Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871

Vol. 6, No. 2, April 2025, Page. 483-500
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4094

Convolutional Neural Network for COVID-19 Detection Using InceptionV3 Transfer Learning

Dhimas Rama Anthony Navy Pratama[1], Yufis Azhar[2]

[1,2]Informatics, Universitas Muhammadiyah Malang, Indonesia

Email: [1]Dhimaspratama1410@gmail.com

Received : Nov 11, 2024; Revised : Nov 26, 2024; Accepted : Dec 31, 2024; Published : Apr 26, 2025

Abstract

The COVID-19 pandemic has underscored the need for rapid and accurate diagnostic methods. Although Reverse Transcription Polymerase Chain Reaction (RT-PCR) is the gold standard for detecting COVID-19, it presents limitations such as high costs, lengthy processing times, and the requirement for specialized personnel. Medical imaging, particularly lung X-rays, offers a viable alternative for COVID-19 detection. This study evaluates five Convolutional Neural Network (CNN) models: a handcrafted CNN, VGG-16, VGG-19, ResNet50, and InceptionV3, with the aim of enhancing classification accuracy between COVID-19 and normal lung images. The dataset, obtained from Kaggle, comprises 13,808 X-ray images, which were balanced using random oversampling to address class imbalance. Data augmentation techniques were applied to improve model generalization and mitigate overfitting. After training the models for 100 epochs, the results revealed that both VGG-19 and InceptionV3 achieved the highest accuracy, each attaining 100%, outperforming the other models. VGG-16 and CNN Handcraft also demonstrated strong performance with an accuracy of 99% and 97%, whereas ResNet50 exhibited the lowest accuracy at 78%. These findings suggest that more complex CNN architectures, such as VGG-19 and InceptionV3, are highly effective in detecting COVID-19 from X-ray images. Future research should explore additional CNN models and employ further model tuning to optimize performance.

Keywords : *Convolutional Neural Network, COVID-19, Image Classification, SARS-CoV-2, Transfer Learning.*

1. INTRODUCTION

The world has faced various pandemics that have had significant impacts on the global population. Past pandemics such as the Black Death in the 14th century, the Spanish Flu in 1918, and the H1N1 influenza outbreak in 2009 [1] have taught us the importance of rapid and effective responses in combating the spread of infectious diseases. These historical events highlight not only the critical need for robust healthcare systems but also the role of technological innovation in mitigating the impacts of pandemics[2], [3]. Each pandemic, has unique characteristics that require specialized approaches for detection and management.

COVID-19, caused by a virus, struck the globe in December 2019, with the first reported cases of pneumonia in Wuhan, Hubei Province, China. Inoculating respiratory samples into human respiratory tract epithelial cells, and the Vero E6 and Huh7 cell lines led to the isolation of a new respiratory virus. Its genome revealed that the virus was a coronavirus related to SARS-CoV, which was subsequently named Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) [4], [5], [6]. This has triggered an unprecedented global health crisis. One of the greatest challenges in the pandemic is to develop methods for the early and accurate detection of COVID-19 infections. Laboratory-based diagnostic tests such as RT-PCR (Reverse Transcription Polymerase Chain Reaction) are accurate, but they have drawbacks including the need for expensive equipment and testing costs, specialized laboratory personnel, relatively long processing times, and high exposure risks [7], [8].

In this situation, medical imaging technologies like Lung X-rays and CT-Scans offer promising alternatives. Medical imaging data can reveal specific patterns indicative of certain disease infections. However, manual analysis can be time-consuming and prone to errors due to differences in individual interpretations. Therefore, an automated solution is needed that can assist in analyzing medical images quickly and accurately. A solution for automation is the use of deep learning. Deep learning, based on artificial neural network technology [9], [10], includes highly effective architectures for image analysis such as Convolutional Neural Networks (CNN).

A Convolutional Neural Network (CNN) comprises main layers of convolution, pooling, non-linear activation, and fully connected layers. The convolutional layer uses primary filters to extract local features from the input image, producing hierarchical feature maps. The pooling layer reduces the spatial dimensions of the feature maps and the number of parameters, making the model more robust to variations in image data. Non-linear activation functions, such as ReLU, introduce non-linearities to capture complex relationships within the data. The fully connected layer classifies the extracted features, serving as the classifier in the CNN, and generates the categorical output from the image [11]. CNNs have demonstrated exceptional performance in pattern recognition and image classification, making this architecture highly beneficial for medical image analysis. However, CNNs achieve good results on large datasets but require substantial computational resources [12], [13].

Transfer learning is a method that can reduce the time and computational resources required [13], [14]. It allows the use of models that have been trained on large datasets with minimal adjustments. One of the frequently used CNN models in transfer learning is InceptionV3. InceptionV3 has a structure that includes input, output, and classification processes. InceptionV3 performs the extraction process and has hidden layers such as convolution, pooling, ReLU, softmax, and fully connected layers that work in a structured manner to organize objects [15]. By leveraging the InceptionV3 model through transfer learning, we can accelerate the training process and enhance accuracy for COVID-19 detection from medical imaging data.

