

Eye Abnormality Automatic Detection using Deep Learning based Model

Audyati Gany^{1,*}^a, Meilan Jimmy Hasugian^{1,2}^b, Erwani Merry Sartika¹^c,
Novie Theresia Br. Pasaribu¹^d and Hannah Georgina¹^e

¹Department of Electrical Engineering, Maranatha Christian University, Surya Sumantri 65, Bandung, Indonesia

²Department of Electrical Engineering, Chung Yuan Christian University, Zhongli, Taiwan

*Corresponding author

Keywords: Ocular Disease, Fundus Retina, Deep Learning, Convolutional Neural Network.

Abstract: Early detection and diagnosis of ocular pathologies would enable to forestall of visual impairment. One challenge that limits the adoption of a computer-aided diagnosis tool by the ophthalmologist is, the sight-threatening rare pathologies such as central retinal artery occlusion or anterior ischemic optic neuropathy and others are usually ignored. The aim of this research is to develop methods for automatic detection of eye abnormality caused by the most common ocular disease along with the rare pathologies. For this purpose, we developed the deep learning-based model trained with Retinal Fundus Multi-disease Image Dataset (RFMiD). This dataset consists of a 1920 fundus retina images captured using three different fundus cameras with 46 conditions annotated through adjudicated consensus of two senior retinal experts. The model is built on the top of some prominent pretrained convolutional neural network (CNN) models. From the experiment, the model could achieve the accuracy level and recall 0.87, whereas precision and F1 score are 0.86, and area under receiver operating characteristic (AUROC) is 0.90. The proposed model built in deep learning structure could be a promising model in automatic classification of ocular disease based on fundus retina images.


1 INTRODUCTION


In the World Report on Vision 2019, (WHO, 2019) stated that approximately 2.2 billion people worldwide have visually impaired, of whom at least 1 billion have a vision impairment that could have been prevented. The world faces considerable challenges in terms of eye care, including inequalities in the coverage and quality of prevention, treatment, and rehabilitation services. Early detection and diagnosis of ocular pathologies would enable to forestall of visual impairment. One challenge that limits the adoption of a computer-aided diagnosis tool by the ophthalmologist is, the sight-threatening rare pathologies such as central retinal artery occlusion or


anterior ischemic optic neuropathy and others are usually ignored.


In the past two decades, many publicly available datasets of color fundus images have been collected with a primary focus on diabetic retinopathy, glaucoma, and age-related macular degeneration, and few other frequent pathologies. There are several researches that use the fundus retina images as the intake in their system (Qummar, et al., 2019) (Soomro, et al., 2019) (Sarki, Ahmed, Wang, & Zhang, 2020). However, most of them applied for diabetic retinopathy disease.


This study is a preliminary research that aims to develop methods in automatic detection of eyes abnormality caused by not only a frequent ocular disease but also along with the rare pathologies. For

^a <https://orcid.org/0000-0002-7389-6667>

^b <https://orcid.org/0000-0003-0759-6663>

^c <https://orcid.org/0000-0003-3720-3584>

^d <https://orcid.org/0000-0001-7774-9675>

^e <https://orcid.org/0000-0002-5885-3503>

this purpose, we developed the deep learning-based model trained with Retinal Fundus Multi-disease Image Dataset (RFMiD).

Therefore, the contribution of this research is its capability in auto detection of eye abnormality for not only most common ocular disease such as diabetic retinopathy, but also for other rare ocular diseases.

2 METHODS AND MATERIALS

2.1 Fundus Retina Datasets

In this research we deployed RFMiD dataset as provided in Retina Image Analysis Challenge (Pachade, et al., 2021). These datasets just published in April 2021. Total images provided during our experiment are 1920 fundus images. These images were captured using three different fundus cameras with 46 conditions annotated through adjudicated consensus of two senior retinal experts. The advantage of this dataset is its wide variety of diseases that appear in clinical settings.

Several diseases/abnormalities that are included in the dataset: diabetic retinopathy (DR), age-related macular degeneration (ARMD), media haze (MH), drusen (DN), myopia (MYA), branch retinal vein occlusion (BRVO), tessellation (TSLN), epiretinal membrane (ERM), laser scar (LS), macular scar (MS), central serous retinopathy (CSR), optic disc cupping (ODC), and many more total 42 types of eye abnormalities. Visualization of these disease is shown Figure 1.

