Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 3, July 2023
DOI : https://doi.org/10.33395/sinkron.v8i3.12721
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Comparison of Residual Network-50 and Convolutional Neural Network Conventional Architecture For Fruit Image Classification

Arie Satia Dharma[1)*], Judah Michael Parluhutan Sitorus[2)], Andreas Hatigoran[3)]
[1,2,3)]Institut Teknologi Del, Laguboti, Indonesia
[1)]ariesatia@del.ac.id, [2)]judahmichael0902@gmail.com, [3)]andreashatigoran@gmail.com

Submitted : July 18, 2023 | **Accepted** : July 28, 2023 | **Published** : July 30, 2023

Abstract: Classification of fruit images using machine learning technology has had a significant impact on human life by enabling accurate recognition of various fruits. With the advancements in technology, machine learning architectures have become increasingly diverse and sophisticated, providing enhanced capabilities for fruit image classification. However, previous studies have primarily focused on classifying fruits at a basic level. Therefore, there is a growing need for the development and application of Fruit Image Classification systems within the community, particularly in the field of agriculture. Such applications can play a pivotal role in leveraging technology to benefit the agricultural sector, empowering users to gain satisfaction and knowledge regarding different fruits through the utilization of these applications. In this study, we employ both a conventional Convolutional Neural Network (CNN) architecture and a Residual Network-50 for fruit image classification. To ensure robust performance evaluation, the dataset is divided into training and testing subsets, with fruits categorized into specific classes. Furthermore, identical preprocessing and optimization techniques are applied to both architectures to maintain consistency and fairness during the evaluation process. The results of our classification experiments on a dataset consisting of 17 different fruit classes reveal that the conventional CNN architecture achieves an impressive accuracy of 0.998 (99%) with a minimal loss of 0.009. On the other hand, the Residual Network-50 demonstrates a slightly lower accuracy of 0.994 (99%) but with a slightly higher loss of 0.02. Despite the higher loss, the Residual Network-50's accuracy remains comparable to that of the conventional architecture, showcasing its potential for fruit image classification. By leveraging the power of machine learning and these advanced architectures, fruit image classification systems can provide valuable insights and assistance to users. They can facilitate informed decision-making in various domains, including agriculture, food production, and consumer education.

Keywords: *Accuracy; Convolutional Neural Network; Fruit Classification; Machine Learning; Residual Network-50*

INTRODUCTION

Classification of image data using machine learning technology has significantly impacted human life. However, with the advancement of technology, there are now numerous machine learning architectures available. Previous research focused on fruit classification using the Convolutional Neural Network (CNN) algorithm but only at the level of fruit categories without comparing other architectures to determine which one is better in fruit classification (Maulana & Rochmawati, 2020). This limited the application of machine learning to classification only, without exploring alternative architectures that might perform better in fruit classification.

Therefore, this study aims to compare different architectures to determine which one is more effective in fruit image classification. The researcher will utilize the Convolutional Neural Network (CNN) algorithm, which is renowned for its excellent image recognition capabilities. CNN has been extensively used in various applications, such as autonomous driving, robotics, drones, security, disease research, and more (Ilahiyah & Nilogiri, 2018). It is considered the leading algorithm for object detection and recognition tasks. CNN can also be applied in Computer Vision tasks to detect objects in images, including categorizing them based on their types.

*Arie Satia Dharma

Previous studies employed the CNN algorithm to evaluate CNN models and classify fruit images into specific classes. However, this study aims to compare architectures specifically for tropical fruits and develop an application that provides information about the fruit's content. The architectures to be investigated are the conventional CNN architecture and the ResNet50 architecture. ResNet50 is chosen as a benchmark because research by (Liang & Zheng, 2020) demonstrated that the Residual Network, utilizing Adam optimization, outperformed other modern architectures such as VGG16, DenseNet121, Xception, and InceptionV3 in X-ray image object detection based on Confusion Matrix evaluation techniques. This indicates that ResNet is not only a modern architecture but also one of the best among other modern architectures.

The Comparative Study of Convolutional Neural Network Architectures for Fundus Classification (Setiawan, 2020) discusses the comparison of various Convolutional Neural Network (CNN) architectures for classifying fundus images. The experimental setup consists of two stages: without optimization and with optimization. In the first stage, VGG19 achieved the highest accuracy of 89.3%, while Gradient Descent, VGG16 achieved an accuracy of 92.31%. RmsProp optimization, ResNet50 achieved the highest accuracy of 88.5%, and with Adam optimization, AlexNet achieved an accuracy of 90.7%. The research findings indicate that the newest architectures are not necessarily the best for image classification. The differentiating factors are the optimization techniques employed and the specific objects used. Therefore, the researcher intends to investigate the comparison between a conventional CNN and Residual Network 50, which is a modern architecture and one of the best architectures for machine learning classification. However, this research will utilize images of fruits as the objects of study. The optimization technique chosen for this study is Adam, which combines the strengths of RmsProp and Gradient Descent optimizations. Adam is a commonly used optimization method that is suitable for large datasets and has been tested extensively for solving optimization problems in machine learning (Kingma & Ba, 2015).

