International Journal of Industrial Engineering & Production Research September 2024 Vol. 35, No. 3: 1-12
DOI: 10.22068/ijiepr.35.3.1935

RESEARCH PAPER

Comparative Analysis of Pre-trained CNN Models on Retinal Diseases Classification

Theodore Alvin Hartanto[1] & Seng Hansun[2*]

Received 7 December 2023; Revised 29 January 2024; Accepted 9 June 2024;

ABSTRACT

One method to diagnose retinal diseases is by using the Optical Coherence Tomography (OCT) scans. Annually, it is estimated that around 30 million OCT scans are performed worldwide. However, the process of analyzing and diagnosing OCT scan results by an ophthalmologist requires a long time so machine learning, especially deep learning, can be utilized to shorten the diagnosis process and speed up the treatment process. In this study, several pre-trained deep learning models are compared, including EfficientNet-B0, ResNet-50V2, Inception-V3, and DenseNet-169. These models will be fine-tuned and trained with a dataset containing OCT scanned images to classify four retinal conditions, namely Choroidal Neovascularization (CNV), Diabetic Macular Edema (DME), Drusen, and Normal. The models that have been trained are then tested to classify the test set and the results are evaluated using a confusion matrix in terms of accuracy, recall, precision, and F1-score. The results show that the model with the best classification results in the batch size of 32 scenario is the ResNet-50V2 model with an accuracy value of 98.24%, precision of 98.25%, recall of 98.24%, and F1-score of 98.24%. While for the batch size of 64, the EfficientNet-B0 model is the model with the best classification results with an accuracy value of 96.59%, precision of 96.84%, recall of 96.59%, and F1-score of 96.59%.

KEYWORDS: Classification; Comparative analysis; Deep learning; Optical coherence tomography; Pre-trained model; Retinal.Introduction.

1. Introduction

Artificial Intelligence (AI) is the simulation of computers and machines to imitate the problem solving and decision-making abilities of the human mind. A difficult and time-consuming task if done by human, can be completed in a shorter time by AI techniques. Therefore, in the current industrial era 4.0, the existence of AI is increasingly needed in various fields of work to increase productivity and efficiency, including in the medical field [1].

Currently, the process of diagnosing retinal diseases is commonly done by using the Optical Coherence Tomography (OCT) scan results. OCT is an imaging technique that is widely used by ophthalmologists to obtain high-resolution images of the retina of the eye. According to Swanson and Fujimoto [2], an estimated of 30 million OCT scans are performed annually. This OCT scan image is used to diagnose various retinal eye diseases suffered by patients, such as Choroidal Neovascularization (CNV), Diabetic Macular Edema (DME), and Drusen. However, the problem is that the process of analyzing and diagnosing OCT results manually by a specialist (ophthalmologist) takes quite a long time [2].

By using Machine Learning (ML) techniques, the process of diagnosing diseases in the patient's retina can be done in a shorter time, so that it will speed up the process of treating patients. Currently, there are many pre-trained ML models that can be used to perform image-based classification. However, each model has a different level of accuracy and training time. Therefore, it is necessary to do a comparison between pre-trained models to find out the model that can classify retinal diseases accurately and efficiently.

In a similar study, Wang et al. [3] discussed the classification of retinal OCT scan results using the ResNet-50 model and produced an accuracy of 96.87%. Another study from Kamble et al. [4] used the Inception-V3 model for the detection of Diabetic Macular Edema (DME) using retinal

* Corresponding author: Seng Hansun
seng.hansun@lecturer.umn.ac.id

1. *Informatics Department, Universitas Multimedia Nusantara, Tangerang, Indonesia*.
2. *Informatics Department, Universitas Multimedia Nusantara, Tangerang, Indonesia*.

OCT and obtained an accuracy of 92.77%. Then in Abirami et al. [5] study, DenseNet-169 model was used in the detection of Choroidal Neovascularization (CNV) disease using retinal OCT and obtained an accuracy of 95.25%.

From these studies, it can be seen that the use of different pre-trained models can result in different levels of accuracy. Therefore, this research goal is to find out the pre-trained model that can classify retinal diseases accurately and efficiently. Some of the pre-trained models that will be used in this research are ResNet-50V2, Inception-V3, DenseNet-169, and EfficientNet-B0.

ResNet-50V2, Inception-V3, and DenseNet-169 were chosen because they performed quite well in similar studies conducted previously [3]–[5]. Meanwhile, EfficientNet-B0 was chosen because according to the research by Tan and Le [6], this model can provide similar or even better performance than other Deep Learning models, but with fewer parameters.

These pre-trained models will learn from a collection of retinal OCT scan images that have been grouped into four category labels, namely Choroidal Neovascularization (CNV), Diabetic Macular Edema (DME), Drusen, and Normal. Fine-tuning will be performed on all the compared models so that each model can specifically classify retinal diseases. In addition, several callbacks are also implemented, such as early stopping, model checkpoint, and reduce learning on plateau in order to obtain optimal training results for all the models being compared.

After the model training process is complete, the model will classify a set of images from the OCT scan on the test set. Then, the classification results carried out by the model will be matched with the actual labels in the dataset by using a confusion matrix. The result of the confusion matrix will show which pre-trained model has the best performance in the classification of retinal diseases.

In this study, we conducted a comparative analysis on four different pre-trained Convolutional Neural Networks (CNN) models for retinal disease classification. Although some of these models had been utilized in previous studies, they were used to detect only one retinal disease with one specific model [3]–[5]. Therefore, an endeavor to use and to compare several models together for several retinal diseases is worth to be explored. Further explanation regarding the data and methods used in this research will be given in Section 2, while the experimental results will be discussed in Section 3.

