JOURNAL MEDICAL INFORMATICS TECHNOLOGY
medinftech.org ISSN 2988-7003 (online)

Glaucoma Detection in Fundus Eye Images using Convolutional Neural Network Method with Visual Geometric Group 16 and Residual Network 50 Architecture

Chandra Nugraha[1*] *and Sri Hadianti*[2]

[1,2] *Universitas Nusa Mandiri Jakarta, Indonesia*

ARTICLE HISTORY

Received: 08 June 23

Final Revision: 10 June 23

Accepted: 20 June 23

Online Publication: 30 June 23

KEYWORDS

Convolutional Neural Network (CNN), Geometric Group-16 (VGG-16), Residual Network-50 (ResNet-50), Glaucoma, Fundus

CORRESPONDING AUTHOR

14002629@nusamandiri.ac.id

DOI

10.37034/medinftech.v1i2.7

A B S T R A C T

Glaucoma is an eye disease usually caused by abnormal eye pressure. One of the causes of abnormal eye pressure is blockage of fluid flow, which if detected too late can lead to blindness. Glaucoma can be identified by examining specific areas on the retina fundus image. The aim of this study is to detect positive and negative glaucoma in fundus images. The image data was obtained from the glaucoma_detection dataset, consisting of 520 images, including 134 glaucoma-infected images and 386 normal images. This study uses the Convolutional Neural Network (CNN) method with Visual Geometric Group-16 (VGG-16) and Residual Network-50 (ResNet-50) architectures. The research and testing results using the VGG-16 architecture obtained an accuracy rate of 78%, while using the ResNet-50 architecture obtained an accuracy rate of 80%.

1. Introduction

The eye is one of the most important organs in the human body. In addition to being an organ, the eye also serves as a sensory organ for vision. As one part of the body, the eye is certainly not immune to disease, whether it is an attack from within or outside the eye. The most common attack on the eye is irritation caused by small objects such as dust or insects that are very small and enter the eye. In addition to irritation, there are also other diseases such as cataracts, styes, myopia, color blindness, presbyopia, glaucoma, and many more [1].

Glaucoma is the second leading cause of blindness worldwide after cataracts. Unlike cataracts, blindness caused by glaucoma is permanent. Based on WHO data from 2010, an estimated 3.2 million people suffer from blindness due to glaucoma. In addition, the rapid development of information and communication technology should be utilized to facilitate health workers in carrying out their duties. That is the benefit expected from this final project, which is to make it easier for health workers to detect glaucoma from retina images of patients [2].

Glaucoma is an eye disease caused by damage to the eye nerves due to an increase in pressure on the eye ball. If not detected early, this disease can cause permanent blindness. Glaucoma is the second leading cause of blindness after cataracts, and generally affects women and Asians [3]. Along with the development of the times, technology is rapidly advancing nowadays. With the development of technology today, it makes it easier for everyone to access anything. One of the technologies that have been discovered is image processing using digital images. Identification performed on an image has been developed for quite some time, one of which is by differentiating the texture of the image. In the texture of an image, it can be differentiated by several factors, including density, uniformity, roughness, and regularity [4].

1.1. Glucoma

Glaucoma is one of the leading causes of blindness in the world, characterized by damage to the eye nerves due to an increase in pressure on the eye ball. In the eye ball, there is a fluid called aqueous humor. Aqueous humor functions as nutrition for the eye, maintains the shape of the eye ball, and maintains the balance of eye ball pressure. Nerve damage is caused by an imbalance between the production of aqueous humor and the outflow of aqueous humor, causing the eye ball pressure in glaucoma patients to be higher than in normal eyes. The difference between normal eyes and glaucoma eyes is shown in Figure 1 [3].

