International Journal of Sciences: Basic and Applied Research (IJSBAR)

ISSN 2307-4531
(Print & Online)

http://gssrr.org/index.php?journal=JournalOfBasicAndApplied

Image Meat Classification Using Deep Learning Approach

I Nyoman Budiastra[a]*, I Wayan Arta Wijaya[b], I Gusti Ngurah Janardana[c]

[a,b,c]Department Electrical Engineering , Udayana University, Bali, Denpasar, 80361, Indonesia

[a]Email: budiastra@unud.ac.id, [b]Email: artawijaya@unud.ac.id, [c]Email: janardana@unud.ac.id

Abstract

In many countries meat is one of the most popular food in the world, but not all kind of meat can be consumed by most people such as pork for muslim and jewish also beef, pork and mutton tend to have several similar characteristics, namely color and texture, this can create confusion in distinguishing the three types of meat when see it by naked eye, therefore we need a system that can help identify the type of meat. In this study the authors used CNN-based Deep Learning approached with pre-defined architecture ResNet50 and ResTet101 combined with transfer learning techniques, the authors also used several types of hyperparameters including learning rate, momentum and epoch to see the effect of these hyperparameter values on the performance of deep learning model, the results of this study show that ResNet50 can outperformed ResNet101 by producing less loss and by 1 % more in F1 score with 95.96% F1 score 95.96% precision and 95.96% recall. In addition the momentum of 0.3 in both architectures has the tendency to produce high loss with a low F1 score, while for a learning rate of 0.0001 on both ResNet50 and ResNet101 architectures also has a tendency to produce high loss and a low F1 score, optimal hyperparameter values range of learning rate is between of 0.001 to 0.01 and 0.6 - 0.9 for momentum.

Keywords: Meat classification; Deep Learning; Transfer learning; ResNet.

----

* Corresponding author.

1. Introduction

Meat is one of the most popular foods in the world and has a lot of mineral and vitamin content which is good for consumption but not all meat can be consumed by people, for example pork which is prohibited for consumption by Muslims and Jewish [1], sometimes meat is very difficult to distinguish by naked eye because the shape, texture and color are very similar, because of that it need machine learning techniques to help to distinguish kind of meat. Machine learning for image recognition has been widely used in various fields, such as breast tumor classification using the image recognition technique of the ensemble method which is a technique that combines several machine learning model methods in conducting image classification [2], Another study is to identify the types of plant diseases using the SVM[3]. Nowdays CNN is widely used in studies to perform several image classifications, some studies have proven that CNN can outperform classic machine learning methods, for example research classify flower types, the study results found that CNN can outperform SVM and KNN [4] Another study is about Tanzanian hand language image processing research which compares the comparison of CNN and SVM, the results of the study conclude that CNN outperforms the performance produced by SVM [5]. In this study the authors used the CNN method with Resnet-50 and Resnet 101 architectures with transfer learning techniques, basically transfer learning is the re-use of a model that has performed well in a certain task to solve a different but related task is known as transfer learning [6], besides using different methods, the author also performs several hyperparameter tunings such as learning rate, momentum and epoch to see the performance provided by the two methods.

2. Data and Methodology

2.1. Data

Data collection is sourced from collected directly at traditional markets in Bali, Indonesia by taking pictures using an iPhone 11 device with a 12 Megapixel camera resolution, the lighting conditions in the room at that time were quite clear. Figure 1 is an example of a mutton image taken using the iPhone 11 cellphone camera before being processed into a dataset

Figure 1: mutton raw image

The process of converting image into a dataset is to cut each photo into small parts with a size of 300 x 300, this is done to get parts of the meat and remove noise that has nothing to do with the meat image. Figure 2 is an illustration of the process of converting raw meat images into a dataset

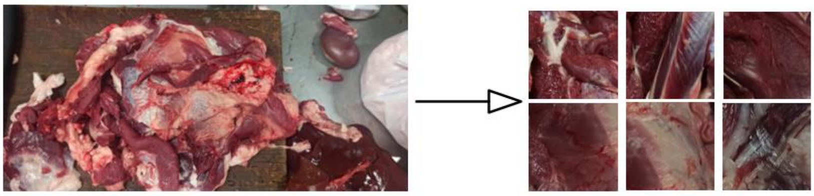


Figure 2: Processing image into dataset

the result of the process of converting raw meat images into datasets produces a number of datasets with a total of 422 dataset images with details of the mutton dataset is 129 images, the pork dataset being 144 images and the beef dataset being 156 images, the finished image dataset will then be divided into 2 parts, the dataset for train and the dataset for validation with details of 75% training and 25% validation, details of the dataset distribution can be seen in Table 1 and an example of datasets can be seen in Figure 3

Table 1: Distribution of Dataset

Dataset	Pork	Mutton	Beef
Train	110	93	121
Validation	34	36	34

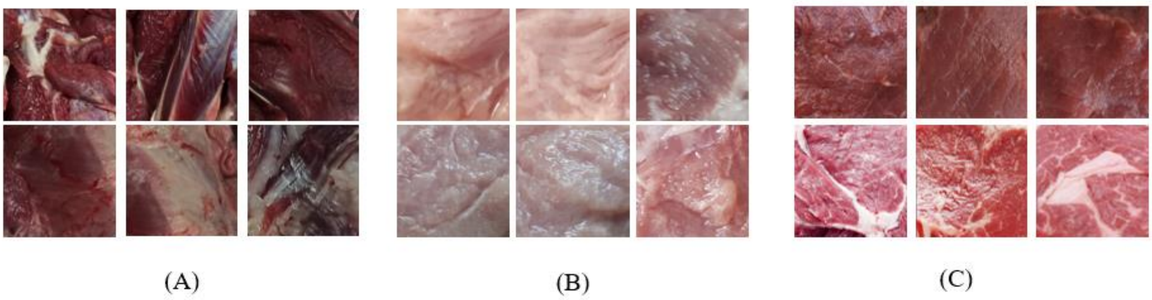


Figure 3: dataset ready to train (A) Mutton (B) Pork (C) Beef

2.2. Convolutional Neural Network

Convolutional Neural Network (CNN) is a supervised learning method in deep learning [7], Structure of CNN consists of 3 main constituent layers, namely convolutional layer, pooling layer, and fully connected layer [8] also architecture CNN usually consists of various levels of network depth, each structure represents a separate feature. At CNN, the convolution process is a process for extracting features contained in an image, the convolution process is simply formulated in equation 1