2. Data and Methods

2.1. Dataset

The dataset used in the study was obtained through Kaggle and provided by Kermany, Zhang, and Goldbaum [7]. It contains retinal OCT scans of adult patients from the Shiley Eye Institute of the University of California San Diego, the California Retinal Research Foundation, Medical Center Ophthalmology Associates, the Shanghai First People's Hospital, and Beijing Tongren Eye Center between 1 July 2013 to 1 March 2017 [8]. The dataset is divided into two groups, namely the training set which contains 4,016 retinal images resulting from OCT scans, and the test set containing 968 retinal images resulting from OCT scans. The existing data have been grouped into four different labels according to the condition of the retina, namely Choroidal Neovascularization (CNV), Diabetic Macular Edema (DME), Drusen, and Normal.

2.2. Retinal diseases

The retina is a sensory membrane that lines the inner surface behind the eyeball [9]. The retina is made up of several layers, including one that contains specialized cells called photoreceptors. There are two types of photoreceptor cells in the human eye, namely rods and cones. Rod photoreceptors play a role in detecting motion, provide black-and-white vision and function well in low light, while cone photoreceptors play a role in central vision, color vision and function well in moderate and bright light.

Like other parts of the body, the retina is also prone to various problems and diseases, such as Choroidal Neovascularization (CNV), Diabetic Macular Edema (DME), and Drusen. One way to diagnose the diseases is by using an Optical Coherence Tomography (OCT) scan which can obtain cross-sectional images of the patient's retina with high resolution [2]. Fig. 1 shows an example of OCT scan on patient with DME.

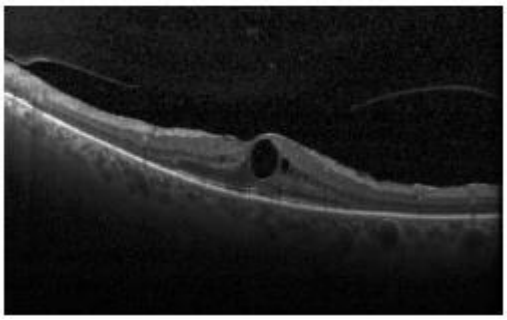


Fig. 1. OCT scan on patient with DME [7]

2.2.1. Choroidal neovascularization (CNV)

Choroidal Neovascularization (CNV) is the formation of new blood vessels in the choroid layer, the layer that contains blood vessels under the retina. These new, abnormal blood vessels leak easily and allow fluid from the blood, and sometimes even red blood cells, to enter the retina. This fluid can directly impair vision because it forms "blisters" on the retina, which are usually flat. Over days to months, this fluid can damage the retina, killing light-sensing cells, called photoreceptors [10].

Some of the consequences that can arise from CNV are macular degeneration, pathological myopia (extreme nearsightedness), ocular histoplasmosis, and severe eye inflammation (uveitis). A person's CNV disease can be diagnosed using an OCT scan. This scanned image can detect up to a small amount of fluid leaking into the retina as a result of the CNV [10].

2.2.2. Diabetic macular edema (DME)

Diabetic Macular Edema (DME) is a complication of diabetes caused by fluid accumulation in the macula that can affect the fovea. The macula is the central part of the retina at the back of the eye and is where vision is sharpest. Vision loss from DME can develop over several months and make it difficult for the sufferer to focus clearly [11].

DME usually occurs in people with diabetes, both type one and type two. Persistently high blood sugar due to poor glucose control over time can weaken the tiny blood vessels in the eye, allowing fluid to leak into the retina and then on to the macula [11]. The diagnosis of DME can be made using an OCT scan. The scan results can detect up to a small amount of fluid and swelling.

2.2.3. Drusen

Drusen are small yellow deposits of fatty proteins (lipids) that build up under the retina. Drusen consist of two different types, namely soft drusen which tend to be large and grouped at one point, and hard drusen which tend to be small and scattered [12]. Hard varieties usually cause no problems and require no treatment. On the other hand, the soft type can injure and cause bleeding in the macular cells.

Drusen does not cause total blindness, but can result in loss of central vision. Central vision allows us to focus on details straight ahead. Several factors that can lead to the emergence of drusen include genetic factors (inherited from the family), smoking, cardiovascular disease, high cholesterol, and can also be caused by age factors (susceptible to occur at the age of 60 years and over) [12]. The appearance of soft drusen can be detected using an OCT scan. The results of the OCT scan will show the presence of a collection of soft drusen that has accumulated on the patient's retina.

2.3. Convolutional neural network (CNN)

Convolutional Neural Network (CNN) is a deep learning architecture that is used for image processing, such as image classification, instance segmentation, semantic segmentation, and synthetic image generation [13]. The initial step is to perform low-level feature extraction from raw input, which is then followed by high-level feature extraction in deeper layers. CNN is designed to process data in array form. In terms of the layer structure contained in CNN, in general CNN consists of three types of layers, namely Convolutional layer, the Pooling layer, and the Fully Connected layer.

2.3.1. EfficientNet-B0

EfficientNet is a CNN architecture published by Google in one of its journals in 2019. This model not only increases accuracy but also increases model accuracy by reducing parameters and Floating Point Operations Per Second (FLOPS) [6]. EfficientNet consists of several models, including B0, B1, B2, B3, B4, B5, B6 and B7. Of these models, B0 is the model that uses the fewest parameters compared to other models. EfficientNet-B0 architecture can be seen in Table 1.

Tab. 1. EfficientNet-B0 architecture [6]

Layer	Operator	Resolution	Channel	Layers
1	Conv 3x3	224x224	32	1
2	MBConv1 3x3	112x112	16	1
3	MBConv6 3x3	112x112	24	2
4	MBConv6 5x5	56x56	40	2
5	MBConv6 3x3	28x28	80	3
6	MBConv6 5x5	14x14	112	3
7	MBConv6 5X5	14x14	192	4
8	MBConv6 3x3	7x7	320	1
9	Conv 1x1 & Pooling & FC	7x7	1280	1

2.3.2. ResNet-50V2

Residual Network (ResNet)-50V2 is a CNN architecture designed by Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun in 2016. The number 50 in ResNet-50V2 shows the layer depth possessed by this model [14]. ResNet-50V2 is a development from the previous version, namely ResNet-50. The difference between ResNet-50V2 and ResNet-50 lies in the use of residual blocks which no longer use post-activation (ResNet-50), but pre-activation (ResNet-50V2), thus allowing the layer to return to the previous layer because it has the same layer shape [15]. Moreover, ResNet-50V2 applies Batch Normalization and ReLU activation to the input before multiplication with the weight matrix (convolution operation). In ResNet-50, this process is performed after the convolution operation [15]. Fig. 2 illustrates the differences between ResNet-V1 and ResNet-V2 unit.

