SITEKIN: Jurnal Sains, Teknologi dan Industri, Vol. 20, No. 1, Desember 2022, pp.62 - 69
ISSN 2407-0939 print/ISSN 2721-2041 online

Bacteria Classification using Image Processing andResidual Neural Network (ResNet)

Dybio Dompu Hot Asih[1], Adnan Purwanto[2], Dwiza Riana[3], Sri Hadianti[4]
[1,2,3,4]Faculty of Information Technology, Magister of Computer Science Study Program,
Nusa Mandiri University
[1,2,3,4]East Jakarta 12450, Indonesia
[1]14002438@nusamandiri.ac.id, [2]14002474@nusamandiri.ac.id, [3]dwiza@nusamandiri.ac.id,
[4]sri.shv@nusamandiri.ac.id

ABSTRACT

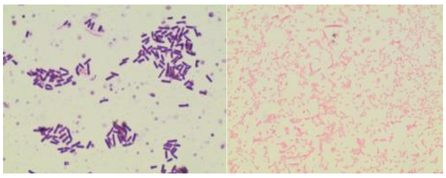
Detection of microorganisms is of particular importance to human health and life, and for the industry in general. For this reason, we want this process to be as fast and precise as possible. We also expect that the automation of this activity (detection of microorganisms) can be widely used in various industries. This article is another attempt to the classification of bacteria that uses a deep learning approach with Residual Neural Network(ResNet) models. The research was conducted by training the ResNet-18,ResNet-34, ResNet50 and ResNet-101 models. The results show that the ResNet-50 and ResNet-101 are the best learning model. It is better to use ResNet-50 than ResNet-101 because of the faster training time. While the results of the research also show that the architecture with the least number of layers is the fastest learning model. ResNet-50 has an accuracy rate of 96.1% with a training time of 451 seconds is the best learning model. ResNet-18 has an accuracy rate of 93.6% with a training time of 185 seconds is the fastest learning model.

Keywords : Deep Learning, Model, Bacteria Classification, ResNet

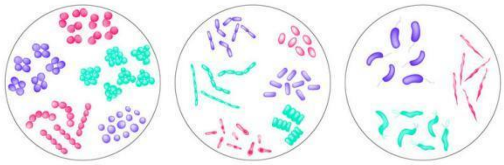
I. INTRODUCTION

The rapid growth of technology has initiated the development of automated system in various fields, including medical[1] and microorganisms. Bacteria can be either bad or good to human body such as lactobacillus, a good bacteria that can improve Symptoms of Irritable Bowel Syndrome or Escherichia coli, a bad bacteria that produces a potent toxin that can harm the lining of the small intestine. There are 2 main methods to classify bacteria.

First, bacteria can be classified based on their cell wall composition and reaction to the Gram stain test. This method can classify bacteria to positive gram as shown in Fig. 1(a) and negative gram as shown in Fig. 1(b). Second, bacteria can be classified by shape such as Cocci, Bacili and Spirilla as shown in Fig 1(c) until Fig 1(e). This shows that there are many types of bacteria and to classify bacteria for diagnosis is going to be rough and need to take time[2].



(a). Gram-Positive (b).Gram-Negative



(c) Cocci (d) Bacili (e) Spirilla

Fig 1. Classification of Bacteria

A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. The pre-processing required in a ConvNet is much lower as compared to other classification algorithms. While in primitive methods filters are hand-engineered, with enough training, ConvNets have the ability to learn these filters/characteristics[3].

The architecture of a ConvNet is analogous to that of the connectivity pattern of Neurons in the Human Brain and was inspired by the organization of the Visual Cortex. Individual neurons respond to stimuli only in a restricted region of the visual field known as the Receptive Field. A collection of such fields overlap to cover the entire visual area[4].

Similar to the Convolutional Layer, the Pooling layer is responsible for reducing the spatial size of the Convolved Feature. This is to decrease the computational power required to process the data through dimensionality reduction. Furthermore, it is useful for extracting dominant features which are rotational and positional invariant, thus maintaining the process of effectively training of the model[5].

