

Experimenting with the Hyperparameter of Six Models for Glaucoma Classification

Muhamad Ilham, Angga Prihantoro, Iqbal Kurniawan Perdana, Rita Magdalena, Sofia Saidah
Faculty of Electrical, Telkom University, Bandung, Indonesia

ARTICLE INFO

Article history:

Received July 03, 2023
Revised July 06, 2023
Published July 13, 2023

Keywords:

Glaucoma;
Early Detection;
CNN;
RIM-ONE DL;
Experiment Hyperparameter

ABSTRACT

Glaucoma is characterized by optic nerve damage, and can potentially lead to blindness, often presenting with no obvious symptoms in most affected individuals. As a result, a large proportion of those affected remain undiagnosed, making early detection crucial for effective treatment. Numerous studies have been conducted to develop glaucoma detection systems. In this particular study, a glaucoma detection system using the CNN method was developed. The contribution of this research is to conduct hyperparameter experiments on AlexNet, Custom Layer, MobileNetV2, EfficientNetV1, InceptionV3, and VGG19 models on the RIM-ONE DL dataset with a total of 933 images that have been augmented. Hyperparameter experiments were conducted to determine the most optimal parameters for each model, specifically testing batch size, learning rate, and optimizer. The batch sizes used were 64, 128, 256, and 512. The learning rates used are 0.1, 0.001, 0.0001, and 0.00001. The optimizers to be tested are Nadam, Adam, and RMSProp. The hyperparameter optimization process yielded the optimal parameters for each model. However, it is important to note that the MobileNetV2, InceptionV1, and VGG19 models exhibited signs of overfitting in the training graph results. Among the models, the custom layer model achieved the highest accuracy of 93%, while InceptionV3 attained the lowest accuracy at 83.5%. Model testing was conducted using data from the Cicendo Eye Hospital, and the RIM-ONE DL testing dataset which totals 200 images. Based on the testing results, it was found that InceptionV3 outperformed the other models in predicting images accurately. Therefore, the study concluded that high accuracy in training does not necessarily indicate superior performance in testing, particularly when limited variation exists in the training dataset.

This work is licensed under a [Creative Commons Attribution-Share Alike 4.0](https://creativecommons.org/licenses/by-sa/4.0/)



Corresponding Author:

Muhamad Ilham, Telecommunication Engineering, Faculty of Electrical Engineering, Telkom University, Indonesia
Email: mhmdilhm@student.telkomuniversity.ac.id

1. INTRODUCTION

Glaucoma is characterized by optic nerve damage, and the potential for blindness, often presents without apparent symptoms in the majority of affected individuals [1][2][3]. According to a study cited from Worldview [4], glaucoma ranks as the second leading cause of blindness worldwide among all eye diseases. The study reveals that primary glaucoma has rendered over 11.1 million individuals bilaterally blind. Most of the studies that have been conducted to evaluate the diagnostic ability of OCT to detect glaucoma have involved patients with existing visual field damage, and few have focused on using this imaging to detect the disease at an early stage. Glaucoma is caused by an increase in intraocular pressure (IOP) at the front of the eye [5][6]. There are two types of glaucoma categorized based on anatomical features: open-angle glaucoma, and angle-closure glaucoma [5][7].

Primary open-angle glaucoma is diagnosed by ruling out other potential ocular or systemic causes, as there are no evident factors directly leading to its development. The condition typically develops gradually, and often goes unnoticed during its initial phases, as there are usually no obvious symptoms while Angle-closure glaucoma presents with noticeable, and severe signs, and symptoms that, if left untreated, can result in permanent blindness within a short period. This condition occurs when the aqueous humor is unable to pass through the pupil into the front chamber of the eye, leading to increased pressure behind the iris. As a result, the iris is forced forward (referred to as iris bombé), obstructing the angle of the front chamber. Acute angle-closure glaucoma can occur when the pupil dilates or when the lens moves forward due to certain triggers [7].

One of the ways to treat glaucoma is to perform early detection to catch the early signs of optic nerve damage. Since glaucoma is the second highest cause of blindness in the world, early detection is crucial [8]. One solution for early detection is to design a glaucoma detection system.

Several studies have been conducted on the early detection of glaucoma. In a research study by C. P. Bragança et al. [9], they investigated the early detection of glaucoma using smartphones integrated with ophthalmoscopes, utilizing the CNN algorithm. Pre-trained models such as Densenet, Mobilenet, InceptionV3, InceptionResnet, Resnet50v2, Resnet101, and Xception were used. The Brazil Glaucoma (BrG) public dataset was divided into 70% training, 30% testing, and 20% of training data for validation. The accuracy of each model was found to be above 80%. In another study by Xiangyu Chen, Yanwu Xu et al. [10], glaucoma detection was performed using a deep convolutional neural network (CNN). The architecture of the network consisted of four convolutional layers, and two fully-connected layers. The ORIGA, and SCES datasets were used for training, and testing the model. Augmentation, dropout, response-normalization layers [11][12], and overlapping-pooling layers were employed to improve the model's performance, and reduce overfitting. The AUC results obtained were 0.831 for the ORIGA dataset, and 0.887 for the SCES dataset. Research conducted by Amer Sallam, Abduguddoos S, et al. [13] focused on the early detection of glaucoma using transfer learning from pre-trained CNN models. Fundus images were used for glaucoma detection, and various pre-trained models were employed. The Large-scale Attention-based Glaucoma (LAG) dataset was utilized for training, and evaluation. The ResNet-152 model exhibited the highest performance with an accuracy rate, precision, and recall of 86.9%. In 2019, Mohammad Norouzifard et al. [14] conducted research to diagnose glaucoma using a comparative method of transfer learning, deep learning, and multilayer neural network techniques. Various classifiers, including MNN, CNN, VGG, InceptionResNet, Xception, and proposed NASnet-based models, were trained using the RIM-ONE V2 dataset. The proposed NASnet-based model achieved the highest test accuracy among these classifiers, which was 90%.

Based on the aforementioned issues, this study aimed to design a glaucoma detection system using the CNN method, and employing six models, namely AlexNet, Custom Layer, MobileNetV2, EfficientNetV1, InceptionV3, and VGG19. The reason why we choose CNN for glaucoma classification is because CNN is capable of processing complex images, and the ability to transfer learning which can improve glaucoma classification performance. In order to achieve optimal results in the study, hyperparameter experiments were conducted for each model. The hyperparameters tested included batch size, learning rate, and optimizer. The batch sizes used in parameter testing are 64, 128, 256, and 512. Learning rates in this parameter test are 0.1, 0.001, 0.0001, and 0.00001. The optimizers that will be tested are Nesterov-accelerated Adaptive Moment Estimation (Nadam), Adaptive Network Estimation (Adam), and Root-Mean-Square Propagation (RMSProp). Propagation (RMSProp). These experiments aimed to identify the best combination of hyperparameters that would maximize the performance of each model. By fine-tuning these parameters, the researchers sought to optimize the performance, and accuracy of the glaucoma classification model in detecting a total of 200 combined testing data from Cicendo Eye Hospital, and data from RIM-ONE DL.

