

Conference Paper

Diabetic Retinopathy Detection Using GoogleNet Architecture of Convolutional Neural Network Through Fundus Images

Amnia Salma *, Alhadi Bustamam, Devvi Sarwinda

Department of Mathematics, Mathematics and Natural Science Faculty, Universitas Indonesia, Depok, Indonesia

*Corresponding author:
E-mail:
amniasalma@sci.ui.ac.id

ABSTRACT

The number of people who have Diabetes is about 422 million in the world. Diabetes is a group of metabolic diseases characterized by elevated levels of blood glucose. The serious damage of blood vessels caused by Diabetes in the tissue at the retina is called Diabetic Retinopathy. Diabetic Retinopathy can cause severe blindness. Early detection can help patients find a suitable treatment and prevent blindness. Ophthalmologists can detect this disease by screening, but this method takes a long time, is very costly, and need professional skills to perform it. In the big data era, many researchers use deep learning models for medical help. One of the models uses image classification. We have designed a tool using image classification to help ophthalmologists detect diabetic retinopathy. In this research, we use image classification to classify Diabetic Retinopathy into two classes which are normal (No DR) and Diabetic Retinopathy. We use 200 datasets of fundus images that we obtain from the Kaggle Database. We used a deep learning model in this research that is one of Convolutional Neural Network architecture called GoogleNet. For training the model we used Python as a programming language with Pytorch library. GoogleNet has a very good performance for image classification and has an accuracy of 88%.

Keywords: Diabetic retinopathy, CNN, GoogleNet, fundus images

Introduction

Based on data from WHO, there are about 422 million people in the world live with diabetes (WHO, 2019). Diabetes is a group of metabolic diseases characterized by elevated levels of blood glucose. One of the complications of Diabetes can make serious damage to blood vessels in the retina called Diabetic Retinopathy. One of the complications of Diabetes patients is DR and it can cause severe blindness (Khojasteh, et al., 2018; Maier et al., 2019). The conventional method to detect DR is complex to do and includes several efforts for patients and doctors. There is an early symptom of DR such as Microaneurysms. The shape of Microaneurysm was a little red spot in the retina (Indumathi & Sathananthavath. 2019; Seoud et al., 2015; Jaya et al., 2015). The other symptom is a hemorrhage or red lesion because of break vessel and hard exudate (Indumathi & Sathananthavath, 2019; Jaya et al., 2015; Joshi & Karule, 2018). Besides that, there is neovascularization as the symptom of DR which is a severe level (Maier et al., 2019). Medical exercise for DR patient's screening is used to detect DR. In fact, it takes a too long time, is very costly, and needs professional skills to perform it.

With technology development, computers can work faster classification once trained. Besides, that computers have the performance to help the doctors in real-time classification (Philip et al, 2007). Significant research has been done by the researcher before on detecting the feature of DR using automated methods with machine learning, such as SVM, random forest, and k-NN as classifiers.

How to cite:

Salma, A., Bustamam, A., & Sarwinda, D. (2021). Diabetic retinopathy detection using GoogleNet architecture of convolutional neural network through fundus images. *1st Bioinformatics and Biodiversity Conference*. NST Proceedings. pages 1-6. doi: 10.11594/ nstp.2021.0701

Convolutional Neural Network (CNN) is a subset of deep learning and have good achievement in image analysis and image interpretation such as handwritten recognition, Chinese handwritten recognition, animal imaging, and medical imaging. In the 1970s, this network architecture was designed to work with images and been useful applications. In 2014, the winner of ILSRVRC and getting 1st place in both detection image and classification is GoogleNet. The error rate of GoogleNet reaches 6,67% in classification images and an ensemble of 6 GoogleNet has 43,9% mAP on ImageNet test (Szegedy et al., 2014).

Automated DR detection through fundus images has such benefits that DR can be detected early and quickly. For an overview of such methods according to Chorage and Khot (2017) use retinal segmentation to extract various features such as hemorrhages to help the program detect DR and help the doctor do suitable treatment for the patient. To enhance the features the researcher used a green filter and CLAHE algorithm. Then to know the presence of DR, SVM is used to detect it.



Figure 1. Normal fundus image



Figure 2. DR fundus image

This refers to Sarwinda et al., (2017), who investigates texture capabilities from fundus images into three classes of Diabetic Retinopathy, Age-related Macular Degradation, and Normal fundus images. The experiment used two types of dataset there are STARE and DIARETDB0. The experiment designed four models for these two types of datasets and used the algorithm Naïve Bayes, SVM, and KNN for the classifiers.

