Multi Label Classification Of Retinal Disease On Fundus Images Using AlexNet And VGG16 Architectures

Reyhansyah Prawira
Department of Mathematics
Universitas Indonesia
Depok, Indonesia
reyhansyah.prawira@sci.ui.ac.id

Alhadi Bustamam
Department of Mathematics
Universitas Indonesia
Depok, Indonesia
alhadi@sci.ui.ac.id

Prasnurzaki Anki
Department of Mathematics
Universitas Indonesia
Depok, Indonesia
prasnurzaki.anki@sci.ui.ac.id

2021 4th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI) | 978-1-6654-0151-7/21/$31.00 ©2021 IEEE | DOI: 10.1109/ISRITI54043.2021.9702817

Abstract— **Diseases of the eye have the potential to cause blindness in sufferers. There have been many types of diseases that exist in the human eye. Some examples of diseases that exist in the eye include Diabetic Retinopathy (DR), Myopia (MA), Optic Disc Cupping (ODC). Fundus images help medical personnel to see what diseases are in the eyes of people with certain diseases. In one fundus image there may be more than one disease in the eye. The research that will be carried out is to find out what diseases are contained in the fundus image by using multi-label classification. The research will be conducted using a deep learning method using the AlexNet and VGG16 architectures which will then be compared between the two models. The data used are fundus images on DR, MA, and ODC diseases as many as 1133 data. The results obtained in this study indicate that the AlexNet model is better than the VGG16 model in performing multi-label classification on fundus images.**

Keywords—deep learning, fundus image, AlexNet, VGG16

I. INTRODUCTION

Diseases of the fundus of the eye greatly affect the function of vision in humans and have the potential to cause blindness. Some common diseases that can affect vision function include Diabetic Retinopathy (DR), Myopia (MA), Optic Disc Cupping (ODC), and several other diseases. To provide effectiveness in treating disease or vision loss due to retinal disease, screening procedures are needed to determine the severity of eye disease. Various techniques have also been applied by experts or doctors to diagnose diseases of the eye, one of which is optical coherence tomography, where the process captures cross-sectional images and fundus photography of the eye [1].

Diabetic Retinopathy (DR) is one of the most common eye diseases found in diabetics [2]. Diabetes Mellitus is a public health problem that is projected to increase to 700 million sufferers by 2045 [3] where Diabetes Mellitus can be characterized by chronic disorders in glucose metabolism [4]. Myopia (MA) or better known as nearsightedness is an increasingly worrying public health problem [5] and it is predicted that by 2050 worldwide, the number of myopia sufferers will increase to 5 billion sufferers [6]. Optic Disc Cupping (ODC) or commonly referred to as glaucoma, is a progressive eye disease that can cause partial or complete blindness due to elevated Intraocular Pressure (IOP) that damages the nervous system [7]. According to WHO, glaucoma is the second dominant cause of vision loss. Until now there is no drug that can prevent the disease, so the detection of glaucoma must be done early before showing alarming symptoms [8].

Retinal fundus images also can be generated using scanner lase ophthalmology (SLO) and optical coherence tomography (OCT), as has been done by Ratheesh K. Meleppat (2021), namely by characterizing lipofuscin and melanolipofuscin granules using optical coherence tomography (OCT) [9]. OCT can also be used as a tool to determine the anatomy and pathology of the retina , and how he way OCT works on the retina is by extracting information from the intensity of light from the retina, so that the amount of reflection generated from the retinal layer can provide information about changes in retinal tissue [10]. Ratheesh K. Meleppat et al (2020), conducted a study by investigating the main organ of the retinal pigment epithelium (RPE) in mice using multicolor confocal fluorescence microscopy (MCFM) so as to produce visualization of the mosaic of RPE cells and the different characteristics of each mouse [11].

Retinal fundus images are also used to detect other diseases as was done by Ratheesh K. Meleppat et al. (2019) i.e. evaluating and comparing performance, different multi-scale Hessian screening was performed on OCTA images on mouse retina [12]. Chung, S. H et al (2020) also conducted a study on retinal disease, namely age-related macular degeneration (AMD) disease by evaluating the relative contribution of AAV delivery efficiency and genome editing level of retinal objects in mice [13].