2. METHODS

This research aims to design a system to classify glaucoma disease using the method in Fig. 1. This method starts by inputting the public dataset, namely RIM-ONE DL [15]. After that preprocessing, at this stage the data is resized, and augmented such as random rotation, zooming, and flips. Furthermore, the hyperparameter experiment process by training by combining each parameter such as batch size, learning rate, and optimizer. The purpose of hyperparameter experimentation is to identify the most optimal parameters to achieve accurate classification. The CNN models used in this research include AlexNet, Custom Layer, MobileNetV2, EfficientNetV1, InceptionV3, and VGG19.

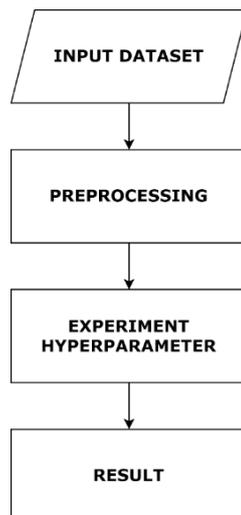


Fig. 1. Proposed method for glaucoma classification

2.1. Dataset

The training data consists of 313 samples categorized into two classes: glaucoma, and normal. This research uses the RIM-ONE DL dataset [15], which is a new version of rim one v1, v2, and v3 by eliminating duplicate images in each of the previous rim one versions. RIM-ONE DL also implements all images to be cropped right around the optic nerve head using the same proportionality criteria, something that was not done in previous versions. The training data consisted of 313 samples categorized into two classes: glaucoma, and normal. Data was collected using different cameras such as Nidek AFC-210 non-mydratic fundus camera with 21.1-megapixel Canon EOS 5D mark II, and non-mydratic Kowa WX 3D stereo. The dataset is taken from various hospitals in Spain such as Hospital Universitario de Canarias (HUC), in Tenerife, Hospital Universitario Miguel Servet (HUMS), in Zaragoza, and Hospital Cl'nico Universitario San Carlos (HCSC), in Madrid. The sample dataset used in this study can be seen in Fig. 2.

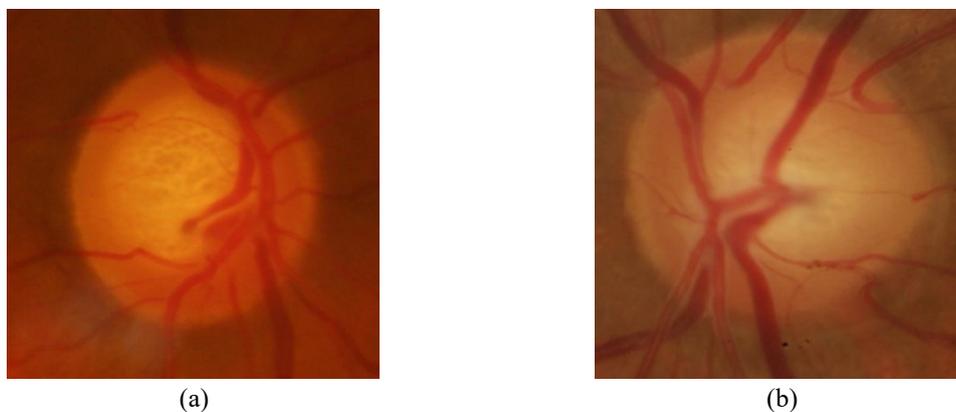


Fig. 2. Sample dataset image of RIM-ONE DL - (a) glaucoma fundus image, (b) normal fundus image

2.2. Preprocessing

Before the training process, the data is resized to 224x224, and converted to RGB format. To prevent overfitting, data augmentation techniques are employed based on the research conducted by A. Diaz-Pinto et al. [16]. The augmentation includes random rotation (-30 to 30 degrees), zooming within a range of 0 to 0.2, as well as horizontal, and vertical flips. As a result of augmentation, the dataset size expands to 933 samples. Then the data is divided into training 80% and testing 20%.

2.3. Convolutional Neural Network

A Convolutional Neural Network (CNN) is an artificial neural network specifically designed to handle structured data, such as images, and audio signals. CNNs have found extensive application in tasks like image classification, object detection, and image segmentation [17]. One of the pioneers of CNN is Yann Lecun, who invented the architecture called LeNet-5 in 1998. LeNet-5 was designed with an input size of 32x32 pixels,

and comprised 5 CNN layers. Its original purpose was to recognize digits in postal codes [18]. In the development of CNN models, LeNet-5 served as the foundation for other architectures. For example, AlexNet adopted convolution layers, pooling layers, and ReLU activation from LeNet-5 [19]. VGGNet also incorporated a convolutional layer concept similar to LeNet-5 [20]. The design of the InceptionNet feature extraction module, which utilizes different spatial scales, was inspired by the feature extraction approach of LeNet-5 [21].

2.3.1. AlexNet

In 2012, Alex Krizhevsky, and colleagues introduced AlexNet, a Convolutional Neural Network (CNN) model that surpassed the LeNet model in terms of depth, and breadth. AlexNet achieved significant success by winning the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) competition [22][23]. This architecture consists of five convolutional layers, three fully connected layers, and three MaxPooling layers [24][25].

Fig. 3 displays the proposed architecture of the AlexNet model used for glaucoma classification. The proposed Alexnet architecture contains eight learned layers: five convolutional layers with activation relays, and three fully connected layers. The outputs of the three fully connected layers are sigmoid classifiers. The first convolutional layer consists of 96 filters, each with a kernel size of 3, and a stride of 4 pixels. The second, and fifth convolutional layers utilize 256 filters, 3 kernel sizes, and a stride of 1 pixel. Similarly, the third, and fourth convolutional layers employ 384 filters, 3 kernel sizes, and a stride of 1 pixel. Moving on to the fully connected layers, the first, and second layers contain 4096 units each, while the third fully connected layer consists of 1000 units.

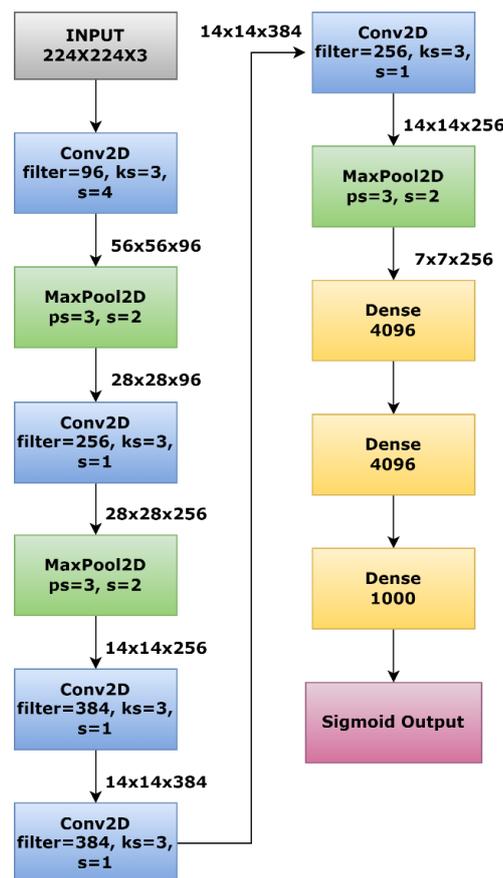


Fig. 3. Alexnet architecture

2.3.2. Custom Layer

Fig. 4 illustrates the proposed architecture of the custom layer model for glaucoma classification research. This architecture consists of seven learning layers: five convolutional layers, and two fully connected layers. The outputs of the two fully connected layers are sigmoid classifiers. The convolutional layers use filters 16, 32, 64, 128, 128, a kernel size with size 5, and ReLU as an activation function. The MaxPooling2D layers use

a pool size of 2, the fully connected layers have 512 units, and a dropout layer with a rate of 0.2 is added based on research [11] to avoid overfitting.