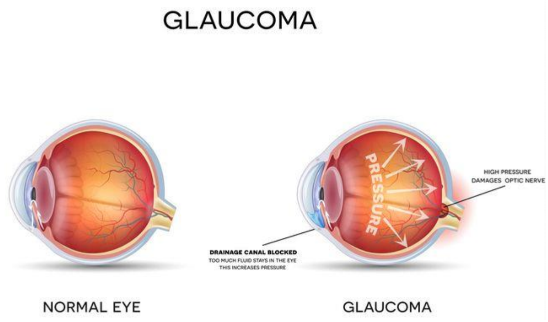


Figure 1. Differences between Normal and Glaucoma Eyes

1.2. Convolutional Neural Network (CNN)

CNN (Convolutional Neural Network) is a variation of Multilayer Perceptron that is inspired by the human neural network. The early research that underlies this discovery was first conducted by Hubel and Wiesel who studied the visual cortex in the cat's visual system. The visual cortex in animals is very powerful in visual processing systems ever existed. Therefore, many researches are inspired by its workings and produce new models, such as Neocognitron, HMAX, and LeNet-5 is shown in Figure 2 [5].

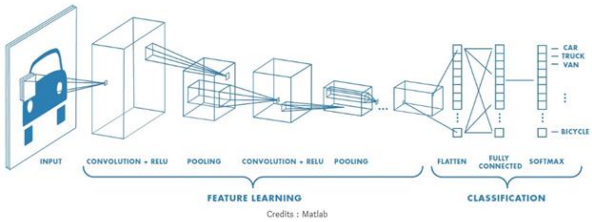


Figure 2. CNN Architecture

The working mechanism of CNN is similar to MLP, but in CNN each neuron is presented in two dimensions, unlike MLP where each neuron is only one-dimensional. Linear operations in CNN use convolution operations, while weights are no longer one-dimensional, but are four-dimensional collections of convolution kernels as shown in Figure 3. The weight dimensions in CNN are: input neuron x output neuron x height x width. Due to the nature of the convolution process, CNN can only be used on data that has a two-dimensional structure such as images and sounds is shown in Figure 3 [6].

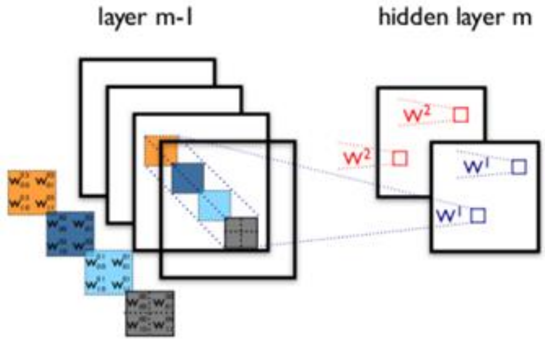


Figure 3. Convolution process in CNN

2. Research Method

The proposed method in this study can be seen in the following Figure 4.

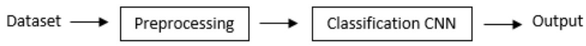
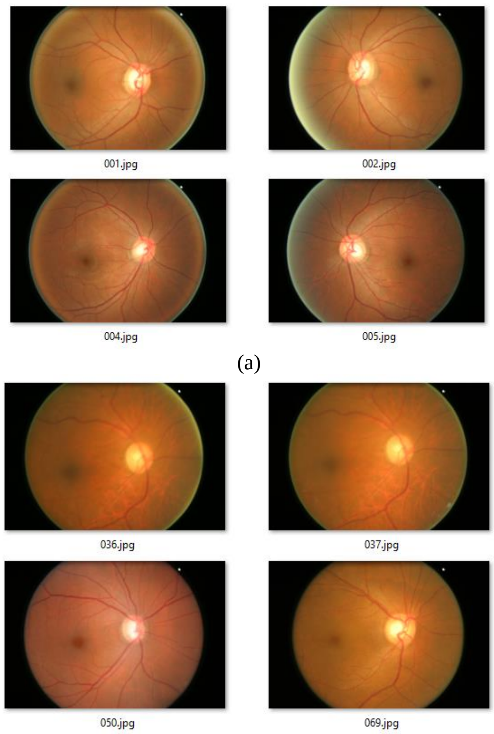


Figure 4. Research Method

2.1. Dateset

In conducting this research, the first step is to collect the glaucoma fundus image dataset obtained from glaucoma detection [7]. It is divided into two labels, the first is glaucoma positive while the second label is glaucoma negative. Then the architecture design starts with determining the network depth, layer arrangement, and selection of the layer type to be used to obtain a model based on the input dataset and the label name index.