The study will ensure uniform treatment of both classification processes, including data splitting, preprocessing, and the number of epochs, to ensure a fair comparison between the two architectures. Previous research by (Yuliani, Aini, & Khasanah, 2020) found that increasing the number of epochs and steps per epoch leads to higher training accuracy. Additionally, a study by (Nashrullah, Wibowo, & Budiman, 2020) investigating the role of epochs in ResNet-50 architecture for pornographic image classification concluded that the number of epochs significantly influences the performance of CNN classification. Increasing the number of epochs allows the model to achieve better performance and generalize the data effectively. Therefore, this research will use a higher number of steps per epoch compared to previous studies, namely 30 epochs.

Another challenge lies in the extensive variation of fruits, making it difficult to identify their composition and types. Although there are thousands of known fruit-bearing plant species, less than 10% are cultivated, and there are only around 300 types of fruits consisting of 61 families and 148 genera, both native and imported, found in Indonesia (Angio & Irawanto, 2019). With such diversity, it becomes challenging to gain comprehensive knowledge about various fruit compositions. The application of machine learning in the field of agriculture and food technology, in general, is currently limited to fruit name recognition and classification. This limitation hinders the development and utilization of technology in the agricultural and food technology sectors. By creating a detailed fruit classification system using the best machine learning architecture, it would greatly benefit the general public. Developing a fruit recognition application based on deep learning technology for Android, which not only identifies fruits but also provides information about their composition, can address these challenges. This solution leverages machine learning algorithms to create a fruit recognition application that offers brief information about different fruits. By implementing this solution, it is expected to facilitate human identification of fruits and support the advancement of machine learning classification systems (Paraijun, Aziza, & Kuswardani, 2022).

To summarize, this study proposes a concise and clear solution to address the aforementioned challenges by developing a fruit recognition application using deep learning technology. The application will utilize machine learning algorithms to recognize fruits and provide brief information about their composition. By implementing this solution, it aims to simplify fruit identification for individuals and contribute to the advancement of machine learning classification systems.

LITERATURE REVIEW

Convolutional Neural Network

This algorithm belongs to the type of neural network capable of processing high-dimensional data and has interconnected neurons. The Convolutional Neural Network is a deep learning algorithm commonly used for image classification, as well as identifying faces, traffic signs, tumors, and other visual data aspects (Arrofiqoh & Harintaka, 2018). CNN consists of three interconnected layers to process input data. The feature extraction layer

e-ISSN : 2541-2019
p-ISSN : 2541-044X

consists of convolution and pooling layers. Both layers enable the processing and learning of input data characteristics. The next layer is the fully connected layer used to classify input data. In the realm of deep learning, CNN is recognized as a beneficial method for discovering relevant features and forming nonlinear hypotheses to enhance model complexity. Due to the high level of complexity, training deep CNN models often requires a significant amount of time. This is why the use of GPUs has become widespread in deep learning, as they expedite the training process.This is the illustration of the algorithm:

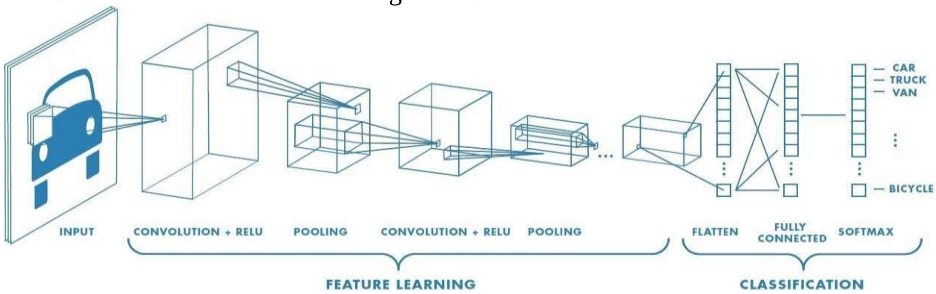

Fig 1. Convolutional Neural Network Architecture

In Figure 1, the initial step in the CNN architecture is the convolution stage. The convolution stage is performed using a kernel with a customized size. The number of kernels used depends on the number of features produced. The next step is the activation function, which uses the ReLU (Rectifier Linear Unit) activation function. After the activation function is completed, the next step goes through the pooling process. This process will be repeated several times until a suitable feature map is obtained to be continued to the fully-connected neural network, and the output class is obtained from the fully-connected network.

Convolution Layer

The Convolution Layer is a part of the CNN architecture stage. Convolution is a mathematical term in which a function is applied repeatedly to the output of another function (Patil & Rane, 2021). The convolution operation applies the output function as a feature map of the input image. Both input and output can be seen as two real-valued arguments. In the Convolutional Layer, neurons are arranged to form a filter with a length and height (pixels). For example, the initial layer in the feature extraction layer is generally convolution. The layer has a size of 5 x 5 x 3, with a length of 5 pixels, a height of 5 pixels, and a thickness of 3 channels from the image. The three filters are then shifted to all parts of the image, and each shift will perform a "dot" operation between the input and the value of the filter, resulting in an output commonly referred to as a feature map.