ResNet is an architecture from CNN which introduces a new concept, namely shortcut connections. The emergence of the concept of shortcut connections that exist in the ResNet-50 architecture is related to the vanishing gradient problem that occurs when efforts to deepen the structure of a network are carried out. ResNet-101 is a convolutional neural network that is 101 layers deep. You can load a pretrained version of the network trained on more than a million images from the ImageNet database[6]. The pretrained network can classify images into 1000 object categories, such as keyboard, mouse, pencil, and many animals. As a result, the network has learned rich feature representations for a wide range of images. The network has an image input size of 224-by-224.

Image Processing is a method for processing images (Image) into digital form for a specific purpose. At first this image processing functioned to improve and improve the quality of an image, but with the development of the times and the emergence of computational sciences it allows humans to retrieve the information contained in an image. The input is an image (image) and the output is an image that has been improved in quality. For example, an image that is less sharp in color, blurred (blurring) and contains noise (eg white spots) requires processing to improve image quality so as to obtain better information[7].

So far the deep learning approach in the analysis of microbiological images taking into account microbial detection and classifification was undertaken and described in several papers.Zielinski et al. [8] developed public available DIBaS (Digital Image of Bacterial Species) data set containing 660 images of 33 different microbes (fungi and bacteria). They used DSIFT, CNN,Fisher Vectors for features extraction, and for classification: SVM, random forest. The final classification result of single microbes was at the level of 97.24 +/- 1.07%. The same data set (DIBaS) was also used by Mohamed et al. [9] in the Bag of Words approach with SVM for 10 different microorganisms (10 * 20 = 200 images). This test ended with a final score of 97% correctness of classification.

With the exponential growth of data and complexity of systems, fast machine learning/artifificial intelligence and computational intelligence techniques are highly required [10]. For this research, Residual Neural Network (ResNet) a popular deep learning model that is widely used for image classification is used to classify 33 species of bacteria dataset from DIBaS dataset (Digital Image of Bacterial Species). Our focus in this research is to compare the accuracy and training time between ResNet models, namely ResNet-18, ResNet-34, ResNet-50 and ResNet-101. We resize images to compromise our hardware limitations without trying to reduce accuracy.

Hypothesis that we propose in this study is:

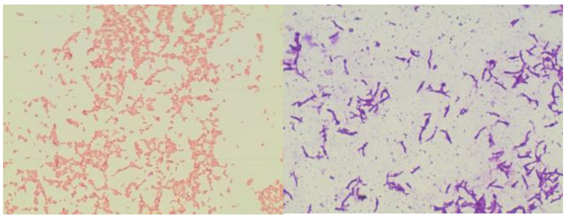
1. H_0= all model of ResNet have equal accuracy
   H_a= at least one model of ResNet is different accuracy
2. H_0= all model of ResNet have equal training time
   H_a= at least one model of ResNet is different training time

II. RESEARCH METHOD

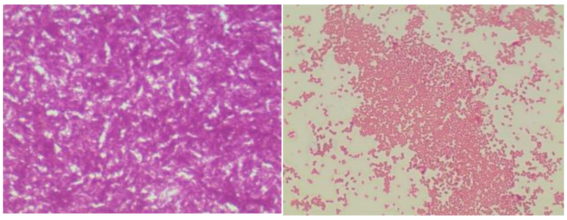
In Fig. 3, we present the research method which we describe as follows

A. Data Set

In this study, we used the DIBAS dataset[8] consisting of 33 classes from 679 images. Example of images shown in Fig 2.



