



Detection of Covid-19 on X-Ray Image of Human Chest Using CNN and Transfer Learning

Jalu Nusantara¹, Faldo Fajri Afrinanto², Wana Salma Labibah³, Zamah Sari⁴, Yufis Azhar⁵
^{1,2,3,4,5}Informatics, Faculty of Engineering, University of Muhammadiyah Malang
jalunusantoro25@gmail.com¹, hafaldo@gmail.com², wanalabibah2@gmail.com³, zamahsari@umm.ac.id⁴,
yufis@umm.ac.id⁵

Abstract

At the end of 2019, a new disease called Coronavirus Disease (COVID-19) originated in Wuhan, China. This disease is caused by respiratory tract infections, ranging from the common cold to serious diseases such as Middle East Respiratory Syndrome (MERS) and Severe Acute Respiratory Syndrome (SARS). In Indonesia, there are tests to detect COVID-19, such as PCR and Rapid Test. This detector takes a long time and is less accurate in producing a diagnosis. This study aims to classify chest X-ray images using CNN and Transfer Learning methods to diagnose COVID-19. The proposed model has 4 scenarios: CNN Handcraft Model, Transfer Learning (VGG 16, VGG 19, and ResNet 50). This model is accompanied by data augmentation and data balancing techniques using undersampling techniques. The dataset used in this study is the "Covid-19 (COVID-19 and Normal) Radiographic Database" with 13,808 data divided into two classes, namely COVID-19 and Normal. Each model built will produce values for accuracy, precision, recall, and confusion matrix. The results of CNN Scenario 1 accuracy is 95%, in Scenario 2 VGG 16 the accuracy is 93%, Scenario 3 VGG 19 is 90% and Scenario 4 ResNet 50 is 80%.

Keywords: COVID-19, Image Classification, Convolutional Neural Network, Transfer Learning

1. Introduction

In the development of the times, various parts of the world have gone through several phases of life. From various fields ranging from technology, education, and health. The health sector itself has gone through various changes, such as increasing technology in the field, the emergence of new diseases, and more efficient screening and disease management. The disease itself has evolved continuously with different conditions. For example, the case of the Spanish flu that occurred in 1918 became a pandemic that continued until the winter of 1919, then there was Severe Acute Respiratory Syndrome (SARS) which had occurred in 2002, and Middle East Respiratory Syndrome (MERS) which had occurred in 2012 [1].

At the end of 2019 in December, a new disease emerged that took the health world by surprise, namely a virus known as Coronavirus Disease (COVID-19) which will continue as a pandemic [2]. COVID-19, which originally originated in the city of Wuhan in China, has spread to almost all parts of the world [3]. The COVID-19 virus in Indonesia entered in March 2020 and has become a pandemic in all parts of the world. As of May 21, 2022, COVID-19 cases have infected 231 countries

with a confirmed case number of 521,920,560 cases and a confirmed death rate of 6,274,323. The USA occupies the highest position, with 81,117,634 cases, followed by India with a total of 43,131,822 cases [4].

The symptoms are almost the same as the flu in general, but COVID-19 reacts more quickly, causing a more severe infection than the flu in general [5]. The period of symptoms that have been caused by this virus is a range of 1-14 days, where transmission of this virus occurs when a person with COVID-19 sneezes or coughs. The transmission of COVID-19 is very fast compared to SARS and MERS with the death rate from SARS (10%) and the death rate from MERS (37%), of course, this percentage is higher than the death rate from COVID-19. From this, the transmission of COVID-19 is very fast and high compared to SARS and MERS [5].

The handling and prevention of COVID-19 have been carried out with various programs implemented in various parts of the world. Including Indonesia, has carried out significant handling and step by step. By providing socialization to the community such as washing hands, avoiding direct contact with sick

people, and using masks to cover the nose and mouth [6]. Giving vaccines to the public is also included in the handling of COVID-19 in Indonesia. This includes handling the COVID-19 virus diagnostic test, which aims to confirm the presence of the COVID-19 virus in a person's body. In Indonesia, there are several tests for the diagnosis of the COVID-19 virus. There are 3 diagnostic tests applied to / for Indonesia, the first is the Rapid Molecular Test (TCM), the second is the Polymerase Chain Reaction (PCR) and the third is the Rapid Test. Of the several diagnostic methods or tests available in Indonesia to detect COVID-19, researchers tried to use lung X-Ray (CT-Scan) images because the results obtained were faster and more accurate than previous tests [7].

The Rapid Molecular Test (TCM) with the COVID-19 examination method uses cartridge-based nucleic acid amplification with a nasopharyngeal swab. the second is the Polymerase Chain Reaction (PCR) using a sample of mucus found in the nasal passages or in the throat passages which takes a long time to get the results, and the third is the Rapid Test using a blood sample and has the disadvantage of being able to produce a 'False Negative' or when the test results appear negative but are positive [8].

Handling COVID-19 does not only lie in mucus or blood but other diagnostic tests support the detection of COVID-19 disease. To ensure this in diagnosing COVID-19 disease, the medical team or doctor performs examinations that can be done, such as the use of CT-Scan photos or chest X-rays [9]. In examination with chest radiography, this is a health facility in the form of chest X-ray images that are most often used and easy to get results [10].

Of the several handlings of COVID-19 in Indonesia, there are several shortcomings in the handling. The drawback of handling COVID-19 in Indonesia is that testing diagnostic tests using the Polymerase Chain Reaction (PCR) method takes a very long time with a period of approximately 6 hours, and sampling certainly requires physical contact which must be avoided [11]. And for Rapid Tests that require fast time within 10-30 minutes, they can produce predictions of COVID-19 but are less than the maximum percentage in predicting COVID-19 [12]. The percentages ranged from 81.8% to 88.1% using whole blood, serum, or plasma samples as well as comparisons on sensitivity for IgM and IgG assays [13].

From these problems, screening and disease detection management must be upgraded to handle this problem. Screening and management of disease detection are what help the decision-making process in the health sector or detect disease. To reduce and control the presence of this COVID-19 disease, a classification process must be required. In classification, CNN can obtain accuracy above 80% and this includes good

performance [14]. This classification uses data in the form of images called chest radiographs, the technique used is a classification that implements chest radiograph images into the machine learning process [15]. Machine Learning includes various algorithms and modeling tools that are used for various data processing tasks both large and small, which have entered most data processing science in recent years [16]. Substantially, machine learning can reduce computational costs, which is one of the most efficient ways to replace repetitive laboratory experiments [17]. In the last decade, machine learning has provided tremendous improvements that have impacted many areas in healthcare, finance, manufacturing, energy, and many more industries [16]. Machine Learning can handle big data so that it can be developed by its users to learn without direction [15].

One of the most important parts of this technology is deep learning. Deep Learning itself is a technique that can be used quickly and efficiently to detect diseases such as COVID-19 or other diseases with fairly good accuracy [18]. Deep learning-based image segmentation and classification is now well established as a powerful tool for predicting these images. It is also seen as providing an attractive and accurate solution for medical imaging and is seen as a key method for future applications in healthcare [19]. This medical image segmentation identifies the pixels of organs or lesions from medical images such as CT or MRI images, is one of the tasks of analyzing medical images that is conveying important information about the shape and volume of these organs. Deep Learning will not only help select and extract features but also build new features. Moreover, it not only diagnoses disease but also measures predictive targets and provides actionable predictive models to help clinicians develop effective treatment plans [20].