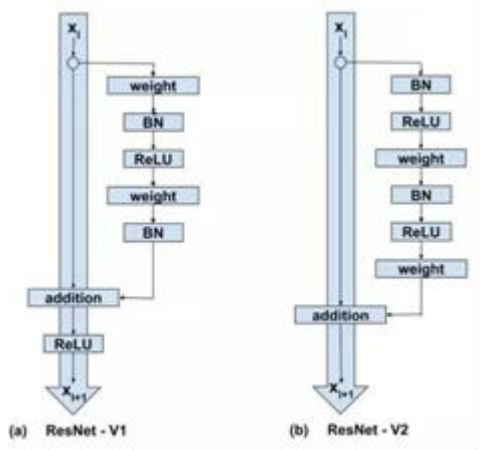


Fig. 2. ResNet-V1 and ResNet-V2 residual unit [15]

2.3.3. DenseNet-169

Densely Connected Convolutional Networks (DenseNet) is a CNN architecture designed by Gao Huang, Zhuang Liu, Laurens van der Maaten, and Kilian Q. Weinberger in 2016 [16]. In DenseNet, each layer gets additional input from all the previous layers and passes the feature-map to all subsequent layers. Therefore, each layer receives "collective knowledge" from all the previous layers. As each layer receives a feature-map of all the previous layers, the network becomes thinner and denser and thus has higher computational and memory efficiency. DenseNet-169 is a DenseNet variant with a layer depth of 169 [16]. The architectural diagram for DenseNet is shown in Fig.3.

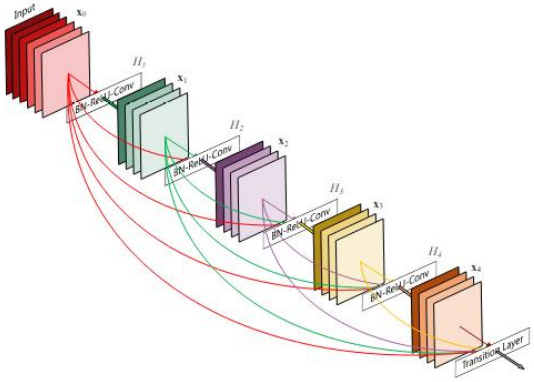


Fig. 3. DenseNet architectural diagram [16]

2.3.4. Inception-V3

Inception-V3 is a CNN architecture developed by Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens, and Zbigniew Wojna in 2016 [17]. Inception-V3 is a development of the previous Inception module with the aim of increasing accuracy in ImageNet classification. In the development of the early versions of Inception modules (Inception-V1 and Inception-V2), this module has been proven to be more computationally efficient, both in terms of the number of parameters generated by the network and the costs incurred (memory and other resources).

Therefore, in the development of Inception-V3, the main focus was to increase accuracy without compromising the efficiency generated by the Inception module. To maintain this efficiency, several techniques were applied in the development of Inception-V3 to optimize the network, such as factorized convolutions, regularization, dimension reduction, and parallelized computations [17]. Fig. 4 depicts the architectural diagram for Inception-V3.

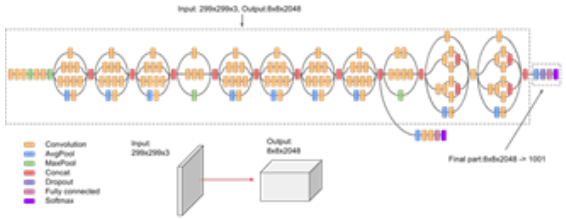


Fig. 4. Inception-V3 architectural diagram [17]

2.4. Confusion matrix

Confusion matrix consists of two dimensions, namely the index of the actual label owned by an object, and the index of the label classified by the model [18]. The data contained in the confusion matrix can be used to evaluate the performance of the resulting classification model. Table 2 shows the conceptual representation of the confusion matrix.

Tab. 2. Confusion matrix concept

Actual	Prediction	
	Positive	Negative
Positive	True Positive	False Negative
Negative	False Positive	True Negative

As shown in Table 2, True Positive (TP) is the value of the classification that is done correctly that an object has a positive label. False Positive (FP) is the value of a classification that is done incorrectly that an object has a positive label, even though the actual label is negative. Then, False Negative (FN) is the value of a classification that is done incorrectly that an object has a negative label, even though the actual label is positive. While True Negative (TN) is the value of the classification that is done correctly that an object has a negative label. The results of the confusion matrix can be used to calculate several metrics, such as accuracy, precision, recall, and F1 score.

Accuracy is the value obtained from the comparison of the correct predictions made by the model with the total predictions made. Precision is the value obtained by comparing the positive

predictions correctly made by the model with all positive predictions made by the model. Recall is the value obtained from the comparison of positive predictions correctly made by the model with all labels that are indeed positive. F1 score is the average value (mean) of precision and recall. The accuracy and capability of the classification generated by the model can be measured by using the F1 Score [19]. Equations (1) to (4) show the accuracy, precision, recall, and F1 score formulation respectively [20], [21].

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

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

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

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

3. Results and Discussion

3.1. Experimental setup and steps

In conducting the research, we used Python v3.9.5 programming language with several packages such as Numpy, Pandas, OS, Kaggle, Tensorflow, Matplotlib, sklearn, Seaborn, and Image. Google Collaboratory was used as the platform for training and testing of pre-trained models.

Before the dataset is processed and trained on the pre-trained model, it is necessary to define several hyperparameters. With references from [3], [4], the hyperparameters that will be used in this study include two batch_size scenarios of 32 and 64, learning_rate of 0.001, and total_epoch of 20. The specified hyperparameters will be the same for all compared models.