The proposed system of our method on convolutional neural networks used GoogleNet. We used fundus images that contain normal fundus images and DR fundus images for training data. The database we obtaining from Kaggle Dataset consists of two hundred images that we used for training the model. We build the model to detect Diabetic Retinopathy and normal fundus images. Diabetic Retinopathy fundus images mean that the eye has symptoms of Diabetic Retinopathy from mild until severe and the normal fundus images or No DR means that the eye is healthy and there are no symptoms of Diabetic Retinopathy.

In this paper, we describe the how to GoogleNet architecture classify image to detect Diabetic Retinopathy. The output of this research classifies DR into two classes there are normal or No DR and DR. Python as the programming language in this research along with Pytorch library. The model that we proposed with GoogleNet has a very good performance and has accuracy 88%.

Material and Methods

In this research, GoogleNet was decided for training and learning the model. After we reviewing the literature about image classification using CNN architecture, GoogleNet has methods such as 1x1 convolution and global average pooling that can make the model getting deeper for learning

the database. 1x1 convolution makes the size of the parameter getting smaller significantly. At the end of the network, the GoogleNet has global average pooling. The layer gains a feature map of 7x7 and averages it to 1x1. This part decreases the number of trainable parameters to zero. In Figure 3 and Fig.4 we explain the difference of total number operation with 1x1 convolution and without 1x1 convolution usage.

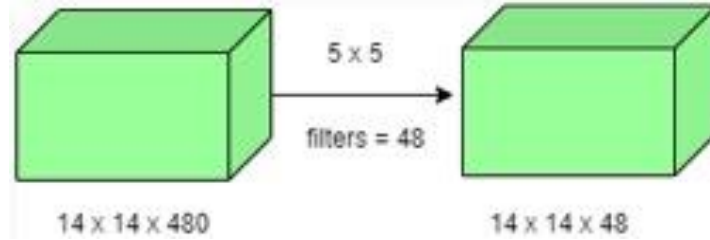


Figure 3. Operation without 1x1 convolution

The operation in Fig. 3 image without 1x1 Convolution is:

$$(14 \times 14 \times 48) \times (5 \times 5 \times 480)$$

Total Number of operations **112.9 M**

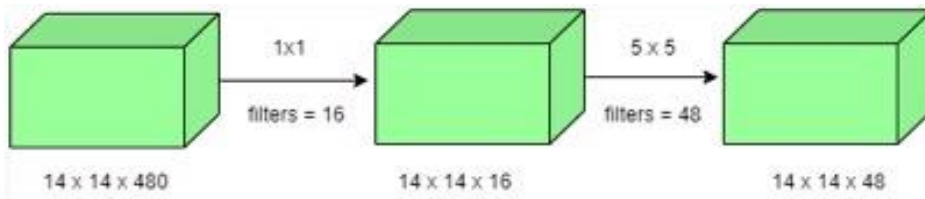


Figure 4. Operation with 1x1 convolution

The operation in Fig. 4 image with 1x1 Convolution is:

$$(14 \times 14 \times 16) \times (1 \times 1 \times 480) + (14 \times 14 \times 48) \times (5 \times 5 \times 16) = 1.5M + 3.8M$$

Total Number of operations **5.3M**

The number operation with 1x1 convolution is much smaller than number operation without 1x1 convolution. By using 1x1 convolution the memory can work efficiently.

In the Alexnet, global average pooling is known as a fully connected layer. The fully connected layers in AlexNet contain the main parameters from architectures and it causes an increase in computational processing. Thus, GoogleNet has global average pooling which has its method to decrease the computation processing and improve the accuracy. On the global average pooling layer, it gains a feature map of 7x7 and averages it to 1x1. Besides that, it decreases the number of trainable parameters to zero.

For fixing convolution size in each layer, GoogleNet has an Inception Module that performed in a parallel way at the input, and the output of these are stacked together to generate final output. Inception modules in the GoogleNet such as 1x1, 3x3, 5x5 convolution, and 3x3 max pooling. The idea of convolution filters of different sizes will handle objects at multiple scales better.