(a)

(b)

Figure 5. Example of dataset, (a) Negative Glaucoma, (b) Positive Glaucoma

2.2. Pre-Processing

Pre-processing is the initial process carried out to obtain the best image by minimizing noise so that the characteristics of the OD and OC images can be easily recognized. In pre-processing, the process of converting RGB images to grayscale images is carried out, with the red channel for the OD part and the green channel for the OC part. This is done to obtain the desired object with maximum results [3].

2.3. CNN Classification

Unlike ordinary classification algorithms, where feature extraction and classification processes are usually carried out separately, the algorithm model in this deep learning field will extract features and classify images in one process. In other words, feature extraction in the CNN algorithm is also involved in learning [8].

2.4. VGG-16 Architecture

The VGG-16 architecture is a convolutional neural network (CNN) that was proposed by the Visual Geometry Group (VGG) at the University of Oxford in 2014. It consists of 16 layers, including 13 convolutional layers and 3 fully connected layers, and has approximately 138 million trainable parameters.

The architecture follows a simple and uniform design principle, where all the convolutional layers have a 3x3 filter size and a stride of 1 pixel, and all the max-pooling layers have a 2x2 filter size and a stride of 2 pixels. The network takes an input image of size 224x224x3 and outputs a vector of probabilities for each of the 1000 classes in the ImageNet dataset is shown in Figure 6.

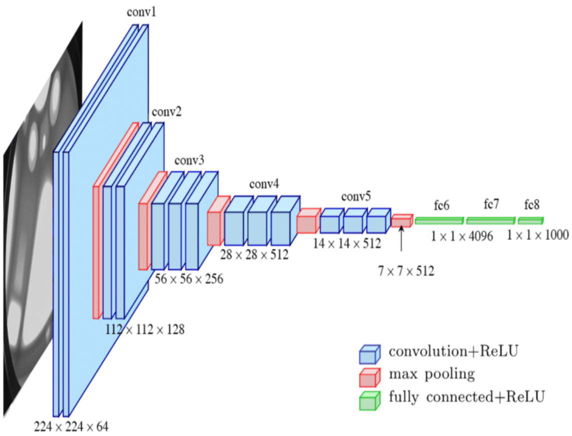


Figure 6. VGG-16 Architecture

2.5. Architecture ResNet50

ResNet50 is a convolutional neural network (CNN) architecture that was proposed by Microsoft Research in 2015. It is a variant of the ResNet family of networks, which are known for their ability to train very deep networks while avoiding the vanishing gradient problem.

ResNet50 consists of 50 layers, including 49 convolutional layers and 1 fully connected layer, and has approximately 23.5 million trainable parameters. The architecture introduces the concept of residual learning, where each residual block contains a shortcut connection that skips one or more layers. This allows the network to learn residual functions instead of attempting to learn the original mapping directly, which can make it easier to optimize deeper networks.

The residual blocks in ResNet50 have different configurations, but they all have a common structure. Each block contains two or three convolutional layers, with batch normalization and ReLU activation between them. The shortcut connection in each block skips over the convolutional layers and adds the input to the output of the last convolutional layer before passing it through the activation function.

The network takes an input image of size 224x224x3 and outputs a vector of probabilities for each of the 1000 classes in the ImageNet dataset dataset is shown in Figure 7.

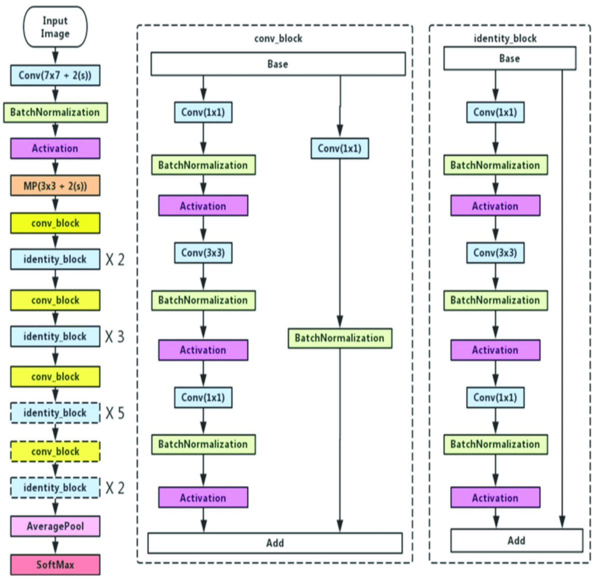


Figure 7. ResNet50 Architecture

3. Result and Discussion

The CNN architecture used in this research is VGG-16, and its implementation can be seen in Figure 8.