The next step is to split data for training and validation. In the training dataset, 75% of the data (3,013 images) will be used for model training, while 25% of the data (1,003 images) will be used for validation. For the test dataset will not be split so that the number remains the same (968 images). The split of the training dataset is done using the validation_split argument available in Tensorflow's "ImageDataGenerator" library. The library also be used to rescale image pixels in the dataset from a range of 0 to 255, to a range of 0 to 1 so that the total loss generated between images with high and low pixel ranges will be more evenly distributed. Table 3 detailed the data splitting results in this study.

Tab. 3. Distribution of the amount of data in the dataset after split

Type	CNV	DME	Drusen	Normal	Amount
Train	754	753	753	753	3013
Valid	250	251	251	251	1003
Test	242	242	242	242	968

After the splitting of the dataset is successful, the next step is to read the images in the dataset. The image being read will be set to a size of 224 x 224 pixels. The reason is that the image input size must be compatible with the image input size used by the pre-trained model in its training [22]. The models EfficientNet-B0, ResNet-50V2, Inception-V3, and DenseNet-169 each use an image input size of 224 x 224 pixels so that the input image size for each model is also 224 x 224 pixels. In addition, setting the image to this size can also speed up computing time. The image reading process will use "ImageDataGenerator" which was initialized in the previous stage for dataset sharing and rescaling. After that, the "flow_from_directory" method is used which will access image data stored in a folder.

The pre-trained models that will be compared, namely ResNet-50V2, DenseNet-169, Inception-V3, and EfficientNet-B0 will be loaded into the Google Collaboratory. The first step to load the pre-trained model is to import the model library. After importing the library, then proceed with declaring the model. The same parameters will be applied to all models in order to see the effect of differences in the pre-trained models used.

The next stage is fine-tuning all pre-trained models so that they can specifically classify retinal diseases using OCT scanned images. In this fine-tuning process, several layers will be added to the pre-trained model, which include flatten layer, dense layer, dropout layer, and also an output layer. The fine-tuned model will then be compiled. Table 4 shows the model structure that has been fine-tuned.

Tab. 4. Model structure after fine-tuning process

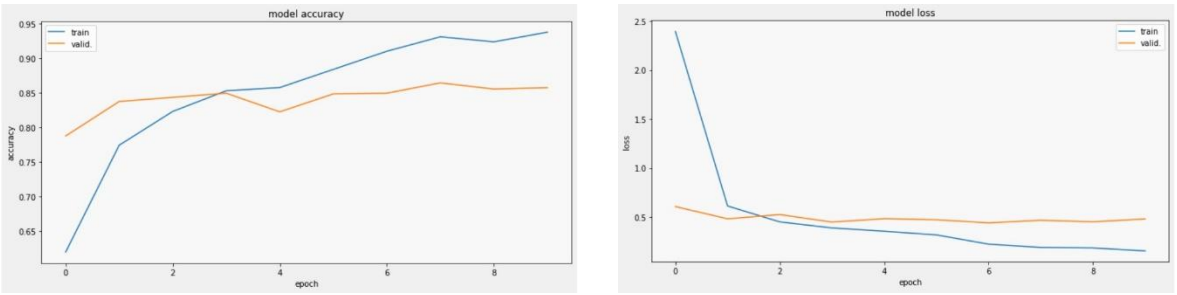
Layer	Description
Base model	Pre-trained model used in this study
Flatten layer	Converting 2-dimensional image input to 1- dimensional
Dense layer 112	Connects all neurons from the flatten layer to 112 neurons belonging to the dense layer
Dropout 25%	Removes 25% of the previous network and prevents overfitting
Output layer	Generate classification based on received input

After all pre-trained models have been fine-tuned, these models will then be trained using the same dataset as shown in Table 3. During the training process, several callbacks will be implemented such as early stopping, checkpoint models, and reduce learning on plateau in order to obtain optimal training results for all pre-trained models being compared.

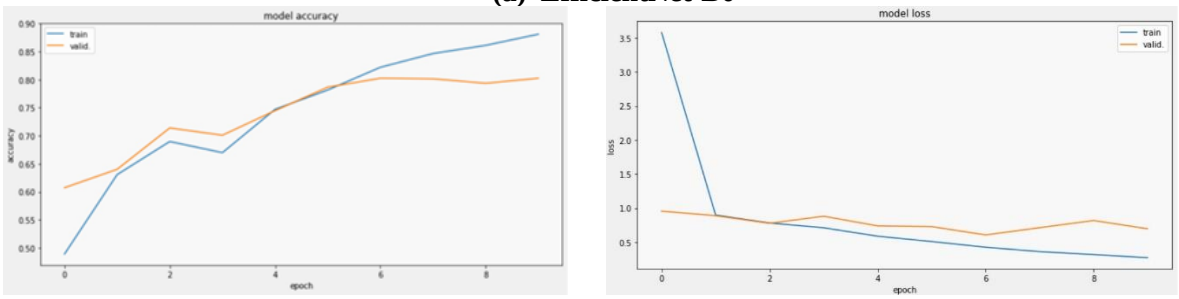
The next step is to evaluate all pre-trained models that have been trained. This evaluation process will be carried out by testing the model to classify the images contained in the test dataset. The results of the classification carried out by the model will be matched with the actual labels in the dataset by using a confusion matrix. The results of the confusion matrix from each model will be used to determine the level of accuracy, precision, recall, and F1 score of the model.