(a) Acinetobacter baumanii (b)Bifidobacterium spp



(c) Lactobacillus casei (d)Neisseria gonorrhoeae

Fig 2. Example of bacteria images

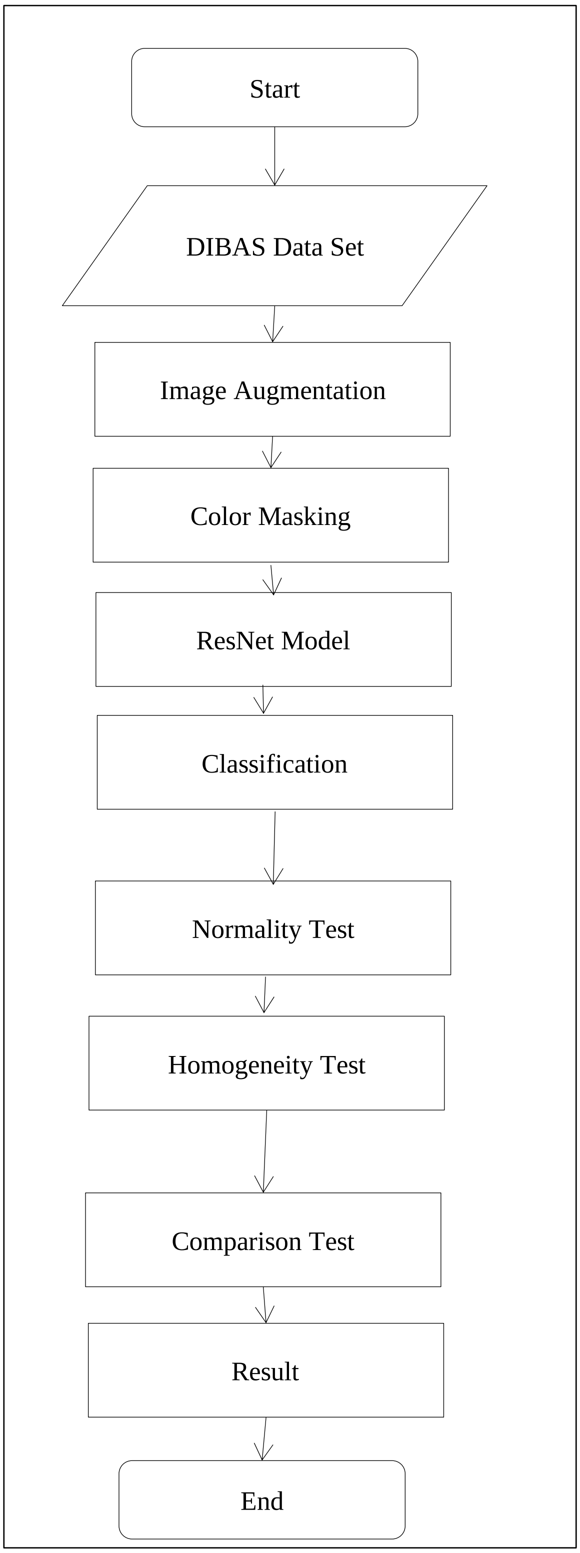


Fig 3.Research Model

B. Image Augmentation

Data augmentation is a commonly used technique for increasing both the size and the diversity of labeled training sets by leveraging input transformations that preserve output labels. In computer vision domain, image augmentations have become a common implicit regularization technique to combat overfitting in deep convolutional neural networks and are ubiquitously used to improve performance. While most deep learning frameworks implement basic image transformations, the list is typically limited to some variations and combinations of flipping, rotating, scaling, and cropping. Moreover, the image processing speed varies in existing tools for image augmentation[11].

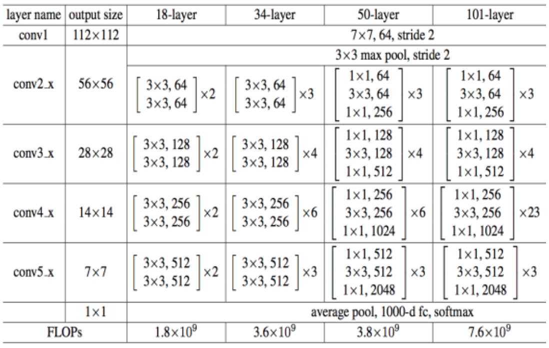
C. Color Masking

HSV segmentation is a separation process object with color selection by value Hue, Saturation, and Value. Hue is attribute that represents pure color. Saturation is an attribute that indicates the effect of white light that affect the degree of color dominance. Value is an attribute that indicates brightness difference in pure color[12].