Convolutional Networks (ConvNets) are currently the most efficient inner model for classifying image data. Convolutional Neural Network (CNN) is a deep learning algorithm that can train large data sets and form 2D images, and combine them with several filters to produce the desired output [21]. CNN is one of the methods in deep learning. Therefore, medical problems, especially those related to image recognition on CT-Scan of the lungs, can be solved using deep learning [22].

CNN also consists of neurons that have bias, activation function, and weight. CNN is divided into 3 layers (Feature Extraction Layer), consisting of weight sharing (reducing computational complexity) and Convolution Layer which works with the sliding window principle, Pooling Layer is used to summarize the information that can be generated by a convolution (reducing dimensions), and which the last to perform the classification using a Fully-Connected Layer [23].

A previous study on the classification of cases of COVID-19, Pneumonia, and Normal from chest X-ray images by applying the transfer learning approach to the VGG-19 architecture conducted by the researchers, yielded several results including accuracy, and precision, and recall of the modeling that had been used. Using a dataset that exposed COVID-19 (1,184 images), Pneumonia (1,294 images), and Normal (1,319 images). The model built in this study obtained an accuracy of 97.11%, an average precision of 97%, and an average recall of 97% on the test dataset. The researcher proposes the original VGG-19 input model with a layer that can take 2562563 images as input. Instead of the last three layers, add a dense layer with three neurons and use softmax as the activation function [24]. From this research, it is suggested to use the VGG 19 model scenario.

In previous research on detecting COVID-19 using Transfer Learning on Convolution Neural Networks such as VGG16, VGG19, MobileNet, DenseNet201, Resnet50. Using a dataset of 6,000 CT-Scan images consisting of COVID-19 (3,000 images) and Normal (3,000 images). The model built in this study obtained different accuracy results, VGG19 (97.56%), VGG16 (97.78%), MobileNet (98.11%), DenseNet201 (97.23%) and ResNet50 (49.17%). Thus, if you want better performance, it is advisable to apply a large dataset [25]. From this research, the ResNet model has a very low accuracy with a relatively low learning epoch.

In a previous study, the detection of lung X-ray images was based on the Convolution Neural Network with 4 scenarios. Using a dataset consisting of 3 classes with a total of 2,905 X-Ray images consisting of COVID-19 (219 images), Normal (1,341), and Viral Pneumonia (1,345). The model built in this study obtained 4 model scenarios with model 1 scenario with a total data of 1,560 images (unbalance) with an average accuracy value of 98.69%, model 2 scenario with a total data of 2,906 images (unbalance) with an average value of 92.96% accuracy. , model 3 scenario with total data of 438 images (balance) with an average value of 97.17% accuracy and model 4 scenario with total data of 657 images (balance) with an average value of accuracy of 91.18% [11]. The experimental results in the performance model can affect the imbalance in the composition of the data.

Of the 3 previous studies, there are several advantages and disadvantages. From these advantages and disadvantages, researchers conducted research on each experiment that had been carried out to get the best performance.

To overcome the conditions in handling problems from each of the previous studies such as the CNN Handcraft Model, it was found to use data augmentation, data balancing, and pooling layers. Manipulating image data

and changing the architecture located at the pooling layer. Previous research has stated that the problem related to the number of unbalanced datasets can be overcome by using undersampling techniques by reducing the datasets in the majority class to be equated with the minority class. And the CNN Handcraft model can solve this problem.

In addition to the CNN Handcraft Model, transfer learning can also participate in the efficiency of prediction in time. Transfer Learning is a method that utilizes a convolutional neural network model that has been previously trained or can be called a pre-trained model. Transfer learning models are typically trained on large datasets such as ImageNet and the weights obtained from trained models can be applied to new datasets. So there is no need to do data training from the beginning and adjustments can be made at the end of the transfer learning model. The reason researchers use transfer learning is that it saves training time, and the performance of neural networks is better in many studies that have been carried out and do not require a lot of data.

This research will build 4 models, namely the CNN handcraft model, Transfer Learning VGG 16, Transfer Learning VGG 19, and Transfer Learning ResNet50. This model is also accompanied by data augmentation and data balancing techniques using undersampling techniques. Augmentation is done by changing the model architecture to suit the dataset, and the undersampling technique used is to handle the imbalance in the amount of data in the class with the aim of equalizing the number of datasets used in each class.

2. Research methods

This research method consists of several stages, namely Research Stages aim to explain the flow of research to evaluation, the dataset used in this study, Architecture Model seeks to explain each scenario model used in this study, and Augmentation Data aims to change the image in such a way, from the human view of the CT-Scan images that think the same, but what is changed is a different image. The scenario test explains the classification, data balancing, and data splitting process.

The stages of this research are listed in Figure 1. This research starts by collecting the dataset which is then carried out in the preprocessing stage which is divided into several preprocessing stages. The first step is to balance the dataset because the dataset used is still imbalanced, so equalization is carried out using the undersampling technique to overcome this problem. After that, the next process is data splitting by dividing the "Covid 19 Radiography Database" dataset into 3 folders consisting of the train, validation, and test data folders. The last stage in the preprocessing process is data augmentation.

1.1. Research Stage

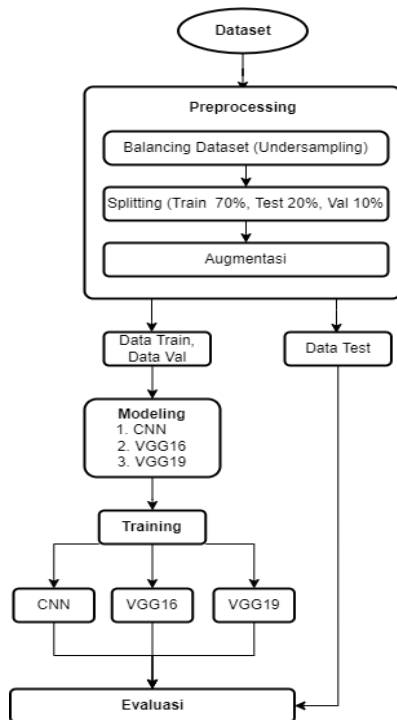


Figure 1. Research Stages

The next step after preprocessing is to process the results of grouping the data. In the data train, the modeling stage is carried out which is used to train the data with the models and parameters that have been built. Meanwhile, in data validation, validation of the train results is carried out with the parameters in the model in it to find out if there is underfitting or overfitting. The model that has been carried out in the model training process can be tested on the test data.

2.2. Dataset

The dataset used in this study is a chest X-ray image of COVID-19. The dataset comes from researchers at Qatar University and Dhaka University in collaboration with medical doctors who have created a database of chest X-ray images for positive cases of COVID-19 along with Normal and Viral Pneumonia Images. Also named a COVID-19 Dataset Award Winner by the Kaggle community. The dataset used for this study consists of two classes of images, namely COVID-19 and Normal, where the dataset consists of 13,808 data onto an image size of 256×256 pixels in PNG format. The following is an example of a dataset from each class, which can be seen in Figure 2 [26].