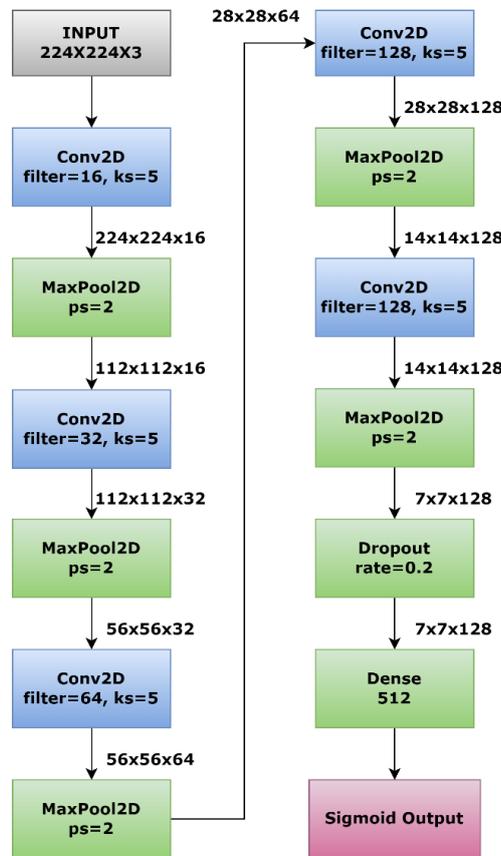


Fig. 4. Custom layer architecture

2.3.3. MobileNetV2

MobileNetV2 incorporates the Depthwise Separable Convolutions technique, which combines Depthwise convolution, and Pointwise convolution. This technique helps reduce the number of parameters, and computational cost significantly, by approximately 18 times compared to regular convolutions. The primary objective of this approach is to enhance portability. To address the issue of information loss in the non-linear layer of the convolution block, MobileNetV2 introduces Linear Bottlenecks. Furthermore, MobileNetV2 introduces a novel structure called Inverted residuals, which aims to maintain information integrity throughout the network [26].

Table 1 provides an overview of the layer order in the MobileNetV2 architecture. The architecture begins with an initial fully convolutional layer consisting of 32 filters. Following that, there are a total of 19 residual bottleneck layers present in the MobileNetV2 architecture [27]. Table 2 show that MobileNetV2 uses ReLU6 activation which can embed the input space into low dimensions [26].

Table 1. MobileNetV2 architecture

Input	Layer	t	c	n	s
$224 \times 224 \times 3$	Conv2D	-	32	1	2
$112 \times 112 \times 32$	Bottleneck	1	16	1	1
$112 \times 112 \times 16$	Bottleneck	6	24	2	2
$56 \times 56 \times 24$	Bottleneck	6	32	3	2
$28 \times 28 \times 32$	Bottleneck	6	64	4	2
$14 \times 14 \times 64$	Bottleneck	6	96	3	1
$14 \times 14 \times 96$	Bottleneck	6	160	3	2
$7 \times 7 \times 160$	Bottleneck	6	320	1	1
$7 \times 7 \times 320$	Conv2D 1x1	-	1280	1	1
$7 \times 7 \times 1280$	avgPool 7x7	-	-	1	-
$1 \times 1 \times 1280$	Conv2D 1x1	-	k	-	-

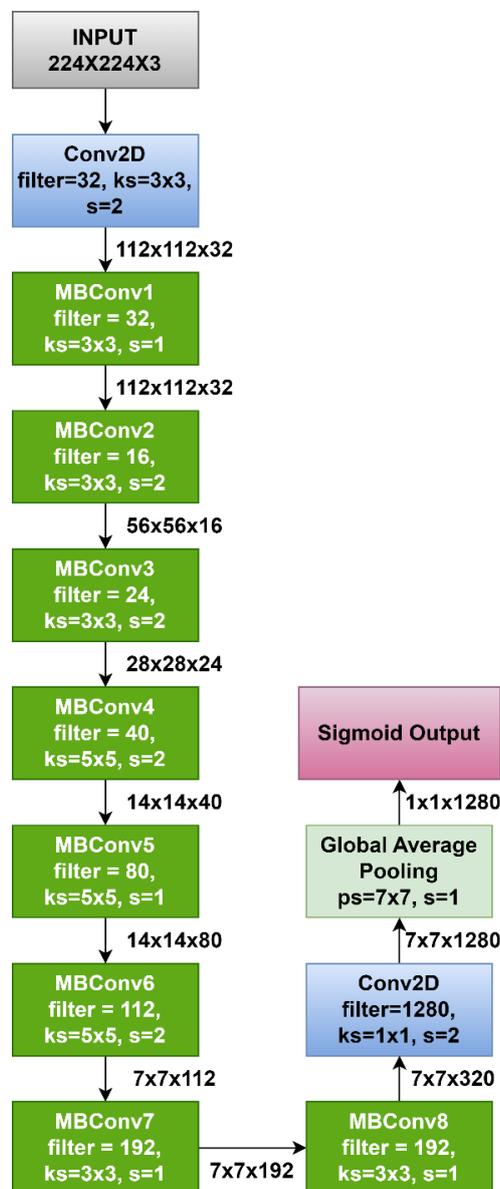
Table 2. Bottleneck residual block transforming from k to k' channels, with stride s , and expansion factor t

Input	Operator	Output
$h \times w \times k$	1×1 conv2d, ReLU6	$h \times w \times (tk)$
$h \times w \times tk$	3×3 dwise $s=s$, ReLU6	$\frac{h}{s} \times \frac{w}{s} \times (tk)$
$\frac{h}{s} \times \frac{w}{s} \times tk$	Linear 1×1 conv2d	$\frac{h}{s} \times \frac{w}{s} \times k'$

2.3.4. EfficientNetV1

EfficientNetV1 is a member of the group of convolutional neural network models known for their efficiency, and effectiveness. It incorporates a scaling technique that adjusts the network's depth, width, and resolution using a compound coefficient. This approach aims to strike the right balance between model size, and performance [28].

