

Computer Aided Diagnosis for Early Detection of Glaucoma Using Convolutional Neural Network (CNN)



Yunendah Nur Fu'adah, Sofia Sa'idah, Inung Wijayanto,
Nur Ibrahim, Syamsul Rizal, and Rita Magdalena

Abstract Glaucoma is an eye disease caused by an increase in eye pressure resulting in damage to the optic nerve and can cause blindness. An ophthalmologist makes the diagnosis of glaucoma by analyzing the fundus images that show the retinal structure. This manual diagnosis requires years of expertise and prone to error. Previous studies have designed a glaucoma CAD system based on Convolutional Neural Network (CNN) and showed promising results. This study proposes the CNN method consisting of three hidden layers that use 3×3 of the filter size with 16, 32, 64 output channels, fully connected layers, and sigmoid activation. The experiment is conducted using the RIMONE R2 fundus images dataset to classify normal and glaucoma conditions. From 455 fundus images, 75% are used as the training data, while the rest is used as the validation data. From the experiment, this study outperforms other previous studies by achieving 91.22% of accuracy. The glaucoma system detection that has been developed in this research, can be helpful for ophthalmologists to establish an initial diagnosis of glaucoma that can reduce the harmful effects of glaucoma.

Keywords Fundus · Glaucoma · CNN · RIMONE R2

1 Introduction

Glaucoma is a multifactorial neurodegenerative disease, which reduces human vision and may cause blindness. The main cause of glaucoma is the Intra Ocular Pressure (IOP). This condition can damage the optic nerve, which made the brain does not get image information from light receptors [1]. The retina and optic nerve changes often occur without symptoms and are not detected by diagnostic tests [2]. Early examination and medical treatment suggested by an ophthalmologist can help reduce the risk of glaucoma [1]. Based on data from the WHO, glaucoma is one of

Y. N. Fu'adah (✉) · S. Sa'idah · I. Wijayanto · N. Ibrahim · S. Rizal · R. Magdalena
School of Electrical Engineering, Telkom University, Bandung, Indonesia
e-mail: yunendah@telkomuniversity.ac.id

the major causes of blindness globally after cataracts. More than 82% of all blind people are aged 50 years and over. In 2013, the number of people (40–80 years old) with glaucoma was 64.3 million and continued to increase until 2020 [3].

Ophthalmologists diagnose Glaucoma by analyzing the retinal structures. They are followed by analyzing ocular parameters such as Cup to Disk Ratio (CDR) and very high Rim to Disk Ratio (RDR). This requires expertise and accuracy. In diagnosing Glaucoma, ophthalmologists use several expensive devices such as the Heidelberg Retinal Tomography (HRT), Optical Coherence Tomography (OCT), Including Confocal Scanning Laser Ophthalmoscopy (CSLO) and relatively limited fundus imaging [4]. Several studies based on fundus image processing have developed for glaucoma detection. Previous studies mentioned that glaucoma disease is detected by localizing and segmenting fundus images to obtain the Optic Nerve Head (ONH) sections. The features are extracted using statistical measurement and classified using K-Nearest Neighbor (K-NN). The method achieves 95.24% accuracy, which tested in a private data set [5]. GeethaRamani et al. [6] did a segmentation process is performed to calculate the Cup to Disc Ratio (CDR) based on image processing. The study obtains an accuracy of 98.7% using a private dataset. In study [7], the feature extraction process is performed using the Empirical Wavelet Transform method, and the Support Vector Machine (SVM) classification method. The method is tested in a private dataset combined with the RIM-ONE dataset. They achieved 98.33% accuracy for the private data set and 81.32% for the RIM-ONE dataset.

In some previous studies, the researches carried out several important stages in designing a glaucoma detection system, such as the preprocessing and segmentation process that must be accurate, proper feature selection and optimization at the classification stage which greatly affect the system performance. Some studies are starting to use the Convolutional Neural Network (CNN) as the main method. CNN has the advantage that it can learn the deterministic features of raw data directly and provide a promising system performance in detecting Glaucoma. Studies by Memon et al. [8] and Bajwa et al. [9] are the examples of study which use CNN as the main method for detecting Glaucoma. They achieve an accuracy of 85% [8] and 87.4% [9], which tested in RIM-ONE [8] and ORIGA [9] dataset. In the previous studies, CNN model was designed using more than three hidden layers and used softmax activation. In this research, CNN model design will be done using three hidden layers and sigmoid activation, which is more effective for classifying two conditions such as normal and glaucoma.