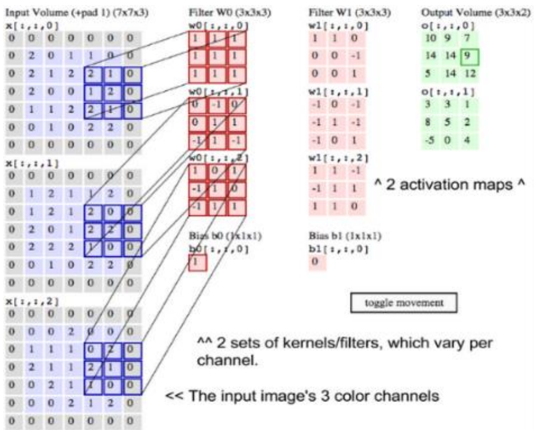

Fig 2. Convolution Layer Process

After the Convolutional Layer produces the values from the convolution process, ReLU (Rectified Linear Unit) is the activation function responsible for normalizing the resulting values (Sitepu & Sigiro, 2021). The following is the mathematical process of the ReLU function equation:

$$ReLU(x) : \{X, IF\ X > 0\}0, otherwise \text{ ....formula(i)}$$

Pooling Layer

The Pooling Layer is located after the Convolution Layer. This layer performs the Max pooling process, which takes the maximum value from each channel of the convolution and ReLU output. The maximum value used by the Pooling Layer input between consecutive convolution layers in the CNN model architecture can progressively

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 3, July 2023
DOI : https://doi.org/10.33395/sinkron.v8i3.12721
e-ISSN : 2541-2019
p-ISSN : 2541-044X

reduce the output volume size on the feature map, thereby reducing the number of parameters and calculations in the network to control overfitting. The following is an illustration of the max-pooling operation:

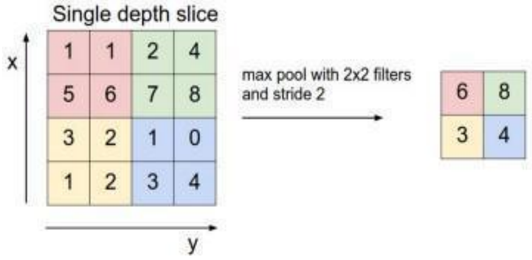

Fig 3. Max-Pooling Operation

Figure 3 shows the process of max pooling. The output of the pooling is a matrix with dimensions that are smaller than the original image. The pooling layer operates on each depth slice of the input volume in turn. In the image, max pooling operation is performed using a 2 x 2 filter size. The input of the process is 4x4 in size, and the maximum value is taken from each set of 4 numbers in the input operation, resulting in a new output size of 2x2.

$$f(x) : Max(0, x) \text{ ....} \textit{formula(ii)}$$

Fully Connected Layer

The Fully-Connected Layer is a layer where each neuron is the result of the activation process from the previous layer connected to neurons in the next layer, similar to a general neural network. This layer is commonly used in MLP (Multi Layer Perceptron) to transform the dimensions of data so that each data can be classified linearly (Sakib, Ahmed, Kabir, & Ahmed, 2018). In this layer, there is a mapping process from the flattened vector result using a dense layer to each respective class. Then, the probability of each class is calculated using the softmax function. The equation for the softmax function is as follows:

$$softmax(x) = \frac{e^{xi}}{\sum_{j=1}^{n} e^{xj}} \text{ ....} \textit{formula(iii)}$$

The difference between a fully-connected layer and a regular convolutional layer is that only certain areas of the input neurons in the convolutional layer are connected, while the fully-connected layer has neurons that are connected overall (Upreti, 2022). However, both layers still perform dot operations, so there is not much difference in their function. The following is the process of the fully-connected layer:

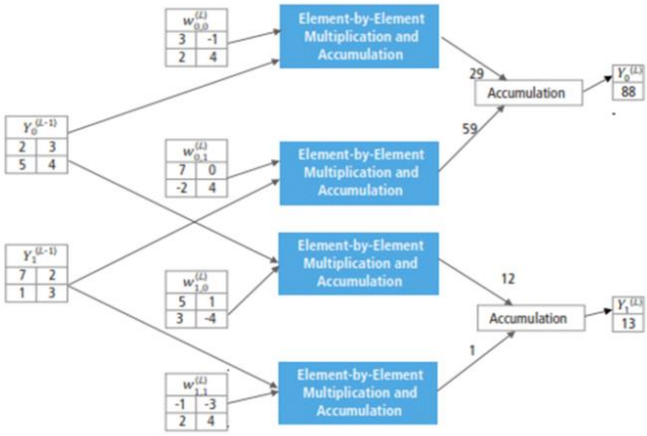

Fig 4. Fully-Connected Layer Process

Residual Network 50

ResNet, short for Residual Networks, is a classic neural network. Residual Networks were introduced by Kaiming in their research entitled "Deep Residual Learning for Image Recognition", which won the ILSVRC competition in 2015 (Shafiq & Gu, 2022) (Yu et al., 2021). Residual Networks were created to facilitate the training process of neural networks that are much deeper than previously trained models (He, Zhang, Ren, & Sun, 2016). This architecture explicitly reformulates its layers with residual function learning that references the input layer, rather than learning unreferenced functions (Wen, Li, & Gao, 2020). The working principle of Residual Networks is that, compared to conventional networks, this architecture builds a deeper network and optimizes the layers to eliminate the problem of vanishing gradients. In Residual Networks, there is a process that distinguishes it from conventional architectures, namely Batch Normalization, where this stage will reduce covariance shift or equalize the distribution of input values because changes in the training process result in constantly changing