Fig. 5 displays the architecture of efficientnetv1, this architecture uses 2 convolutional layers with 32, and 1280 filters, 1 pooling layer, and 8 MBConv (Mobile Inverted Residual Bottleneck) layers, each MBConv layer uses 3x3, and 5x5 kernel sizes include 1x1 pointwise convolutional, a depthwise separable convolution layer, and another 1x1 pointwise convolutional layer. The utilization of MBConv helps reduce computational complexity while maintaining model performance [28].

**Fig. 5.** EfficientNetV1 architecture

2.3.5. InceptionV3

InceptionV3 was invented by C.Szegedy in 2015, it is designed to address the challenge of increasing model depth without significantly escalating computational complexity by employing the concept of the "Inception Module" [21][29]. Fig. 6 shows the architecture of inceptionnetv3, this model uses 5 convolutional layers, and 2 pooling layers, and 11 inception modules which are the main components of the inceptionv3 model. Each inception module consists of 4 branches, which contain a total of 6 convolutional layers having various kernel sizes (1x1, 3x3, 5x5), 1 pooling layer, and 1 concatenation layer at the end of the inception module which serves to combine the output of several different branches.

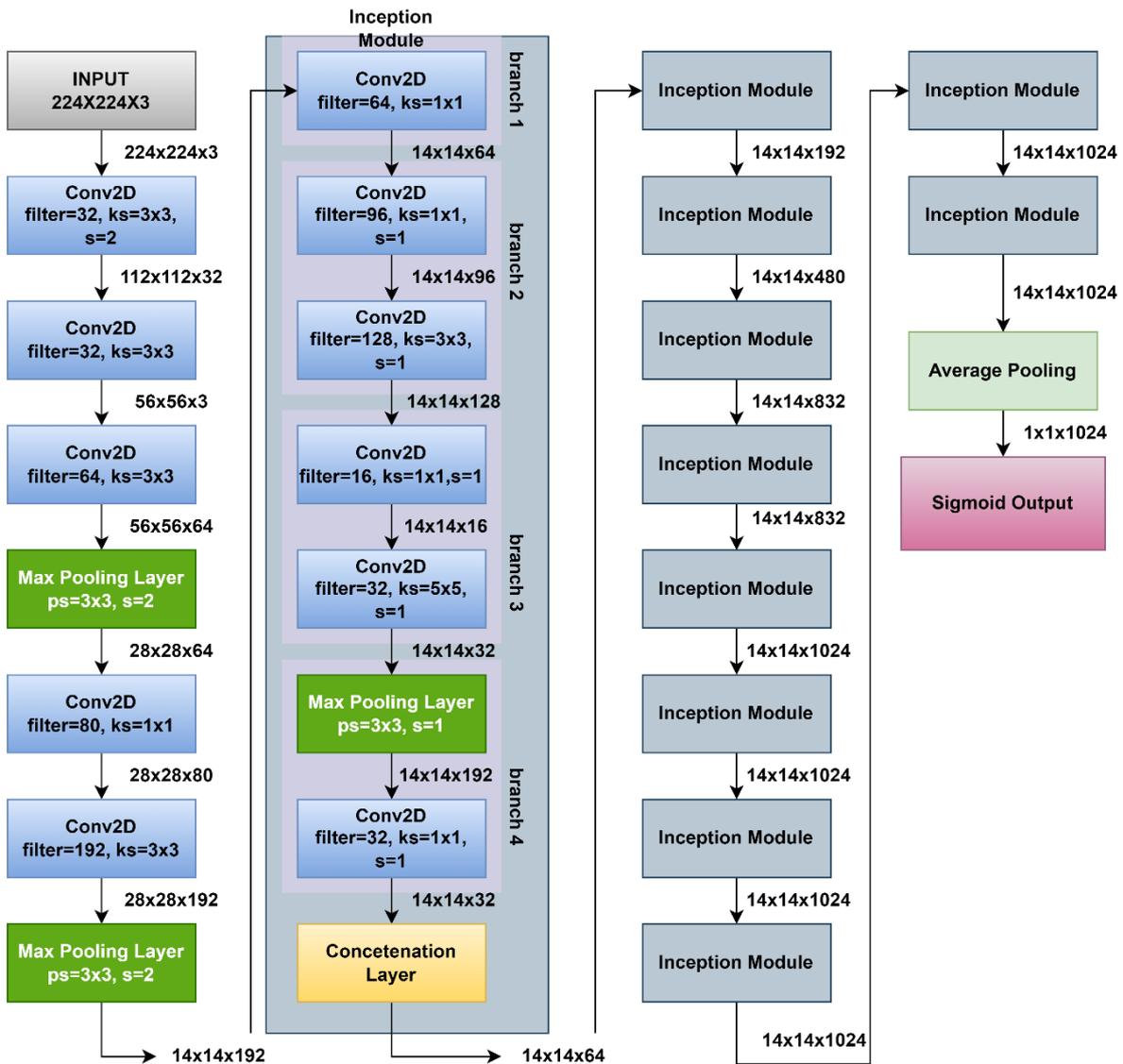


Fig. 6. InceptionV3 architecture

2.3.6. VGG19

VGG19 is a variant of the VGGNet architecture introduced by Karen Simonyan, and Andrew Z in 2014. VGG19 achieves high accuracy by utilizing a deep architecture, however, this model has large number of parameters which demands significant computational resources [20][29]. Fig. 7 illustrates the architecture of the VGG19 model, which consist of a total of 19 layers. This architecture comprises 16 convolutional layers, 2 fully connected layers, and 1 sigmoid activation function. This model utilizes 3x3 convolutional filters to learn local patterns, max pooling layers with a pooling size of 2x2 for downsampling, and a flatten layer followed by 2 dense layers as fully connected layers.