2 Method

This study designed a detection system to classify glaucoma and normal condition based on image processing. The system used the CNN method with three hidden layers, which used 3×3 filter size of each layers and 16, 32, 64 output channels

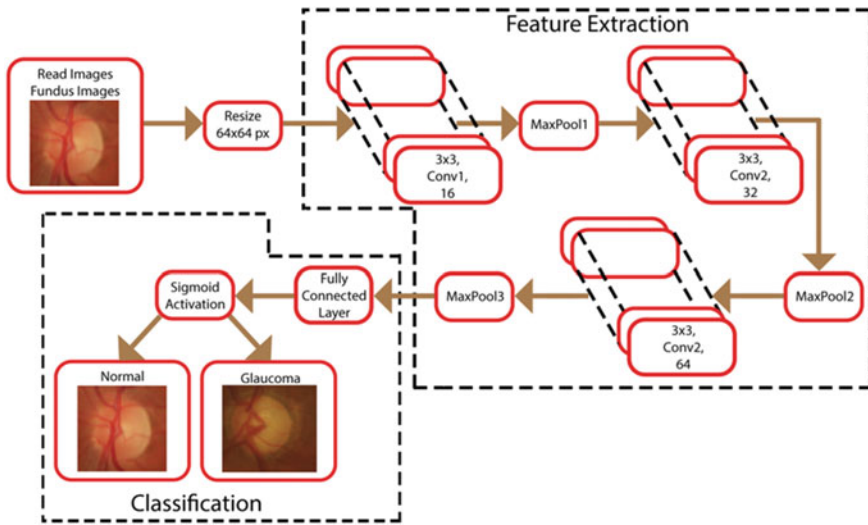


Fig. 1 The proposed system model of CNN for glaucoma detection

respectively. Furthermore, fully connected layers and sigmoid activation used to classify normal and glaucoma conditions. In general, the CNN model proposed in this study is shown in Fig. 1.

2.1 Dataset

This study uses RIM-ONE R2 dataset. The dataset has 455 fundus images, consist of 255 fundus images in normal condition and 200 fundus images in glaucoma condition. The distribution of training data and validation data in this study are 75% and 25%, respectively. Thus, the training data used are 341 fundus images, while the validation data are 114 fundus images. The data is then resized with a size of 64×64 to be processed at the next stage.

2.2 Convolutional Neural Network

Convolutional Neural Network (CNN) is one of the Deep Neural Networks, which implemented for image recognition [10]. CNN is a development of the Multilayer Perceptron (MLP). MLP accepts one-dimensional input data and propagates the data on the network to produce output. The CNN data propagated on the network are two-dimensional [11]. Thus, it can only be used on data that has a

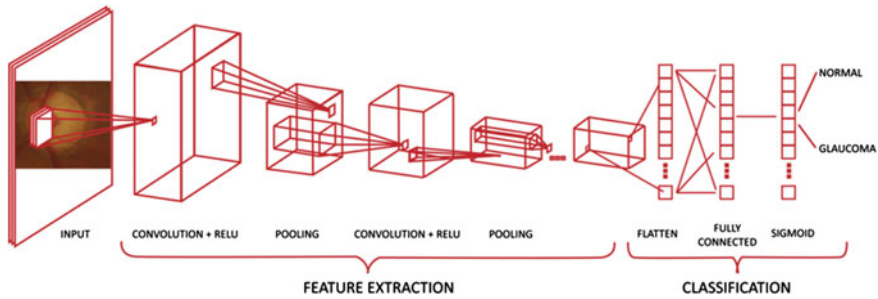


Fig. 2 The architecture of CNN

two-dimensional structure such as image data. In general, CNN consists of a Feature Extraction Layer and Classification Layer, which shown by the CNN architecture in Fig. 2.

