed image database: The Messidor database. *Image Analysis and Stereology*, *33*(3), 231–234. https://doi.org/10.5566/ias.1155

Duh, E. J., Sun, J. K., & Stitt, A. W. (2017). Diabetic retinopathy: current understanding, mechanisms, and treatment strategies. *JCI Insight*, *2*(14), 1–13. https://doi.org/10.1172/jci.insight.93751

G Alaslani, M., & A. Elrefaei, L. (2018). Convolutional Neural Network Based Feature Extraction for IRIS Recognition. *International Journal of Computer Science and Information Technology*, *10*(2), 65–78. https://doi.org/10.5121/ijcsit.2018.10206

Han, X., & Li, Y. L. (2015). The Application of Convolution Neural Networks in Handwritten Numeral Recognition. *International Journal of Database Theory and Application*, *8*(3), 367–376. https://doi.org/10.14257/ijdta.2015.8.3.32

M, S., K, Laskshmi, D., Madel, M., & Kurakula, K. (2018). Convolutional Neural Network For Supervised Image Classification. *International Journal of Pure and Applied Mathematics*, *119*(14), 77–83.

Majaj, N. J., & Pelli, D. G. (2018). Deep learning-Using machine learning to study biological vision. *Journal of Vision*, *18*(13), 1–13. https://doi.org/10.1167/18.13.2

Noronha, K., & Nayak, K. P. (2012). A review of fundus image analysis for the automated detection of diabetic retinopathy. *Journal of Medical Imaging and Health Informatics*, *2*(3), 258–265.

https://doi.org/10.1166/jmihi.2012.1098

Pratt, H., Coenen, F., Broadbent, D. M., Harding, S. P., & Zheng, Y. (2016). Convolutional Neural Networks for Diabetic Retinopathy. *Procedia Computer Science*, *90*(July), 200–205. https://doi.org/10.1016/j.procs.2016.07.014

Regina Lourdhu Suganthi, S., Hanumanthappa, M., & Kavitha, S. (2018). Image classification using Deep learning. *ICSNS 2018 - Proceedings of IEEE International Conference on Soft-Computing and Network Security*, *7*, 614–617. https://doi.org/10.1109/ICSNS.2018.8573655

Sisodia, D. S., Nair, S., & Khobragade, P. (2017). Diabetic retinal fundus images: Preprocessing and feature extraction for early detection of Diabetic Retinopathy. *Biomedical and Pharmacology Journal*, *10*(2), 615–626. https://doi.org/10.13005/bpj/1148

Visa, S., Ramsay, B., Ralescu, A., & Van Der Knaap, E. (2011). Confusion matrix-based feature selection. *CEUR Workshop Proceedings*, *710*, 120–127.

Vogt, M. (2019). *An Overview of Machine Learning and its Applications*. (January), 178–202. https://doi.org/10.1007/978-3-658-23751-6_17

Wang, W., & Lo, A. C. Y. (2018). Diabetic retinopathy: Pathophysiology and treatments. *International Journal of Molecular Sciences*, *19*(6). https://doi.org/10.3390/ijms19061816

Xu, K., Feng, D., & Mi, H. (2017). Deep convolutional neural network-based early automated detection of diabetic retinopathy using fundus image. *Molecules*, *22*(12). https://doi.org/10.3390/molecules22122054