Research [16] focused on analyzing the performance comparison of X-ray lung image classification algorithms for COVID-19 detection, employing three main methods: K-nearest neighbors (KNN), Support Vector Machine (SVM), and Convolutional Neural Network (CNN). The dataset, sourced from Kaggle, included X-ray images of normal lungs, lungs infected with COVID-19, and viral pneumonia. This study utilized five-fold cross-validation to test the performance of these algorithms. The results indicated that CNN outperformed KNN and SVM across all performance metrics, including accuracy, precision, recall, and F1 score, despite having a longer execution time. This research opens opportunities for the use of CNNs in future medical applications, although they require greater computational resources.

However, this study [16] did not explore other Convolutional Neural Network (CNN) architectures. It provides a foundation that CNNs are a more robust approach compared to KNN and SVM but still necessitates further testing on more complex CNN architectures to enhance accuracy.

In Research [17], the use of Convolutional Neural Network (CNN) models VGG19 and ResNet50 was examined for the classification of X-ray lung images in COVID-19 cases, focusing on modifications to the dropout regularization values and the number of classification layers. The dataset consisted of 21,165 images, subsequently divided into 10% for testing and 90% for training and validation. A five-fold cross-validation methodology was used to assess the performance of the models, with results showing that ResNet50 outperformed VGG19, achieving an accuracy of 94.4%. This study underscores the effectiveness of optimized models in medical diagnostics and suggests that enhancements to the dataset and the application of techniques like dropout can counteract overfitting, thereby increasing the accuracy and reliability of the models in medical applications.

This study [17] successfully addressed the limitations of previous research that had not explored various Convolutional Neural Network (CNN) models. By utilizing the ResNet50 and VGG19 models, researchers were able to compare the performance across more complex CNN architectures. The use

P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4094

of augmentation techniques also helped improve model accuracy and addressed weaknesses from previous studies.

In Research [18], a hybrid neural network combining the Xception and ResNet50V2 models was developed for the classification of X-ray lung images, identifying normal, pneumonia, and COVID-19 cases. Utilizing two open-source datasets, this study tackled the challenge of imbalanced datasets by dividing the training set into eight sequential phases, ensuring nearly equal data representation from each class in every phase. This network was trained using transfer learning techniques with ImageNet pre-trained weights and optimized with specified training parameters, achieving an average accuracy of 99.50% for COVID-19. This research suggests that this method can be applied to situations with highly imbalanced datasets and hopes that larger data in the future could further enhance accuracy.

Although this study [18]successfully improved the accuracy of models from previous research, the minority class in this study, COVID-19, still requires improvement, despite the overall high accuracy. The datasets used in this research remain imbalanced and unresolved.

In Research [19], the focus was on classifying X-ray lung images to detect COVID-19 using Convolutional Neural Network (CNN) methods and Transfer Learning. By leveraging four different scenarios: CNN Handcraft Model, VGG16, VGG19, and ResNet50, this study implemented data augmentation techniques and data balancing using undersampling. The dataset consisted of 13,808 X-ray lung images, divided into two classes: COVID-19 and Normal. Results showed that the first scenario with the CNN Handcraft model, achieving an accuracy score of 95%, outperformed the others. This research highlights the importance of augmentation and data-balancing techniques in enhancing the accuracy of medical image classification models. For further development in this study, the use of oversampling methods for data balancing and the exploration of other transfer learning architectures such as InceptionV3 are recommended.

This study [19] used undersampling techniques to address the issue of dataset imbalance noted in previous research. It also conducted evaluations of several Convolutional Neural Network (CNN) models to provide a clearer understanding of the advantages and limitations of each CNN model.

This research will develop five models: CNN Handcraft, Transfer Learning VGG16, Transfer Learning VGG19, Transfer Learning ResNet50, and Transfer Learning InceptionV3. The goal is to develop more accurate COVID-19 detection models. Based on the findings and shortcomings of previous research, this study focuses on refining the model architectures to maximize accuracy results. Previous research has shown that CNNs outperform other algorithms; however, issues with dataset imbalance still affect the outcomes of the models.

Therefore, this research will utilize data augmentation and rebalancing techniques using oversampling to ensure that each class in the dataset is evenly represented. Additionally, model tuning will be conducted to minimize overfitting and stabilize the training process. The inclusion of the InceptionV3 model aims to explore more modern models for comparison with other CNN architectures.

2. METHOD

Vol. 6, No. 2, April 2025, Page. 483-500
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4094