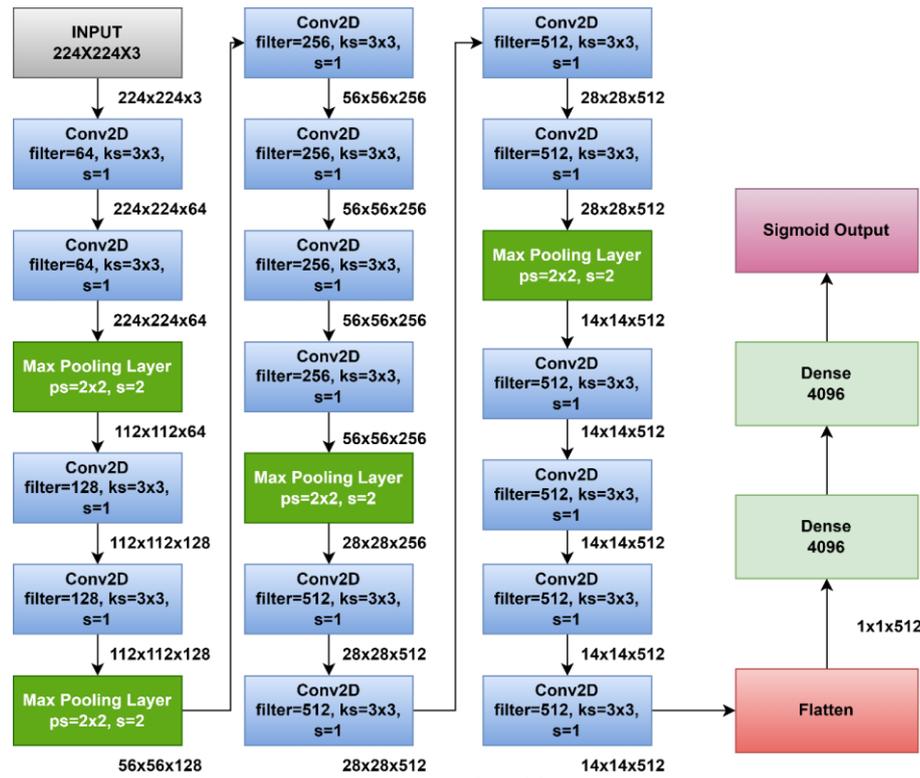


Fig. 7. VGG19 architecture

2.4. Experiment Hyperparameter

In this research, the parameters tested are batch size (64, 128, 256, 512), optimizer (Adam, Nadam, RMSprop), and learning rate (0.01, 0.001, 0.0001, 0.00001) with 50 epochs. Batch size refers to the number of training data samples used in each iteration, high batch size will make the network take a long time to reach convergence (no more accuracy improvement), and if the batch size is too low it will cause the network to bounce around without achieving acceptable performance [30][31][32]. Learning rate is a parameter that controls the speed of the algorithm in learning patterns from a given dataset, using a learning rate that is too low can cause slow convergence, and using a high learning rate causes faster convergence, and results in an unstable loss function [33][34]. Optimizers are fundamental algorithms that allow the machine to learn from its experience, different optimizers will affect the speed of convergence, stability, and generalization [35]. Experiments were conducted by combining each parameter, and comparing the results one by one in the form of classification reports to get the best parameters for each model.

2.5. Classification Report

To assess the effectiveness of a model in accurately identifying images within each class, performance metrics such as accuracy, precision, recall, and f1-score can be utilized. These metric can be computed using the provided formulas in the equations (1) to (4).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP + TN}{TP + FN} \quad (3)$$

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

In the context of model predictions, True Positive (TP) refers to the correct identification of positive data, while True Negative (TN) denotes the accurate prediction of negative data. False Positive (FP) indicates the

incorrect labeling of negative data as positive, whereas False Negative (*FN*) signifies the incorrect labeling of positive data as negative [36][37]. Equation (1) is a parameter to measure the performance of the model in classifying all data correctly. Equation (2) is a parameter to measure the *TP* ratio with all data predicted as positive data. Equation (3) is a parameter to measure the ratio of *TP* to all data that is actually positive data. Equation (4) is a parameter to measure the harmonic mean of the precision, and recall parameters.

3. RESULTS AND DISCUSSION

This research utilizes the RIM-ONE DL dataset [15] which consists of 933 segmented fundus images including 348 glaucoma fundus images, and 585 normal fundus images. The dataset is divided into training data, and validation data, which comprises 746 fundus images, and validation data which comprises 187 fundus images. The input data used for training, and validation is sized at $224 \times 224 \times 3$. Hyperparameter experiments are conducted by combining each parameter. After performing the hyperparameter experiments, the best results of each model will be presented in the form of accuracy charts, classification reports, and confusion matrices. After obtaining the optimal parameters for each model, the model will then be tested using testing data in the form of a combination of RIM-ONE DL data, and glaucoma fundus image data from Cicendo Eye Hospital, Bandung. The testing results will be displayed in the confusion matrix as a reference to find the best model.

3.1. Optimal Hyperparameter

After conducting the experiment with hyperparameters by combining each parameter, the best results from each model were selected. The selection process involved comparing the shape of the graph, and the classification report for each hyperparameter combination. The selected graphs are those that do not show signs of overfitting or are best among other graphs, and have a high accuracy value in the classification report of each model. Fig. 8 displays the best training graph for each model.

Additionally, Table 3 showcases the most optimal hyperparameter for each model. In Fig. 8 parts (c), (d), (f) show signs of overfitting, overfitting is a condition where the model performs the training process well but when the model is tested with validation data shows results that are not as good as the training results which will cause the model to have high accuracy but not have good performance when tested using new data.

Table 3. Optimal hyperparameter for each model

Models	Hyperparameters			Description
	Optimizer	Batch Size	Learning Rate	
AlexNet	Adam	64	0.00001	Good fit
Custom Layer	Adam	256	0.001	Good fit
MobileNetV2	Adam	128	0.01	Overfitting
EfficientNetV1	Adam	256	0.01	Overfitting
InceptionV3	Nadam	64	0.001	Good fit
VGG19	Adam	64	0.01	Overfitting

3.2. Classification Report

Table 4 presents the training results obtained using the most optimal parameters for each model. Based on the results, the model with custom layer achieves the highest accuracy value of 93%, and the model that gets the lowest accuracy is InceptionV3 with 83.5% accuracy.

Table 4. Best classification report of each model

Model	Evaluation Matrix				
	Class	Precision	Recall	F1-score	Accuracy
AlexNet	Glaucoma	89%	85%	87%	90%
	Normal	92%	94%	93%	
Custom Layer	Glaucoma	95%	87%	90%	93%
	Normal	94%	98%	96%	
MobileNetV2	Glaucoma	86%	83%	85%	88%
	Normal	90%	91%	91%	
EfficientNetV1	Glaucoma	93%	76%	84%	87.5%
	Normal	87%	97%	91%	
InceptionV3	Glaucoma	80%	79%	80%	83.5%
	Normal	87%	88%	87%	
VGG19	Glaucoma	85%	86%	86%	88.5%
	Normal	91%	90%	91%	