The feature extraction layer consists of convolutional layers and pooling layers. Convolutional layers convert images with convolution processes to produce feature maps that show the original image's unique characteristics. Convolutional layers operate differently from other neural network layers that use connection weights. The convolutional layers use convolutional filters to produce a feature map [10].

Furthermore, the convolutional layer's activation process uses Rectified Linear Units (ReLU) activation to increase the training stage on the neural network so it can minimize errors and saturation. In this study, the ReLU activation function is used at each hidden layer of the neural network. The ReLU activation function is shown in (1) [12].

$$f(x) = \begin{cases} x, & \dots x > 0 \\ 0, & \dots x \leq 0 \end{cases} \quad (1)$$

The ReLU activation function changes the negative pixel value on the image to 0 on the feature map. The pooling layer on the feature extraction layer serves to reduce the size of the layer. There are two types of pooling methods, namely maximum pooling, which sees the maximum value and mean pooling that looks for the average value. An illustration of the pooling process can be seen in Fig. 3 [10].

Based on Fig. 3, it can be seen that the Pooling Layer reduces the size of the image, the image which was originally sized 4×4 to 2×2 without losing significant information. To avoid overfitting and help generalize at the training stage, this study used a dropout at the last hidden layer. The dropout consists of setting to zero the output of each hidden neuron according to the probability value used. Hence when the neurons in CNN are dropped out, the neurons do not contribute to the forward pass stage and do not participate in the back propagation stage [13].

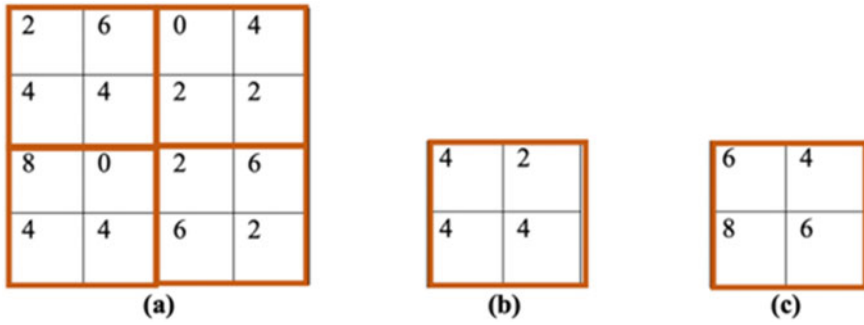


Fig. 3 Illustrate of pooling layer, **a** the 4×4 pixel input image, **b** mean pooling, **c** maximum pooling

2.3 Classification Layer

The classification stage consists of flattening stages to change feature maps, which are multidimensional arrays into one-dimensional arrays [14]. The fully connected layers emit vectors K , where K is the number of classes that can be predicted by the network. This study used two classes, namely normal and glaucoma. In the last stage, the sigmoid activation function is used following equation, which showed in (2) [15]. It can be seen that the sigmoid activation function transform the input value x into the range of 0–1.

$$S(x) = \frac{1}{1 + e^{-x}} \tag{2}$$

Table 1 shows the summary of the proposed CNN model as well as the output of each layer that affects the image size.

Table 1 Details of CNN model proposed

| Layer type | Output shape | Parameter | Layer type | Output shape | Parameter |
|-------------|--------------|-----------|-------------|--------------|-----------|
| Input image | 64,64,3 | 0 | Convolution | 16,16,64 | 18,496 |
| Convolution | 64,64,16 | 448 | ReLU | 16,16,64 | 0 |
| ReLU | 64,64,16 | 0 | Max pooling | 8,8,64 | 0 |
| Max pooling | 32,32,16 | 0 | Dropout | 8,8,64 | 0 |
| Convolution | 32,32,32 | 4640 | Flatten | 4096 | 0 |
| ReLU | 32,32,32 | 0 | Dense | 2 | 8194 |
| Max pooling | 16,16,32 | 0 | Sigmoid | 2 | 0 |