2.1. Research Stages

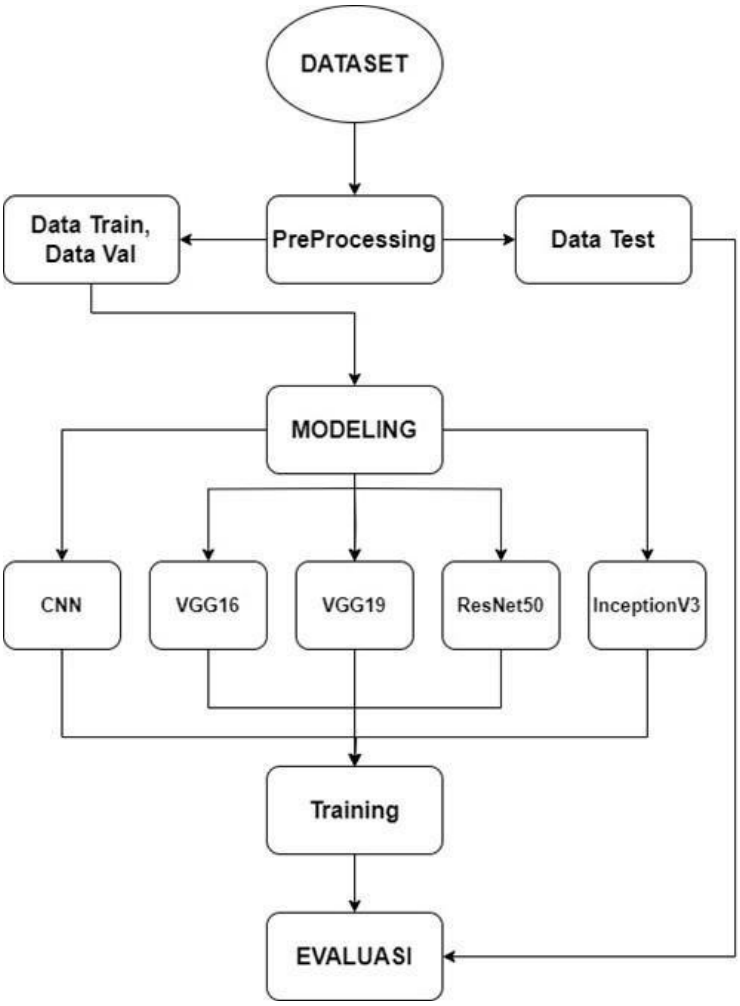


Figure 1. Research Stage

The research method used consists of five stages: dataset collection, preprocessing, classification model formulation, model training, and evaluation of each model. Figure 1 illustrates the flow of the research methods used. During the preprocessing stage, data balancing is achieved using Oversampling Techniques, splitting the data into 70% Train Data, 20% Test Data, and 10% Validation Data, followed by data augmentation. In the training phase, modelling is conducted to train the data using the developed model and parameters. Meanwhile, validation data is used to validate the training results with model parameters to detect any occurrences of underfitting or overfitting.

2.2. Dataset

The dataset used in this research consists of COVID-19 X-ray lung images. The "COVID-19 Radiography Database" from Kaggle includes images from researchers at Qatar University and Dhaka University who collaborated with physicians to create a database containing images of X-ray lung scans positive for COVID-19, as well as normal lung images, Tuberculosis, and Viral pneumonia. The data utilized in this study comprises two classes: COVID-19 and Normal, containing 13,808 image data with a resolution of 256x256 pixels in PNG format. Examples from this dataset can be seen in Figure 2.

Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871

Vol. 6, No. 2, April 2025, Page. 483-500
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4094

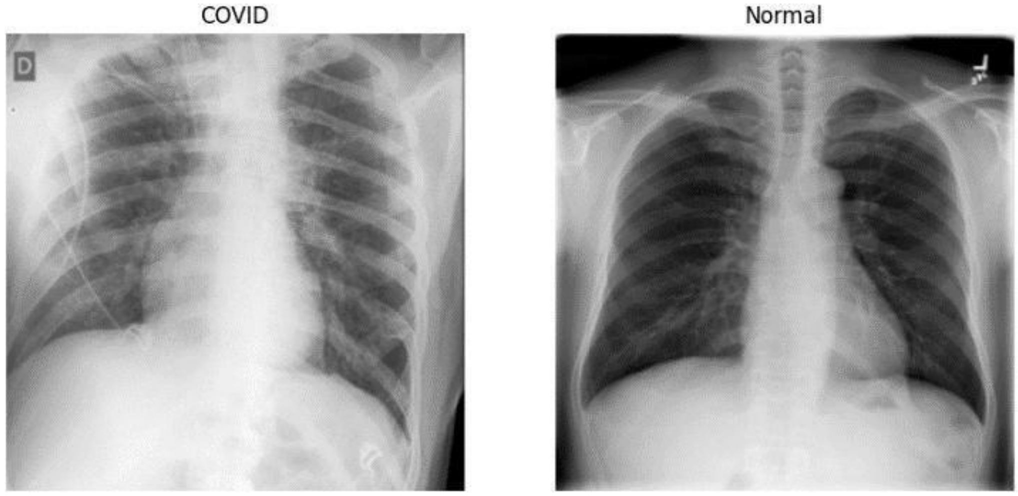


Figure 2. Sample Data of COVID-19 and Normal X-ray Images

2.3. Model Architecture

This study used five models for comparative testing. The original image size of 256x256 pixels was resized to 150x150 to focus the images that could help indicate the presence of COVID-19 and to accelerate the computational process so as to differentiate the two classes of images consisting of lungs infected with COVID-19 and normal lungs.

The first model, illustrated in Figure 3, applies three convolutional layers and three pooling layers using the max pooling technique with a 2x2 filter size. Subsequently, three more convolutional layers are used, where the layers have different filters sized 128, 64, and 32. The kernels in these convolutional layers are 3x3 and followed by ReLU activation. This is followed by a Fully Connected layer consisting of Flatten and Dense layers using ReLU activation, followed by a Dropout layer with a rate of 0.5 to address overfitting during training. The sequence ends with a dense layer using Softmax activation.