The widespread use of artificial intelligence technology has created opportunities to develop deep learning applications to help detect diseases of the retina. In recent times, deep learning methods have been widely applied in diagnosing eye diseases based on fundus images [14]. Deep learning is widely used to classify diseases on retinal fundus images. Many researchers have devised different techniques for classifying diseases present on retinal images [15]. Deep learning models can provide better performance compared to traditional methods where the feature design process is still done manually, whereas deep learning models can optimize features automatically and thoroughly.

Prior to the advent of deep learning models, computer-aided diagnosis (CAD) was applied to the diagnosis of fundal diseases [16]. In 2004, Zhang et al. Conducted research using machine learning methods to detect and classify exudate and cotton spots on color fundus images [17]. Prasnurzaki et al conducted research on disease classification in the medical field by comparing several machine learning models and producing an accuracy of 97% using decision tree model [18]. The use of deep learning in assisting in the diagnosis of retinal disease has prompted researchers to undertake other studies in recent years [19]. In a study on classification using a deep learning model in diabetic retinopathy, Alhadi Bustamam et al succeeded in modeling with an accuracy of 97% using the ResNet-50 architecture [20].

In this study, we performed a multi-label classification to find out what types of disease from the three diseases DR, MYA, and ODC on fundus images. Deep learning model used in this research using AlexNet architecture, VGG16, and a combination of the two architectures. In this classification, it will be seen in the fundus image what type of disease is infected on the retina and it is possible that in one fundus there are three of these diseases at once.

II. MATERIALS

In this paper we used the fundus image dataset from website https://riadd.grand-challenge.org/download-all-classes/ by taking image who have Diabetic Retinopathy (DR), Myopia (MA), and Optic Disc Cupping (ODC). The amount of data used is 1133 and splitted into training and testing data with a ratio of 70:30 and 80:20 and the validation data used is by using test data that has been divided from the two comparison ratios of train data and test data. The modeling process is carried out by using the python programming language on google collaboratory.

The steps we took in conducting this research can be seen in the flowchart below:

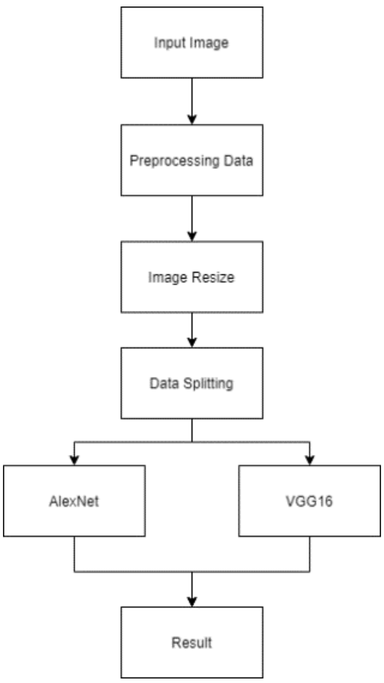

Fig 3. Modelling Flowchart

A. *Preprocessing*

Preprocessing is done by changing the pixel image size on the fundus to 227 x 227 pixels. The image data is compressed to a size of 227 x 227 to facilitate the learning process carried out by the deep learning model [22]. The data taken in the dataset is only fundus image data that has Diabetic Retinopathy (DR), Myopia (MA), and Optic Disc Cupping (ODC) diseases. The data is then split to become train data and test data with a ratio of 70:30 and 80:20.

B. *AlexNet Architecture (2)*

AlexNet is a CNN architecture designed by Krizhevsky et al. using 8 layers, of which 5 are convolutional layers and 3 are fully-connected layers. The architecture developed by Krizhevsky et al. it won test set error rates of 35.7% and 17.0% in the 2010 ImageNet Large Scale Visual Recognition Challenge (ILSVRC) competition [21]. An illustration of the architectural form on Alexnet itself can be seen in the Fig 2 and Table 1 below.