D. ResNet Model

In this study, we used 4 ResNet models namely ResNet-18, ResNet-34, ResNet-50, ResNet-101. ResNet was developed by Kaiming He et al. [11] in 2016. A residual learning method was proposed to train deeper networks that are practically difficult to train. Network layers were reformulated to learn residual functions with reference to the layer inputs[13].

The difference between each model on ResNet can be seen in Fig. 4.



layer name	output size	18-layer	34-layer	50-layer	101-layer
conv1	112×112	7×7, 64, stride 2			
conv2_x	56×56	3×3 max pool, stride 2			
		$\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$
conv3_x	28×28	$\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times4$
conv4_x	14×14	$\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix}\times6$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times6$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times23$
conv5_x	7×7	$\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$
	1×1	average pool, 1000-d fc, softmax			
FLOPs		1.8×10^9	3.6×10^9	3.8×10^9	7.6×10^9

Fig 4. ResNet Model

E. Classification

Classification is a task that requires the use of machine learning algorithms that learn how to assign a class label to examples from the problem domain[14]. Output from this task for this reseach are accuracy and training time.

F. Normality Test

Before conducting a comparison test of the four models, we perform data normality testing using Shapiro-Wilk test. Data normality testing is a common practice before the statistical method. Normality test is one part of the test requirements for data analysis or commonly called classical assumptions. The purpose of the normality test is to find out whether the data distribution follows or approaches the normal distribution. The standard normal distribution is the most important continuous probability distribution has a bell-shaped density curve described by its mean and SD and extreme values in the data set have no significant impact on the mean value. If a continuous data is follow normal distribution then 68.2%, 95.4%, and 99.7% observations are lie between mean ± 1 SD, mean ± 2 SD, and mean ± 3 SD, respectively[15].

For a data normality test, the hypothesis are as follows:

H_0: Data follow a normal distribution.

H_a: Data do not follow a normal distribution.

G. Homogeneity Test

Purpose homogeneous test is to find out homogeneous for a variance across groups [16]. We perform homogeneous testing using Levene's test For homogeneous test, the hypothesis are as follows :

H_0: Variance across group is not homogeneous.

H_a: Variance across group is homogeneous.

H. Comparison Test

In this research, we compared 4 models, so we used ANOVA if the data were normal and homogeneous. Meanwhile, if the data is not normal or homogeneous, we use the Kruskal-Wallis rank sum test[24]. In addition, ANOVA requires data with a sufficient number of samples [17]. In this research, each model was tested 30 times.

III. RESULTS

A. Image Augmentation

Each image we resize to 224 X 224 pixels. The dataset must be increased by Image Augmentation, this process will double the size from 20 files each to 40 files by using flipping and rotation techniques. Example result of flipping and rotation can be seen in Fig.5.

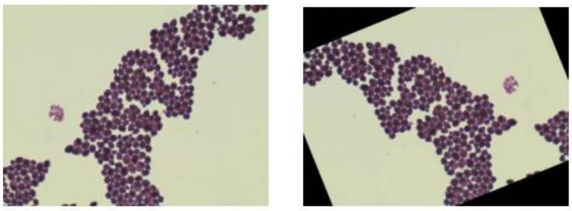


Fig 5. Example of flipping and rotation

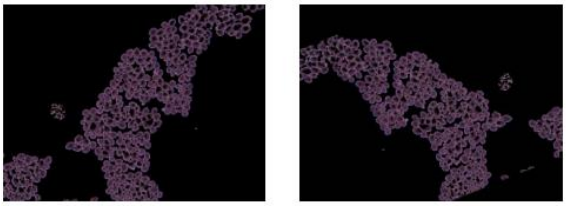


Fig 6. Example of Color Masking

B. Color Masking

The background color is an unnecessary feature, so the second stage is to deal with background color via Image Processing. HSV color-space is the method we used because the default colors in the RGB color-space are coded using the three channels, it is more difficult to scope the background color. So, the background was removed in this process. Example result of color masking can be seen in Fig. 6.