3.2. Experimental results

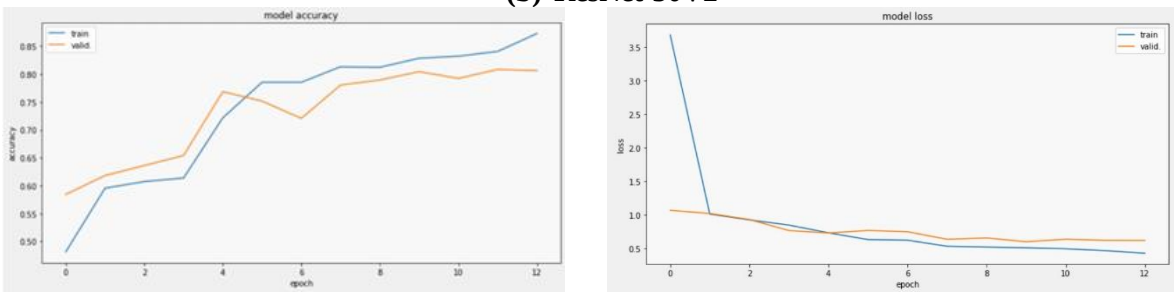
This section will describe the results of the training and the classification performed by each pre-trained model in two test scenarios, viz. by using batch size of 32 and 64 for each pre-trained model. The test results of each model in form of accuracy and loss plots found during training are shown in Figs. 5 and 6. Meanwhile, the model classification results in the form of a confusion matrix are shown in Figs. 7 and 8. The overall performance value of each model is shown in Table 5.



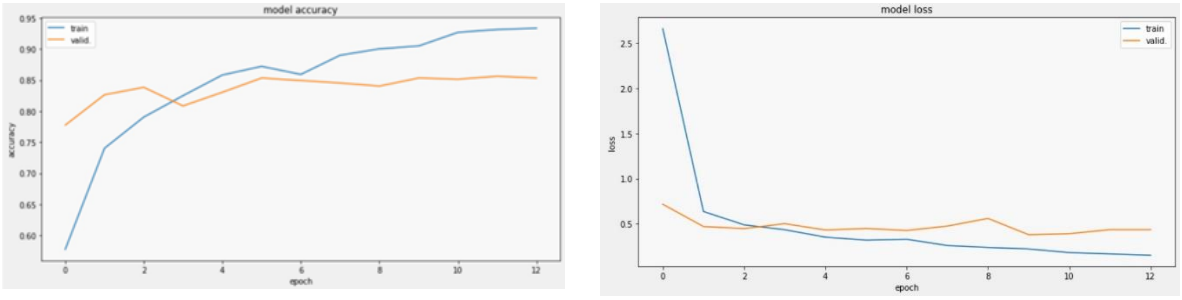
(a) EfficientNet-B0



(b) ResNet-50V2

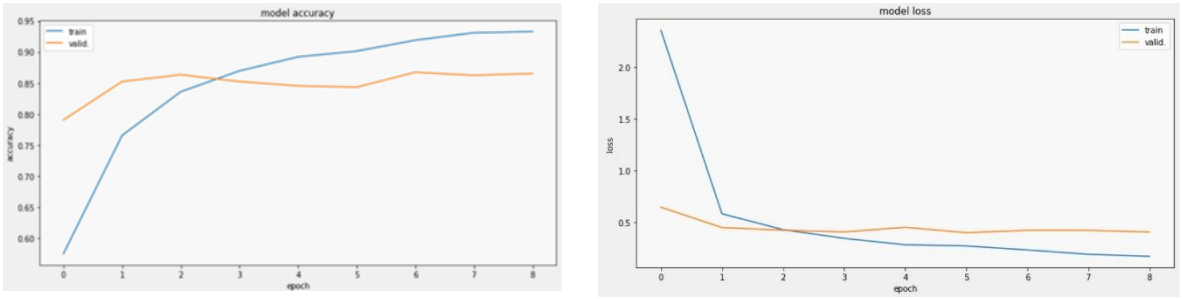


(c) Inception-V3

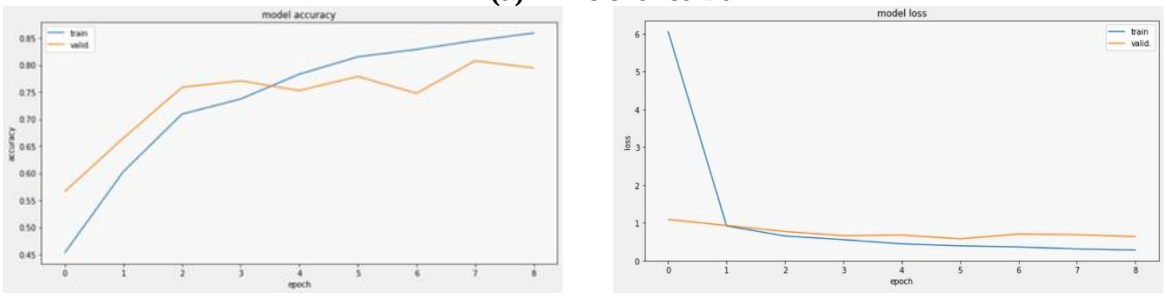


(d) DenseNet-169

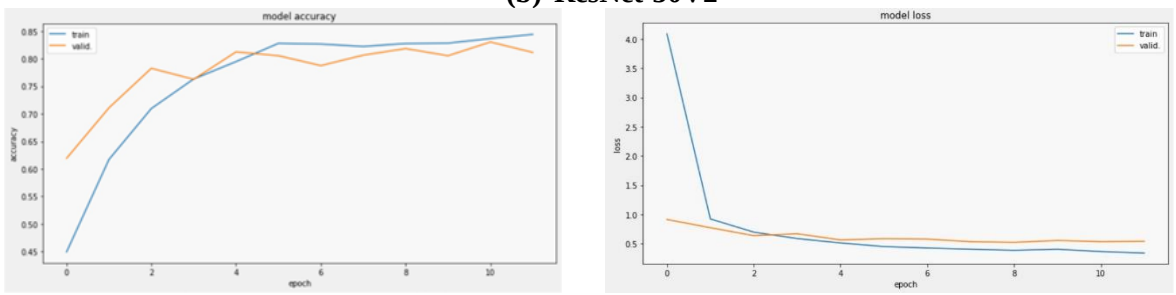
Fig. 5. Accuracy and loss plot of each compared model on batch size 32