2.3. Model Architecture

This study proposes 4 models which will later be used as a comparative test. The first model the researcher uses is the CNN model, which uses an input layer of 150 150 pixels, starting from 224×224 pixels. The pixel sizes changed into 150×150 aims to focus the image of

objects that can help in marking the presence of COVID-19 disease or not, and also to speed up computing so that it can distinguish 2 image classes consisting of images infected with COVID-19 and Normal.

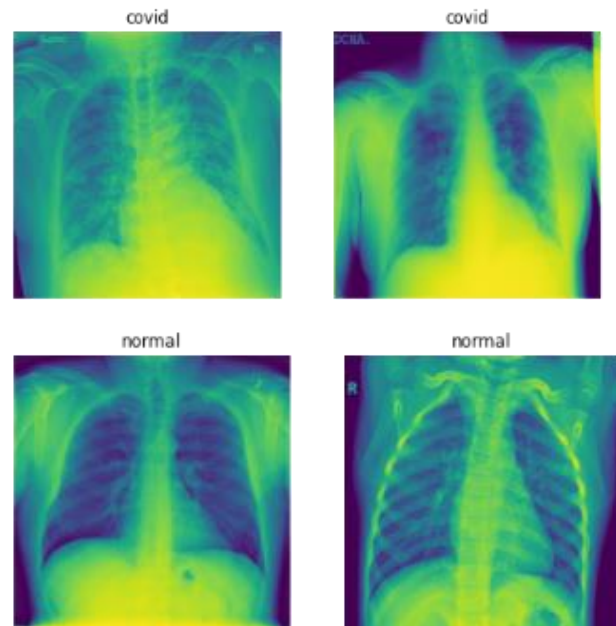


Figure 2. Dataset Chest X-ray Covid and Normal

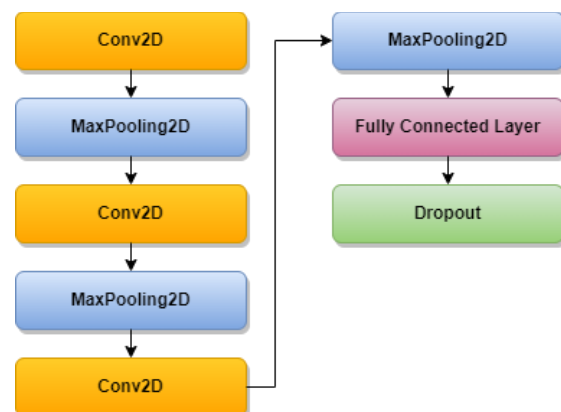


Figure 3. CNN Model Architecture

The first model architecture is shown in Figure 3, where this model applies three convolutional layers and three pooling layers by implementing max pooling with a filter of 2×2 . Apart from the pooling layer, this research was also built using a convolutional layer 3 times, each of which has a filter of 128, 64 and 32, the kernel from the convolutional is 3×3 and is followed by relu activation. Then, after this stage, it is followed by a fully connected layer in which there is a flatten layer, a dense layer using relu activation followed by a dropout layer of 0.5 which aims to deal with overfitting during the training process. Finally, there is a dense layer using activation softmax.

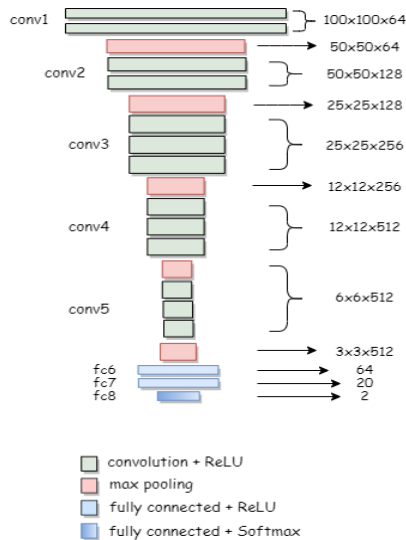


Figure 4. VGG16 Model Architecture

The second model architecture in this study uses transfer learning Visual Geometry Group 16 (VGG16). VGG16 is a deep convolutional neural network architecture with a total of 16 convolutional layers [27]. As seen in Figure 4, the model consists of 13 convolutional layers followed by 3 fully connected layers. The difference between each convolution layer is in the filter. The first 2 convolution layers have a filter number of 100 x 100 with a depth of 64. Layer 3,4 convolution has a filter of 50 x 50 with a depth of 128. Next, layer 5,6,7 convolution has a filter of 25 x 25 and a depth of 256. After that, the convolution 8,9,10 layer has a 12 x 12 filter and a 512 depth. The last layer convolution 11,12,13 has a 6 x 6 filter and the depth is the same as the previous layer.

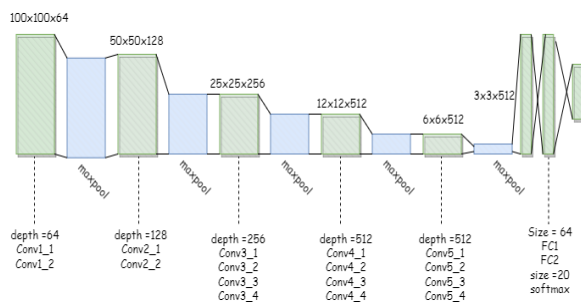


Figure 5. VGG 19 Model Architecture

The third model architecture used in this research is using Visual Geometry Group 19 (VGG19). VGG19 is a deep convolutional neural network architecture that has 19 layers [28]. As shown in Figure 5, this model consists of 16 convolutional layers, 3 fully connected layers, and 5 max-pooling layers. The difference between each convolution layer is in the filter and depth.

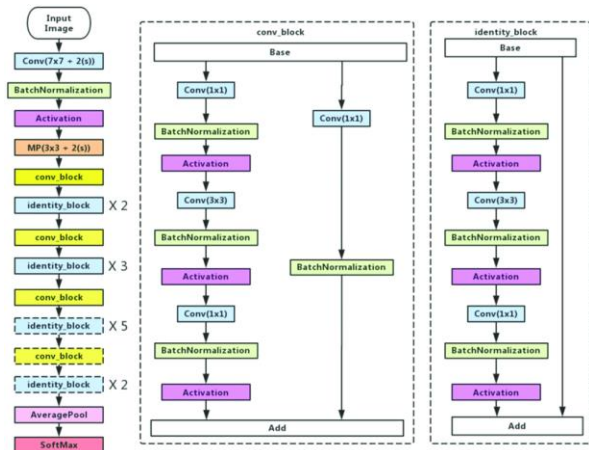


Figure 6. ResNet50 Model Architecture

Figure 6 shows the architecture of the ResNet-50 model used in this study. ResNet-50 is a convolutional neural network that has a depth of 50 layers. From Figure 5 it can be seen that there are 3 image blocks, on the left is the ResNet-50 architecture, while the middle is a convolution block that changes the input dimensions, and on the right is an identity block that will not change the input dimensions [29].