$$h(x) = f(x) * g(x) \quad (1)$$

from equation 1, $h(x)$ is the output matrix resulting from the convolution or it can be called a feature map, $f(x)$

is the matrix of the input image while $g(x)$ is the filter in the convolution process, after the image convolution process is complete, it produces an output in the form of a feature map, the feature map will be reduced using the pooling layer after going through the convolution and pooling processes the calculated value of the previous layer will be forwarded to the fully connected layer for the prediction process to produce class output. The general CNN architecture is shown in Figure 4.

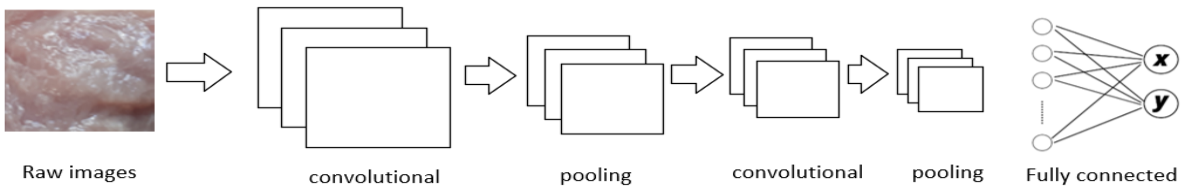


Figure 4: CNN architecture in general [9]

2.3. Resnet Architecture

Deep residual network or commonly called ResNet is one of the CNN structures proposed by He in 2016 [10]. ResNet is one of the CNN architectures that introduces a new concept, namely shortcut connections. The emergence of the concept of shortcut connections that exist in the architecture ResNet has a relationship with the vanishing gradient problem that occurs when trying deepen the structure of a network is done however deepen a network with the aim of improving its performance can not be done just by piling up layers. The deeper a network can bring up a new problem, namely vanishing gradient, vanishing gradient can make the gradient very small which results in a decrease in performance or accuracy [10].

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2				
conv2_x	56×56	3×3 max pool, stride 2				
		[3×3, 64; 3×3, 64] ×2	[3×3, 64; 3×3, 64] ×3	[1×1, 64; 3×3, 64; 1×1, 256] ×3	[1×1, 64; 3×3, 64; 1×1, 256] ×3	[1×1, 64; 3×3, 64; 1×1, 256] ×3
conv3_x	28×28	[3×3, 128; 3×3, 128] ×2	[3×3, 128; 3×3, 128] ×4	[1×1, 128; 3×3, 128; 1×1, 512] ×4	[1×1, 128; 3×3, 128; 1×1, 512] ×4	[1×1, 128; 3×3, 128; 1×1, 512] ×8
conv4_x	14×14	[3×3, 256; 3×3, 256] ×2	[3×3, 256; 3×3, 256] ×6	[1×1, 256; 3×3, 256; 1×1, 1024] ×6	[1×1, 256; 3×3, 256; 1×1, 1024] ×23	[1×1, 256; 3×3, 256; 1×1, 1024] ×36
conv5_x	7×7	[3×3, 512; 3×3, 512] ×2	[3×3, 512; 3×3, 512] ×3	[1×1, 512; 3×3, 512; 1×1, 2048] ×3	[1×1, 512; 3×3, 512; 1×1, 2048] ×3	[1×1, 512; 3×3, 512; 1×1, 2048] ×3
	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	11.3×10^9

Figure 5: ResNet Architecture

Therefore ResNet introduces a new concept, namely the shortcut connection and in this concept the features that were input from the previous layer are also used as input for the output of that layer. This concept is carried out as a solution to minimize the loss of important features during the convolution process. Overall ResNet consists of 5 stages of the convolution process then continued with average pooling and ended with a fully connected layer as a prediction layer There is a difference between the ResNet-50 architecture used in this study and original architecture [10] This is because there are a few modifications made to according to the needs of the classification system. Modifications made include:

• Changed the number of outputs on the fully connected layer from 1000 classes to 3 classes.

• Added 64 hidden layer in fully connected layer before output layer

In addition, this study also uses transfer learning techniques combined with the Resnet architecture, Transfer learning itself is a process of reuse from pre-trained models that have been trained with large-scale datasets by researchers before [11] and the application of transfer learning in this study was by freezing the top layer on the ResNet architecture which had been previously trained using Imagenet data

2.4. Training and Analysis

Training and analysis stage to discuss the flow of this research, the training and analysis stage in general can be seen in Figure 6