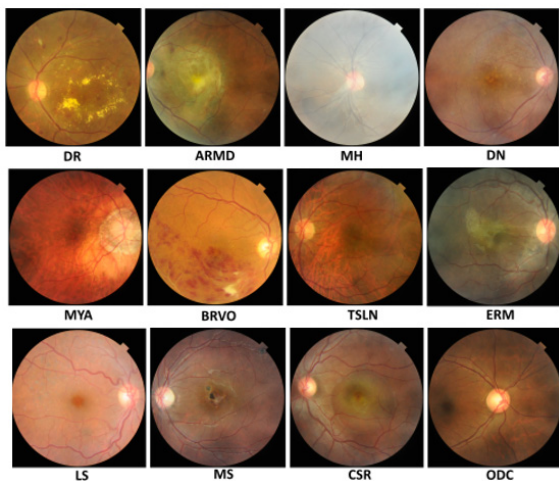


Figure 1: Fundus retina images for several eye diseases

The data distribution between the normal eyes and eyes with abnormalities is presented in the diagram below.

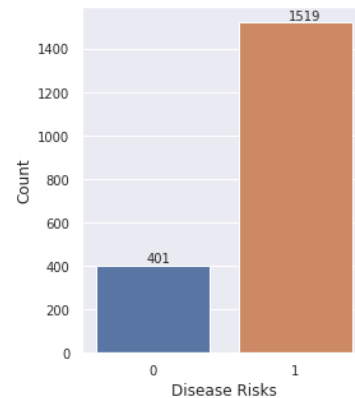


Figure 2: Distribution between normal and abnormal eyes.

Around 20.9% (401 images) are normal eyes and the rest of it are from eyes with several diseases.

2.2 Method

The usage of deep learning in computer vision problem especially in biomedical imaging analysis become more compelling (Liu, et al., 2018), (Tajbakhsh, Shin, Gurudu, & Hurst, 2016) (Roy, et al., 2020). In this paper, we report a deep-network based model to develop a system for detecting eye abnormality automatically. The model was built on the top of some prominent pretrained convolutional neural network (CNN) e.g., VGG-16 and Inception-v3. These networks were trained by using ImageNet dataset of over 14 million images belonging to 1000 classes. Therefore, it's very suitable for common computer vision problem. However, for fundus retina images, we deployed the network in a transfer learning fashion and without or with image-augmentation to improve the generalization of the model in detecting eyes abnormality.

The basic structure of VGG16 model (Simonyan & Zisserman, 2015) is depicted in Figure 3. This model consists of 16 layers, with 13 of convolutional layers and 3 fully connected (FC) layers.

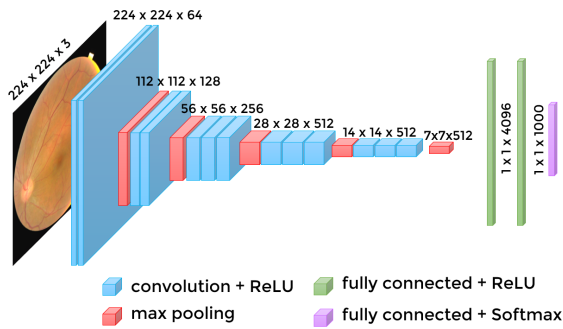


Figure 3: The basic structure of VGG16 model.

This fully connected layers parts of the model was adjusted for the purpose of this research. The complete configuration of the model, without and with image-augmentation process is reported in Table 1 and Table 2 respectively.

Table 1: The configuration of the fully-connected layers in VGG16, without image-augmentation.

Layers	Model A1	Model A2
FC 1	1024, ReLU	1024, ReLU
Drop out	n.a.	0.5
FC 2	512, ReLU	512, ReLU
Drop out	0.5	0.5
Output	1, sigmoid	1, sigmoid

We deployed Rectifier Linear Unit (ReLU) as the activation function in FC1 and FC2, and sigmoid activation function at the output, for binary classification system.

Table 2: The configuration of the fully-connected layers in VGG16, with image-augmentation.

Layers	Model B1	Model B2	Model B3	Model B4
FC 1	512, ReLU	512, ReLU	512, ReLU	512, ReLU
Drop out	0.3	n.a.	0.3	n.a.
FC 2	512, ReLU	512, ReLU	512, ReLU	512, ReLU
Drop out	n.a.	0.3	0.3	n.a.
Output	1, sigmoid	1, sigmoid	1, sigmoid	1, sigmoid

To avoid overfitting during the training, drop-out technique was used as the regularization for the deep learning model.