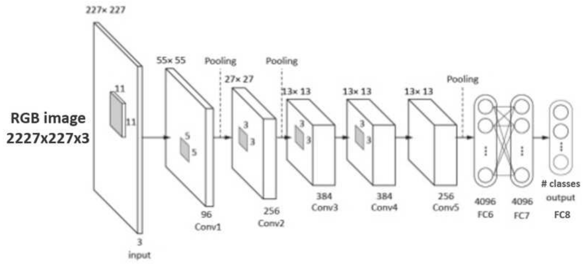

Fig 4. AlexNet Architecture

Table 1. AlexNet Architecture

Layer	Output Shape
Conv-1	55 x 55 x 96
Pooling-1	27 x 27 x 96
Conv-2	27 x 27 x 256
Pooling-2	13 x 13 x 256
Conv-3	13 x 13 x 384
Conv-4	13 x 13 x 384
Conv-5	13 x 13 x 384
Pooling-3	6 x 6 x 256
Fully_Connected-1	1 x 1 x 4096
Fully_Connected-2	1 x 1 x 4096
Fully_Connected-3	1 x 1 x 3

C. *VGG16 Architecture*

The VGG16 architecture developed by K. Simonyan managed to achieve an accuracy of 92.7% using the ImageNet dataset in the ILSVRC competition in 2014. This architecture is a development of the AlexNet architecture by replacing large kernel-sized filters using a 3 x 3 kernel size [23]. An overview of the VGG16 architecture can be seen in Fig 3 and Table 2.

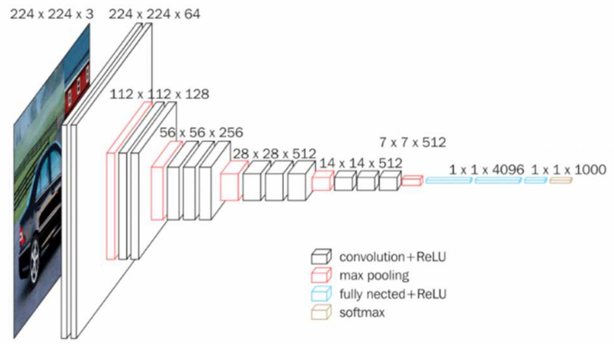

Fig 5. VGG16 Architecture

Table 2. VGG16 Architecture

Layer	Output Shape
Conv-1	227 x 227 x 64
Conv-2	227 x 227 x 64
Pooling-1	113 x 113 x 64
Conv-3	113 x 113 x 128
Conv-4	113 x 113 x 128
Pooling-2	56 x 56 x 128
Conv-5	56 x 56 x 256
Conv-6	56 x 56 x 256
Conv-7	56 x 56 x 256
Pooling-3	28 x 28 x 256
Conv-8	28 x 28 x 512
Conv-9	28 x 28 x 512
Conv-10	28 x 28 x 512
Pooling-4	14 x 14 x 512
Conv-11	14 x 14 x 512
Conv-12	14 x 14 x 512
Conv-13	14 x 14 x 512
Pooling-5	13 x 13 x 512
Fully_Connected-1	1 x 1 x 256
Fully_Connected-2	1 x 1 x 128
Fully_Connected-3	1 x 1 x 3

D. Adam Optimizer

Diederik P. Kingma proposed an efficient stochastic optimization method requiring only first-order gradients with little memory called Adam [24]. The optimizer is designed to suit non-stationary targets and problems with very noisy gradients. The weight updates are performed as:

$$w_t = w_{t-1} - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon}$$

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t}$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t}$$

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2$$

E. Output Program Analysis

Confusion matrix can be used to solve multiclass classification problems as well as binary classification whose use is very popular in machine learning or deep learning research [25]. A confusion matrix is a performance measure for analyzing programs in machine learning or deep learning problems where the output can be of more than one class. There are four terms that represent the results of the classification process in the confusion matrix, namely True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN).

Table 3. Confusion Matrix

Actual Value	**Recognize Value**	
	Positive	**Negative**
Positive	TP	FN
Negative	FP	TN

From Table 3 it can build the formula for accuracy as seen below:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} * 100\%$$

III. RESULT

In this study, 1133 data were classified on fundus image disease, where the diseases affected were diabetic retinopathy (DR), myopia (MA), and Optic Disc Cupping (ODC). The research was conducted using two architectural models, namely AlexNet and VGG16. The fundus image used in our research is resized to 227 x 227 pixels as shown in Fig 6 before training.