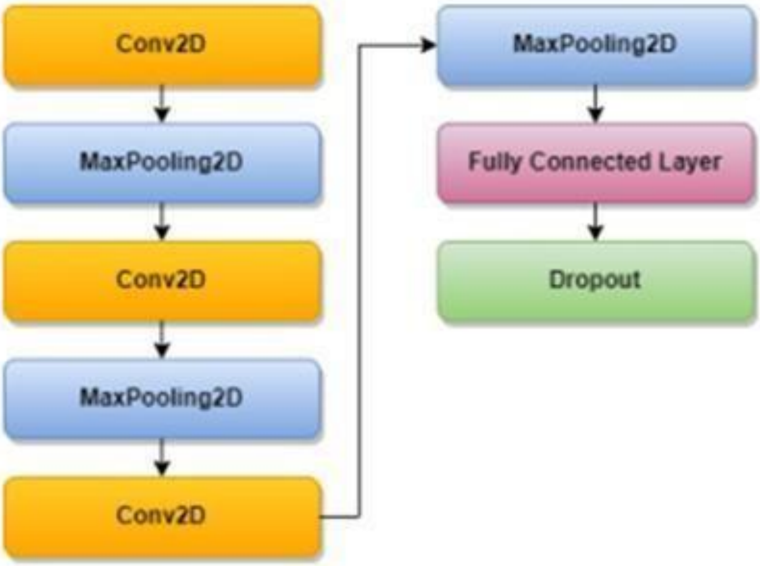


Figure 3. CNN Model Architecture

Vol. 6, No. 2, April 2025, Page. 483-500
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4094

The second model uses Transfer Learning with VGG-16, illustrated in Figure 4. VGG-16 is a convolutional model that utilizes convolutional layers with a 3x3 convolutional filter size and has 16 layers, including 13 convolutional layers and 3 fully connected layers [20]. This model begins with a series of convolutional layers, where the first layer (conv1) takes an input image of 224x224 and applies 64 3x3 filters with ReLU activation, followed by a second convolutional layer (conv2) which reduces the spatial dimension to 112x112 while increasing the depth to 128 filters. The process continues with deeper convolutional layers (conv3, conv4, and conv5) that gradually increase the number of filters to 256 and 512, while the spatial size continues to decrease.

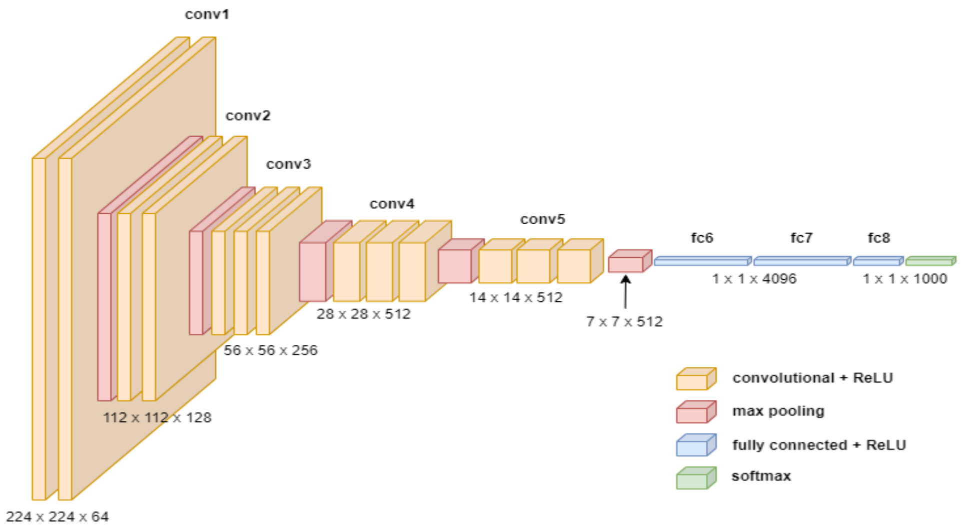


Figure 4. VGG16 Model Architecture [src. Towardsdatascience]

Third model used transfer learning with the VGG-19 architecture, as shown in Figure 5. Similar to VGG-16, VGG-19 included additional layers to enhance its capability to extract features from images [21]. It started with an input image size of 224x224x3 and comprised five blocks, each consisting of several convolutional layers activated by ReLU. The first two layers used 64 3x3 filters followed by max pooling to reduce spatial dimensions. The second and third blocks each had two and four convolutional layers with 128 and 256 filters, respectively, followed by max pooling. The fourth and fifth blocks contained four convolutional layers with 512 filters, followed by max pooling. After the convolution and pooling, the extracted features were flattened and passed through three fully connected layers. The first two layers each had 4096 units, and the third was an output layer with 1000 units using softmax activation for classification.