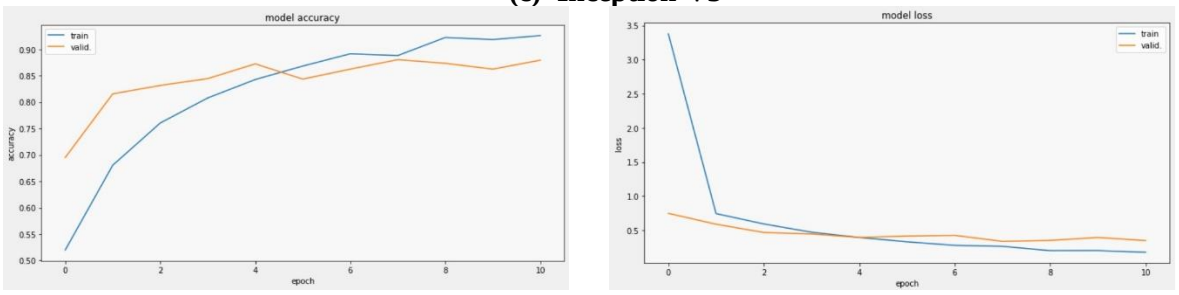
(a) EfficientNet-B0



(b) ResNet-50V2

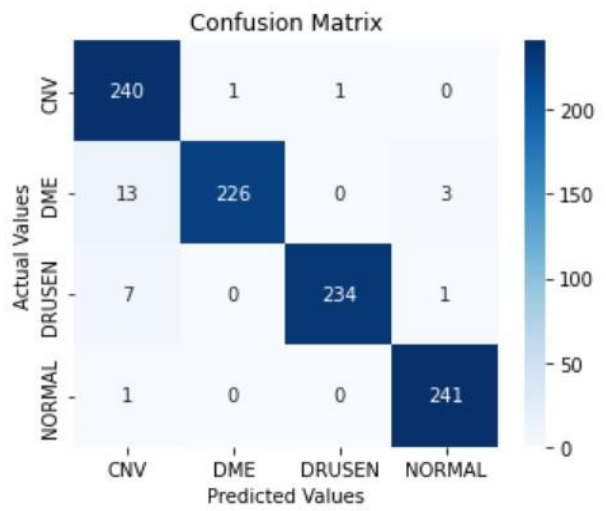


(c) Inception-V3

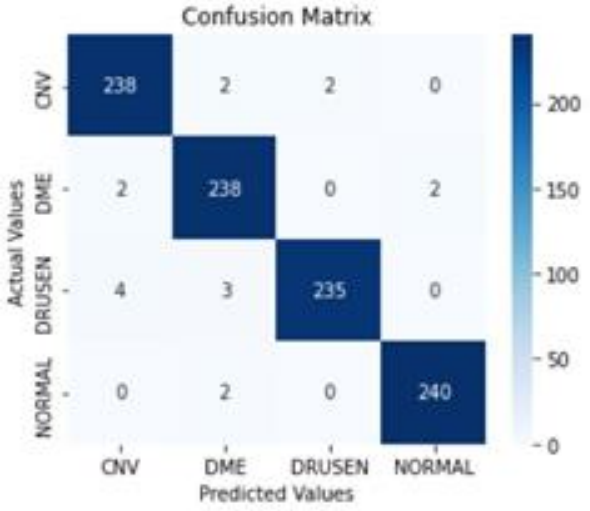


(d) DenseNet-169

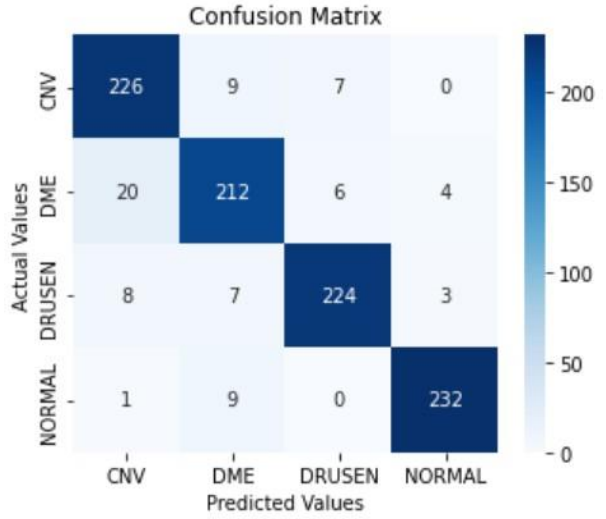
Fig. 6. Accuracy and loss plot of each compared model on batch size 64