3 System Performance

This study uses accuracy, recall, precision, and f1 scores to measure the performance. The calculation is shown in (3), (4), (5) and (6) [9]. True Positive (TP) shows the exact glaucoma data detected as glaucoma, True Negative (TN) shows normal detected as normal data, False Positive (FP) shows glaucoma detected normal data while False Negative (FN) shows normal detected glaucoma data.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \tag{3}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{4}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{5}$$

$$F1 \dots \text{Score} = 2 \times \frac{\text{recall} \times \text{precision}}{\text{recall} + \text{precision}} \tag{6}$$

4 Result and Discussion

The RIM-ONE R2 dataset used in this study consist of 455 fundus images for glaucoma and normal conditions. The number of training data used is 341 fundus images. On the other hand, the amount of test data used is 114 fundus images for 67 normal condition images and 47 glaucoma conditions. For glaucoma and normal conditions, fundus images are trained using the CNN model with Adam optimizer learning rate 0.001, and loss binary cross-entropy. Performance parameters measured in this study are accuracy, recall, precision, F1 score, and loss. The model of accuracy and the loss of the proposed model is shown in Fig. 4.

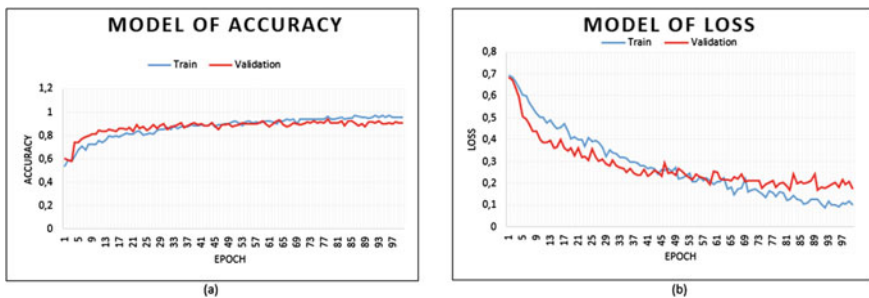


Fig. 4 a Accuracy of the proposed model, b loss of the proposed model

Fig. 5 Confusion matrix of validation data

| | | | |
|------|----------|-----------|--------|
| TRUE | Glaucoma | 61 | 4 |
| | Normal | 6 | 43 |
| | | Glaucoma | Normal |
| | | Predicted | |

Based on the experiment conducted in this study, the accuracy increases for each iteration (epoch), and the difference in accuracy between the training data and the validation data accuracy is not much different, as shown in Fig. 4a. Based on these results, it can be concluded that there is no overfitting of the system designed using the proposed model. Figure 4b shows a decrease in the value of the loss at each iteration (epoch) with a value that is not much difference between the value of loss training and loss validation. It can be concluded that learning errors that occur both for training data and test data achieved are minimum so that the model can recognize normal conditions and glaucoma conditions with the best accuracy performance of 91.22%, and loss of 0.1758.

Based on the Confusion Matrix shown in Fig. 5, it can be seen that from 114 validation data used, 104 data are successfully classified according to their class. Other parameters used to evaluate system performance are precision, recall, and F1-score, which have a range of values from 0 to 1 (a value of 1 indicates no error). Based on the data shown in Table 2, the value of system performance parameters is closed to 1. This condition shows that the CNN model can classify glaucoma and normal conditions with high accuracy and minimal missclassification.

Compared to some previous studies, which used RIM ONE data set and Convolution Neural Network with softmax activation as the main method to classify normal and glaucoma conditions, the proposed CNN method with sigmoid activation outperforms the previous studies. The sigmoid activation that used in the classification layer has proven to be more precise in classifying two conditions.

Table 2 Performance of CNN model proposed

| Class | Precision | Recall | F1-Score | No of images |
|----------|-----------|--------|----------|--------------|
| Glaucoma | 0.91 | 0.94 | 0.92 | 65 |
| Normal | 0.91 | 0.88 | 0.90 | 49 |
| Total | 0.91 | 0.91 | 0.91 | 114 |

5 Conclusion

In this research, a computer-aided diagnose system for the early detection of glaucoma based on digital image processing is designed. The CNN model used in this study consists of three hidden layers. Each of them is using 3×3 filter sizes with 16, 32, and 64 channel outputs, a fully connected layer, and sigmoid activation. The experiment showed that the proposed CNN model is able to classify raw fundus image datasets directly into glaucoma conditions and normal conditions with an accuracy of 91.22%, loss of 0.1758, and the value of precision, recall, an f1-score average of each amounted to 0.91. In further research, a system for the classification of glaucoma conditions can be developed based on its severity.