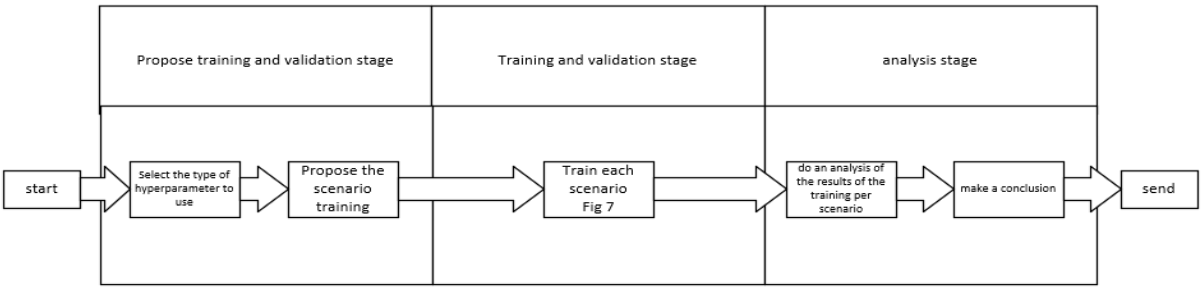


Figure 6: Training and Analysis Flowchart

After preparing the dataset, the next step is to choose the scenario that will be used in this study, scenario consists of by tuning several hyperparameters such as learning rate, momentum and epoch, the value of learning rate, momentum and epoch used can be seen in Table 2

Table 2: Distribution of Hyperparameter

Learning rate	momentum	epoch
0.1	0.3	50
0.001	0.6	75
0.0001	0.9	100

After setting the type of hyperparameter, the scenario training process will try each type of hyperparameter as shown in Figure 7

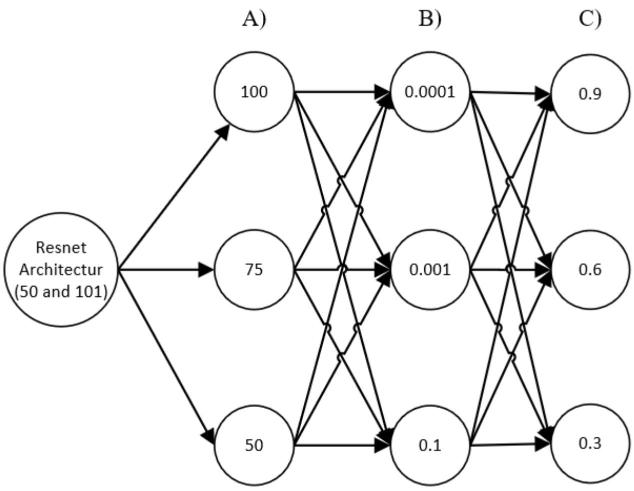


Figure 7: (A) epoch (B) Learning rate (C) Momentum

after the scenario training process is complete, the next stage is the analysis process, the analysis process performs a loss and accuracy graph analysis which is the output of the training process other metrics are used such as F1 score, precision, recall based on confusion matrix on each scenario

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

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

$$F\ score(\%) = \frac{2\ recall\ x\ precison}{recall + precision} \tag{4}$$

Calculation of f1 score, precision and recall based on TP, FN, FP and TP data obtained from the Confusion Matrix Table. True Positive (TP) is the number of correct predictions in the positive class, whereas False Negative (FN) means the number of predictions which is false in the negative class, False Negative (FP) is the number of wrong predictions in the positive class whereas for True Negative (TN) means the number of predictions which is true in the negative class. Illustration of the Confusion Matrix can be seen in Figure 8.

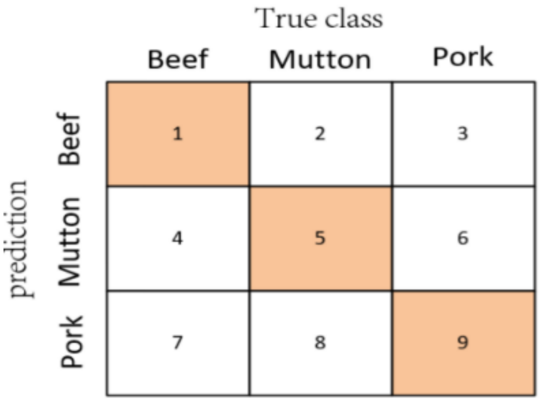
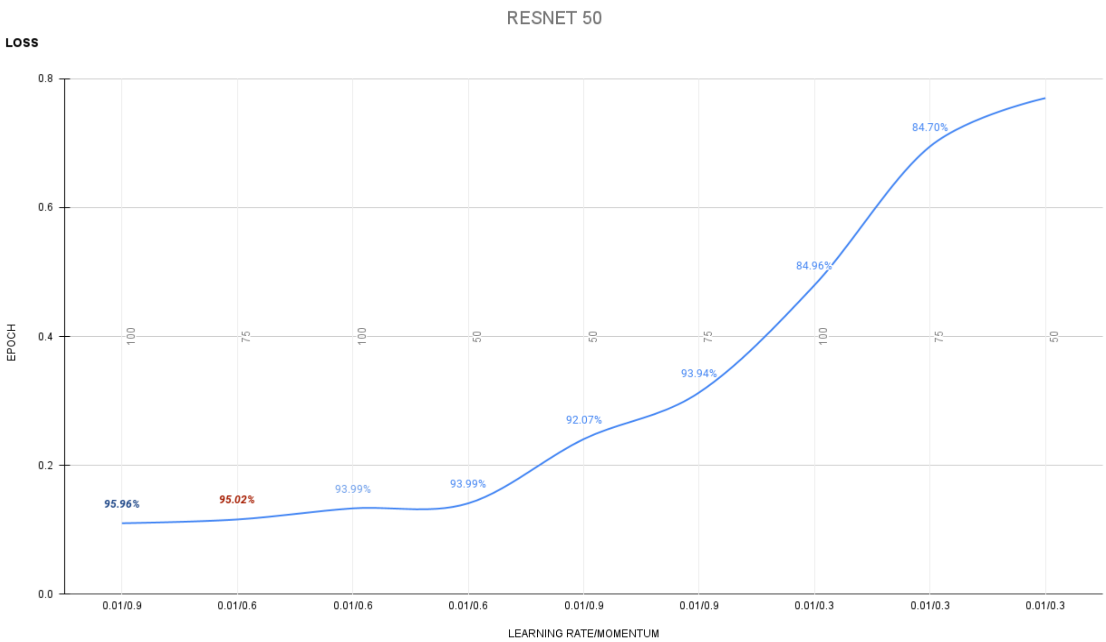