*Arie Satia Dharma

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 3, July 2023
DOI : https://doi.org/10.33395/sinkron.v8i3.12721
e-ISSN : 2541-2019
p-ISSN : 2541-044X

values. Batch normalization is a technique for normalizing activations in the deepest layers of neural networks. This technique continuously improves activations to have a mean of zero and a standard deviation of one, allowing for a larger gradient step and faster processing.
mini batch mean formula:

$$\mu\ B \leftarrow \frac{1}{m}\sum_{i=1}^{m} \quad xi \text{ ....}\textit{formula(iv)}$$

mini batch varian formula:

$$\sigma_B^2 \leftarrow \frac{1}{m}\sum_{i=1}^{m} \quad (xi - \mu B)^2 \text{ ....}\textit{formula(v)}$$

annotation:
m = column number of pixels
μB = average value of the batch
σB = standard deviation of minibatches

Normalization input

$$\hat{x}i \leftarrow \frac{xi-\mu B}{\sqrt{\sigma_B^2+\epsilon}} \text{ ....}\textit{formual(vi)}$$

Keterangan :
$\hat{x}$ = zero-centered and input normalization
ϵ = additional value ,usually 10^{-8}

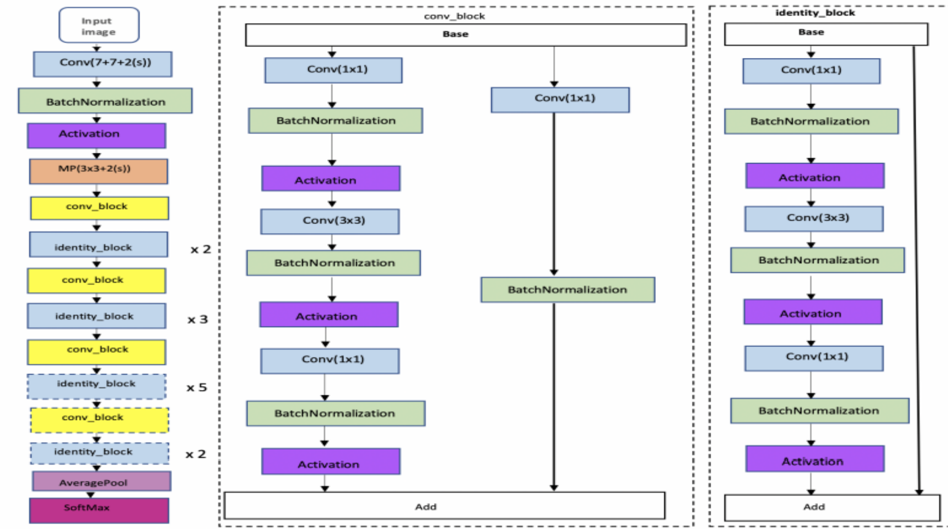

***Fig 5.* Architecture Residual Network 50**

Initially, an image with a size of 255 x 255 pixels from the input layer will be convolved with a kernel size of 7x7 and stride 2. The results of the convolution will get output for the feature map and will be normalized by Batch Normalization. Batch Normalization is useful for reducing covariance shifts or equalizing the distribution of input values because changes in the training process produce values that are always changing.Batch Normalization goes through several processes, namely mini batch mean, mini batch variant, normalization and scale shift.After the process is complete, it will then enter the activation process which uses the ReLu function to make the results of feature extraction non-linear. The output value of ReLU will be reduced at the maxpool layer first before continuing with the next convolution process. Between the next to the fifth convolution stage, the feature extraction process will be carried out by a combination of conv-block and identity-block. After the extraction process has been completed, then the feature map will be forwarded to the fully connected layer and then a classification decision will be made assisted by the softmax function. From the analysis carried out on the architectures of Conventional CNN and Residual Network 50 apart from the optimization carried out, the researcher has an initial hypothesis where Residual Network 50 has a better architecture than Conventional CNN because it has a structure that can manage the deepest network from training properly (Wu, Ying, Zhou, Pan, & Cui, 2023).

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 3, July 2023
DOI : https://doi.org/10.33395/sinkron.v8i3.12721
e-ISSN : 2541-2019
p-ISSN : 2541-044X