The basic structure of Inception-v3 network (Szegedy, Vanhoucke, Ioffe, Shlens, & Wojna, 2016) is illustrated in Figure 4. This model contains 42 layers. The detail structure in Stem, Inception A,

Reduction A, Inception B, Reduction B, and Inception C blocks are depicted in Figure 5-10. This model employs the Inception network (Szegedy, et al., 2015).

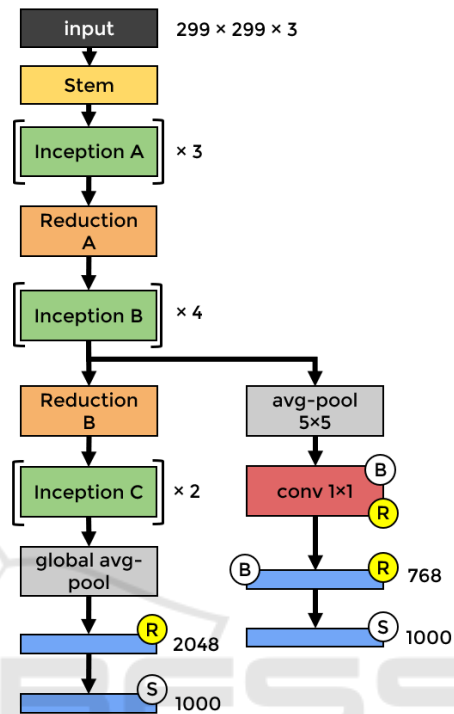


Figure 4: The basic structure of Inception-v3 model.

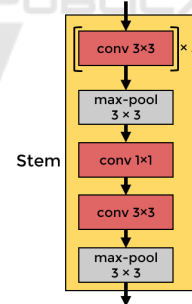


Figure 5: The Stem block in Inception-v3 model.

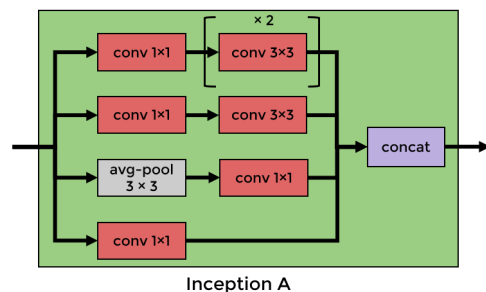


Figure 6: The Inception A block in Inception-v3 model.

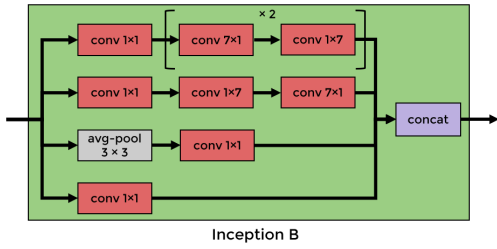


Figure 7: The Inception B block in Inception-v3 model.

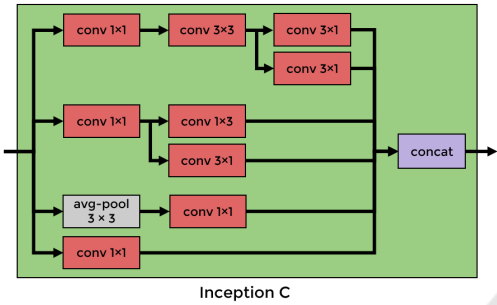


Figure 8: The Inception C block in Inception-v3 model.

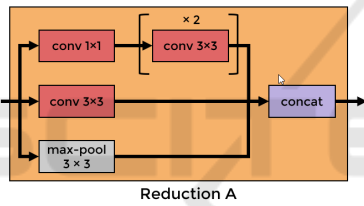


Figure 9: The Reduction A block in Inception-v3 model.

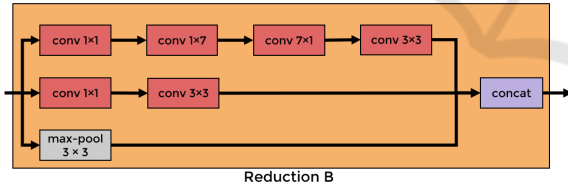


Figure 10: The Reduction B block in Inception-v3 model.

The fully connected layers parts of this model was modified in order to be applicable in our purpose. The complete configuration of the model, without and with image-augmentation process is described in Table 3 and Table 4, respectively.

Table 3: The configuration of the fully-connected layers in Inception-v3, without image-augmentation.