Figure 8: Confusion Metric Table

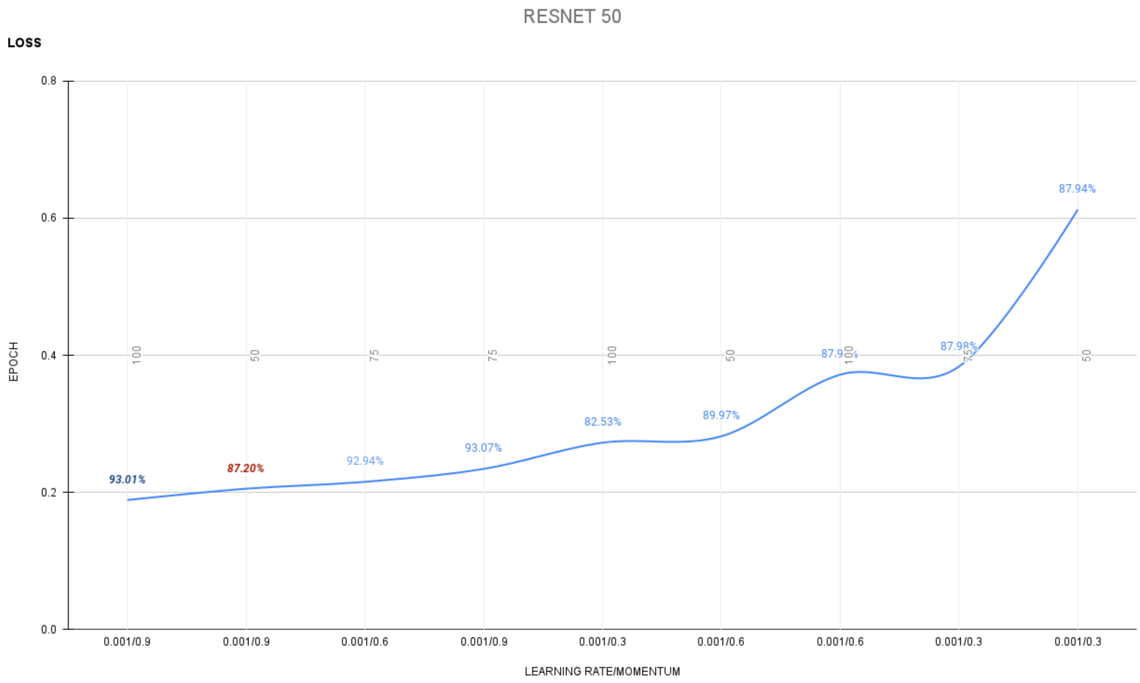
3. Result And Discussion

The analysis in this study is divided into 3 parts, ResNet50 analysis , ResNet101 analysis and Summary, each analysis uses 2 types of metrics namely loss and F1 score, the model with the lowest loss will be selected because the model will be able to generalize the new image [12]

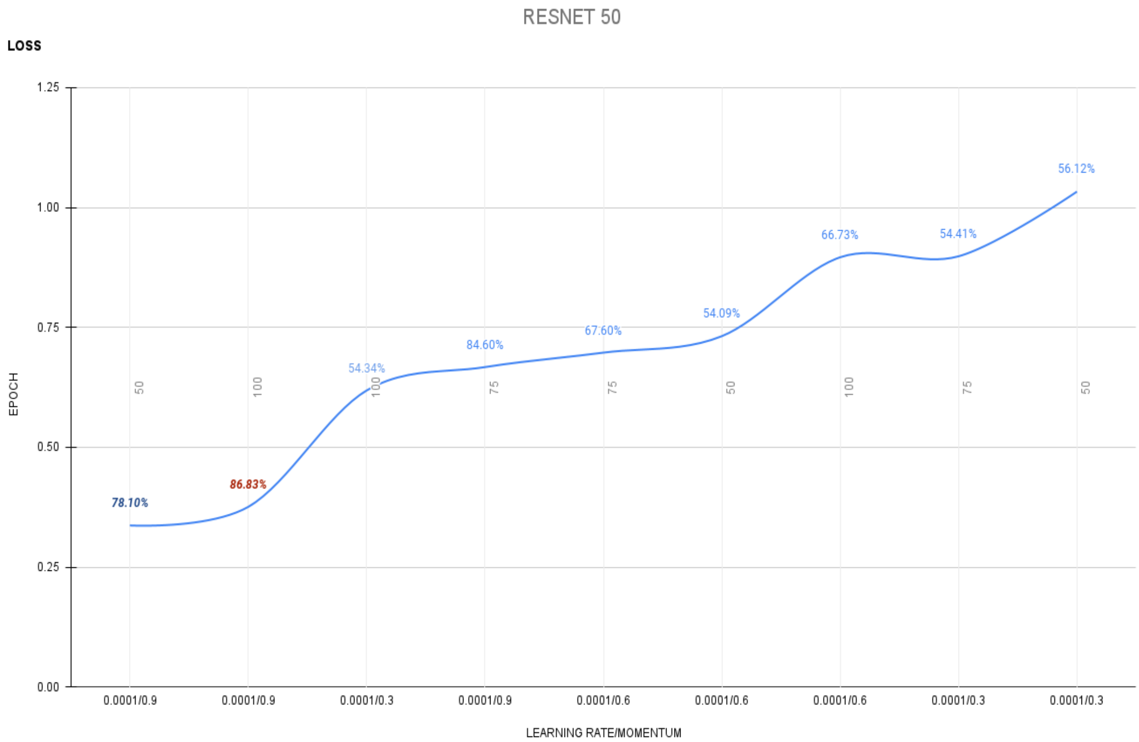
3.1 ResNet50 Analysis



(A)