Model: "vgg16"

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	[(None, 224, 224, 3)]	0
block1_conv1 (Conv2D)	(None, 224, 224, 64)	1792
block1_conv2 (Conv2D)	(None, 224, 224, 64)	36928
block1_pool (MaxPooling2D)	(None, 112, 112, 64)	0
block2_conv1 (Conv2D)	(None, 112, 112, 128)	73856
block2_conv2 (Conv2D)	(None, 112, 112, 128)	147584
block2_pool (MaxPooling2D)	(None, 56, 56, 128)	0
block3_conv1 (Conv2D)	(None, 56, 56, 256)	295168
block3_conv2 (Conv2D)	(None, 56, 56, 256)	590080
block3_conv3 (Conv2D)	(None, 56, 56, 256)	590080
block3_pool (MaxPooling2D)	(None, 28, 28, 256)	0
block4_conv1 (Conv2D)	(None, 28, 28, 512)	1180160
block4_conv2 (Conv2D)	(None, 28, 28, 512)	2359808
block4_conv3 (Conv2D)	(None, 28, 28, 512)	2359808
block4_pool (MaxPooling2D)	(None, 14, 14, 512)	0
block5_conv1 (Conv2D)	(None, 14, 14, 512)	2359808

Figure 8. Implementation of VGG-16 Architecture

The CNN architecture used in this study is ResNet50, its application can be seen in Figure 9.

Model: "resnet50"

Layer (type)	Output Shape	Param #	Connected to
input_1 (InputLayer)	[(None, 224, 224, 3)]	0	[]
conv1_pad (ZeroPadding2D)	(None, 230, 230, 3)	0	['input_1[0][0]']
conv1_conv (Conv2D)	(None, 112, 112, 64)	9472	['conv1_pad[0][0]']
conv1_bn (BatchNormalization)	(None, 112, 112, 64)	256	['conv1_conv[0][0]']
conv1_relu (Activation)	(None, 112, 112, 64)	0	['conv1_bn[0][0]']
pool1_pad (ZeroPadding2D)	(None, 114, 114, 64)	0	['conv1_relu[0][0]']

Figure 9. Application of ResNet50 Architecture

In the pre-processing stage, the image size was resized to 224x224 pixels. With uniform image size, it can make it easier for computation and recognition stages. Next, we move on to the data distribution which can be seen in Table 1.

Table 1. Distribution of Train, Test, and Validation Data

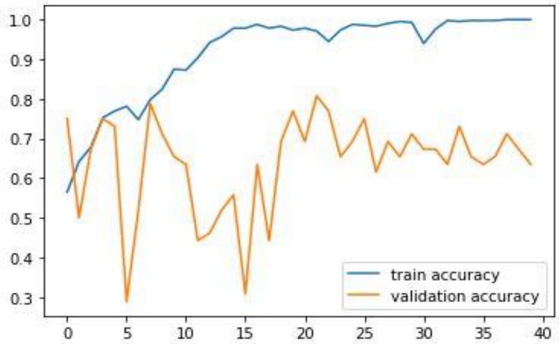
Dataset			Total
Train	Test	Valid	
416	52	52	520

Based on Table 1, the training and validation data were used in the process of training, tuning, and evaluating the CNN model, while the test data were used to test the performance of the trained model. After obtaining images with a uniform size and distribution of image data, we move on to the classification stage. In this stage, parameters such as those seen in Table 2 are used.

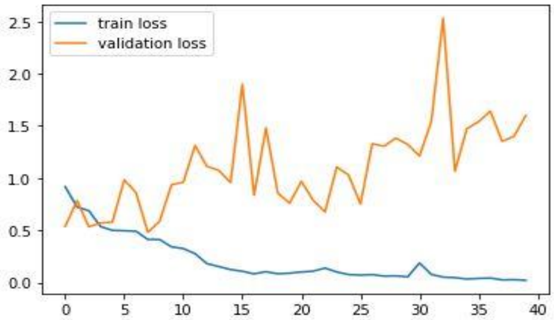
Table 2. Parameters

Size (pixel)	ePoch	Batch Size	Optimizer
224 x 224	20	32	ADAM

Based on the parameters used, which are VGG-16 and ResNet50 architectures with the applied model, the results obtained can be seen in the following Figure10 and Figure 11.