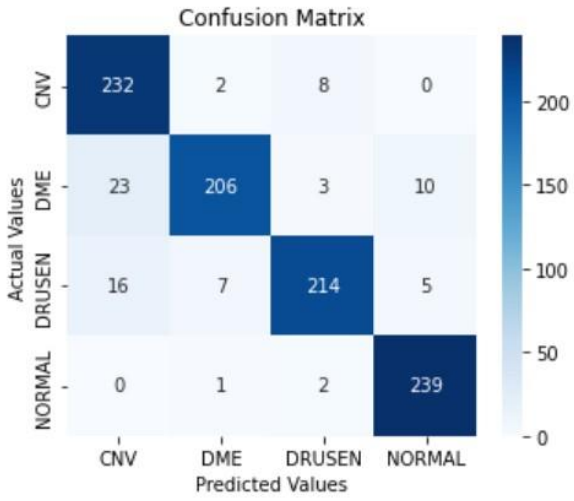
(a) EfficientNet-B0



(b) ResNet-50V2

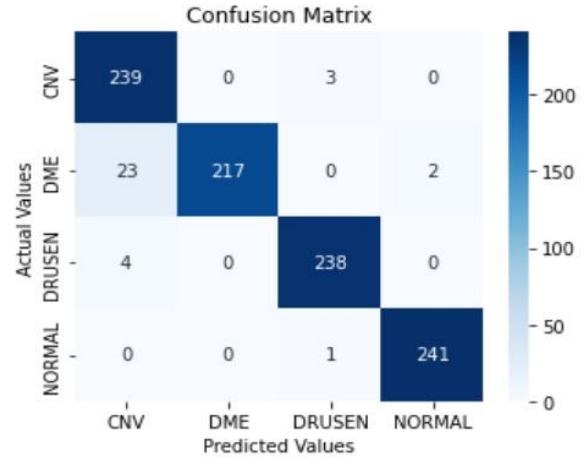


(c) Inception-V3

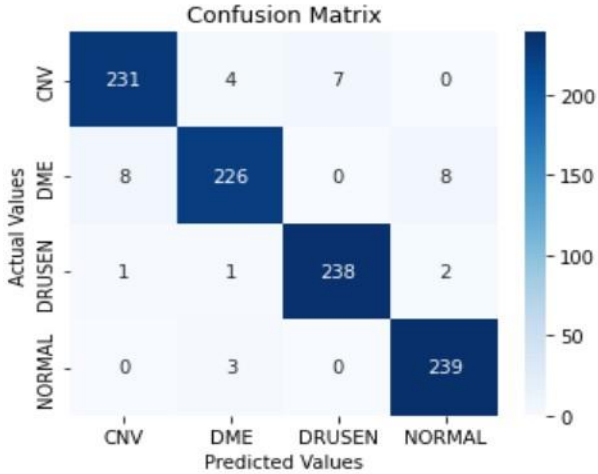


(d) DenseNet-169

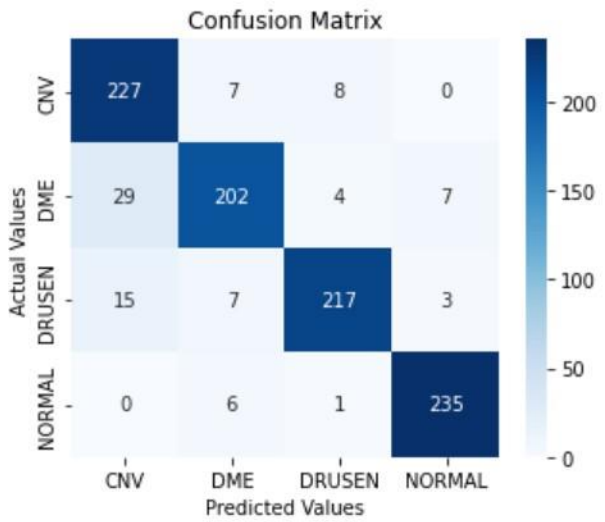
Fig. 7. Confusion matrices of each compared model on batch size 32



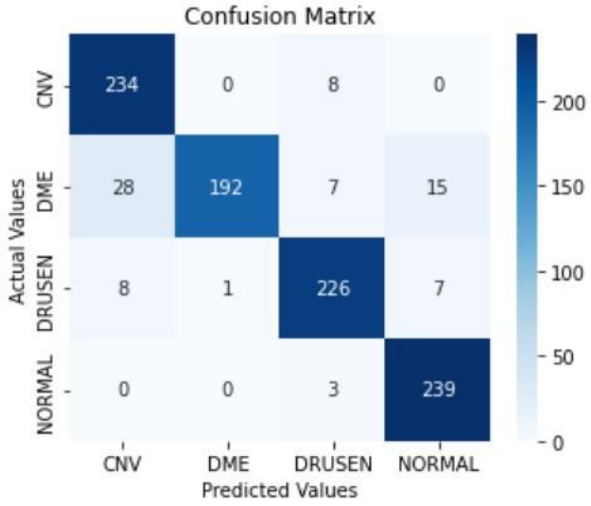
(a) EfficientNet-B0



(b) ResNet-50V2



(c) Inception-V3



(d) DenseNet-169

Fig. 8. Confusion matrices of each compared model on batch size 64

Tab. 5. Overall model performance with batch size of 32 and 64

Model	Batch Size	Accuracy	Precision	Recall	F1 score
EfficientNet-B0	32	97.21%	97.36%	97.21%	97.21%
	64	96.59%	96.84%	96.59%	96.59%
ResNet-50V2	32	98.24%	98.25%	98.24%	98.24%
	64	96.49%	96.49%	96.49%	96.48%
Inception-V3	32	92.36%	92.42%	92.36%	92.36%
	64	91.01%	91.26%	91.01%	91.01%
DenseNet-169	32	92.05%	92.34%	92.05%	92.01%
	64	92.05%	92.59%	92.05%	91.93%

The results show that the model with the highest performance value in the two scenarios tested overall is ResNet-50V2 with a batch size of 32. Meanwhile, the Inception-V3 model with a batch size of 64 has the lowest performance value compared to other models. Through a more in-depth study of each model, several factors have the potential to influence these results, namely the batch size used, as well as the characteristics and suitability of the model to the dataset.

Based on the results of the model performance in Table 5, it can be seen that each model has a better performance value when using a batch size of 32 compared to 64. This result is also supported by research in [23] regarding the effect of batch size on model classification results. The results of this study indicate that the use of batch size with the right size can produce better model performance. Using a batch size that is too large can result in overfitting and poor generalization of the model. On the other hand, the use of a batch size with a size that is too small can result in the model not being able to properly classify datasets with large sizes and high complexity. Therefore, it is proven that the use of different batch sizes can affect the performance of the resulting model.

The next factor is the characteristics and suitability of each model to the dataset. The main characteristic of ResNet is the application of residual blocks aimed at increasing accuracy. The application of residual block allows ResNet-50V2 to be able to perform identity mapping, which is to combine the output of one residual block with the next residual block. In addition, the application of pre-activation as a development on ResNet-50V2 also makes all inputs normalized first before being accepted by each residual block. The combination of identity mapping and pre-activation can improve the final output so that ResNet-50V2 has good performance. In contrast to ResNet, the characteristics of the Inception-V3 model emphasize width scaling, where there are four operations that run in parallel on each layer, including 1 x 1 convolution layer, 3 x 3 convolution layer, 5 x 5 convolution layer, and max pooling. Each operation will extract features from the input image and then combine these features before being passed on to the next layer. This set of layers will determine the feature with the greatest value or weight of an image. From the

results of the performance value, it can be seen that the Inception model actually has a pretty good value. However, the characteristics of the Inception model are less effective for classifying diseases of the eye retina using OCT scanned images when compared to the other three models. From this study results, several key observations can be taken for decision making process. Firstly, among the four CNN models experimented, the ResNet-50V2 is recommended to be used in classifying eye retinal diseases. Secondly, rather than using a larger batch size demanding bigger computational power, the batch size of 32 is proven to be more effective and accurate in almost all scenarios conducted.