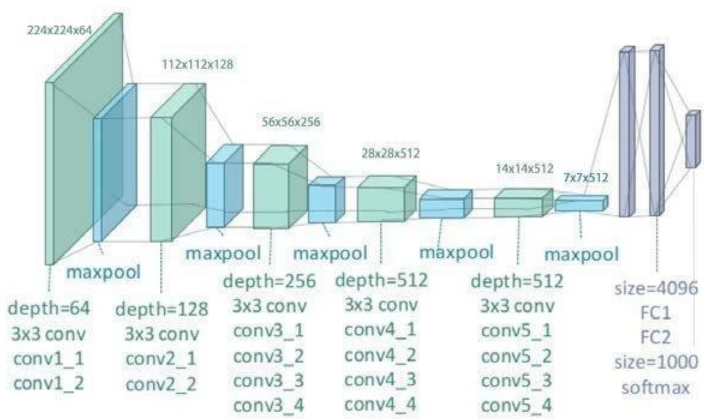


Figure 5. VGG19 Model Architecture [src. Pemrogramanmatlab]

Vol. 6, No. 2, April 2025, Page. 483-500
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4094

The fourth model used transfer learning with ResNet50, illustrated in Figure 6. ResNet50 introduced the concept of Shortcut Connections, which prevented the system from losing significant information during the training process [22]. It began with a Zero Padding layer followed by a standard convolutional layer, then equipped with Batch Normalization and activated by ReLU for training stabilization. A max pool layer was used to reduce spatial dimensions before entering a series of Identity Blocks and Convolutional Blocks that further altered dimensions. The process continued with Average Pooling that flattened features into a more manageable form and concluded with a Fully Connected layer that classified the output.

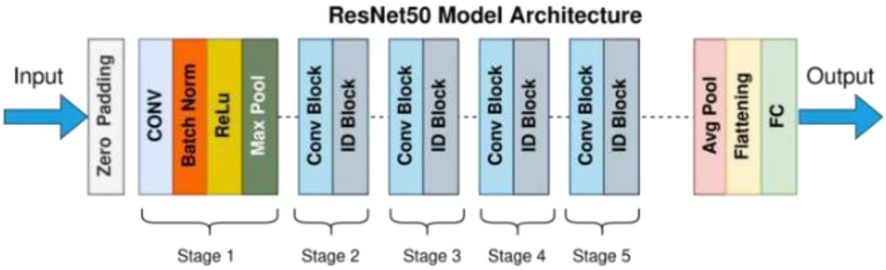


Figure 6. ResNet50 Model Architecture [src. Towardsdatascience]

The fifth model used InceptionV3, as depicted in Figure 7. InceptionV3 performed extraction processes and featured hidden layers such as convolution, pooling, ReLU activation, softmax, and fully connected layers [23]. It began with several convolutional and max pooling layers for initial dimension reduction, followed by a series of Inception modules. Each Inception module used a combination of convolutional layers with different filter sizes (1x1, 3x3, 5x5) and pooling layers in parallel to capture features at various scales within a single network layer. This was followed by additional layers for normalization and further dimension reduction, culminating in the final layers where the model used average pooling followed by softmax for classification.

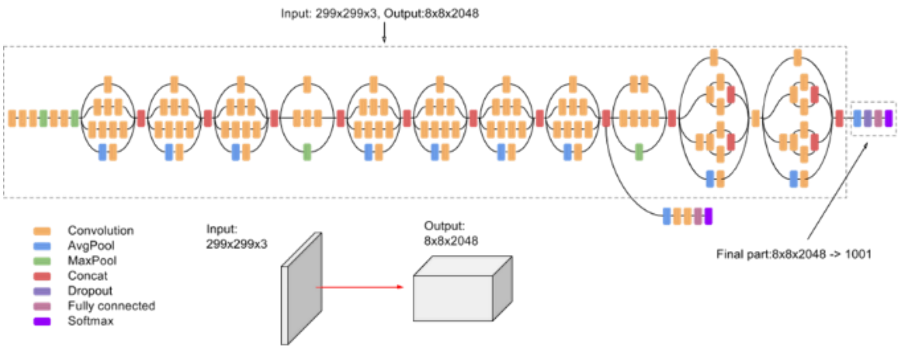


Figure 7. InceptionV3 Model Architecture [src. Paperspace]

2.4. Data Augmentation

Data augmentation is a technique that can reduce overfitting by minimally increasing the dataset size [24]. Data augmentation is crucial for enhancing the performance, durability, and ability of machine learning models to generalize effectively[25]. The following image augmentation parameters were used: rotation (rotation_range = 20), width shift (width_shift_range = 0.10), height shift (height_shift_range = 0.10), scaling (rescale = 1/255), shear (shear_range = 0.1), zoom (zoom_range = 0.1), horizontal flip (horizontal_flip = True), and fill mode (fill_mode = 'nearest').

P-ISSN: 2723-3863
https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871
DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4094

2.5. Model Testing

In this research, classification was performed on two classes: COVID X-ray and normal X-ray. balancing data is fundamental in research to guarantee that machine learning models are precise, equitable, and able to consistently deliver trustworthy predictions across various classes [26]. Random oversampling was used to balance the data between classes. Initially, there were 13,808 images, split into 3,616 COVID-19 X-ray images and 10,192 normal X-ray images. After applying oversampling, the total number of images equalized to 20,384, with each class containing 10,192 images.

Subsequently, the data was divided into three types: 70% train data, 10% validation data, and 20% test data from the total data in each class. In Table 1, the data used from the dataset was unbalanced, where data counts were not equal across classes. Random Oversampling (ROS) was used to balance and even out data across each class. Oversampling is a method of generating minority class data equal to the majority class data [27], [28]. Random Oversampling (ROS) involves randomly assigning data from the minority class to the training data. This process was repeated until the minority class data matched the majority class data in quantity [29], [30].