2.4. Data Augmentation

Augmentation Data aims to change the image in such a way, from the human view of the CT-Scan image that thinks the same, but what is changed is a different image [30]. The advantage of using data augmentation techniques can be to improve the performance of the model [31]. The following augmentation parameters used in this study are *rotation_range* = 20, *width_shift_range* = 0.10, *height_shift_range* = 0.10, *rescale* = 1/255, *shear_range* = 0.1, *zoom_range* = 0.1, *horizontal_flip* = True, *fill_mode* = 'nearest'.

2.5. Testing Scenario

In this study, two classes will be classified, namely X-Ray image Covid and X-Ray image Normal. Next, the undersampling method is applied to the dataset so that the amount of data in the two classes are balanced.

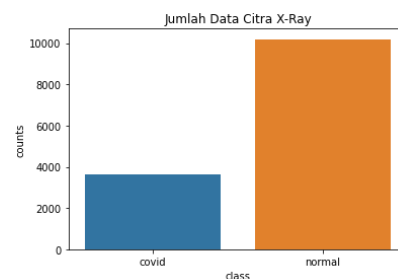


Figure 7. Count of datasets before undersampling

In Figure 7 is the amount of data in each class before undersampling with a total of 13.808 images with each

COVID-19 image totaling 3.616 and Normal images 10.192 images.

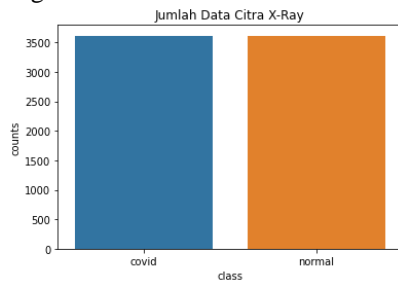


Figure 8. Count of datasets after undersampling

Figure 8 is the amount of data in each class after undersampling with a total of 7,232 images with each Covid-19 image totaling 3,616 and Normal images at 3.616 images.

Then the dataset is divided into three types of data with a division of 70% train data, 10% validation data, and 20% test data from the total data in each class. The number of datasets before and after the augmentation process is the same, this is because the augmentation process does not add new images, but only changes each existing image into various shapes according to the augmentation parameters. Table 1 is the amount of train, validation, and test data before the augmentation and undersampling process is carried out. The dataset shown in Table 1 is the original dataset that the researcher downloaded from the official website link kaggle [26]. This dataset has unbalanced data, which is a condition where one data is not balanced with other data. Therefore, the researcher balances the data using the undersampling technique, which is balancing the datasets whose conditions are uneven and reducing the datasets from the majority class to the number of minority classes.

Table 1. Amount of data before augmentation & undersampling

Kelas	Training Images	Validation Images	Testing Images	Total Images
Covid	2531	361	724	3.616
Normal	7.134	1019	2039	10.192

Then after carrying out the undersampling technique, the results are in Table 2, by equating the amount of data in all classes by following the amount of data in the minority class. The following is the train, validation, and test data after the augmentation and undersampling process.

Table 2. Amount of data after augmentation and undersampling

Kelas	Training Images	Validation Images	Testing Images	Total Images
Covid	2531	361	724	3.616
Normal	2531	361	724	3.616

Furthermore, the dataset will be tested with 4 scenarios. Scenario 1 uses the CNN handcraft model, scenario 2

uses the VGG 16 model, scenario 3 uses the VGG 19 model and scenario 4 uses the ResNet50 model.

3. Results and Discussion

The results of this study are the stages carried out based on the arrangement of the test scenarios. Then test the 4 scenarios of the proposed model and the performance results will be compared. The performance parameters used are in the form of a classification report that provides accuracy, precision, and recall values for each model scenario. And the researcher uses the confusion matrix as information on the comparison of the prediction results of the classification that has been done by the model with the actual classification results.

Before the testing is carried out on the 4 model scenarios, the dataset will enter the preprocessing stage which has been described in the test scenario. This is followed by the creation and training of models. The model will be trained with 100 iterations (epochs).

3.1. Scenario 1 (CNN Handcraft Model)

The dataset was tested using scenario model 1 of the Convolutional Neural Network (CNN). In the training process that has been carried out with the CNN model, a plot graph shows the accuracy graph and the loss graph listed in Figure 9 and Figure 10. Based on the two plot graphs, the accuracy value is 0.95 and the loss value is 0.09.

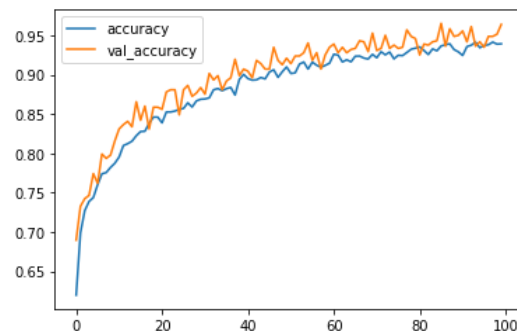


Figure 9. CNN model accuracy graphic

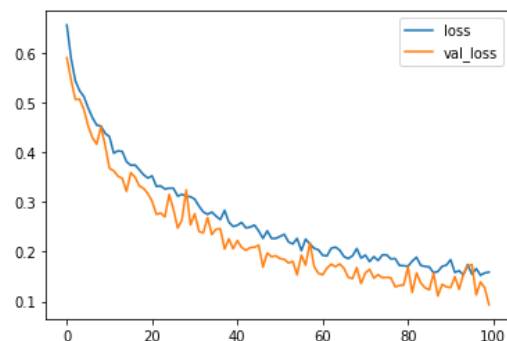


Figure 10. CNN model loss graphic

After knowing the plot graph, then proceed with the process of evaluating the model using the confusion matrix as shown in Figure 11.

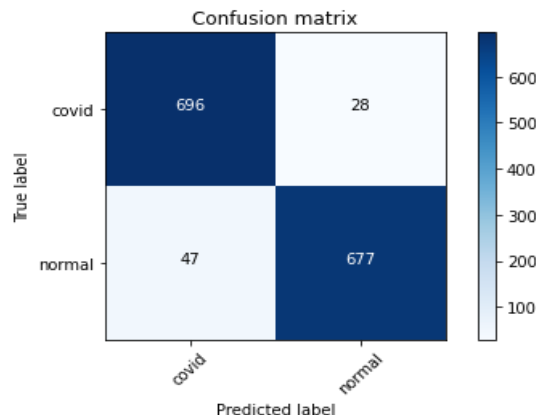


Figure 11. Confusion Matrix CNN Model

Figure 11 is the result of the CNN confusion matrix model. In the picture, it can be seen that in the covid class there are 696 image data that are predicted to be correct and 28 image data that are predicted to be incorrect by the model. And in the normal class, there are 677 image data that are predicted to be correct and 47 images that are predicted to be wrong by the model.

Then the CNN model is tested with test data in which each class displays 5 images to predict. From the average prediction results, the researcher displays 2 images of the average predictions seen in Figure 12 showing that the average image can be predicted correctly in the covid class and achieves an average accuracy of 0.99 and 0.95 as well as a long prediction time. not too far away, namely 0.032 seconds and 0.033 seconds.