(B)



(C)

Figure 9: ResNet 50 (A) learning rate 0.01 (B) learning rate 0.001 (C) learning rate 0.0001

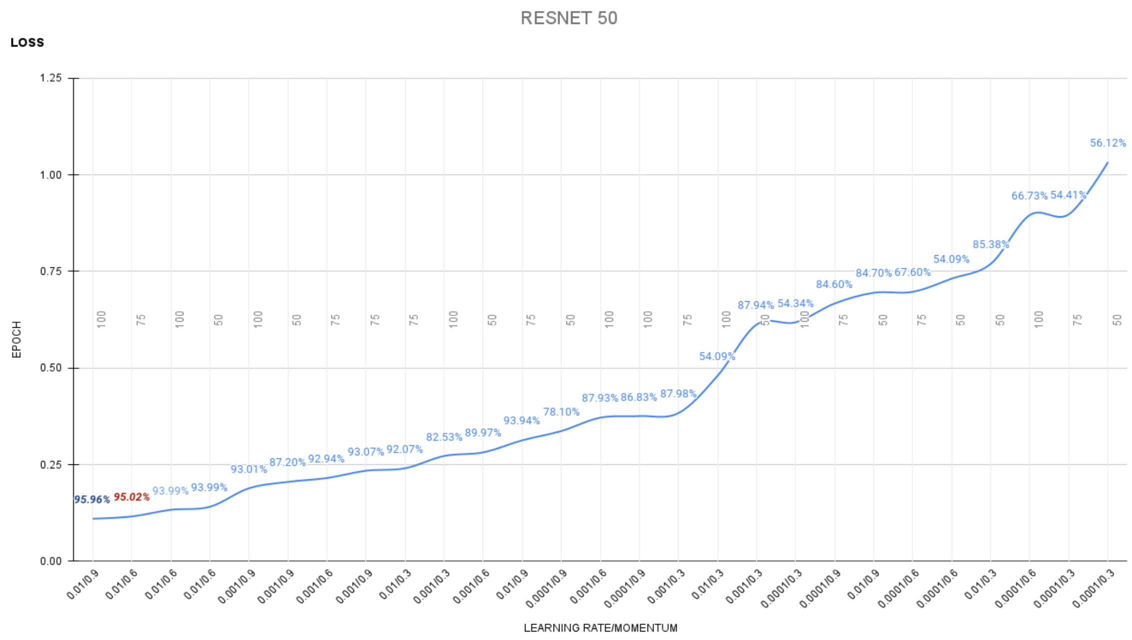


Figure 10: Overall Graph of ResNet 101 Scenario

Figure 9 consists of 3 graphs A, B and C, those graph images generated by trial for each scenario on ResNet50, it can be seen that from the three graphs above, momentum 0.3 tends to have a high loss with low F1 scores, while momentum 0.9 in the three graphs above has the lowest loss with high F1 scores, while the best learning rate values are in the range of 0.001 and 0.1.

Figure 10 is the overall graph of the ResNet50 scenario, from the graph above the best performance for ResNet 50 is the lowest loss (blue with bold font on the Figure 10) 0.110296231150675 with 95.96% of F1 score ,95.96% of precision and 95.96% of recall and has hyperparameter configuration such as learning rate of 0.01, momentum of 0.9 and epoch 100 generated loss of 0.110296231150675 and 95.96% of F1 score, 95.96% of precision and 95.96% of recall, the detail of accuracy graph and loss graph can be seen on Figure 11