Table 1. Data Quantity Before Oversampling

CLASS	TRAIN DATA	VALIDATION DATA	TEST DATA	TOTAL DATA
COVID	2.531	361	724	3.616
NORMAL	7.124	1.019	2.049	10.192

After Random Oversampling (ROS), as shown in Table 2, the data was evened out with the majority class and then divided into train, validation, and test data.

Table 2. Data Quantity After Oversampling

CLASS	TRAIN DATA	VALIDATION DATA	TEST DATA	TOTAL DATA
COVID	7.124	1.019	2.049	10.192
NORMAL	7.124	1.019	2.049	10.192

the dataset that had undergone data balancing was trained on five models. The first model used CNN Handcraft, the second model used VGG-16, the third model used VGG-19, the fourth model utilized ResNet50, and the fifth model used InceptionV3.

3. RESULT

The testing phase in this study involved training the five proposed models. Model development and training for each model utilized 100 iterations (epochs). Subsequently, results were presented through classification outcomes, testing, and confusion matrices for each tested model.

3.1. Training Results with the CNN Handcraft Model

The results from training the Convolutional Neural Network (CNN) model are depicted in the graphs shown in Figures 8 and 9, which illustrate accuracy and loss across 100 epochs. Based on the graphs, the model achieved an accuracy of 0.95 and a loss value of 0.09.

Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 6, No. 2, April 2025, Page. 483-500
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4094

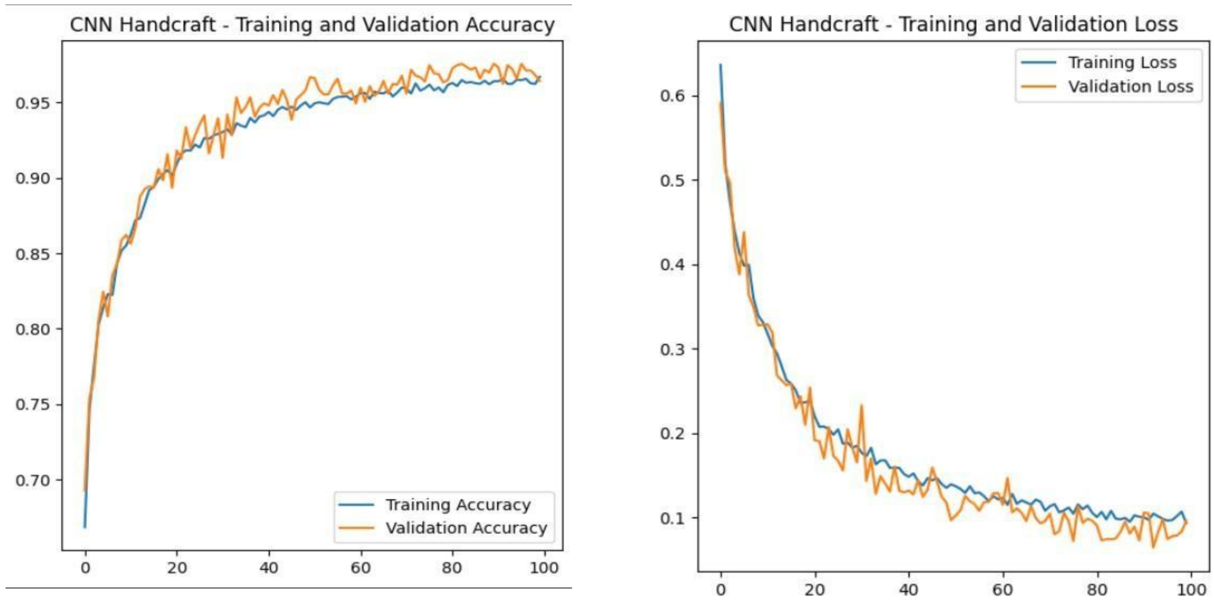


Figure 8 & 9 Accuracy and Loss Graph of the CNN Handcraft Model

Following the graph analysis, the model evaluation process continued with a confusion matrix, the results of which are displayed in Table 4. It shows that for the COVID class, 2006 data points were correctly predicted while 43 were mispredicted. Similarly, in the Normal class, 1955 data points were correctly predicted and 94 were mispredicted. Subsequently, the model was tested with test data to predict images. Figure 10 shows an image correctly predicted as the COVID class with an accuracy of 0.99 and a processing time of 0.273 seconds. Further testing involved predictions with the normal class. Figure 11 shows an image correctly predicted as the Normal class with an accuracy of 1.00 and a processing time of 0.057 seconds.