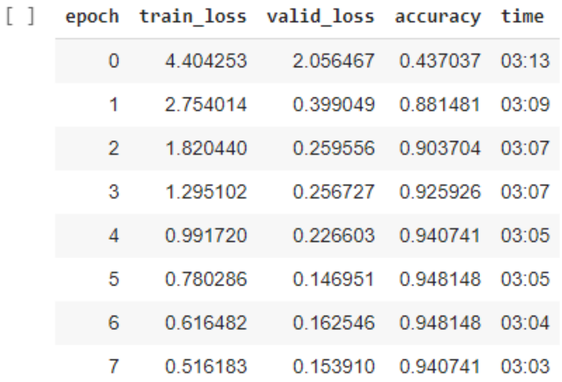
C. ResNet Model

Now, the whole dataset is ready, next is to build a Residual network (ResNet) by using Fastai with Pytorch as backend. Before the training session, dataset was split into 2 parts: Training set and Test set with 80:20 ratio for the model training and evaluating. This approach used 8 epoch, 17 step per pochs, batch size at 32.

IV. DISCUSSION

A. Classification

The classification results in one test will appear as shown in the Fig 7.



```
[ ]  epoch  train_loss  valid_loss  accuracy  time
         0    4.404253    2.056467  0.437037  03:13
         1    2.754014    0.399049  0.881481  03:09
         2    1.820440    0.259556  0.903704  03:07
         3    1.295102    0.256727  0.925926  03:07
         4    0.991720    0.226603  0.940741  03:05
         5    0.780286    0.146951  0.948148  03:05
         6    0.616482    0.162546  0.948148  03:04
         7    0.516183    0.153910  0.940741  03:03
```

Fig7. Example output result classification

Result of accuracy classification each model after 30 times test shown in table 1 and table 2.

Table 1. Result Accuracy Test

No	Resnet-18	ResNet-34	Resnet-50	ResNet-101
1	0.955556	0.948148	0.97037	0.955556
2	0.948148	0.918519	0.97037	0.925926
3	0.955556	0.918519	0.97037	0.962963
4	0.933333	0.911111	0.962963	0.962963
5	0.948148	0.925926	0.955556	0.985185
6	0.955556	0.918519	0.977778	0.97037
7	0.933333	0.948148	0.962963	0.955556
8	0.933333	0.918519	0.962963	0.977778
9	0.933333	0.918519	0.977778	0.962963
10	0.933333	0.925926	0.955556	0.948148

No	Resnet-18	ResNet-34	Resnet-50	ResNet-101
11	0.918519	0.933333	0.962963	0.985185
12	0.933333	0.933333	0.97037	0.985185
13	0.955556	0.925926	0.97037	0.977778
14	0.940741	0.911111	0.962963	0.940741
15	0.933333	0.903704	0.97037	0.948148
16	0.940741	0.948148	0.955556	0.955556
17	0.925926	0.911111	0.955556	0.977778
18	0.933333	0.955556	0.948148	0.940741
19	0.955556	0.925926	0.97037	0.97037
20	0.925926	0.948148	0.940741	0.955556
21	0.925926	0.933333	0.948148	0.955556
22	0.940741	0.948148	0.985185	0.97037
23	0.918519	0.962963	0.962963	0.955556
24	0.911111	0.918519	0.955556	0.962963
25	0.940741	0.948148	0.940741	0.962963
26	0.911111	0.955556	0.955556	0.925926
27	0.918519	0.955556	0.955556	0.97037
28	0.918519	0.940741	0.955556	0.97037
29	0.955556	0.955556	0.962963	0.940741
30	0.940741	0.940741	0.940741	0.948148

Descriptive statistical test results are shown in the table 2.