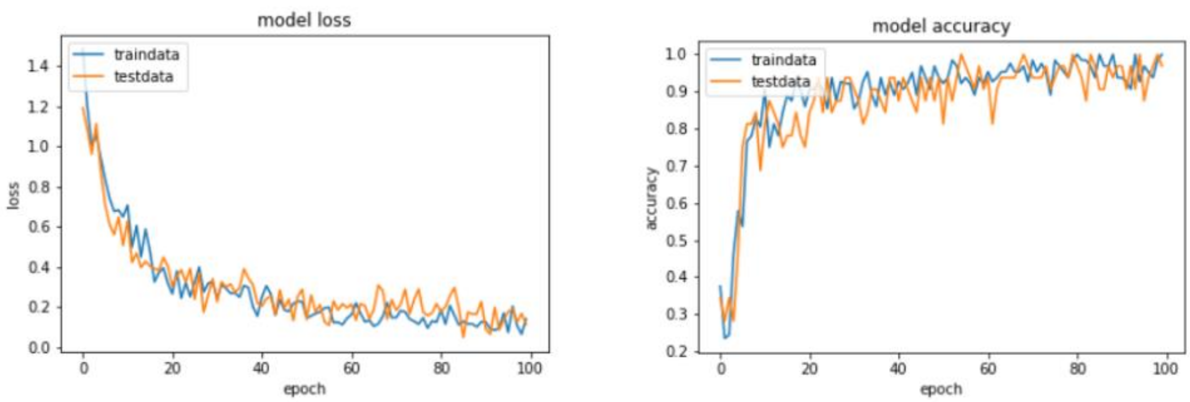
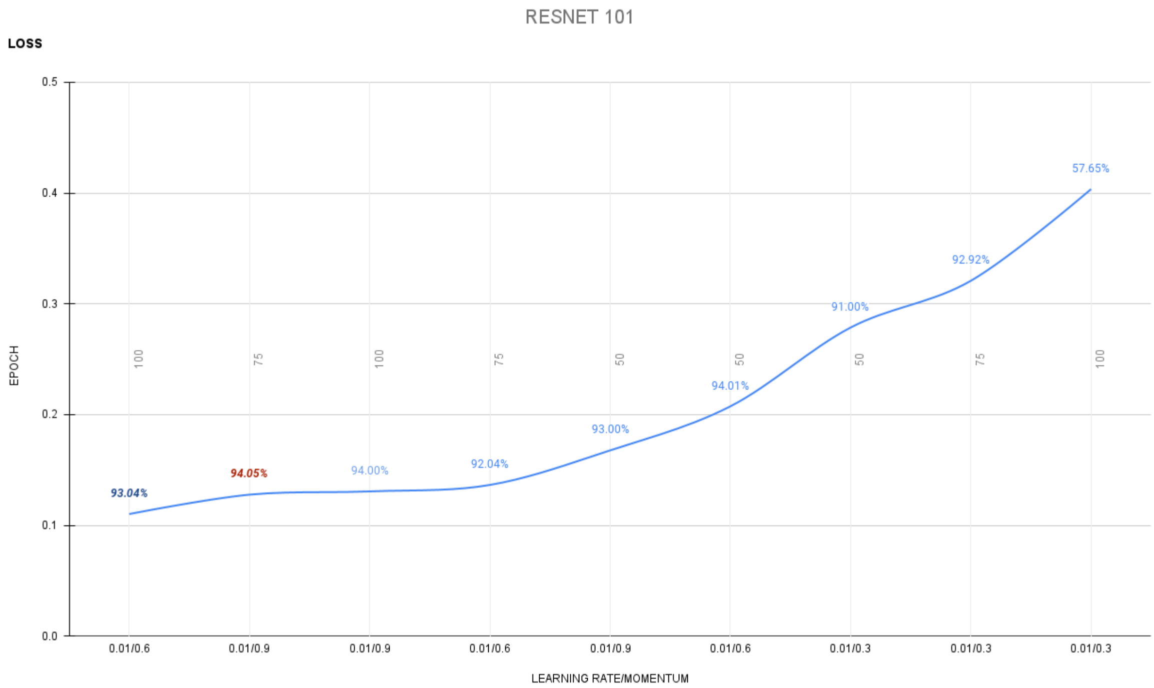
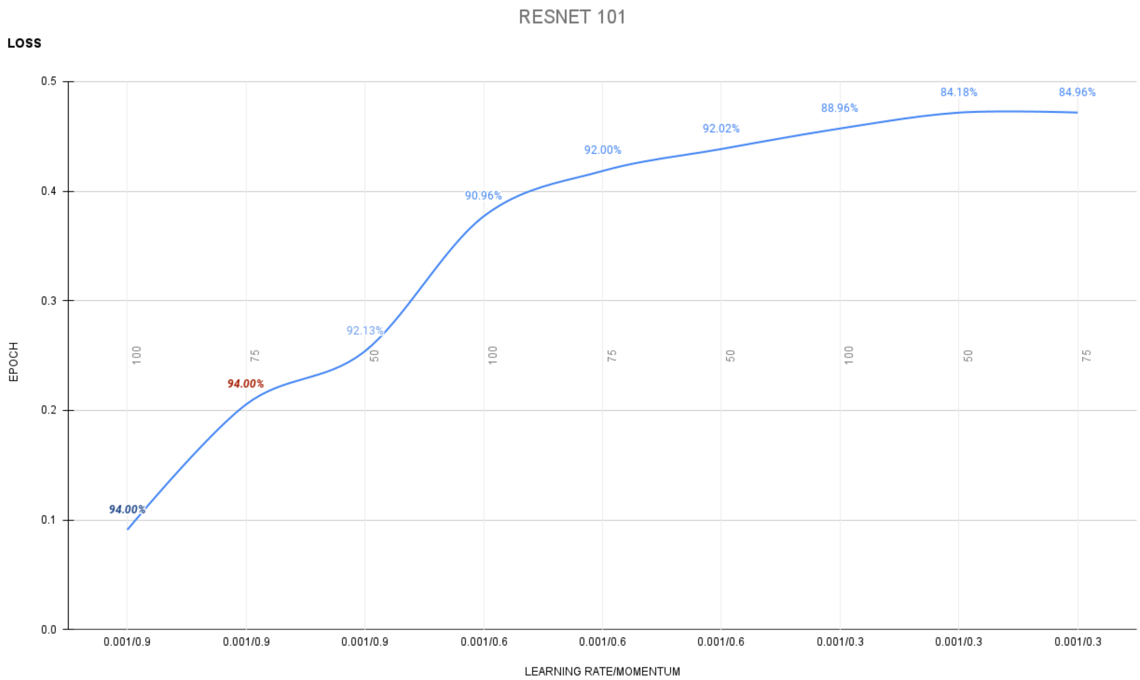