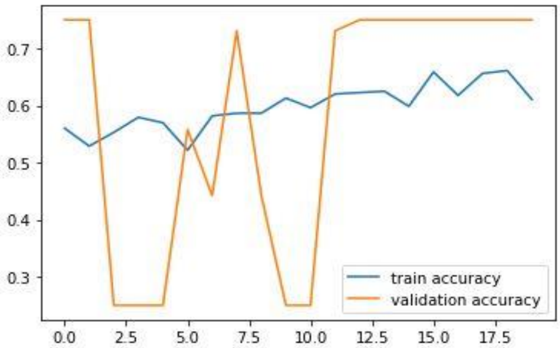


(a)

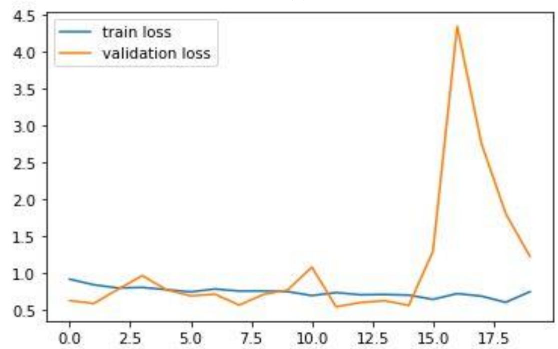


(b)

Figure 10. VGG-16 Architecture, (a) Accuracy Graph, (b) Loss Graph



(a)



(b)

Figure 11. ResNet50 Architecture, (a) Accuracy Graph, (b) Loss Graph

Based on Figure 10 and Figure11, that can be explained that the classification of eye diseases using Convolutional Neural Network (CNN) with VGG-16 architecture resulted in an accuracy of 78%. Meanwhile, using the ResNet50 architecture resulted in an accuracy of 80%. To determine the classification performance, evaluation of classification metrics and confusion matrix was carried out as can be seen in Figures 12 and 13.

The statement is saying that in the experiment, two different CNN architectures, VGG-16 and ResNet50, were used to classify eye diseases. The accuracy of the classification was measured for both architectures, and the results were shown in Figures 10 and 11. The accuracy of VGG-16 was found to be 78%, while that of ResNet50 was 80%. This indicates that ResNet50 performed better than VGG-16 in classifying the eye diseases.

To further evaluate the performance of the classification, classification metrics and confusion matrix were calculated and presented in Figures 12 and 13. The classification metrics provide information about the precision, recall, and F1 score for each class, while the confusion matrix shows the number of true positives, true negatives, false positives, and false negatives for each class. These metrics can help to analyze the performance of the model in more detail and identify the areas that need improvement.

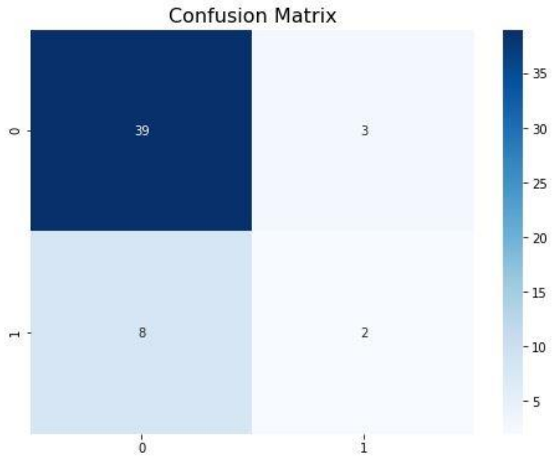
	precision	recall	f1-score	support
0	0.83	0.93	0.88	42
1	0.40	0.20	0.27	10
accuracy			0.79	52
macro avg	0.61	0.56	0.57	52
weighted avg	0.75	0.79	0.76	52

(a)