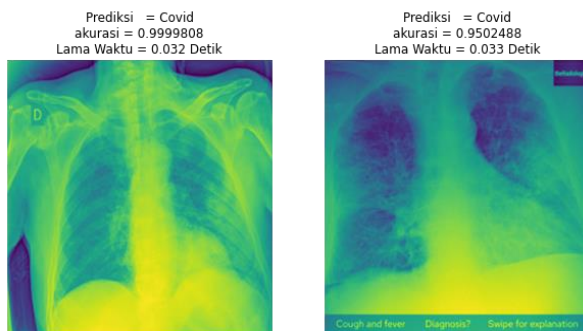


Figure 12. Test results for Covid images with the CNN Model

Then the test data test is continued in the normal class which also displays 5 predicted images and 2 images are taken the average is shown in Figure 13 shows that the image can be predicted correctly and achieves an average accuracy of 0.99 and the prediction time is not too far away. namely 0.033 seconds and 0.034 seconds.

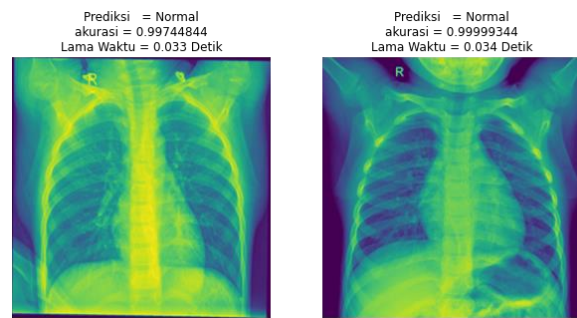


Figure 13. Test results Normal image testing with CNN Model

3.2. Scenario 2 (VGG16 Model)

The dataset was tested using the model from VGG 16. In the training process that was carried out with the VGG 16 model, a plot graph was obtained showing the accuracy graph and the loss graph listed in Figure 14 and Figure 15. Based on both of the plot graphs, the val_acc value of 0,93 was obtained. and val_loss value 0,16.

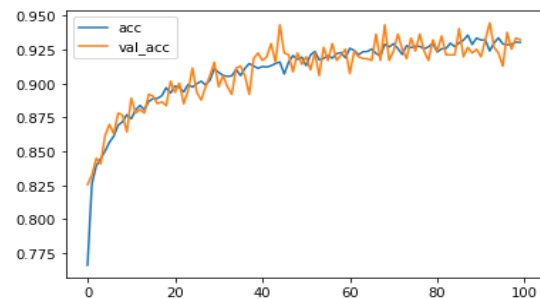


Figure 14. VGG16 Model accuracy graph

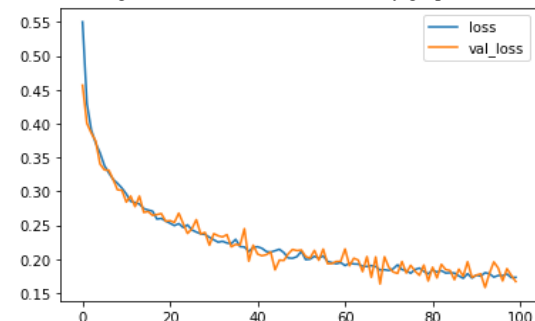


Figure 15. VGG16 Model loss graph

After knowing the plot graph, then proceed with the process of evaluating the model using the confusion matrix as shown in Figure 16.

Figure 16 is the result of the confusion matrix Model VGG 16, it can be concluded that in the covid class there are 664 image data that are predicted to be correct and 60 image data are predicted to be incorrect by the model. And in the normal class there are 688 image data that are predicted to be correct and 36 images that are predicted to be wrong by the model.

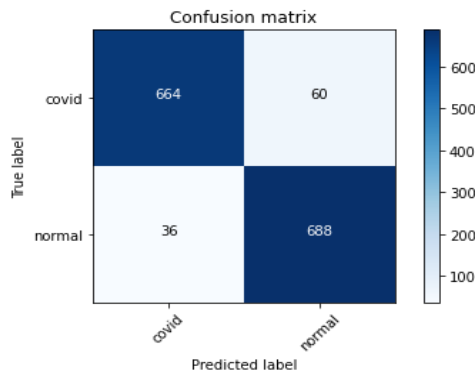


Figure 16. Confusion Matrix VGG 16 Model

Then the VGG 16 model was tested with test data by performing 5 predicted images and 2 images taken. The average prediction shown in Figure 17 shows that the image can be predicted correctly in the covid class and achieves an average accuracy of 0,99 and a prediction time of 0,08 seconds and 0,04 seconds.

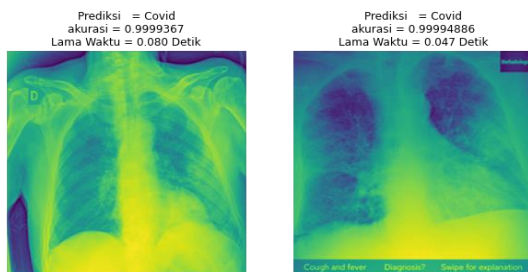


Figure 17. Test results for Covid images with the VGG16 model

Then the test data test was continued in the normal class which also displayed 5 predicted images and 2 images were taken, the average of which is shown in Figure 18 showing that the image can be predicted correctly and achieves an average accuracy of 1.0 and a prediction time of 0,047 seconds and 0,046 seconds.

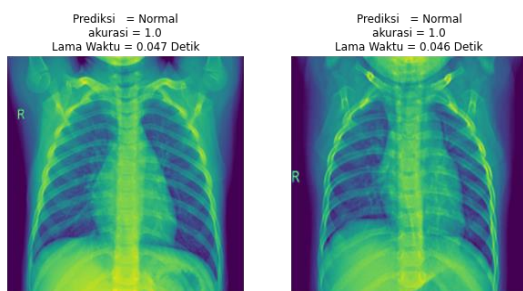


Figure 18. Test Results Normal image testing with VGG16 Model

3.3. Scenario 3 (VGG 19 Model)

The dataset is tested using the model from VGG 19. In the training process that has been carried out with the VGG 19 model, a plot graph shows the accuracy graph and loss graph listed in Figure 19 and Figure 20. Based on the two plot graphs, the val_acc value is 0,90 and val_loss value 0,21.