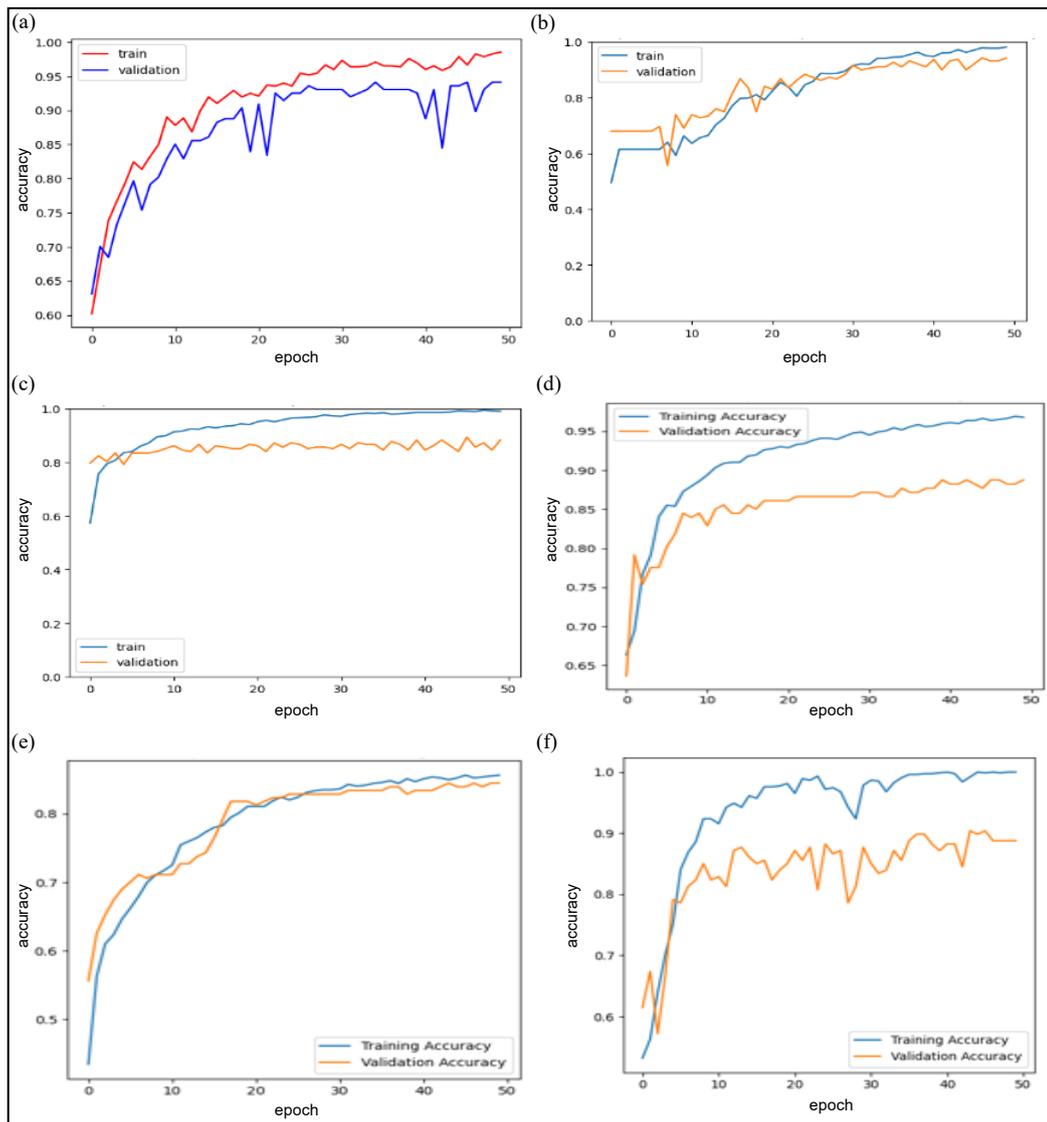


Fig. 8. Training, and validation graphs of the 6 generated models. (a) AlexNet; (b) Custom Layer; (c) MobileNetV2; (d) EfficientNetV1; (e) InceptionV3; (f) VGG19

3.3. Testing

After conducting hyperparameter experimentation, and obtaining the optimal parameters for each model, the next step is model testing. The testing data comprises primary data obtained from the Cicendo Eye Hospital, Bandung, West Java, as well as testing data from RIM-ONE DL [15]. The testing data includes 100 normal fundus images, and 100 glaucoma fundus images. Table 5 displays the distribution of testing data between Cicendo Eye Hospital, and RIM-ONE DL.

Table 5. The distribution of testing data

Class	Dataset	
	Cicendo	RIM-ONE DL
Glaucoma	56	44
Normal	-	100

According to Fig. 9, the models that perform the best in image classification are InceptionV,3 and VGG19. InceptionV3 achieves a successful classification rate of 53% for glaucoma, and 85% for normal images, while VGG19 achieves a successful classification rate of 55% for glaucoma, and 77% for normal.

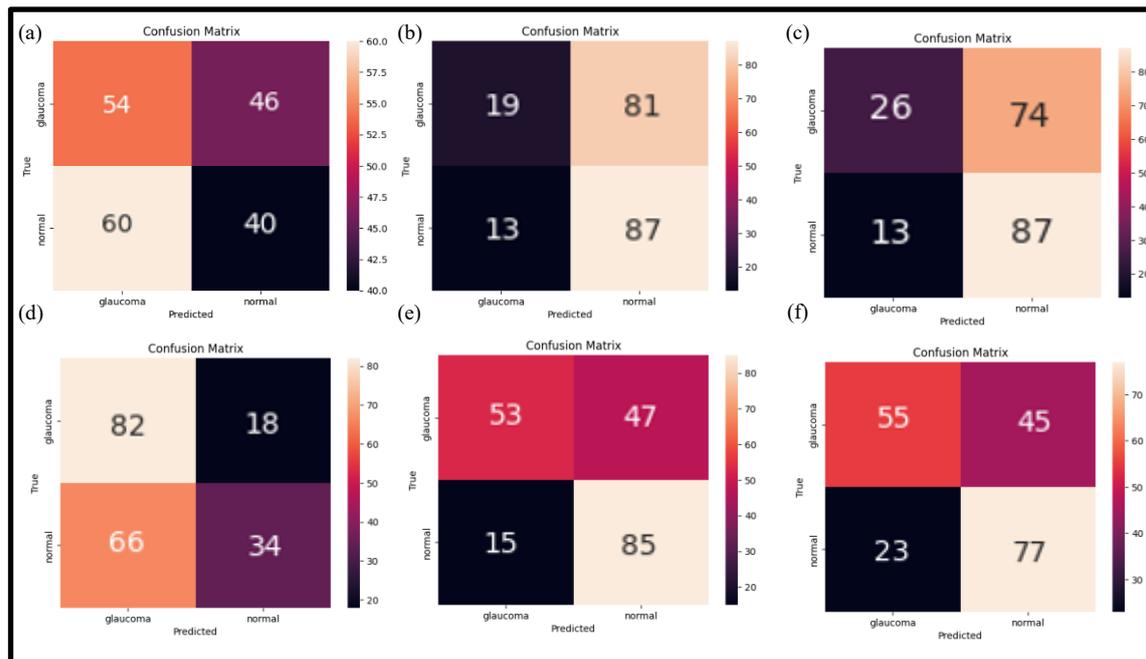


Fig. 9. Confusion matrix of the 6 generated models. (a) AlexNet; (b) Custom Layer; (c) MobileNetV2; (d) EfficientNetV1; (e) InceptionV3; (f) VGG19

3.4. Discussion

In this study, [Table 4](#) reveals that the Custom Layer model achieves the highest average accuracy of 93%, as shown in [Fig. 8](#) part b, which demonstrates a training graph without overfitting, however, the high accuracy of the model does not necessarily translate to accurate results in the confusion matrix using the test data. This discrepancy may arise from the similarity between glaucoma, and normal classes, leading to misclassification during the detection process. Consequently, the model may be deemed unreliable in detecting glaucoma disease. To address this issue, several approaches can be considered. Firstly, collecting a more diverse range of data can provide greater variation. Additionally, implementing feature engineering techniques such as image segmentation on the dataset can enhance object representation, and eliminate unnecessary features [38][39]. Furthermore, employing other data augmentation methods can augment the training data with more samples, improving data generalization, and enhancing the differentiation between classes [40].