METHOD

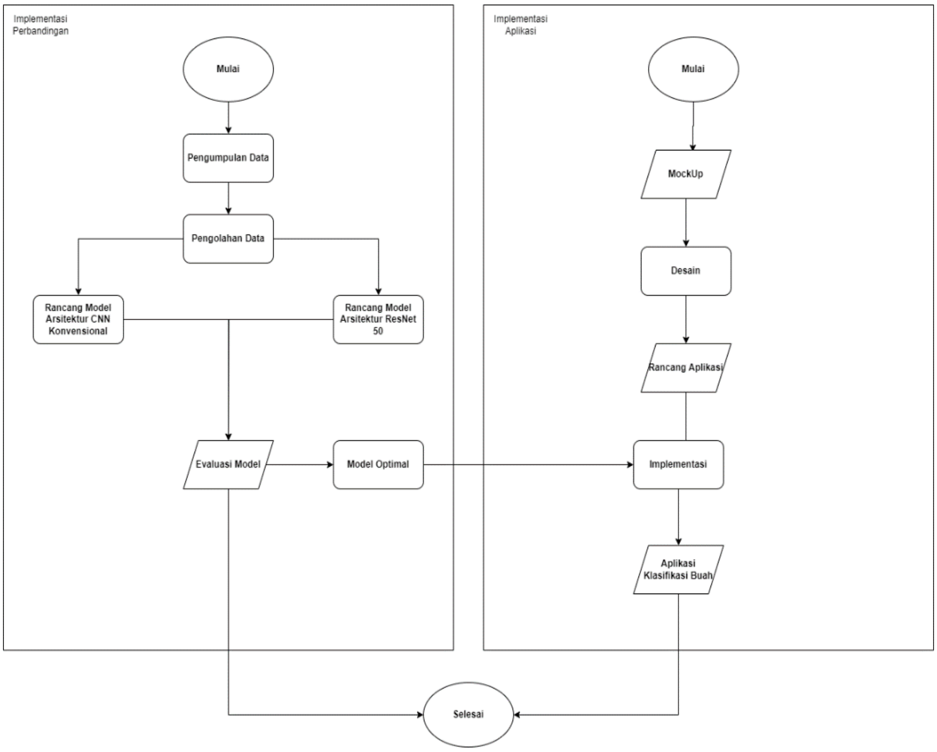

Fig 6. Research Flow

The figure above illustrates the research design in making fruit classification applications. This stage consists of two parts, namely for Comparison and Application Implementation. The architectural comparison design process will begin by observing previous research related to image classification using the Convolutional Neural Network architecture and also the Residual Network-50. After analyzing the previous research, the researcher looked for fruit datasets that would be used as architectural model training data and would classify fruit datasets that belong to tropical fruit. From the 7995 fruit dataset images that have been collected, the augmentation will be used to analyzed what will be part of the preprocessing to make the model better in doing learning. The preprocessing carried out on the data for the first training is to normalize each fruit in the dataset in size 255 x 255, after normalization, the fruit data will be sharpened with the sharpen function, then the image will be smoothed so that the color is even, then the fruit dataset Rotate function will be given so that the fruit can be recognized not only in an upright state. The processed dataset will become material for the model as learning from the model architecture. After that the data will be splitted by a comparison of the train data and test data of 60:40, 70:30, 80:20, 90:10 for each class, then the design model will be made based on the type of architecture. Residual Network-50 Architecture and Conventional CNN will be trained using the data and then will be tried to classify fruit. After the fruit is classified, the results of the model evaluation will be taken. Accuracy and f-measure will be taken from both architectures to see which architecture is the best in classifying fruit images.

RESULT

Splitting Datasets

The datasets used consists of 16 types of fruits with a total of 7995 image data, the distribution of datasets of train data and test data is divided by a ratio of 60:40, 70:30, 80:20, and 90:10. Types of fruit and distribution of the dataset into training data and test data as shown in Table 1.

Table 1 Distribution of Fruit Datasets

No.	Nama Buah	Source Data	Ratio of Splitting Datasets							
			60	: 40	70	: 30	80	: 20	90	: 10
1	Mangosteen	300	180	120	210	90	240	60	270	30
2	Papaya	492	295	197	344	148	394	98	443	49

*Arie Satia Dharma

3	Kiwi	466	280	186	326	140	373	93	419	47
4	Leci	490	294	196	343	147	392	98	441	49
5	Banana	450	270	180	315	135	360	90	405	45
6	Pineapple	490	294	196	343	147	392	98	441	49
7	Orange	499	299	200	349	150	399	100	449	50
8	Avocado	427	256	171	299	128	342	85	384	43
9	Guava	490	294	196	343	147	392	98	441	49
10	Passion Fruit	510	306	204	357	153	408	102	459	51
11	Mango	492	295	197	344	148	394	98	443	49
12	Coconut	490	294	196	343	147	392	98	441	49
13	Corn	450	270	180	315	135	360	90	405	45
14	Watermelon	475	285	190	332	143	380	95	427	48
15	Limes	490	294	196	343	147	392	98	441	49
16	Rambutan	492	295	197	344	148	394	98	443	49
17	Pomegranate	492	295	197	344	148	394	98	443	49

Preprocessing Datasets

After the data that has been divided is ready to enter the preprocessing stage, the dataset will be processed to be normalized so that the size of the image data is uniform, besides that, data smoothing is carried out using the Gaussian blur function, then the dataset will be image enhancement by sharpening the colors of the dataset. The following are the results of preprocessing image.

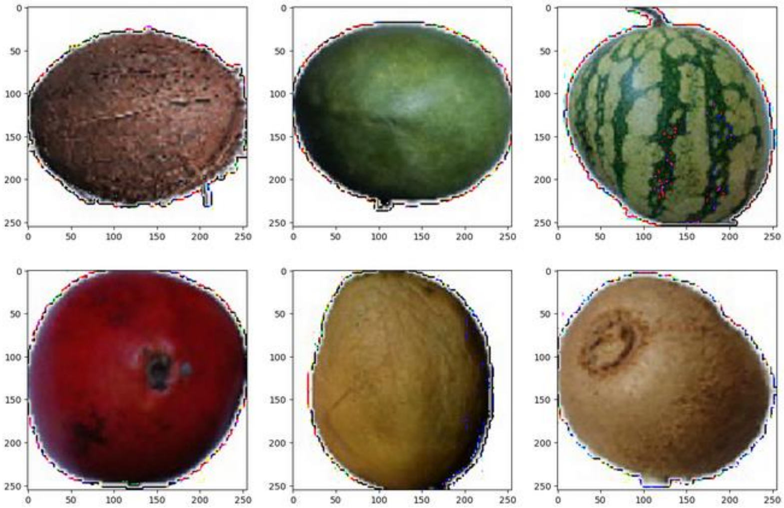

Fig 7. Fruits Image

Convolutional Neural Network Train Models

We built the CNN architecture for the training model using 3 convolution layers, 2 max pooling layers, and 1 fully connected layer. In this architecture we define a number of parameters used such as kernel size, number of filters, and stride to form the activation shape and activation size as shown in the table 2.