Table 2. Descriptive Statistic Accuracy Test

No	Model	Count	Mean
1	ResNet-18	30	0.936±0.0138
2	ResNet-34	30	0.934±0.0168
3	ResNet-50	30	0.961±0.0111
4	ResNet-101	30	0.960±0.0160

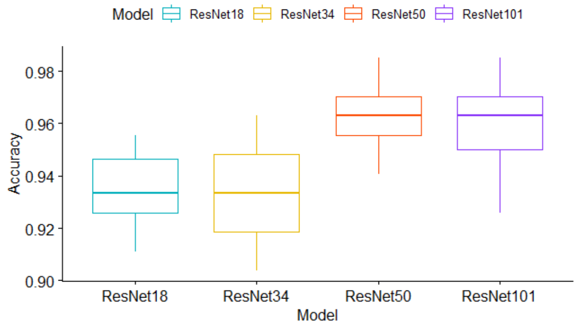


Fig 8. Box plot of Accuracy Test

Box plot of the accuracy testing shown on Fig8. The results show that the architecture with the highest number of layers is more accurate.

Result of training time classification shown in table 2.

Table 3. Result Training Time Test

No	Resnet-18	ResNet-34	Resnet-50	ResNet-101
1	196	287	480	789
2	178	241	481	704
3	185	243	466	703
4	196	243	464	754
5	175	243	468	714
6	184	245	467	772
7	196	248	465	770
8	196	252	464	777
9	185	246	536	770
10	173	256	507	773
11	197	258	530	776
12	182	254	516	771
13	176	251	475	787
14	175	262	469	813
15	185	261	476	812
16	196	265	475	813
17	175	263	470	709
18	183	272	427	624
19	197	265	428	815
20	174	263	387	709
21	183	256	387	623
22	208	258	426	793
23	174	265	390	626
24	190	258	429	714
25	195	256	405	769
26	174	269	425	620
27	183	263	408	714
28	196	266	402	773
29	174	271	428	618
30	183	261	384	714

Descriptive statistical test results are shown in the table 4.

Table 4. Descriptive Statistic of Training Time

No	Model	Count	Mean
1	ResNet-18	30	185±9.73
2	ResNet-34	30	258±10.4
3	ResNet-50	30	451±42.6
4	ResNet-101	30	737±63.0

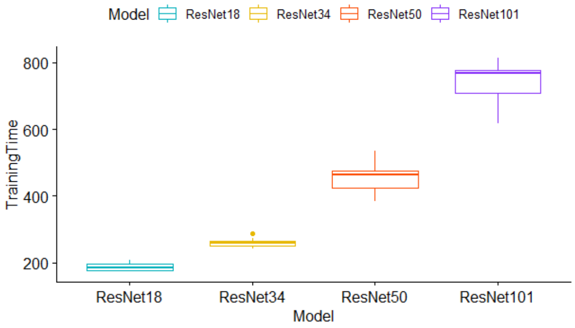


Fig 9. Box plot of Training Time

Box plot of the training time testing shown on Fig9. The results show that the architecture with the lowest number of layers is faster learner.

B. Normality Test

Table5. Shapiro-Wilk Test of Accuracy

	W	**Asymp. Sig**
Residual	0.98285	0.1309

Using R Studio software, rejection and acceptance of hypothesis are seen from Asymp. Sig. (2-tailed) value. Because the number of samples is less than 50, the test results used are the Shapiro-Wilk test. If the Asymp. Sig. (2-tailed) value of Shapiro-Wilk test is greater than 0.05, the data are normal. If it is below 0.05, the data significantly deviate from a normal distribution. Based on table5, the result shows that the Asymp. Sig. (2-tailed) value of accuracy are 0.1309, thus the null hypothesis is accepted. The conclusion from the Shapiro-Wilk test shows that the data of accuracy follow a normal distribution.

Table 6. Shapiro-Wilk Test of Training Time

	W	**Asymp. Sig**
Residual	0.92457	4.481e-06

Based on table 6, the result shows that the Asymp. Sig. (2-tailed) value of training time are below 0.05, thus the null hypothesis is rejected. The conclusion from the Shapiro-Wilk test shows that the data of training time do not follow a normal distribution.