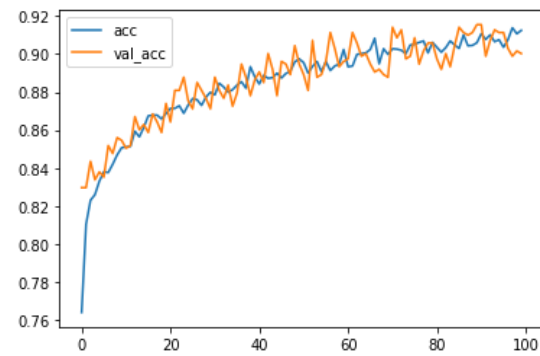


Figure 19. VGG19 Model accuracy graph

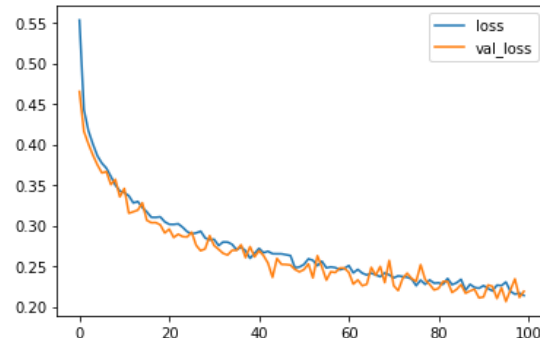


Figure 20. VGG19 Model loss graph

After knowing the plot graph, then proceed with the process of evaluating the model using the confusion matrix as shown in Figure 21.

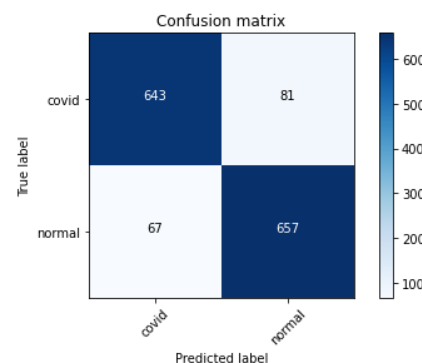


Figure 21. Confusion Matrix VGG 19 Model

Figure 21 is the result of the confusion matrix Model VGG 19, it can be concluded that in the covid class there are 643 image data that are predicted to be correct and 81 image data are predicted to be incorrect by the model. And in the normal class there are 657 image data that are predicted to be correct and 67 images that are predicted to be incorrect by the model.

Then the VGG 16 model was tested with test data by performing 5 predicted images and 2 images of the average prediction were taken. Figure 22 shows that the image can be predicted correctly in the covid class and achieves an average accuracy of 0,99 and an average prediction time. The average covid class reaches 0,054 seconds.

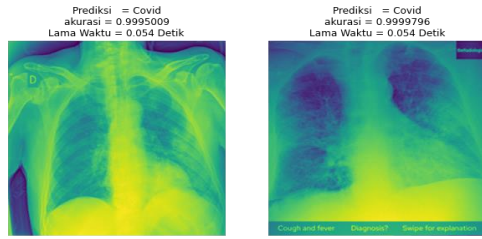


Figure 22. Test results for Covid images with the VGG19 model

Then the test data test was continued in the normal class which also displayed 5 predicted images and 2 images were taken the average of which is shown in Figure 23 showing that the image can be predicted correctly and achieves an average accuracy of 1.0 with a long image prediction time of 0,055 seconds and 0,056 second.

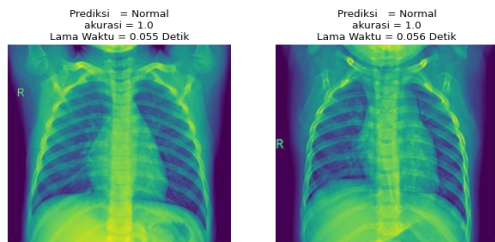


Figure 23. Test Results Testing normal images with the VGG19 Model

3.4. Scenario 4 (Model ResNet 50)

The dataset is tested using the model from ResNet 50. In the training process that has been carried out with the ResNet 50 model, a plot graph shows the accuracy graph and the loss graph listed in Figure 24 and Figure 25. Based on both of the plot graphs, the validation accuracy value is 0.80 and the value of validation loss is 0.44.

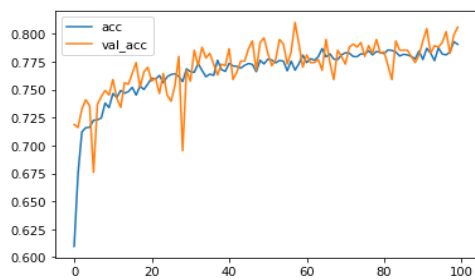


Figure 24. ResNet50 Model accuracy graph

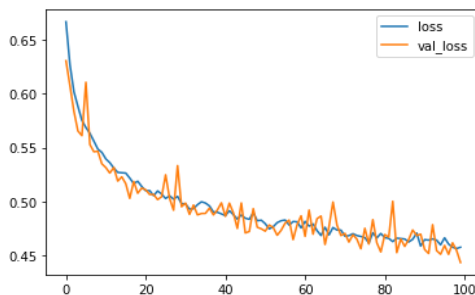


Figure 25. ResNet50 Model loss graph

After knowing the plot graph, then proceed with the process of evaluating the model using the confusion matrix as shown in Figure 26.

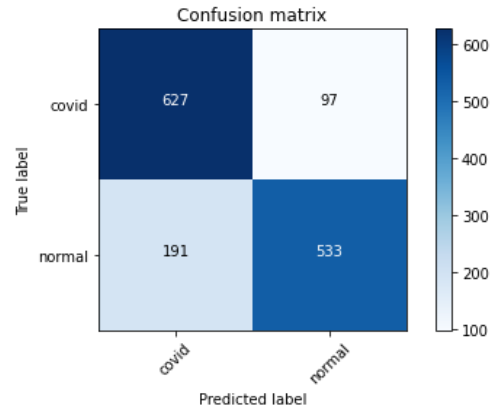


Figure 26. Confusion Matrix of ResNet50 Model

Figure 26 is the result of the ResNet 50 model confusion matrix, it can be concluded that in the covid class there are 627 image data that are predicted to be correct and 97 image data that are predicted to be incorrect by the model. And in the normal class, there are 533 image data that are predicted to be correct and 191 images are predicted to be wrong by the model.

Then the ResNet 50 model was tested with test data by carrying out 5 predicted images and 2 images of the predicted average shown in Figure 27 showing that the image can be predicted correctly in the covid class and achieving an average accuracy of 0.86 and 0.97 and the length of the image prediction time is 0.053 seconds and 0.077 seconds.

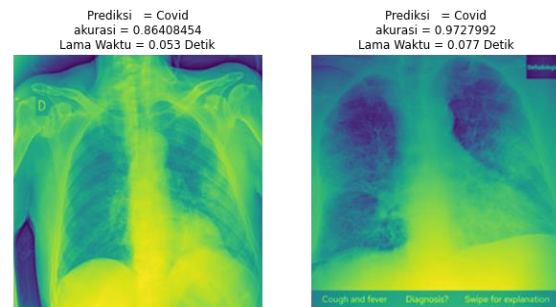


Figure 27. Test results Covid image testing with Resnet50 Model

Then the test data test was continued in the normal class which also displayed 5 predicted images and 2 images were taken the average in the normal class shown in Figure 28 shows that the image can be predicted correctly and achieves an average accuracy of 0.99 and a long prediction time images of 0.053 seconds and 0.077 seconds.