References

1. Ahmad H, Yamin A, Shakeel A, Gillani SO, Ansari U (2014) Detection of glaucoma using retinal fundus images. In: 2014 International conference on robotics and emerging allied technologies in engineering (iCREATE), pp 321–324. IEEE. <https://doi.org/10.1109/iCREATE.2014.6828388>
2. Carrillo J, Bautista L, Villamizar J, Rueda J, Sanchez M, Rueda D (2019) Glaucoma detection using fundus images of the eye. In: 22nd Symposium image signal processing artificial vision STSIVA 2019—conference proceedings, pp 1–4. <https://doi.org/10.1109/STSIVA.2019.8730250>
3. Nucci C, Osbornec NN, Bagetta G, Cerulli L (2008) Glaucoma: an open-window to neurodegeneration and neuroprotection. Preface Prog Brain Res 173:173. [https://doi.org/10.1016/S0079-6123\(08\)01145-X](https://doi.org/10.1016/S0079-6123(08)01145-X)
4. Kim M, Han JC, Hyun SH, Janssens O, Van Hoecke S, Kee C, De Neve W (2019) Medinoid: computer-aided diagnosis and localization of glaucoma using deep learning. Appl Sci 9:2357–2362. <https://doi.org/10.3390/app9153064>
5. Septiarini A, Khairina DM, Kridalaksana AH, Hamdani H (2018) Automatic glaucoma detection method applying a statistical approach to fundus images. Healthc Inform Res 24:53–60. <https://doi.org/10.4258/hir.2018.24.1.53>
6. GeethaRamani R, Dhanapackiam C (2014) Automatic localization and segmentation of optic disc in retinal fundus images through image processing techniques. In: International conference on recent trends in information technology automatic
7. Maheshwari S, Pachori RB, Acharya UR (2017) Automated diagnosis of glaucoma using empirical wavelet transform and correntropy features extracted from fundus images. IEEE J Biomed Heal Inf 21:803–813. <https://doi.org/10.1109/JBHI.2016.2544961>
8. Memon I, Ursani AA, Bohyo MA, Chandio R (2019) Automated diagnosis of glaucoma using deep learning architecture. Eng Sci Technol Int Res J 3. <https://doi.org/10.2174/1574362414666190906112910>
9. Bajwa MN, Malik MI, Siddiqui SA, Dengel A, Shafait F, Neumeier W, Ahmed S (2019) Correction to: two-stage framework for optic disc localization and glaucoma classification in retinal fundus images using deep learning. BMC Med Inf Decis Mak 19(136). <https://doi.org/10.1186/s12911-019-0842-8>. BMC Med Inform Decis Mak 19:1–16. <https://doi.org/10.1186/s12911-019-0876-y>
10. Kim P (2017) MATLAB deep learning. Apress, Berkeley, CA. <https://doi.org/10.1007/978-1-4842-2845-6>

11. Noeman A, Handayani D (2020) Detection of Mad Lazim Harfi Musyba images uses convolutional neural network. IOP Conf Ser Mater Sci Eng 771. <https://doi.org/10.1088/1757-899X/771/1/012030>
12. Kim P (2017) MATLAB deep learning: with machine learning, neural networks and artificial intelligence. Apress, Berkeley, CA
13. Wang L, Wen M, Wu Y (2019) Glaucoma detection based on deep convolutional neural network. Acta Microsc 28:331–338
14. Yamashita R, Nishio M, Do RKG, Togashi K (2018) Convolutional neural networks: an overview and application in radiology. Insights Imaging 9:611–629. <https://doi.org/10.1007/s13244-018-0639-9>
15. Feng J, Lu S (2011) Performance analysis of various activation functions in artificial neural networks. Int J Artif Intell Expert Syst 1. <https://doi.org/10.1088/1742-6596/1237/2/022030>