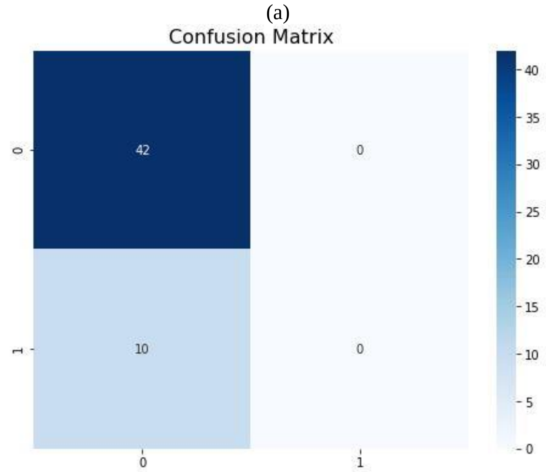
	precision	recall	f1-score	support
0	0.81	1.00	0.89	42
1	0.00	0.00	0.00	10
accuracy			0.81	52
macro avg	0.40	0.50	0.45	52
weighted avg	0.65	0.81	0.72	52

(b)

Figure 12. Classification Metrics, (a) VGG-16 Architecture, (b) ResNet50 Architecture



(a)



(b)

Figure 13. Confusion Matrix, (a) Arsitektur VGG-16, and (b) Arsitektur ResNet50

4. Conclusion

Based on the results of the research and testing on the classification of fundus glaucoma images using CNN (convolutional neural network) with VGG-16 and ResNet50 architectures with two identity labels in fundus images, the accuracy level for VGG-16 architecture was 78% and for ResNet50 architecture was 80%. Therefore, with these accuracy values, it can be concluded that glaucoma detection using CNN method with ResNet50 architecture has a better accuracy value.

References

[1] A. R. Tobias, D. S., & Widiarti, “Deteksi Glaukoma pada Citra Fundus Retina dengan Metode K-Nearest Neighbor,” Semin. Nas. Ilmu Komput., 2016.

[2] A. Mustofa, H. Tjandrasa, and B. Amaliah, “Deteksi Penyakit Glaukoma pada Citra Fundus Retina Mata Menggunakan Adaptive Thresholding dan Support Vector Machine,” J. Tek. ITS, vol. 5, no. 2, Oct. 2016, doi: 10.12962/j23373539.v5i2.18929.

[3] M. R. A. Romadhoni, R. Magdalena, and R. Y. N. Fuadah, “Klasifikasi Mata Glaukoma dan Mata Normal Menggunakan Metode Support Support Vector Machine,” e-Proceeding Eng., vol. Vol.7, No., no. ISSN : 2355-9365, p. 768, 2020.

[4] Y. G. K, I. Santoso, and R. R. Isnanto, "Klasifikasi Citra Dengan Matriks Ko-Okurensi Aras Keabuan (Gray Level Co-occurrence Matrix-GLCM) Pada Lima Kelas Biji-Bijian," Undergrad. thesis, no. Jurusan Teknik Elektro Fakultas Teknik Undip, 2011.

[5] M. Zufar and B. Setiyono, "Convolutional Neural Networks Untuk Pengenalan Wajah Secara Real-time," J. Sains dan Seni ITS, vol. 5, no. 2, 2016.

[6] F. N. Cahya, N. Hardi, D. Riana, and S. Hadiyanti, "Klasifikasi Penyakit Mata Menggunakan Convolutional Neural Network (CNN)," SISTEMASI, vol. 10, no. 3, p. 618, Sep. 2021, doi: 10.32520/stmsi.v10i3.1248.

[7] M. Douider, "glaucoma_détection." https://www.kaggle.com/meryemdouider/glaucoma-dtection.

[8] F. F. Maulana and N. Rochmawati, "Klasifikasi Citra Buah Menggunakan Convolutional Neural Network," J. Informatics Comput. Sci., vol. Volume 01, no. JINACS, pp. 104–108, 2019.Riana, A. N. Hidayanto, S. Hadianti, and D. Napitupulu, "Integrative Factors of E-Health Laboratory Adoption: A Case of Indonesia," Future Internet, vol. 13, no. 2, p. 26, Jan. 2021, doi: 10.3390/fi13020026.