C. Homogeneity Test

Table 7. Levene's Test of Accuracy

	df	**F Value**	**Sign**
Group	3	2.6713	0.05073

Using R Studio software, rejection and acceptance of hypothesis are seen from Sig. value. If the Sig. value of Levene's test is greater than 0.05, the variance across group is homogeneous. If it is below 0.05, variance across group is not homogeneous. Based on table 7, the result shows that the Sig value are 0.05073, thus the null hypothesis is rejected. The conclusion from the Levene's test shows that the variance across group of accuracy is homogeneous.

Table 8. Levene's Test of Training Time

	df	**F Value**	**Sign**
Group	3	15.161	2.178e-08

Based on table 8, the result shows that the Sig value are below 0.05, thus the null hypothesis is accepted. The conclusion from the Levene's test shows that the variance across group of TrainingTime is not homogeneous.

D. Comparative Test

Based on the results of the normality and homogeneity tests above, the accuracy can use Anova while the training time uses the Kruskal-Wallis rank sum test.

Hypothesis for accuracy :

H_0= all model of ResNet have equal accuracy

H_a=at least one model of ResNet is different accuracy

Table 9. Anova Test of Accuracy

	Df	**Sum Sq**	**Mean Sq**	**F value**	**Pr(>F)**
Model	3	0.02045	0.006815	32.02	3.74e-15
Residual	116	0.02469	0.000213		

Result of Anova Test for Accuracy shown in table 9. As the Pr(>F) value is less than the significance level 0.001, thus the null hypothesis is rejected.The conclusion from Anova Testshows that there are significant differences between the models. To find out which model is different accuracy, we use Tukey HSD.

Table 10. Result Tukey HSD of Accuracy

	Diff	**Lower**	**Upper**	**P Adj**
ResNet18-ResNet101	-0.0244444000	-0.034263360	-0.01462544	0.0000000
ResNet34-ResNet101	-0.0266666000	-0.036485560	-0.01684764	0.0000000
ResNet50-ResNet101	0.0009876667	-0.008831294	0.01080663	0.9936535
ResNet34-ResNet18	-0.0022222000	-0.012041160	0.00759676	0.9349462
ResNet50-ResNet18	0.0254320667	0.015613106	0.03525103	0.0000000
ResNet50-ResNet34	0.0276542667	0.017835306	0.03747323	0.0000000

Result of Tukey Honest Significant Differences (HSD) show in table 10. If the P Adj value of Tukey HSD test is greater than 0.05, there are not different between two models. If it is below 0.05, there are different between two models.It can be seen from the table 5 that ResNet-50 and ResNet-101 have P Adjusted value is 0.9936535. It means that ResNet-50 and ResNet-101 have equal accuracy. ResNet-34 and ResNet-18 have P Adjusted value is 0.9349462. It means that ResNet-34 and ResNet-18 have equal accuracy. Another row in Table 10 show that another comparison model does not have P Adjusted value greater than 0.05.

Hypothesis for training time :

H_0= all model of ResNet have equal training time

Ha=at least one model of ResNet is different training time

Table 11. Kruskal-Wallis rank sum test of Training Time

	Df	**Chi-Squared**	**P Value**
Kruskal-Wallis	3	111.61	< 2.2e-16

Result of Kruskal-Wallis rank sum test of Training Time shown in table 11. As the P value is less than the significance level 0.05, thus the null hypothesis is rejected. The conclusion from Kruskal-Wallis rank sum test shows that there are significant differences between the models for Training Time.To find out which model is different accuracy, we use Pairwise Wilcox Test.

Table 12. Pairwise Wilcox Test of Training Time

	ResNet101	ResNet18	ResNet34
ResNet18	3e-11	-	-
ResNet34	3e-11	3e-11	-
ResNet50	3e-11	3e-11	3e-11

.Result of Pairwise Wilcox Test show in table 12. If the p value of Pairwise Wilcox Testt is greater than 0.05, there are not different between two models. If it is below 0.05, there are different between two models.It can be seen from the table 10 that every models have p value below 0.05 so every model significant difference from another models.