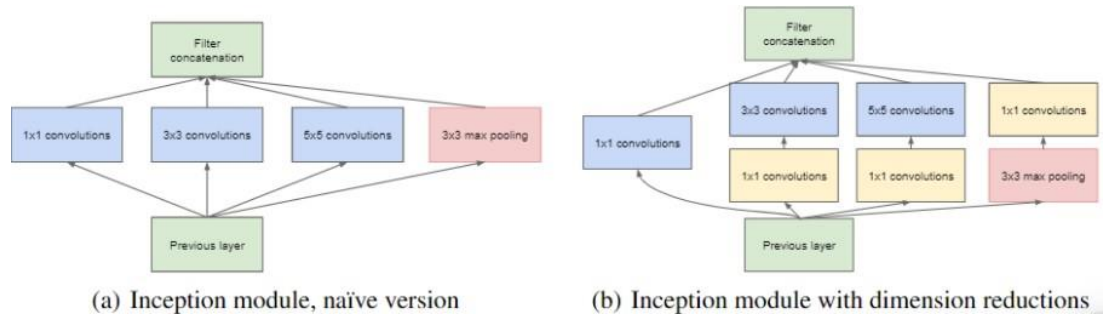


Figure 5. Inception module

The inception architecture used some intermediate classifier branches in the middle of the architecture, these branches are used only during training. These branches consist of a 5×5 average pooling layer with a stride of 3, then 1×1 convolution with 128 filters, two fully connected layers of 1024 outputs and 1000 outputs, and a softmax classification layer. The generated loss of these layers added to total loss with a weight of 0.3. These layers help in combating the gradient vanishing problem and provide regularization too.

The architecture has 22 layers deep to perform the model. It was designed to keep computational efficiency in the process. It can perform in the low computational resources. This architecture has two auxiliary classifier layers that are connected to the output of inception and inception layers. The architecture is shown in Figure 6.

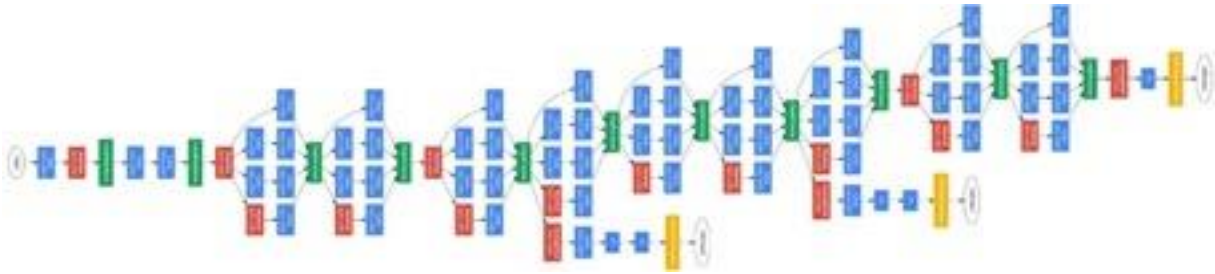


Figure 6. GoogleNet architecture

GoogleNet gains the image size of 224×224 with RGB color channels. All the convolutions inside the architecture use Rectified Linear Units (ReLU) as activations function. The ReLU gains the pixels of the image dataset and if the pixels have a positive value, the value is kept and if the value is negative it will be converted as zero.

Dataset, hardware, and software

The dataset used in the model for training was obtained by the Kaggle dataset and open access at the website (<https://www.kaggle.com>) it contains 200 images which 161 DR fundus images and 39 normal fundus images, the approximate size is 9M pixels per image and scale of fundus images. We resized the image because it is too large and built the program through Google Colaboratory and python programming language along with Pytorch library to perform the model. Then we run the program on a high-end GPU in the Google Colaboratory. We used Lenovo Yoga type 370 Signature Edition with Processor Intel Core i7 Gen 7th and RAM 8 GB.

Image preprocessing

Preprocessing image from a huge part of CNN (Convolutional Neural Network). The dataset contained images from a patient who comes from the Asia Pacific area and has many characteristics. The images were high resolution and need large memory for saving the data. The images are shown in Figure 1 and Figure 2. Then the dataset was resized to 224x224 pixels which fit size on GoogleNet input. We used 200 datasets. There are 39 normal fundus images and 161 DR fundus images. The size of each image is 3152 x 3000 or 9M pixels per image with the type in .jpg format. The images are resized to fit GoogleNet architecture properly. The type color of images is RGB images because RGB is complex to process and requires a lot of time, a greater number of features can be extracted from such images.

Training data

The GoogleNet Architecture is trained only in two normal classes and DR. The normal class indicates there is no disease and the DR class indicates there is an infected eye containing the disease. The model was pre-trained on 120 datasets until it reached a significant level. This was needed 120 minutes for training with 200 epochs. After 200 epochs of training the images, the result getting significantly increase from 67% accuracy to 88% accuracy. The model trained on the full dataset of 200 images for 40 epochs. Our model suffers from over-fitting, especially in a small dataset of normal fundus images. To solve the overfitting, we set the input parameter architecture of GoogleNet from customizing to default GoogleNet architecture that we obtained from the Pytorch library. The GoogleNet was trained using adam optimizer. A low learning rate of 0,001 was used for 40 epochs to stabilize the weight of each image input.