4. CONCLUSION

This paper discusses hyperparameter experiments to find optimal parameters for each of the AlexNet, Custom Layer, MobileNetV2, EfficientNetV1, InceptionV3, and VGG19 models which will later be used for glaucoma detection systems using deep learning with CNN algorithms. Hyperparameter experiments were conducted on the RIM-ONE DL dataset which has been augmented with a total of 933 images. The hyperparameter experiment results show that the Custom Layer model achieves the highest accuracy of 93% using the Adam optimizer with a learning rate of 0.001, and a batch size of 256. Notably, the Custom Layer model exhibits training graph performance that avoids overfitting. Following the Custom Layer model, the AlexNet, VGG19, MobileNetV2, EfficientNetV1, and InceptionV3 models achieved 90%, 88.5%, 88%, 87.5%, and 83.5% accuracy, respectively. Although the six models show high average accuracy, it should be noted that the performance on the testing dataset of 200 images consisting of Cicendo Hospital data, and RIM-ONE DL dataset is not necessarily in line with the accuracy where in the testing data the best performance is actually achieved by the model that has the lowest accuracy, namely InceptionV3 with batch size 64, learning rate 0.001, and using Nadam optimizer. The difference between high training accuracy, and poor testing results can be caused by several factors such as limited training data, and a lack of variety in the datasets used so that the training data, and testing data experience a mismatch. Furthermore, in future research to get better model results it is recommended to use a much larger dataset. It would be helpful to use datasets from different collection methods with different tools to make the dataset more varied, and consider using other types of augmentation. A larger, and more diverse dataset will minimize the possibility of misclassification in the testing process.

Acknowledgments

The authors would like to express their gratitude to the Department of Telecommunication Engineering, Telkom University, Cicendo Eye Hospital, and all parties involved in the completion of this paper.

REFERENCES

- [1] S. F. Mohammadi, G. S. Anari, C. Alinia, E. Ashrafi, R. Daneshvar, and A. Sommer, "Is screening for glaucoma necessary? A policy guide and analysis," *J. Ophthalmic Vis. Res.*, vol. 9, no. 1, pp. 3–6, 2014, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4074467/>.
- [2] R. N. Weinreb, T. Aung, and F. A. Medeiros, "The pathophysiology and treatment of glaucoma: A review," *Jama*, vol. 311, no. 18, pp. 1901–1911, 2014, <https://doi.org/10.1001/jama.2014.3192>.
- [3] A. J. Tatham, F. A. Medeiros, L. M. Zangwill, and R. N. Weinreb, "Strategies to improve early diagnosis in glaucoma," *Prog. Brain Res.*, vol. 221, pp. 103–133, 2015, <https://doi.org/10.1016/bs.pbr.2015.03.001>.
- [4] H. A. Quigley and A. T. Broman, "The number of people with glaucoma worldwide in 2010 and 2020," *British journal of ophthalmology*, vol. 90, no. 3, pp. 262–267, 2006, <https://doi.org/10.1136/bjo.2005.081224>.
- [5] D. A. Lee and E. J. Higginbotham, "Glaucoma and its treatment: A Review," *American Journal of Health-System Pharmacy*, vol. 62, no. 7, pp. 691–699, 2005, <https://doi.org/10.1093/ajhp/62.7.691>.
- [6] C. Cook and P. Foster, "Epidemiology of glaucoma: What's new?," *Can. J. Ophthalmol.*, vol. 47, no. 3, pp. 223–226, 2012, <https://doi.org/10.1016/j.cjco.2012.02.003>.
- [7] B. E. Prum *et al.*, "Primary angle closure preferred practice pattern® guidelines," *Ophthalmology*, vol. 123, no. 1, pp. P1–P40, 2016, <https://doi.org/10.1016/j.ophtha.2015.10.049>.
- [8] F. Topouzis and E. Anastasopoulos, "Glaucoma—The Importance of Early Detection and Early Treatment," *US Ophthalmic Rev.*, vol. 2, p. 12, 2007, <https://doi.org/10.17925/USOR.2007.02.00.12>.
- [9] C. P. Bragança, J. M. Torres, C. P. D. A. Soares, and L. O. Macedo, "Detection of Glaucoma on Fundus Images Using Deep Learning on a New Image Set Obtained with a Smartphone and Handheld Ophthalmoscope," *Healthcare*, vol. 10, no. 12, p. 2345, 2022, <https://doi.org/10.3390/healthcare10122345>.
- [10] X. Chen, Y. Xu, D. W. Kee Wong, T. Y. Wong, and J. Liu, "Glaucoma detection based on deep convolutional neural network," *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 715–718, 2015, <https://doi.org/10.1109/EMBC.2015.7318462>.
- [11] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, 2014, <https://www.jmlr.org/papers/v15/srivastava14a.html>.
- [12] S. Singh and S. Krishnan, "Filter Response Normalization Layer: Eliminating Batch Dependence in the Training of Deep Neural Networks," *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 11234–11243, 2020, <https://doi.org/10.1109/CVPR42600.2020.01125>.
- [13] A. Sallam *et al.*, "Early Detection of Glaucoma using Transfer Learning from Pre-trained CNN Models," *2021 International Conference of Technology, Science and Administration (ICTSA)*, pp. 1–5, 2021, <https://doi.org/10.1109/ICTSA52017.2021.9406522>.
- [14] M. Norouzifard, A. Nemat, A. M. Abdul-Rahman, H. GholamHosseini, and R. Klette, "A Comparison of Transfer Learning Techniques, Deep Convolutional Neural Network and Multilayer Neural Network Methods for the Diagnosis of Glaucomatous Optic Neuropathy," in *New Trends in Computer Technologies and Applications: 23rd International Computer Symposium*, pp. 627–635, 2018, https://doi.org/10.1007/978-981-13-9190-3_69.
- [15] F. Fumero, T. D.-Alema, J. Sign, S. Alayo, R. Arna, and D. Angel-Pereir, "Rim-One D1: A Unified Retinal Image Database for Assessing Glaucoma Using Deep Learning," *Image Anal. Stereol.*, vol. 39, no. 3, pp. 161–167, 2020, <https://doi.org/10.5566/ias.2346>.
- [16] A. D.-Pinto, S. Morales, V. Naranjo, T. Köhler, J. M. Mossi, and A. Navea, "CNNs for automatic glaucoma assessment using fundus images: An extensive validation," *Biomed. Eng. Online*, vol. 18, no. 1, pp. 1–19, 2019, <https://doi.org/10.1186/s12938-019-0649-y>.
- [17] A. Patil and M. Rane, "Convolutional Neural Networks: An Overview and Its Applications in Pattern Recognition," *Smart Innov. Syst. Technol.*, vol. 195, pp. 21–30, 2021, https://doi.org/10.1007/978-981-15-7078-0_3.
- [18] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," in *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998, <https://doi.org/10.1109/5.726791>.
- [19] B. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017, <https://doi.org/10.1145/3065386>.
- [20] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.*, pp. 1–14, 2015, <https://ora.ox.ac.uk/objects/uuid:60713f18-a6d1-4d97-8f45-b60ad8aebbce>.
- [21] C. Szegedy *et al.*, "Going deeper with convolutions," *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–9, 2015, <https://doi.org/10.1109/cvpr.2015.7298594>.
- [22] T. Tsutsui and T. Matsuzawa, "Virtual Metrology Model Robustness Against Chamber Condition Variation Using Deep Learning," in *IEEE Transactions on Semiconductor Manufacturing*, vol. 32, no. 4, pp. 428–433, Nov. 2019, <https://doi.org/10.1109/TSM.2019.2931328>.
- [23] O. Russakovsky *et al.*, "ImageNet Large Scale Visual Recognition Challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, 2015, <https://doi.org/10.1007/s11263-015-0816-y>.