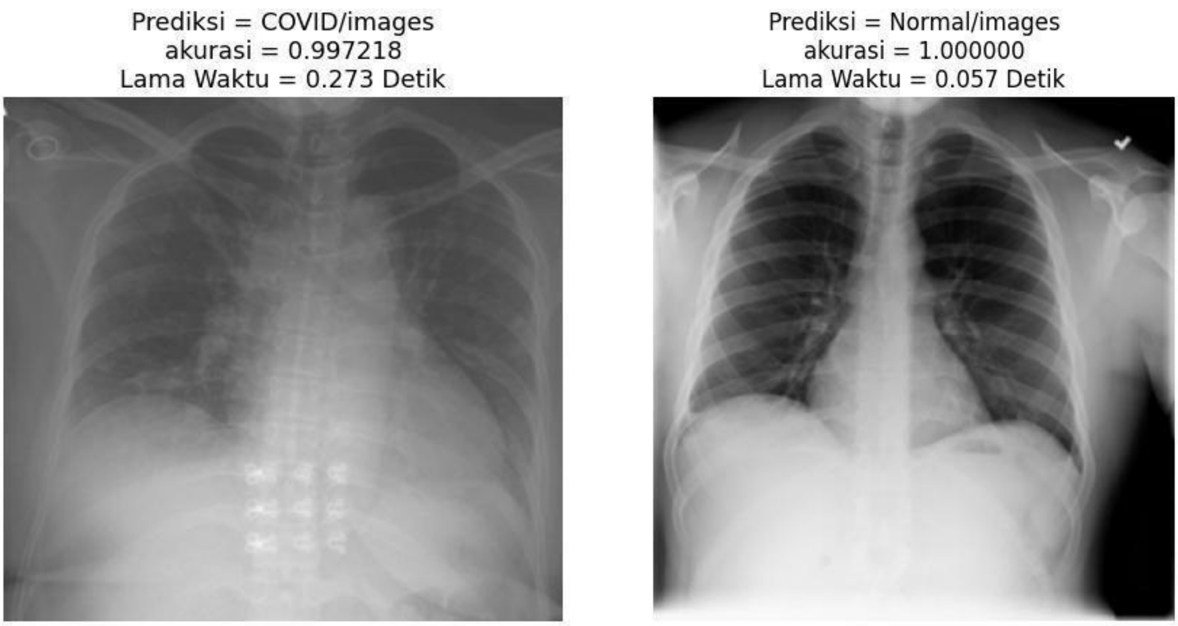


Figure 10 & 11 Covid and Normal Detection Result of the CNN Handcraft Model

Jurnal Teknik Informatika (JUTIF) Vol. 6, No. 2, April 2025, Page. 483-500
P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4094

3.2. Training Results with the VGG - 16 Model

The results from training the VGG-16 model are depicted in the graphs shown in Figures 12 and 13, which illustrate accuracy and loss over 100 epochs. The graphs indicated an accuracy of 0.99 and a loss value of 0.03.

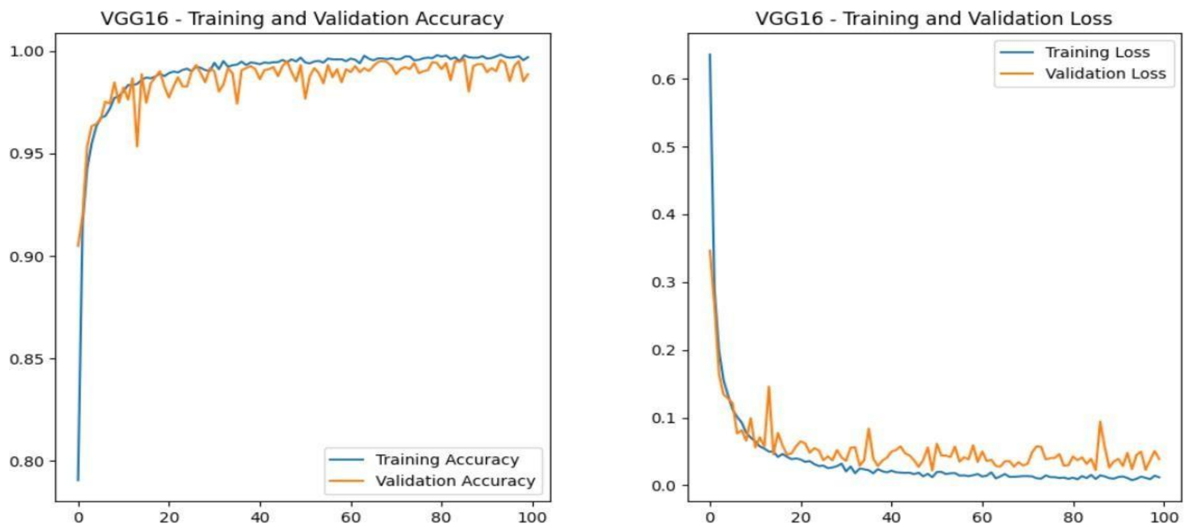


Figure 12 & 13 Accuracy and Loss Graph of the VGG -16 Model

After analyzing the graph plots, the model evaluation process continued with a confusion matrix, the results of which are displayed in Table 4. It shows that for the COVID class, 2047 data points were correctly predicted while 2 were mispredicted. Similarly, in the Normal class, 2013 data points were correctly predicted and 36 were mispredicted. Subsequently, the model was tested with test data to predict images. Figure 14 shows an image correctly predicted as the COVID class with an accuracy of 0.99 and a processing time of 0.945 seconds. Further testing involved predictions with the normal class. Figure 15 shows an image correctly predicted as the Normal class with an accuracy of 1.00 and a processing time of 0.060 seconds.