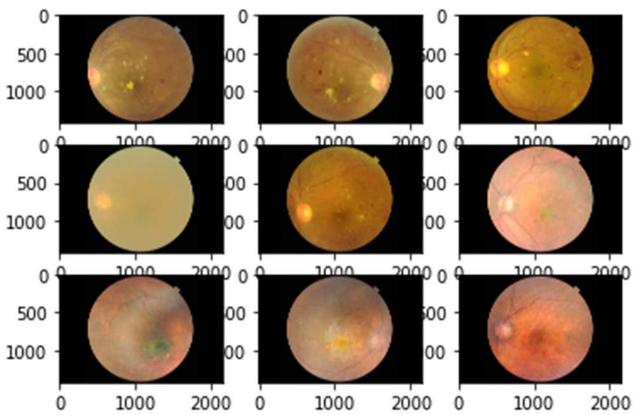

Fig 6. Resized Data Image

The data used is split into training data and testing data with a ratio of 70:30 in the first experiment, then using a ratio of 80:20 in the second experiment. In the AlexNet model, running the model uses 200 epochs, while the VGG16 model uses 100 epochs. The flow of the model that we did in this research can be seen in Fig 7 below where we get 4 results from each model that we run.

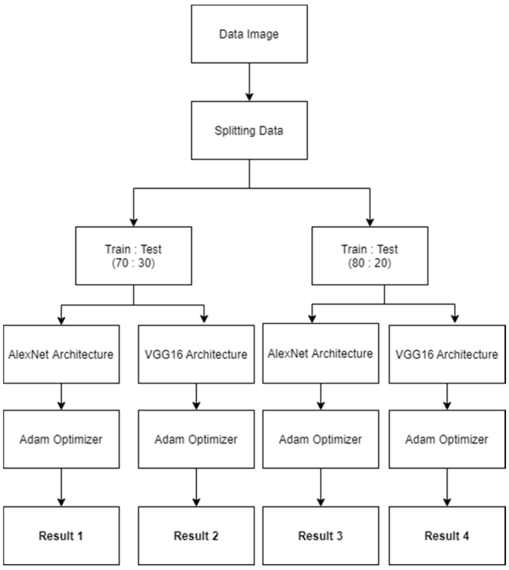

Fig 7. Modelling Flow

The results of the model seen from the following flowchart are presented in table 4 below:

Table 4. Result

Architecture	Optimizer	Train:Test	Accuracy
AlexNet	Adam	70:30	95%
		80:20	94%
VGG16		70:30	89%
		80:20	88%

From the following experiment, the difference in the ratio of the training data and test data as well as the difference in the optimizer does not show significant results. It is hoped that the development of the variation of parameters tested in the program in the future is expected to be more varied and varied according to the references in the research reference, the more program trials with various parameter variations the possibility of getting better program output results [26]. In the results provided by the AlexNet architecture, better results were obtained compared to the VGG16 architecture. A graph that displays the results of accuracy and loss performed on the two models can be seen in the following figure:

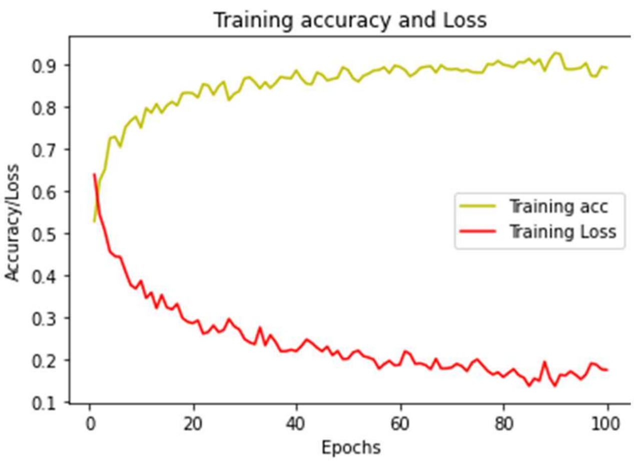

Fig 8. Training Accuracy and Loss AlexNet

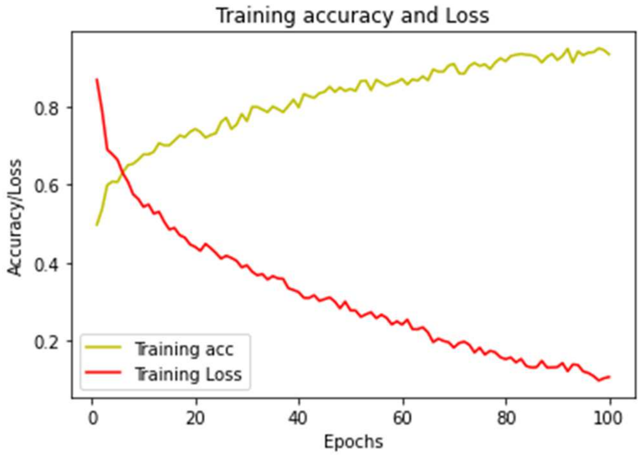

Fig 9. Training Accuracy and Loss VGG16

IV. CONCLUSION

There are a lot of diseases in the retina of the eye, and the detection of these diseases can be seen on the fundus image. In one fundus image there may be several diseases in it. In this era, machine learning and deep learning technology can be very helpful in detecting what disease is in the fundus image. In this study, we carried out a deep learning model using AlexNet and VGG16 architecture to determine what diseases are contained in a single fundus image by limiting only Diabetic Retinopathy, Myopia, and Optical Disk Cupping diseases.

The results given in this study indicate accuracy. Based on the various trials that have been carried out in this study, it can be concluded that the trial using the AlexNet model architecture with Adam optimizer at a train:test ratio of 70:30 which produces an acuuracy of 95%, is the best result obtained in this study.

ACKNOWLEDGEMENTS

This research was supported by the Penelitian Tahun Jamak Penelitian Terapan Bidang Kesehatan Kementerian Pendidikan, Kebudayaan, Riset dan Teknologi research grant from the University of Indonesia with contract number NKB-705/UN2.RST/HKP.05.00/2021. The authors deliver a huge appreciation to colleagues from the Kemendikbudristek and Data Science Centre Department at the Faculty of Mathematics and Natural Sciences who advanced expertise and insights to cultivate this research in numerous ways.

REFERENCES

[1] Chea, N., & Nam, Y. (2021). Classification of fundus images based on deep learning for detecting eye diseases. Computers, Materials, & Continua, 67(1), 411-426. doi:http://dx.doi.org/10.32604/cmc.2021.013390

[2] Alicia J. Jenkins, Mugdha V. Joglekar, Anandwardhan A. Hardikar, Anthony C. Keech, David N. O'Neal, S. Andrzej, Januszewski, Biomarkers in diabetic retinopathy, Rev. Diabet. Stud.: Reg. Dev. Stud. 12 (1–2) (2015) 159.

[3] International Diabetes Federation. International diabetes federation diabetes atlas, ninth ed.https://www.diabetesatlas.org/en/.

[4] Anqi Shan, Xi Chen, Xueli Yang, Baoqun Yao, Fengchao Liang, Ze Yang, Fangchao Liu, Song Chen, Xiaochang Yan, Jianfeng Huang, Shaoye Bo, Nai-Jun Tang, Dongfeng Gu, Hua Yan,Association between long-term exposure to fine particulate matter and diabetic retinopathy among diabetic patients: A national cross-sectional study in China,Environment International,Volume 154,2021,106568,ISSN 0160-4120, https://doi.org/10.1016/j.envint.2021.106568.

[5] Dolgin E. The myopia boom. Nature 2015;519(7543):276.

[6] Holden BA, Fricke TR, Wilson DA, et al. Global Prevalence of Myopia and High Myopia and Temporal Trends from 2000 through 2050. Ophthalmology 2016;123(5):1036–42.

[7] R. Lim, I. Goldberg, Glaucoma in the twenty-first century, in: The Glaucoma Book, Springer, 2010, pp. 3–21.

[8] C.G. de Moraes, J.M. Liebmann, F.A. Medeiros, R.N. Weinreb, Management of advanced glaucoma: characterization and monitoring, Survey Ophthalmol. 61 (5)(2016) 597–615.

[9] Meleppat, R. K., Ronning, K. E., Karlen, S. J., Burns, M. E., Pugh, E. N., & Zawadzki, R. J. (2021). In vivo multimodal retinal imaging of disease-related pigmentary changes in retinal pigment epithelium. Scientific Reports, 11(1). doi:10.1038/s41598-021-95320-z

[10] Meleppat RK, Zhang P, Ju MJ, Manna SK, Jian Y, Pugh EN, Zawadzki RJ. Directional optical coherence tomography reveals melanin concentration-dependent scattering properties of retinal pigment epithelium. J Biomed Opt. 2019 Jun;24(6):1-10. doi: 10.1117/1.JBO.24.6.066011

[11] Meleppat, R. K., Ronning, K. E., Karlen, S. J., Kothandath, K. K., Burns, M. E., Pugh, E. N., & Zawadzki, R. J. (2020). In Situ Morphologic and Spectral Characterization of Retinal Pigment Epithelium Organelles in Mice Using Multicolor Confocal Fluorescence Imaging. Investigative Opthalmology & Visual Science, 61(13), 1. doi:10.1167/iovs.61.13.1

[12] Ratheesh K. Meleppat, Eric B. Miller, Suman K. Manna, Pengfei Zhang, Edward N. Pugh Jr., and Robert J. Zawadzki "Multiscale Hessian filtering for enhancement of OCT angiography images", Proc. SPIE 10858, Ophthalmic Technologies XXIX, 108581K (4 April 2019); https://doi.org/10.1117/12.2511044.