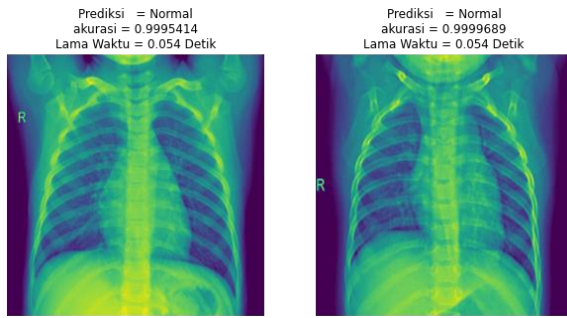


Figure 28. Test results Normal image testing with Resnet50 Model

3.5 Model Performance Comparison

After training the model, the 4 model scenarios were tested for performance using classification reports and Confusion Matrix. The following is a comparison of the performance results of the 4 model scenarios in the classification report, which can be seen in Table 3.

Table 3. Performance results from each scenario using a classification report

Scenario	Class	Precision	Recall	Accuracy
Scenario 1	Covid	94%	96%	95%
	Normal	96%	94%	
Scenario 2	Covid	95%	92%	93%
	Normal	92%	95%	
Scenario 3	Covid	91%	89%	90%
	Normal	89%	91%	
Scenario 4	Covid	77%	87%	80%
	Normal	85%	74%	

Then the comparison of the performance results of the 4 model scenarios in the confusion matrix, can be seen in Table 4.

Table 4. Performance results of each scenario using a Confusion Matrix

Scenario	Covid		Normal	
	Correct Prediction	Incorrect Prediction	Correct Prediction	Incorrect Prediction
Scenario 1	696 images	28 images	677 images	47 images
Scenario 2	664 images	60 images	688 images	36 images
Scenario 3	643 images	81 images	657 images	67 images
Scenario 4	627 images	97 images	533 images	191 images

After 4 scenarios, the performance is tested using classification reports and Confusion Matrix. As seen in Table 3 and Table 4, Scenario 1 with the CNN model gets the best performance results among other scenarios.

The results of the confusion matrix in each scenario can be different due to differences in architectural models and complexity in each scenario. In scenario 4 the normal image prediction results have a higher error rate than the other scenarios, this can be because the model

architecture implemented in scenario 4 does not match the conditions of the dataset used.

3.6 Comparison of the Best Model Performance with Previous Research

After experimenting with several different test scenarios, the next step is to compare the performance of the best model with the results obtained from previous studies. In Figure 29, the classification report in Scenario 1 CNN obtained an accuracy of 95% with a precision of 94% in the Covid class and 96% in the Normal class.

	precision	recall	f1-score	support
0	0.94	0.96	0.95	724
1	0.96	0.94	0.95	724
accuracy			0.95	1448
macro avg	0.95	0.95	0.95	1448
weighted avg	0.95	0.95	0.95	1448

Figure 29. Scenario 1 Classification Report (CNN Model)

Based on Table 5, this study resulted in the model in Scenario 1 as the best model, but this model experienced a decrease in accuracy from the previous study by 3%.

Table 5. Accuracy Results of Each Scenario Model

Model	Accuracy
CNN Purposed [24]	97.11%
CNN ResNet50 [25]	49.17%
Best proposed scenario model with ResNet50	95%

3.7 Testing

The CNN Scenario 1 model became the best scenario model which was then retested using the dataset used. In Figure 30, the results obtained show that there are 5 image data that came out as predictions for the Covid class correctly and achieved an average accuracy of 90.8% with an average time duration of 0.1736 seconds.

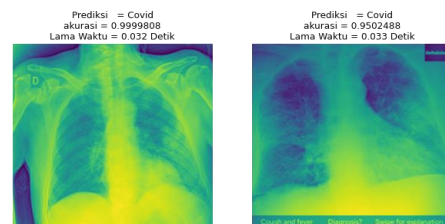


Figure 30. Example of an image of the test results in the Covid Class

Then in Figure 31, the results obtained show that 5 image data came out as predictions in the Normal Class correctly and achieved an average accuracy of 99.2% with an average time duration of 0.0326 seconds.

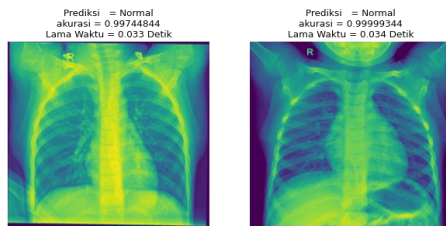


Figure 30. Example of an image of the test results in the Normal class

4. Conclusion

Based on research that has been carried out by implementing four models, it shows that the use of the Convolutional Neural Network (CNN) model gets 95% accuracy with an average precision and recall of 95% which is superior to the VGG16, VGG19, and ResNet50 models. From the results of this study, it can be concluded that the good or bad performance of a model is caused by the number of layers, data conditions, kernel size in the input layer, and data balance. The dataset in this study uses undersampling because the difference in the amount of data from the covid and normal classes is too far from 3:10 to 3:3. A balanced dataset affects the accuracy gain which can increase the accuracy of the model.

For further development of this research, the researcher provides suggestions for balancing data using the oversampling method on the dataset. And the researchers suggest experimenting on each layer in the CNN model architecture and also experimenting with other transfer learning such as InceptionV3.

References

- [1] Gerald Aldian Wijaya, "Mengulas Riwayat Pandemi Dunia," *Bem.Fk.Ui.Ac.Id.* 2020, [Online]. Available: <http://bem.fk.ui.ac.id/mengulas-riwayat-pandemi-dunia/>.
- [2] R. N. Putri, "Indonesia dalam Menghadapi Pandemi Covid-19," *J. Ilm. Univ. Batanghari Jambi*, vol. 20, no. 2, p. 705, 2020, doi: 10.33087/jiubj.v20i2.1010.
- [3] S. Sukesih, U. Usman, S. Budi, and D. N. A. Sari, "Pengetahuan Dan Sikap Mahasiswa Kesehatan Tentang Pencegahan Covid-19 Di Indonesia," *J. Ilmu Keperawatan dan Kebidanan*, vol. 11, no. 2, p. 258, 2020, doi: 10.26751/jikk.v11i2.835.
- [4] World Health Organization, "WHO Coronavirus Disease (COVID-19) Dashboard With Vaccination Data | WHO Coronavirus (COVID-19) Dashboard With Vaccination Data," *World Health Organization*. pp. 1–5, 2021, [Online]. Available: <https://covid19.who.int/%0Ahttps://covid19.who.int/%0Ahttps://covid19.who.int/region/searo/country/bd>.
- [5] L. Amalia, I. Irwan, and F. Hiola, "Analisis Gejala Klinis Dan Peningkatan Kekebalan Tubuh Untuk Mencegah Penyakit Covid-19," *Jambura J. Heal. Sci. Res.*, vol. 2, no. 2, pp. 71–76, 2020, doi: 10.35971/jjhsr.v2i2.6134.
- [6] J. Moudy and R. A. Syakurah, "Pengetahuan terkait usaha pencegahan Coronavirus Disease (COVID-19) di Indonesia," *Higeia J. Public Heal. Res. Dev.*, vol. 4, no. 3, pp. 333–346, 2020, doi: doi.org/10.29100/jipi.v6i2.2102.
- [7] Z. S. Ahmed, N. M. S. Surameery, R. D. Rashid, S. Q. Salih, and H. K. Abdulla, "CNN-based Transfer Learning for Covid-