Figure 10: Loss and Accuracy Graph ResNet50

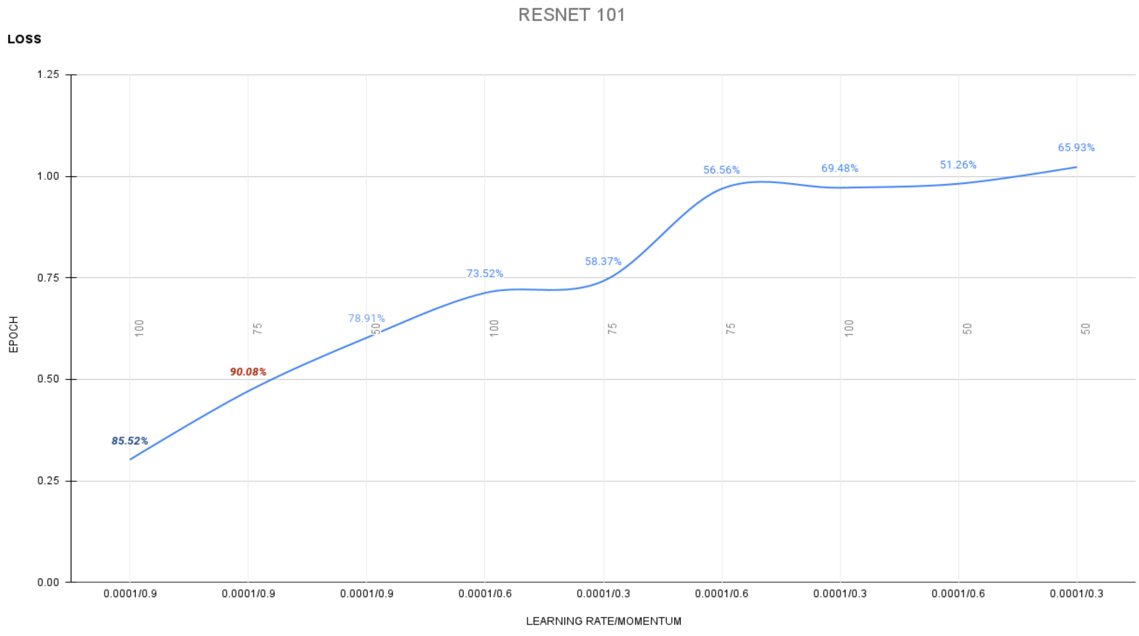
3.2 ResNet101 Analysis



(A)



(B)



(C)

Figure 11: ResNet101 (A) learning rate 0.01 (B) learning rate 0.001 (C) learning rate 0.0001

Figure 12 is the same as in Figure 9 which represents 3 graph images for each scenario that has been tried, from Figure 12 it can be seen that momentum 0.3 has a high tendency of loss and a low F1 score, the same as ResNet 50, momentum with a value of 0.9 have a tendency to get a low loss and a high F1 score

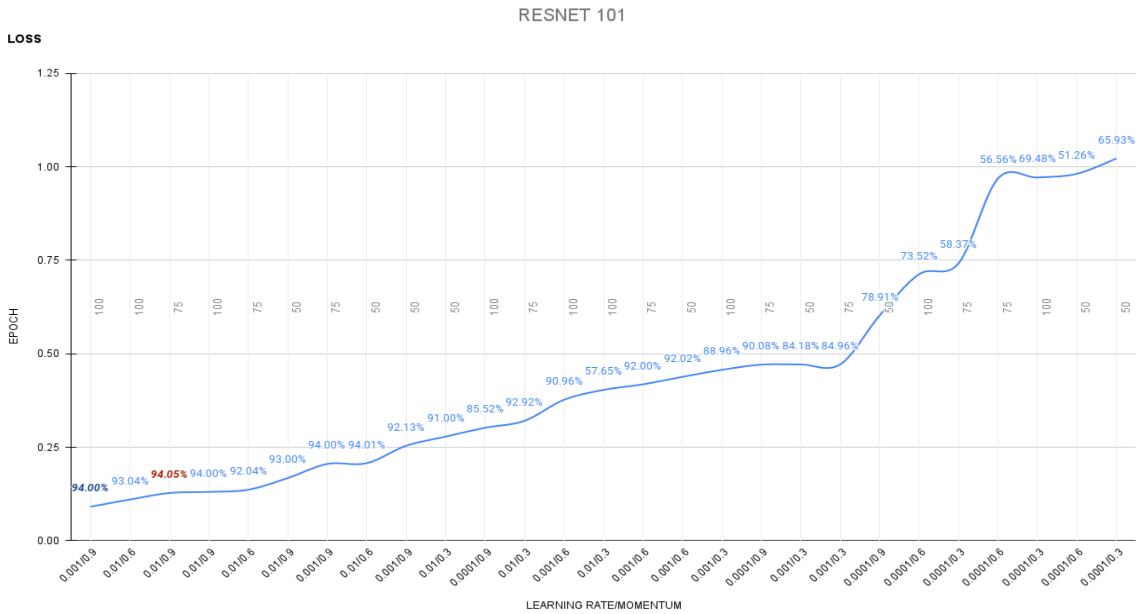


Figure 12: Overall Graph of ResNet101 Scenario

Figure 13 is an overall graph of the scenario that has been tested on ResNet101, from the graph it can be seen that there are 2 interesting scenario, the first scenario (blue with bold font on the Figure 13) has the lowest loss 0.110326200723648 but have F1 score 94.00%, recall 94.00% precision 94.03% while scenario 2 (red with bold font on the figure 13) have the highest F1 score with 94.05% F1 score, 94.05% Recall and 94.24% Precision but have loss of 0.130744844675064, from the those two output scenario it can be seen that scenario 2 have F1 score, precision and recall higher than the first scenario but the second scenario have lowest loss which mean more able to generalize [12] to the new image so that the first scenario is choosen which have loss 0.110326200723648, F1 score 94.00%, recall 94.00% precision 94.03%. The detail of accuracy graph and loss graph of scenario 1 can be seen on Figure 14