[13] Chung, S. H., Mollhoff, I. N., Nguyen, U., Nguyen, A., Stucka, N., Tieu, E., … Yiu, G. (2020). Factors Impacting Efficacy of AAV-mediated CRISPR-based Genome Editing for Treatment of Choroidal Neovascularization. Molecular Therapy - Methods & Clinical Development. doi:10.1016/j.omtm.2020.01.006.

[14] Li, T., Bo, W., Hu, C., Kang, H., Liu, H., Wang, K., & Fu, H. (2021). Applications of deep learning in fundus images: A review. Medical Image Analysis, 69, 101971. doi:10.1016/j.media.2021.101971

[15] Mateen, M., Nasrullah, Sun, S., & Huang, Z. (2019). Fundus image classification using VGG-19 architecture with PCA and SVD. Symmetry, 11(1). doi:http://dx.doi.org/10.3390/sym11010001

[16] J. Wang, L. Yang, Z. Huo, W. He and J. Luo, "Multi-Label Classification of Fundus Images With EfficientNet," in IEEE Access, vol. 8, pp. 212499-212508, 2020, doi: 10.1109/ACCESS.2020.3040275.

[17] Zhang XH, Chutatape O, Ieee (2004) Detection and classification of bright lesions in color fundus images. Icip: 2004 international conference on image processing 1–5: 139–142

[18] Prasnurzaki Anki, Alhadi Bustamam, Rinaldi Anwar Buyung, Looking for the link between the causes of the COVID-19 disease using the multi-model application, Commun. Math. Biol. Neurosci., 2021 (2021), Article ID 75

[19] Pan X, Jin K, Cao J, Liu Z, Wu J, You K, Lu Y, Xu Y, Su Z, Jiang J, Yao K, Ye J. Multi-label classification of retinal lesions in diabetic retinopathy for automatic analysis of fundus fluorescein angiography based on deep learning. Graefes Arch Clin Exp Ophthalmol. 2020 Apr;258(4):779-785. doi: 10.1007/s00417-019-04575-w. Epub 2020 Jan 14. PMID: 31932886.

[20] Alhadi Bustamam, Devvi Sarwinda, Radifa H. Paradisa, Andi Arus Victor, Anggun Rama Yudantha, Titin Siswantining, Evaluation of convolutional neural network variants for diagnosis of diabetic retinopathy, Commun. Math. Biol. Neurosci., 2021 (2021), Article ID 42.

[21] Krizhevsky, Alex; Sutskever, Ilya; Hinton, Geoffrey E. (2017-05-24). "ImageNet classification with deep convolutional neural networks" (PDF). Communications of the ACM. 60 (6): 84–90. doi:10.1145/3065386. ISSN 0001-0782. S2CID 195908774

[22] Lakhani, P. (2020). The Importance of Image Resolution in Building Deep Learning Models for Medical Imaging. Radiology: Artificial Intelligence, 2(1), e190177. doi:10.1148/ryai.2019190177.

[23] Simonyan, Karen & Zisserman, Andrew. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv 1409.1556.

[24] Kingma, Diederik & Ba, Jimmy. (2014). Adam: A Method for Stochastic Optimization. International Conference on Learning Representations.

[25] Kulkarni, A., Chong, D., & Batarseh, F. A. (2020). Foundations of data imbalance and solutions for a data democracy. Data Democracy, 83–106. doi:10.1016/b978-0-12-818366-3.00005-8

[26] Prasnurzaki Anki, Alhadi Bustamam, Rinaldi Anwar Buyung, "Comparative Analysis of Performance between Multimodal Implementation of Chatbot Based on News Classification Data Using Categories," Electronics, vol. 10, no. 21, p. 2696, Nov. 2021.