Table 2 CNN Model Parameters

Layer	Number of Filters	Activation Shape	Activation Size
Input Image	-	(100,100,3)	30000
Conv2d(f=7,s=1)	32	(100,100,32)	320000
MaxPool(f=2,s=2)	-	(50,50,32)	80000
Conv2d(f=3,s=1)	64	(24,24,64)	36864
MaxPool(f=2,s=2)	-	(12,12,64)	9216
Conv2d(f=3,s=1)	128	(5,5,128)	3200

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 3, July 2023
DOI : https://doi.org/10.33395/sinkron.v8i3.12721
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Flatten	-	(3200,1)	3200
Dense	-	(100,1)	100
Softmax	-	(17,1)	17

For hyper parameter we using Adam optimizer, learning rate of 0.001, ReLu activation function, 1000 epoch, 30 batch, and dropout of 0.2. We train ratio of splitting datasets using the same model to get accuracy values as shown in the table 3.

Table 3 Accuracy Ratio of Splitting Datasets

Model Evaluate	60:40	70:30	80:20	90:10
Loss	0.0295	0.0221	0.0049	0.0205
Accuracy	0.9961	0.9941	0.9980	0.9960

The results of the evaluation of the architectural model from the distribution of dataset in the table 3 shown that optimal ratio of splitting data is 80:20 that obtained an accuracy of 0.998 (99%) with a loss of 0.0049. Then we will use the ratio 80:20 for evaluation of CNN architecture that uses 3 times the convolution process and also 2 times does the pooling. The results of the classification report using the confusion matrix obtained the following results.

Table 4 Classification Report CNN Conventional

	Precision	Recall	F1-Score	Support
0	1.00	1.00	1.00	86
1	1.00	1.00	1.00	90
2	1.00	1.00	1.00	98
3	1.00	1.00	1.00	90
4	1.00	1.00	1.00	98
5	0.99	1.00	1.00	94
6	1.00	0.99	0.99	98
7	1.00	1.00	1.00	98
8	1.00	1.00	1.00	99
9	0.98	0.95	1.00	60
10	1.00	1.00	1.00	100
11	1.00	1.00	1.00	99
12	1.00	0.99	0.99	102
13	1.00	1.00	1.00	98
14	1.00	0.99	0.99	99
15	1.00	1.00	1.00	99
16	1.00	1.00	1.00	95
Accuracy			0.998	1603

$$Accuracy = \frac{1.00 + 1.00 + 1.00 + 1.00 + 1.00 + 0.99 + 1.00 + 1.00 + 1.00 + 1.00 + 1.00 + 0.99 + 1.00 + 0.99 + 1.00 + 1.00}{17} = 0.998\ (99\%)$$

*Arie Satia Dharma

e-ISSN : 2541-2019
p-ISSN : 2541-044X

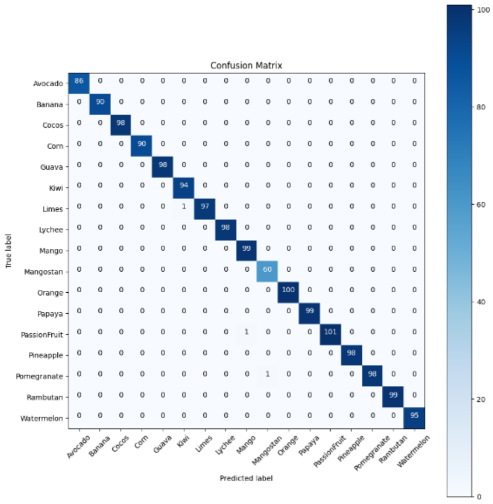

***Fig.* 8 Confusion Matrix CNN Conventional Table**

From the results of the data above, no more than 1 fruit was missed in the recognition of fruit by Conventional CNN. This can happen because the training process is carried out as many as 1000 epochs which makes the algorithm better at recognizing the fruit itself. From the results above, precision can be seen from how the predicted label matches the true label and recall, seen from how well the model can remember the fruits that have been classified in the table above. It can be seen how the predicted label matches the true label of certain fruits.

Residual Network-50 Train Models

We built the Restnet50 architecture for the training model using 5 iteration that consist convolution layers, batch normalization, identity block, max pooling layers, and fully connected layer. We train ratio of splitting datasets using the same model to get accuracy values as shown in the table 3.

Table 5 Accuracy Ratio of Splitting Datasets RestNet50

Model Evaluate	60:40	70:30	80:20	90:10
Loss	0.0179	0.0162	0.0210	0.0178
Accuracy	0.9922	0.9941	0.9943	0.9902

The results of the evaluation of the architectural model from the distribution of dataset in the table 5 shown that optimal ratio of splitting data is 80:20 that obtained an accuracy from the test results with test data are 99% with the loss obtained is 0.02. Then we will use the ratio 80:20 for evaluation of RestNet50 architecture that use 5 iteration. From the graph it can also be seen that the validation of accuracy is also high but there are also many significant losses, this can occur because the Residual Network shifts covariance and also reduces unnecessary features and parameters. The Residual Network 50 architecture uses 5 stages where at each stage there is a convolutional block and an identity block in which there are Batch Normalization, activation and convolution processes. In this study using epochs of 1000 epochs in order to provide the best information for architecture and architecture can learn better in image classification. From the model training process, the accuracy is 99% (0.994) with a loss of 0.02 from the test data.