Layers	Model C1
FC 1	4096, ReLU
Drop out	0.5
FC 2	512, ReLU
Drop out	0.3
Output	1, sigmoid

Table 4: The configuration of the fully-connected layers in Inception-v3, with image-augmentation.

Layers	Model D1	Model D2	Model D3	Model D4	Model D5
FC 1	512, ReLU	512, ReLU	512, ReLU	512, ReLU	1024, ReLU
Drop out	n.a.	n.a.	0.3	0.3	n.a.
FC 2	512, ReLU	512, ReLU	512, ReLU	512, ReLU	512, ReLU
Drop out	n.a.	0.3	n.a.	0.3	0.3
Output	1, sigmoid	1, sigmoid	1, sigmoid	1, sigmoid	1, sigmoid

From the total 1920 fundus images, is split into 1305 for training images, 231 images for validation data, and 384 for testing data. Stratified sampling method is used for splitting the images, on order to maintain the similar distribution for each class (normal and abnormal eyes).

In order to improve the ability of the model to preserve the generalization principles, we applied data augmentation process. The parameter for this augmentation is described in Table 5.

Table 5: Parameters for image-augmentation process.

Parameter	Value
rotation angle	30
width shift	0.2
height shift	0.2
range	
shear range	0.1
horizontal flip	True

2.3 Evaluation Metrics

The performance of the model is measured in terms of the ability of the model to correctly detect the eyes with abnormality and the normal eyes as true positive (TP) and true negative (TN) respectively. And misjudge the normal eye to be abnormal as well as abnormal eye classified as normal as false positive (FP) and false negative (FN) respectively.

The metrics for evaluation are precision (Pre), recall (Re), accuracy (Acc), and F1 score.

$$\text{Pre} = \text{TP} / (\text{TP} + \text{FP}) \quad (1)$$

$$\text{Re} = \text{TP} / (\text{TP} + \text{FN}) \quad (2)$$

$$\text{Acc} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN}) \quad (3)$$

$$\text{F1-score} = 2 * \text{Pre} * \text{Re} / (\text{Pre} + \text{Re}) \quad (4)$$

And another metrics that are prevalent in binary classification is area under receiver operating curve (AUROC) or sometimes called as AUC for

simplicity. That is the area under the curve between false positive rate and true positive rate.

3 RESULTS AND DISCUSSION

The result for each model based on VGG16 structure is reported in Table 6. It can be seen in the table that model A1 outperformed all other models based on VGG16 network in every evaluation metric.

Table 6: Model performance based on VGG16 network.

Model	Acc	Pre	Re	F1	AUC
A1	0.87	0.86	0.87	0.86	0.87
A2	0.85	0.84	0.85	0.84	0.84
B1	0.86	0.85	0.86	0.85	0.84
B2	0.84	0.83	0.84	0.82	0.83
B3	0.83	0.81	0.83	0.81	0.83
B4	0.84	0.83	0.84	0.83	0.84

The comparison of ROC curve for each model is presented in Figure 11. Figure 11 shows that the AUC of model A1 is more superior than the other models.

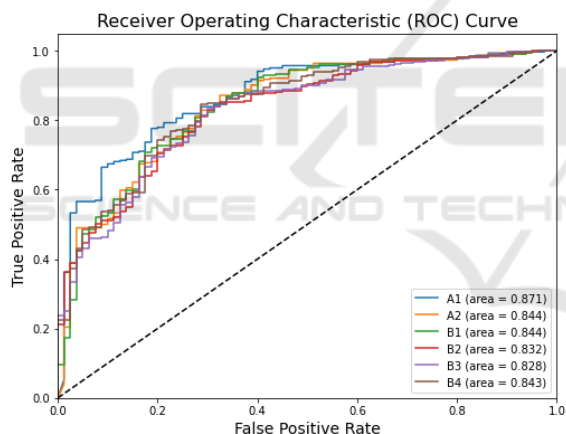


Figure 11: The comparison of ROC curve result for each model based on VGG16 structure.

Table 7 presents all metrics for each model based on Inception-v3 structure.

Table 7: Model performance based on Inception-v3 network.

Model	Acc	Pre	Re	F1	AUC
C1	0.85	0.84	0.85	0.84	0.89
D1	0.86	0.85	0.86	0.85	0.90
D2	0.85	0.85	0.85	0.85	0.90
D3	0.86	0.85	0.86	0.85	0.90
D4	0.86	0.85	0.86	0.85	0.89
D5	0.83	0.84	0.83	0.84	0.89

In Table 7, model D1 and D3 shows a same result. Both surpass the other models in every metric.