- [24] W. Yu *et al.*, “Visualizing and Comparing AlexNet and VGG using Deconvolutional Layers,” *International Conference on Machine Learning*, vol. 48, pp. 1–7, 2016, <https://icmlviz.github.io/icmlviz2016/assets/papers/4.pdf>
- [25] R. V. Moorthy, A. G. Kannan, and R. Balasubramanian, “Comparative analysis on deep convolutional neural network for brain tumor data set,” *Int. J. Health Sci.*, vol. 6, no. S1, pp. 1917–1930, 2022, <https://doi.org/10.53730/ijhs.v6nS1.4949>.
- [26] K. Dong, C. Zhou, Y. Ruan, and Y. Li, “MobileNetV2 Model for Image Classification,” *2020 2nd International Conference on Information Technology and Computer Application (ITCA)*, pp. 476–480, 2020, <https://doi.org/10.1109/ITCA52113.2020.00106>.
- [27] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, “Mobilenetv2: Inverted residuals and linear bottlenecks,” In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4510–4520, 2018, <https://doi.org/10.1109/CVPR.2018.00474>.
- [28] M. Tan and Q. Le, “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks,” *Can. J. Emerg. Med.*, vol. 15, no. 3, p. 190, 2013, <https://doi.org/10.2310/8000.2013.131108>.
- [29] L. Alzubaidi *et al.*, “Review of deep learning: concepts, CNN architectures, challenges, applications, future directions,” *Journal of Big Data*, vol. 8, no. 1, Jan. 2021, <https://doi.org/10.1186/s40537-021-00444-8>.
- [30] I. Kandel and M. Castelli, “The effect of batch size on the generalizability of the convolutional neural networks on a histopathology dataset,” *ICT Express*, vol. 6, no. 4, pp. 312–315, 2020, <https://doi.org/10.1016/j.ict.2020.04.010>.
- [31] Z. Yao *et al.*, “Large batch size training of neural networks with adversarial training and second-order information,” *arXiv preprint arXiv:1810.01021*, 2018, <https://arxiv.org/abs/1810.01021>.
- [32] P. M. Radiuk, “Impact of Training Set Batch Size on the Performance of Convolutional Neural Networks for Diverse Datasets,” *Inf. Technol. Manag. Sci.*, vol. 20, no. 1, pp. 20–24, 2018, <https://doi.org/10.1515/itms-2017-0003>.
- [33] X. Fang, H. Luo, and J. Tang, “Structural damage detection using neural network with learning rate improvement,” *Comput. Struct.*, vol. 83, no. 25–26, pp. 2150–2161, Sep. 2005, <https://doi.org/10.1016/j.compstruc.2005.02.029>.
- [34] Y. Wu *et al.*, “Demystifying Learning Rate Policies for High Accuracy Training of Deep Neural Networks,” *2019 IEEE International Conference on Big Data (Big Data)*, pp. 1971–1980, 2019, <https://doi.org/10.1109/BigData47090.2019.9006104>.
- [35] R. Zaheer and H. Shaziya, “A Study of the Optimization Algorithms in Deep Learning,” *2019 Third International Conference on Inventive Systems and Control (ICISC)*, pp. 536–539, 2019, <https://doi.org/10.1109/ICISC44355.2019.9036442>.
- [36] M. Sokolova and G. Lapalme, “A systematic analysis of performance measures for classification tasks,” *Inf. Process. Manag.*, vol. 45, no. 4, pp. 427–437, 2009, <https://doi.org/10.1016/j.ipm.2009.03.002>.
- [37] M. Hossin and M. N. Sulaiman, “A Review on Evaluation Metrics for Data Classification Evaluations,” *Int. J. Data Min. Knowl. Manag. Process*, vol. 5, no. 2, pp. 1–11, 2015, <https://doi.org/10.5121/ijdkp.2015.5201>.
- [38] S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz, and D. Terzopoulos, “Image Segmentation Using Deep Learning: A Survey,” in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 7, pp. 3523–3542, 2022, <https://doi.org/10.1109/TPAMI.2021.3059968>.
- [39] Y. Yu *et al.*, “Techniques and Challenges of Image Segmentation: A Review,” *Electron.*, vol. 12, no. 5, p. 1199, 2023, <https://doi.org/10.3390/electronics12051199>.
- [40] C. Shorten and T. M. Khoshgoftaar, “A survey on Image Data Augmentation for Deep Learning,” *Journal of Big Data*, vol. 6, no. 1, pp. 1–48, 2019, <https://doi.org/10.1186/s40537-019-0197-0>.

BIOGRAPHY OF AUTHORS



Muhamad Ilham is a final year-year Telecommunication Engineer student at the Faculty of Electrical Engineering at Telkom University, Bandung, Indonesia. His current research interest is image processing. Email: mhmdilhm@student.telkomuniversity.ac.id.



Angga Prihantoro is a final year-year Telecommunication Engineer student at the Faculty of Electrical Engineering at Telkom University, Bandung, Indonesia. His current research interest is computer vision. Email: anggaprihan@student.telkomuniversity.ac.id.



Iqbal Kurniawan Perdana is a final year-year Telecommunication Engineer student at the Faculty of Electrical Engineering at Telkom University, Bandung, Indonesia. His current research interest is mobile development. Email: iqbalperdana@student.telkomuniversity.ac.id.



Rita Magdalena is currently is currently a Lecturer is School of Electrical Engineering, Telkom university. Her research interest in information signal processing especially in image processing, audio processing and biomedical engineering. Email: ritamagdalenastudent.telkomuniversity.ac.id.



Sofia Saidah received the B.S and M.S degree from Telecommunication Engineering, Telkom Institute of Technology, Bandung , Indonesia in 2012 and 2014 respectively. She is currently a Lecturer in School of Electrical Engineering Telkom University. Her research interest includes, image processing, audio processing, biomedical engineering, steganography and watermarking. Email: sofiasaidahsfi@telkomuniversity.ac.id.