4.Conclusion

Based on the results of the research that has been done, the fine-tuning model with the best classification results in the batch size 32 test scenario is the ResNet-50V2 model with an accuracy value of 98.24%, precision of 98.25%, recall of 98.24%, and F1-score of 98.24%. While in the batch size 64 test scenario, the EfficientNet-B0 model is the model with the best classification results with an accuracy value of 96.59%, precision of 96.84%, recall of 96.59%, and F1-score of 96.59%. Thus, we conclude that the best pre-trained CNN model with the results of eye retinal disease classification among EfficientNet-B0, ResNet-50V2, Inception-V3, and DenseNet-169 is ResNet-50V2 with a batch size of 32. This selection is based on the consideration of the best model performance values (accuracy, precision, recall, F1 score) between the different scenarios being tested.

In further research, comparison with other emerging pre-trained models, such as Xception, AmoebaNet, NASNet, can be conducted. Further experiments to check other hyperparameters setting influence to the pre-trained models performance results for eye retinal disease classification can also be done.

5. Acknowledgement

The authors would like to acknowledge the support given by Universitas Multimedia Nusantara during this study.

6. Conflicts of Interest

The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] L. Gurbeta Pokvic, L. Spahic, and A. Badnjevic, "Implementation of Industry 4.0 in Transformation of Medical Device Maintenance Systems," in *Handbook of Research on Integrating Industry 4.0 in Business and Manufacturing*, (2020), pp. 512-532.

[2] E. A. Swanson and J. G. Fujimoto, "The ecosystem that powered the translation of OCT from fundamental research to clinical and commercial impact [Invited]," *Biomed. Opt. Express*, Vol. 8, No. 3, (2017), p. 1638.

[3] J. Wang *et al.*, "Deep learning for quality assessment of retinal OCT images," *Biomed. Opt. Express*, Vol. 10, No. 12, (2019), p. 6057.

[4] R. M. Kamble *et al.*, "Automated Diabetic Macular Edema (DME) Analysis using Fine Tuning with Inception-Resnet-v2 on OCT Images," in *2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, (2018), pp. 442-446.

[5] M. S. Abirami, B. Vennila, K. Suganthi, S. Kawatra, and A. Vaishnava, "Detection of Choroidal Neovascularization (CNV) in Retina OCT Images Using VGG16 and DenseNet CNN," *Wirel. Pers. Commun.*, (2021).

[6] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," (2019).

[7] D. Kermany, K. Zhang, and M. Goldbaum, "Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification," *Mendeley Data, V2*, (2018).

[8] D. S. Kermany *et al.*, "Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning," *Cell*, Vol. 172, No. 5, (2018), pp. 1122-1131.e9.

[9] R. H. Masland, "The Neuronal Organization of the Retina," *Neuron*, Vol. 76, No. 2, (2012), pp. 266-280.

[10] S. L. Baxter *et al.*, "Risk of Choroidal Neovascularization among the Uveitides,"

Am. J. Ophthalmol., Vol. 156, No. 3, (2013), pp. 468-477.e2.

[11] G. E. Lang, "Diabetic Macular Edema," *Ophthalmologica*, Vol. 227, No. Suppl. 1, (2012), pp. 21-29.

[12] S. Hamann, L. Malmqvist, and F. Costello, "Optic disc drusen: understanding an old problem from a new perspective," *Acta Ophthalmol.*, Vol. 96, No. 7, (2018), pp. 673-684.

[13] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in *2017 International Conference on Engineering and Technology (ICET)*, (2017), pp. 1-6.

[14] S. Hansun, A. Argha, H. Alinejad-Rokny, S.-T. Liaw, B. G. Celler, and G. B. Marks, "Revisiting Transfer Learning Method for Tuberculosis Diagnosis," in *2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, (2023), pp. 1-4.

[15] K. He, X. Zhang, S. Ren, and J. Sun, "Identity Mappings in Deep Residual Networks," in *Lecture Notes in Computer Science, vol 9908*, Springer, Cham, (2016), pp. 630-645.

[16] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jul. (2017), pp. 2261-2269.

[17] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (2016), pp. 2818-2826.

[18] X. Deng, Q. Liu, Y. Deng, and S. Mahadevan, "An improved method to construct basic probability assignment based on the confusion matrix for classification problem," *Inf. Sci. (Ny).*, Vol. 340-341, (2016), pp. 250-261.

[19] K. P. Shung, "Accuracy, Precision, Recall or F1?," *Towards Data Science*, (2018). https://towardsdatascience.com/accuracy-precision-recall-or-f1-331fb37c5cb9.

[20] S. Hansun, A. Suryadibrata, R. Nurhasanah, and J. Fitra, "Tweets Sentiment on PPKM Policy as a COVID-19 Response in Indonesia," *Indian J. Comput. Sci. Eng.*, Vol. 13, No. 1, (2022), pp. 51-58.

[21] M. Karmelia, M. Widjaja, and S. Hansun, "Candlestick Pattern Classification Using Feedforward Neural Network," *Int. J. Adv. Soft Comput. its Appl.*, Vol. 14, No. 2, (2022), pp. 80-95.

[22] O. Russakovsky *et al.*, "ImageNet Large Scale Visual Recognition Challenge," *Int. J. Comput. Vis.*, Vol. 115, No. 3, (2015), pp. 211-252.

[23] K. Shen, "Effect of batch size on training dynamics," *Medium*, (2018). https://medium.com/mini-distill/effect-of-batch-size-on-training-dynamics-21c14f7a716e.

Follow this article at the following site:

Theodore Alvin Hartanto & Seng Hansun. Comparative Analysis of Pre-trained CNN Models on Retinal Diseases Classification. IJIEPR 2024; 35 (3) :1-12
URL: http://ijiepr.iust.ac.ir/article-1-1935-en.html