The ROC curve comparison between all models based on Inception-v3 is shown in Figure 12.

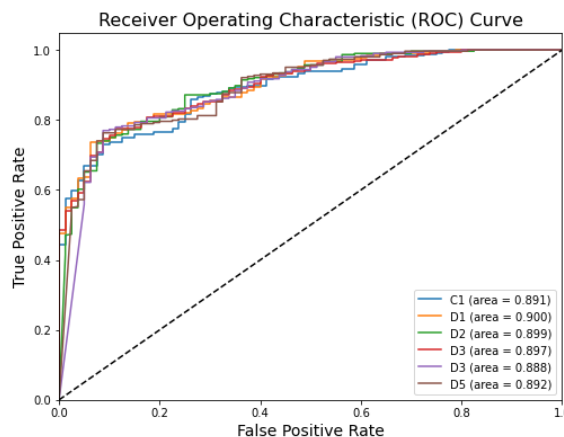


Figure 12: The Reduction B block in Inception-v3 model.

Figure 12 shows that the AUC of each model is different not too significant. Therefore, it can be said, the variation of fully connected layers do not affect the results essentially.

VGG16 structure is very simple compared to Inception-v3. The number of neurons in FC layers layer should be able to accommodate the features as the output of the convolutional layer. Therefore, the sudden change in numbers will affect the performance. This causes the model A1 outperforms more than the other model. The image augmentation process seems not to help very much to improve the performance.

On the other hand, Inception-v3 structure is already complex. The variation of neuron numbers in FC layers would not affect significantly. As it is seen in Figure 12, all models have quite similar ROC. However, the image augmentation gives more impact to the performance.

To the best of our knowledge, since these datasets were launched in April 2021, we haven't found yet any reported publication that used the same dataset in their model. Therefore, at this current moment we couldn't provide any comparison from others' model that are trained with the same dataset.

4 CONCLUSIONS

From the experiment it can be concluded that VGG16 based model is not sensitive in image augmentation but to the fully connected layers. On the other hand, Inception-v3 based model is more impacted by image

augmentation. However, by analysing the ROC curve, both structures are still promising in development of eye abnormality detection in further research.

REFERENCES

- Liu, J., Pan, Y., Li, M., Chen, Z., Tang, L., Lu, C., & Wang, J. (2018, Mar). Applications of deep learning to MRI images: a survey. *Big Data Mining and Analytics*, 1(1).
- Pachade, S., Porwal, P., Thulkar, D., Kokare, M., Deshmukh, G., Sahasrabudhe, V., . . . Mériaudeau, F. (2021). Retinal fundus multi-disease image dataset (RFMiD): a dataset for multi-disease detection research. *Data*, 6(2). doi:https://doi.org/10.3390/data6020014
- Qummar, S., Khan, F. G., Shah, S., Khan, A., Shamsheerband, S., Rehman, Z. U., . . . Jadoon, W. (2019). A deep learning ensemble approach for diabetic retinopathy detection. *IEEE Access*, 7, 150530-150539.
- Roy, S., Menapace, W., Oei, S., Luijten, B., Fini, E., Saltori, C., . . . Demi, L. (2020, Agt). Deep learning for classification and localization of COVID-19 markers in point-of-care lung ultrasound. *IEEE Trans. on Medical Imaging*, 39(8).
- Sarki, R., Ahmed, K., Wang, H., & Zhang, Y. (2020). Automatic detection of diabetic eye disease through deep learning using fundus images: a survey. *IEEE Access*, 151133 - 151149.
- Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. *International Conference on Learning Representations (ICLR)*. San Diego.
- Soomro, T. A., Afifi, A. J., Zheng, L., Soomro, S., Gao, J., Hellwich, O., & Paul, M. (2019). Deep learning models for retinal blood vessels segmentation: a review. *IEEE Access*, 71696 - 71717.
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., . . . Rabinovich, A. (2015). Going deeper with convolutions. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (pp. 2818-2826).
- Tajbakhsh, N., Shin, J. Y., Gurudu, S. R., & Hurst, R. T. (2016, May). Convolutional neural networks for medical image analysis: full training or fine tuning. *IEEE Trans. on Medical Imaging*, 35(5).
- WHO. (2019). *World report on vision*. World Health Organization.