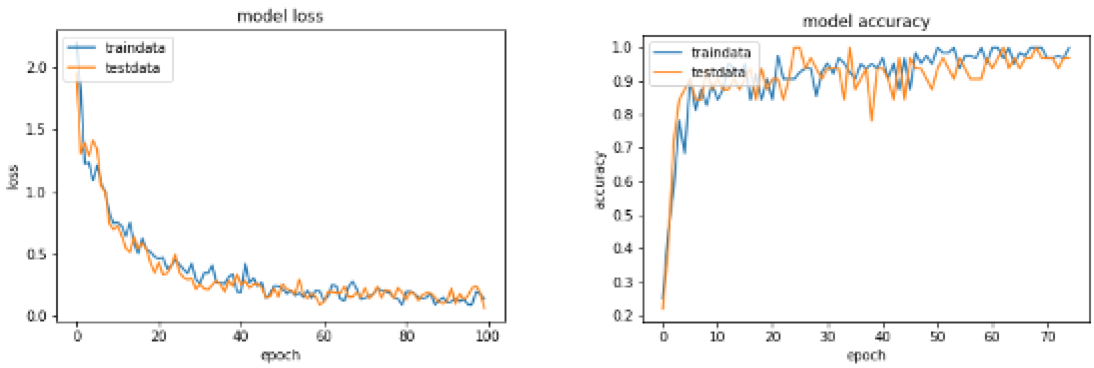


Figure 13: Loss and Accuracy Graph ResNet50

3.3 Summary

After doing all scenario trials for each ResNet50 and ResNet101 architecture, a summary can be obtained that can be compared to find the best, the best scenario results for each architecture with its hyperparameters can be seen in Table 3

Table 3: Summary of Trials

Arch	Loss	F1-Score	Precission	Recall	LR	Momentum	Epoch
ResNet50	0.110296231150675	95.96%	95.96%	95.96%	0.01	0.9	100
ResNet101	0.110326200723648	94.00%	94.00%	94.03%	0.001	0.9	100

From Table 3 data can be seen that ResNet50 gets a lower Loss and a higher F1 score than ResNet101 in the same scenario tested

4. Conclusion

After conducting trial and error of the 2 architectures ResNet50 and ResNet101 with the same scenario it can be seen that the momentum of 0.3 in both architectures has the tendency to produce high loss with a low f1 score, while for a learning rate of 0.0001 on both ResNet50 and ResNet101 architectures also has a tendency to produce high loss and a low f1 score, optimal hyperparameter values range in the range of 0.001 to 0.01 for learning rate and 0.6 - 0.9 for momentum. After comparing the two architectures by trial and error, ResNet50 is the best because it produces the lowest loss and highest F1score, precision, with each value for loss 0.110296231150675 and 95.96% for F1Score, 95.96% precision and 95.96% recall

References

[1] M. A. Pangestu and H. Bunyamin, "Analisis Performa dan Pengembangan Sistem Deteksi Ras Anjing pada Gambar dengan Menggunakan Pre-Trained CNN Model," *J. Tek. Inform. dan Sist. Inf.*, vol. 4, pp. 337–344, 2018.

[2] A. S. Assiri, S. Nazir, and S. A. Velastin, "Breast Tumor Classification Using an Ensemble Machine Learning Method," *J. Imaging*, vol. 6, no. 6, 2020, doi: 10.3390/JIMAGING6060039.

[3] B. K. R, B. K, hreyas M. S, onu S, A. H, and A. B. J, "SVM Based Plant Diseases Detection using Image Processing," *Int. J. Comput. Sci. Eng.*, vol. 7, no. 5, pp. 1263–1266, 2019, doi: 10.26438/ijcse/v7i5.12631266.

[4] S. Ghosh, A. Singh, Kavita, N. Z. Jhanjhi, M. Masud, and S. Aljahdali, "SVM and KNN Based CNN Architectures for Plant Classification," *Comput. Mater. Contin.*, vol. 71, no. 2, pp. 4257–4274, 2022, doi: 10.32604/cmc.2022.023414.

[5] K. Myagila and H. Kilavo, "A Comparative Study on Performance of SVM and CNN in Tanzania Sign Language Translation Using Image Recognition," *Appl. Artif. Intell.*, vol. 36, no. 1, 2022, doi: 10.1080/08839514.2021.2005297.

[6] P. Hridayami, I. K. G. D. Putra, and K. S. Wibawa, "Fish species recognition using VGG16 deep convolutional neural network," *J. Comput. Sci. Eng.*, vol. 13, no. 3, pp. 124–130, 2019, doi: 10.5626/JCSE.2019.13.3.124.

[7] A. Viebke and S. Pllana, "The Potential of the Intel (R) Xeon Phi for Supervised Deep Learning," in *2015 IEEE 17th International Conference on High Performance Computing and Communications, 2015 IEEE 7th International Symposium on Cyberspace Safety and Security, and 2015 IEEE 12th International Conference on Embedded Software and Systems*, 2015, pp. 758–765, doi: 10.1109/HPCC-CSS-ICESS.2015.45.

[8] K. Zhou, L. Zhuo, Z. Geng, J. Zhang, and X. G. Li, "Convolutional Neural Networks Based Pornographic Image Classification," in *2016 IEEE Second International Conference on Multimedia Big Data (BigMM)*, 2016, pp. 206–209, doi: 10.1109/BigMM.2016.29.

[9] C. B. S. Maior, J. M. M. Santana, I. D. Lins, and M. J. C. Moura, "Convolutional neural network model based on radiological images to support COVID-19 diagnosis: Evaluating database biases," *PLoS One*,

vol. 16, no. 3, pp. 1–25, 2021, doi: 10.1371/journal.pone.0247839.

[10] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-Decem, pp. 770–778, 2016, doi: 10.1109/CVPR.2016.90.

[11] N. ALHAWAS and Z. TÜFEKCİ, "The Effectiveness of Transfer Learning and Fine-Tuning Approach for Automated Mango Variety Classification," *Eur. J. Sci. Technol.*, no. 34, pp. 344–353, 2022, doi: 10.31590/ejosat.1082217.

[12] A. Fitrianto and B. Sartono, "Image Classification Modelling of Beef and Pork Using Convolutional Neural Network," *Int. J. Sci. Basic Appl. Res. Int. J. Sci. Basic Appl. Res.*, vol. 57, no. 2, pp. 26–38, 2021, [Online]. Available: http://gssrr.org/index.php?journal=JournalOfBasicAndApplied.