	Precision	Recall	F1-Score	Support
0	1.00	1.00	1.00	86
1	1.00	1.00	1.00	90
2	0.97	1.00	0.98	98
3	1.00	1.00	1.00	90
4	1.00	1.00	1.00	98
5	1.00	1.00	1.00	94
6	1.00	1.00	1.00	98

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 3, July 2023
DOI : https://doi.org/10.33395/sinkron.v8i3.12721
e-ISSN : 2541-2019
p-ISSN : 2541-044X

7	0.98	1.00	0.99	98
8	1.00	1.00	1.00	99
9	1.00	1.00	1.00	60
10	1.00	1.00	1.00	100
11	1.00	1.00	1.00	99
12	1.00	0.94	0.97	102
13	1.00	1.00	1.00	98
14	1.00	0.97	0.98	99
15	1.00	1.00	1.00	99
16	0.96	1.00	0.98	95
Accuracy			0.99	1603

Table 3 Classification Report Residual Network 50

$$Accuracy = \frac{1.00 + 1.00 + 0.98 + 1.00 + 1.00 + 1.00 + 1.00 + 0.99 + 1.00 + 1.00 + 1.00 + 1.00 + 0.97 + 1.00 + 0.98 + 1.00 + 0.98}{17}$$

$$= 0.994\ (99\%)$$

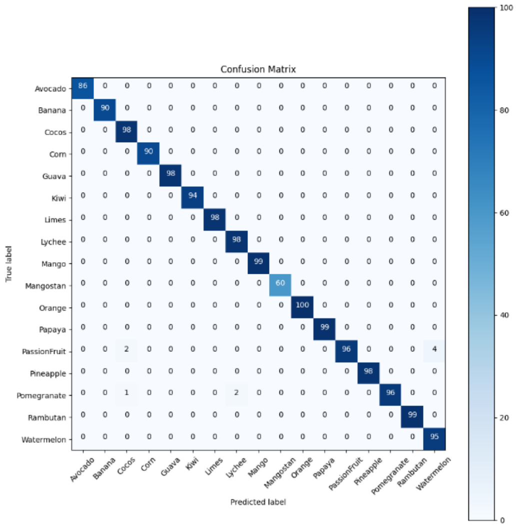

Fig. 10 Confusion Matrix Residual Network Table

From the results of the Confusion Matrix table above, it can be analyzed the results where the Residual Network 50 model has more errors than the Conventional Convolutional Neural Network, this is due to the large number of losses that occur as seen from the graph of the model training results.

DISCUSSIONS

Errors in predicting experienced by each architecture can occur due to the different number of parameters of the two architectures. The number of parameters possessed by the Conventional CNN architecture is more than ResNet-50, this is because the concept of Conventional CNN follows all convolution processes without any reduction of features while Residual Network, there is a process that distinguishes it from conventional architecture, namely Residual Blocks in which there are Batches Normalization which at this stage will reduce covariance shifts or equalize the value distribution of the input because changes in the training process produce values that are always changing where the technique for normalizing activation is in the deepest block layer of neural networks. This technique continuously refines the activation to a zero mean and unit standard deviation, which allows for a larger gradient step. And the long loss is due to the characteristics and design of the ResNet architectural model itself where the original information from the input will be forwarded to the output layer via the Skip Connection line which allows the network to correct errors that occur in the middle layers of the network.

*Arie Satia Dharma

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 3, July 2023
DOI : https://doi.org/10.33395/sinkron.v8i3.12721
e-ISSN : 2541-2019
p-ISSN : 2541-044X

When viewed from the graphs of the two architectures, the conventional CNN architecture graph has better stability than ResNet-50, this can happen because of the Batch Normalization. In addition, other factors may occur because the dataset used uses fruit images taken from different perspectives, so the model can be wrong in predicting fruit class.

CONCLUSION

The results of the comparison of the Conventional Convolutional Neural Network and Residual Network 50 architectures in the classification of subtropical fruit images show that the Conventional Convolutional Neural Network method is better applied, because it produces the highest accuracy value compared to the accuracy results of the Residual Network 50 architecture, even though the accuracy results obtained are the same but there is a significant difference in loss. The result of the highest accuracy of the Convolutional Neural Network is 0.99% (0.998) with a loss of 0.009 and the highest accuracy of the Residual Network 50 99% (0.994) with a loss of 0.02. From the results of the classification carried out on 17 classes of tropical fruit types, obtained information that can be maintained as expected from the results of fruit classification. This is evidenced by the results of the evaluation of the data test which showed a result of 0.99 for both architectures. The models that have been created for each Convolutional Neural Network Conventional and Residual Network 50 architectures have very good prediction accuracy for each fruit class supported by good preprocessing. good for training algorithms to learn to recognize fruits. The model built has considered implementing aspects of the image such as converting the image into color in this implementation, namely RGB, data normalization, adjusting brightness and sharpness as well as smoothing of image data. The implementation is carried out using hardware technology to accommodate models on Android. , the implementation was successful in giving the user the opportunity to take pictures or take from the image gallery of the fruit that you want to classify and then display the information data on the fruit content. The initial hypothesis from the researchers was not proven correct, this was due to processes carried out outside as well as from architectural performance such as a comparison of the number the parameters of both architectures, the preprocessing technique and the optimization technique applied proved that the Conventional CNN architecture has the best accuracy than the Residual Network-50 as evidenced by the Confusion Matrix evaluation method.