Data Augmentation

The original images were only used for training the network once. After that, we did data-augmentation to improve the localization ability of the network. We augmented images with random rotation from 0-90 degrees, random images to flip horizontally and vertically. The ample of termite nest was collected in September 2016 from Pananjung Pangandaran Nature

Results and Discussion

All of the images we obtained from the Kaggle dataset were saved for validation purposes with a total number of 200. The time-consuming validation network took 108 seconds. For the two classes of Normal fundus images and DR fundus images, we defined three indicators such specificity, sensitivity, and accuracy. Sensitivity is the number of patients correctly identified as having DR out of the true total amount with DR. Specificity is the number of patients correctly identified as not having DR out of the true total amount not having DR. Accuracy is the total number of patients with a correct classification. The resulting training of our model achieved 52% specificity, 75% sensitivity, and 88% accuracy.

The research shows that the DR detection problem can be solved using the GoogleNet model. GoogleNet has shown good signs of learning the feature required to detect fundus images, accurately classifying the mainly of DR and No DR fundus images. High accuracy, specificity, and sensitivity show that GoogleNet is suitable CNN architecture for DR detection through fundus images.

The advantages of using GoogleNet for detecting DR through fundus images are trained the model with efficient memory and low computational to perform. By using Google Colaboratory the performance can increase quickly because of GPU usage. The trained of our model makes a fast diagnosis and correct detection for the patient.

The model has no problem detecting the DR images. It is just because we have a large number of DR fundus images in the dataset. While in the training the model required to classify the images at the extreme ends of the scale was significantly less. Then problems come from unbalancing data so the specificity achieved low than sensitivity and accuracy. In the future scope we have plans to classify DR into five classes such every level on DR can detect correctly and CNN trained to focus

not only detect normal and DR fundus images but also the classify of DR level from mild, moderate, severe, and proliferative. Also, we have plans to obtain datasets from real Indonesia screening setting such as Indonesia hospitals, particularly in eye treatment. The ongoing development of CNN combines with the support algorithm allows much deeper networks to learn the data getting better.

Conclusion

In conclusion, we have shown that CNN has the potential training to identify the feature of DR through fundus images. This paper represents a new model for the detection of diabetic retinopathy disease. Using GoogleNet architecture from CNN, the result gives very good accuracy for 88% besides that achieves 52% specificity, 75% sensitivity. In this paper, we start the model to implement deep learning through fundus images. Besides that, this system was created to help ophthalmologists detect diabetic retinopathy they not focusing on replacing the doctors.

Acknowledgment

The authors wish to thank Faculty of Mathematics and Natural Science, Universitas Indonesia. This work was supported in part by a grant from Publikasi Terindeks Internasional Sains Teknologi dan Kesehatan (PUTI SAINTEKES Q4) 2020 Universitas Indonesia.

References

- Chorage, S., & Khot, S. S. (2017). *Detection of diabetic retinopathy and cataract by vessel extraction from fundus images. International conference on electronics, communication and aerospace technology.*
- Indumathi, G., & Sathananthavath, V. (2019). Microaneurysms Detection for Early Diagnosis of Diabetic Retinopathy Using Shape and Steerable Gaussian Features. *Telemedicine Technologies*, 57-69.
- Jaya, T., Dheeba, J., & Singh, N. A. (2015). Detection of Hard Exudates in Colour Fundus Images Using Fuzzy Support Vector Machine-Based Expert System. *Journal of Digital Imaging*, 28(6), 761-768.
- Joshi, S., & Karule, P. T. (2018). A review on exudates detection methods for diabetic retinopathy. *Jurnal Biomedicine & Pharmacotherapy*, 97 (2018), 1454-1460
- Khojasteh, P., Aliahmad, B., & Kumar, D. K. (2018). Fundus images analysis using deep features for detection of exudates, hemorrhages and microaneurysms. *BMC Ophthalmology*, 275-288.
- Maier, A., Syben, C., Lasser, T., & Riess, C. (2019). A gentle introduction to deep learning in medical image processing. *ScienceDirect*, 1-16.
- Philip, S., Fleming, A. D., Goatman, K. A., Fonseca, S., McNamee, P., Scotland, G. S., Prescott, G. J., Sharp, P. F., & Olson, J. A. (2007). The efficacy of automated disease/no disease grading for diabetic retinopathy in a systematic screening programme. *British Journal Optalmology*, 91(11),1512-1217.
- Sarwinda, D., Bustamam, A., & Arymurthy, A. M. (2017). *Fundus Image Texture Features Analysis in Diabetic Retinopathy Diagnosis. International Conference on Sensing Technology.*
- Seoud, L., Hurtut, T., Chelbi, J., Cheriet, F., Langlois, J. M. P. (2015). *Red lesion detection using diabetic retinopathy screening. IEEE transactions on Medical Imaging.*
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2014). *Going Deeper with Convolution. ILSVRC.*
- WHO. Diabetes. 30 Oktober 2018 <https://www.who.int/news-room/fact-sheets/detail/diabetes>. diakses pada 21 Desember 2019.