- 19 Diagnosis," in *2021 International Conference on Information Technology (ICIT)*, Jul. 2021, pp. 296–301, doi: 10.1109/ICIT52682.2021.9491126.
- [8] F. R. Makarim, "Mengenal 3 Jenis Tes Corona yang Digunakan di Indonesia," *Tes Diagnosis Virus Corona*. pp. 2–3, 2020, [Online]. Available: <https://www.halodoc.com/mengenal-jenis-tes-corona-yang-digunakan-di-indonesia>.
- [9] D. L. Smith, J. P. Grenier, C. Batte, and B. Spieler, "A characteristic chest radiographic pattern in the setting of the covid-19 pandemic," *Radiol. Cardiothorac. Imaging*, vol. 2, no. 5, 2020, doi: 10.1148/ryct.2020200280.
- [10] D. E. Litmanovich, M. Chung, R. R. Kirkbride, G. Kicska, and J. P. Kanne, "Review of Chest Radiograph Findings of COVID-19 Pneumonia and Suggested Reporting Language," *J. Thorac. Imaging*, vol. 35, no. 6, pp. 354–360, 2020, doi: 10.1097/RTI.0000000000000541.
- [11] Bambang Pilu Hartato, "Penerapan Convolutional Neural Network pada Citra Rontgen Paru-Paru untuk Deteksi SARS-CoV-2," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 4, pp. 747–759, 2021, doi: 10.29207/resti.v5i4.3153.
- [12] G. Cota *et al.*, "Diagnostic performance of commercially available COVID-19 serology tests in Brazil," *Int. J. Infect. Dis.*, vol. 101, pp. 382–390, 2020, doi: 10.1016/j.ijid.2020.10.008.
- [13] E. Mohit, Z. Rostami, and H. Vahidi, "A comparative review of immunoassays for COVID-19 detection," *Expert Review of Clinical Immunology*, vol. 17, no. 6. Taylor & Francis, pp. 573–599, 2021, doi: 10.1080/1744666X.2021.1908886.
- [14] S. A. Widiarto, W. A. Saputra, and A. R. Dewi, "Klasifikasi citra x-ray toraks dengan menggunakan contrast limited adaptive histogram equalization dan convolutional neural network (studi kasus: pneumonia)," vol. 06, pp. 348–359, 2021, doi: doi.org/10.29100/jipi.v6i2.2102.
- [15] M. Batta, "Machine Learning Algorithms - A Review," *Int. J. Sci. Res. (IJ)*, vol. 9, no. 1, pp. 381–undefined, 2020, doi: 10.21275/ART20203995.
- [16] G. Carleo *et al.*, "Machine learning and the physical sciences," *Rev. Mod. Phys.*, vol. 91, no. 4, p. 45002, 2019, doi: 10.1103/RevModPhys.91.045002.
- [17] J. Wei *et al.*, "Machine learning in materials science," *InfoMat*, vol. 1, no. 3, pp. 338–358, 2019, doi: 10.1002/inf2.12028.
- [18] D. Singh, V. Kumar, Vaishali, and M. Kaur, "Classification of COVID-19 patients from chest CT images using multi-objective differential evolution-based convolutional neural networks," *Eur. J. Clin. Microbiol. Infect. Dis.*, vol. 39, no. 7, pp. 1379–1389, 2020, doi: 10.1007/s10096-020-03901-z.
- [19] M. I. Razzak, S. Naz, and A. Zaib, "Deep learning for medical image processing: Overview, challenges and the future," *Lect. Notes Comput. Vis. Biomech.*, vol. 26, pp. 323–350, 2018, doi: 10.1007/978-3-319-65981-7_12.
- [20] M. H. Hesamian, W. Jia, X. He, and P. Kennedy, "Deep Learning Techniques for Medical Image Segmentation: Achievements and Challenges," *J. Digit. Imaging*, vol. 32, no. 4, pp. 582–596, 2019, doi: 10.1007/s10278-019-00227-x.
- [21] N. Jmour, S. Zayen, and A. Abdelkrim, "Convolutional neural networks for image classification," in *2018 International Conference on Advanced Systems and Electric Technologies, IC_ASET 2018*, 2018, pp. 397–402, doi: 10.1109/ASET.2018.8379889.
- [22] M. R. Fauzi, P. Eosina, and D. Primasari, "Deteksi Coronavirus Disease Pada X-Ray Dan CT-Scan Paru Menggunakan Convolutional Neural Network," *JUSS (Jurnal Sains dan Sist. Informasi)*, vol. 3, no. 2, pp. 17–27, 2021, doi: 10.22437/juss.v3i2.10888.
- [23] A. Yanuar, "Fully-Connected Layer CNN dan Implementasinya – Universitas Gadjah Mada Menara Ilmu Machine Learning." 2018, [Online]. Available: <https://machinelearning.mipa.ugm.ac.id/2018/06/25/fully-connected-layer-cnn-dan-implementasinya/>.
- [24] S. Chakraborty, S. Paul, and K. M. A. Hasan, "A Transfer Learning-Based Approach with Deep CNN for COVID-19-

- and Pneumonia-Affected Chest X-ray Image Classification,” *SN Comput. Sci.*, vol. 3, no. 1, pp. 1–10, 2022, doi: 10.1007/s42979-021-00881-5.
- [25] Buyut Khoirul Umri and V. Delica, “Penerapan transfer learning pada convolutional neural networks dalam deteksi covid-19,” *Jnanaloka*, pp. 9–17, 2021, doi: 10.36802/jnanaloka.2021.v2-no2-9-17.
- [26] T. Rahman, “COVID-19 Radiography Database | Kaggle,” *Kaggle*, vol. 4, no. March. p. 2021, 2021, [Online]. Available: <https://www.kaggle.com/tawsifurrahman/covid19-radiography-database/activity%0Ahttps://www.kaggle.com/tawsifurrahman/covid19-radiography-database>.
- [27] R. Rismiyati and A. Luthfiarta, “VGG16 Transfer Learning Architecture for Salak Fruit Quality Classification,” *Telematika*, vol. 18, no. 1, p. 37, 2021, doi: 10.31315/telematika.v18i1.4025.
- [28] LTDC Team, “What is the VGG-19 neural network_ - Quora.” [Online]. Available: <https://www.quora.com/What-is-the-VGG-19-neural-network>.
- [29] Q. Ji, J. Huang, W. He, and Y. Sun, “Optimized deep convolutional neural networks for identification of macular diseases from optical coherence tomography images,” *Algorithms*, vol. 12, no. 3, pp. 1–12, 2019, doi: 10.3390/a12030051.
- [30] K. H. Mahmud, Adiwijaya, and S. Al Faraby, “Klasifikasi Citra Multi-Kelas Menggunakan Convolutional Neural Network,” *e-Proceeding Eng.*, vol. 6, no. 1, pp. 2127–2136, 2019.
- [31] L. Perez and J. Wang, “The Effectiveness of Data Augmentation in Image Classification using Deep Learning,” 2017, [Online]. Available: <http://arxiv.org/abs/1712.04621>.
-