REFERENCES

Angio, M. H., & Irawanto, R. (2019). PENDATAAN JENIS BUAH LOKAL INDONESIA KOLEKSI KEBUN RAYA PURWODADI. *Jambura Edu Biosfer Journal*, *1*(2). https://doi.org/10.34312/jebj.v1i2.2476

Arrofiqoh, E. N., & Harintaka, H. (2018). IMPLEMENTASI METODE CONVOLUTIONAL NEURAL NETWORK UNTUK KLASIFIKASI TANAMAN PADA CITRA RESOLUSI TINGGI. *GEOMATIKA*, *24*(2). https://doi.org/10.24895/jig.2018.24-2.810

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *2016-December*. https://doi.org/10.1109/CVPR.2016.90

Ilahiyah, S., & Nilogiri, A. (2018). Implementasi Deep Learning Pada Identifikasi Jenis Tumbuhan Berdasarkan Citra Daun Menggunakan Convolutional Neural Network. *JUSTINDO (Jurnal Sistem Dan Teknologi Informasi Indonesia)*, *3*(2).

Kingma, D. P., & Ba, J. L. (2015). Adam: A method for stochastic optimization. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*.

Liang, G., & Zheng, L. (2020). A transfer learning method with deep residual network for pediatric pneumonia diagnosis. *Computer Methods and Programs in Biomedicine*, *187*. https://doi.org/10.1016/j.cmpb.2019.06.023

Maulana, F. F., & Rochmawati, N. (2020). Klasifikasi Citra Buah Menggunakan Convolutional Neural Network. *Journal of Informatics and Computer Science (JINACS)*, *1*(02). https://doi.org/10.26740/jinacs.v1n02.p104-108

Nashrullah, F., Wibowo, S. A., & Budiman, D. G. (2020). Investigasi Parameter Epoch Pada Arsitektur ResNet-50 Untuk Klasifikasi Pornografi. *COMPLETE* , *1*(1).

Paraijun, F., Aziza, R. N., & Kuswardani, D. (2022). Implementasi Algoritma Convolutional Neural Network Dalam Mengklasifikasi Kesegaran Buah Berdasarkan Citra Buah. *KILAT*, *11*(1). https://doi.org/10.33322/kilat.v10i2.1458

Patil, A., & Rane, M. (2021). Convolutional Neural Networks: An Overview and Its Applications in Pattern Recognition. *Smart Innovation, Systems and Technologies*, *195*. https://doi.org/10.1007/978-981-15-7078-0_3

Sakib, S., Ahmed, N., Kabir, A. J., & Ahmed, H. (2018). An Overview of Convolutional Neural Network: Its Architecture and Applications. *Preprints 2018*, (February).

Setiawan, W. (2020). PERBANDINGAN ARSITEKTUR CONVOLUTIONAL NEURAL NETWORK UNTUK KLASIFIKASI FUNDUS. *Jurnal Simantec*, *7*(2). https://doi.org/10.21107/simantec.v7i2.6551

*Arie Satia Dharma

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 3, July 2023
DOI : https://doi.org/10.33395/sinkron.v8i3.12721

e-ISSN : 2541-2019
p-ISSN : 2541-044X

Shafiq, M., & Gu, Z. (2022). Deep Residual Learning for Image Recognition: A Survey. *Applied Sciences (Switzerland)*, Vol. 12. https://doi.org/10.3390/app12188972

Sitepu, A. C., & Sigiro, M. (2021). Analisis Fungsi Aktivasi ReLu Dan Sigmoid Menggunakan Optimizer SGD Dengan Representasi MSE Pada Model Backpropagation. *Jurnal Teknik Informatika Komputer Universal*, *1*(1).

Upreti, A. (2022). Convolutional Neural Network (CNN): A comprehensive overview. *International Journal of Multidisciplinary Research and Growth Evaluation*. https://doi.org/10.54660/anfo.2022.3.4.18

Wen, L., Li, X., & Gao, L. (2020). A transfer convolutional neural network for fault diagnosis based on ResNet-50. *Neural Computing and Applications*, *32*(10). https://doi.org/10.1007/s00521-019-04097-w

Wu, D., Ying, Y., Zhou, M., Pan, J., & Cui, D. (2023). Improved ResNet-50 deep learning algorithm for identifying chicken gender. *Computers and Electronics in Agriculture*, *205*. https://doi.org/10.1016/j.compag.2023.107622

Yu, X., Kang, C., Guttery, D. S., Kadry, S., Chen, Y., & Zhang, Y. D. (2021). ResNet-SCDA-50 for Breast Abnormality Classification. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *18*(1). https://doi.org/10.1109/TCBB.2020.2986544

Yuliani, E., Aini, A. N., & Khasanah, C. U. (2020). Perbandingan Jumlah Epoch Dan Steps Per Epoch Pada Convolutional Neural Network Untuk Meningkatkan Akurasi Dalam Klasifikasi Gambar. *Jurnal Informa : Jurnal Penelitian Dan Pengabdian Masyarakat*, *5*(3). https://doi.org/10.46808/informa.v5i3.140

*Arie Satia Dharma