V. CONCLUSSION

This paper showed that for accuracy, ResNet-18 and ResNet-34 have a same result. ResNet-50 and ResNet-101 have a same accuracy result. From the result testing, it found that ResNet-18 is the fastest model for training and ResNet-101 is the slowest. The next research that can be developed is to compare ResNet with Xception, Inception, MobileNet and DenseNet etc.

REFERENCES

[1] A. Wibisono, J. Rachmad, and E. Anderson, “Deep Learning and Classic Machine Learning Approach for Automatic Bone Age Assessment,” *2019 4th Asia-Pacific Conf. Intell. Robot Syst.*, pp. 235–240, 2019.

[2] C. Rujichan, “Bacteria Classification using Image Processing and Deep Convolutional Neural Network,” 2019.

[3] I. W. S. E. P, A. Y. Wijaya, and R. Soelaiman, “Klasifikasi Citra Menggunakan Convolutional Neural Network (Cnn) pada Caltech 101,” vol. 5, no. 1, 2016.

[4] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, “A Survey of the Recent Architectures of Deep Convolutional Neural Networks 1 Introduction,” pp. 1–70.

[5] D. Joshi, “Approaching Bio Cellular Classification for Malaria Infected Cells Using Machine Learning and then Deep Learning to compare & analyze K-Nearest Neighbours and Deep CNNs,” 2017.

[6] S. B. Griesemer and G. Van Slyke, “Assessment of Sample Pooling for Clinical SARS-CoV-2 Testing,” no. January, 2021.

[7] A. Kour, “A Review on Image Processing,” vol. 4, no. 1, pp. 270–275, 2013.

[8] B. Zieli, A. Plichta, K. Misztal, P. Spurek, M. Brzychczy-włoch, and D. Ocho, “Deep learning approach to bacterial colony classification,” 2017.

[9] B. A. Mohamed and H. M. Afify, “Automated classification of Bacterial Images extracted from Digital Microscope via Bag of Words Model,” no. 1, pp. 1–4.

[10] I. Transactions and O. N. Systems, “Efficient and Rapid Machine Learning Algorithms for Big Data and Dynamic Varying Systems,” pp. 1–2, 2017.

[11] A. Buslaev, V. I. Iglovikov, E. Khvedchenya, A. Parinov, M. Druzhinin, and A. A. Kalinin, “Albumentations: Fast and flexible image augmentations,” *Inf.*, vol. 11, no. 2, 2020, doi: 10.3390/info11020125.

[12] S. R. G. B. Grayscale, “Perbaikan Hasil Segmentasi Hsv Pada Citra Digital Menggunakan Metode Segmentasi Rgb Grayscale,” *Edu Komputika J.*, vol. 6, no. 1, pp. 32–37, 2019, doi: 10.15294/edukomputika.v6i1.23025.

[13] F. Ramzan *et al.*, “A Deep Learning Approach for Automated Diagnosis and Multi-Class Classification of Alzheimer’s Disease Stages Using Resting-State fMRI and Residual Neural Networks,” *J. Med. Syst.*, vol. 44, no. 2, 2020, doi: 10.1007/s10916-019-1475-2.

[14] S. B. Kotsiantis, I. D. Zaharakis, and P. E. Pintelas, “Machine learning: A review of classification and combining techniques,” *Artif. Intell. Rev.*, vol. 26, no. 3, pp. 159–190, 2006, doi: 10.1007/s10462-007-9052-3.

[15] J. Feys, “Nonparametric Tests for the Interaction in Two-way Factorial Designs Using R,” vol. 8,

no. 2008, pp. 367–378, 2016.

[16] A. Calle-Saldarriaga, H. Laniado, and F. Zuluaga, “Homogeneity Test for Functional Data based on Data-Depth Plots,” pp. 1–25, 2020.

[17] Ö. Karadağ and S. Aktaş, “Optimal sample size determination for the ANOVA designs,” *Int. J. Appl. Math. Stat.*, vol. 25, no. 1, pp. 127–134, 2012.