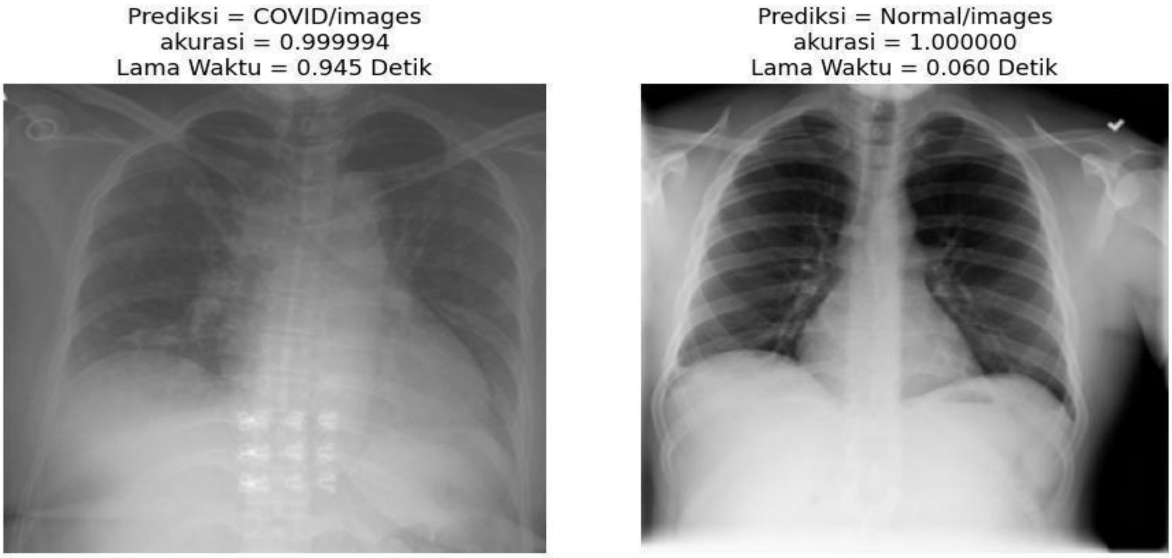
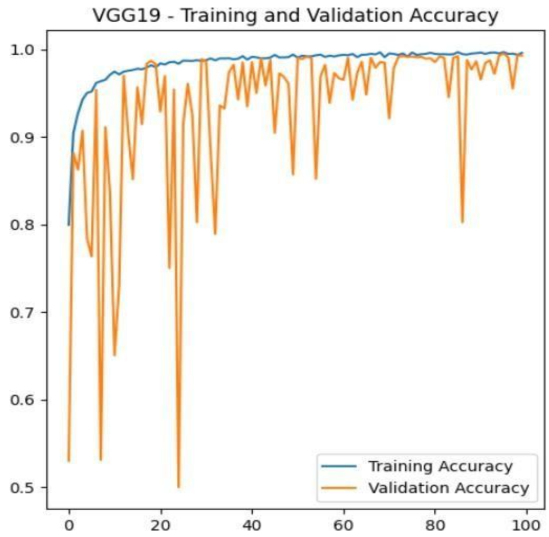


Figure 14 & 15 Covid and Normal Detection Result of the VGG - 16 Model

Jurnal Teknik Informatika (JUTIF) Vol. 6, No. 2, April 2025, Page. 483-500
P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4094

3.3. Training Results with the VGG - 19 Model

The results from training the VGG-19 model are depicted in the graphs shown in Figures 16 and 17, which illustrate accuracy and loss over 100 epochs. The graphs indicated an accuracy of 0.99 and a loss value of 0.03.



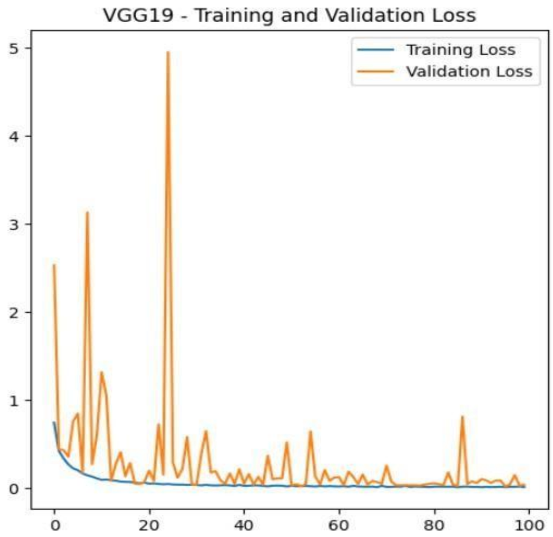


Figure 16 & 17 Accuracy and Loss Graph of the VGG -19 Model

After analyzing the graph plots, the model evaluation process continued with a confusion matrix, the results of which are displayed in Table 4. It shows that for the COVID class, 2034 data points were correctly predicted while 15 were mispredicted. Similarly, in the Normal class, 2047 data points were correctly predicted and 2 were mispredicted. Subsequently, the model underwent testing with test data to predict images. Figure 18 shows an image correctly predicted as the COVID class with an accuracy of 0.99 and a processing time of 0.214 seconds. Further testing involved predictions with the normal class. Figure 19 shows an image correctly predicted as the Normal class with an accuracy of 1.00 and a processing time of 0.061 seconds.

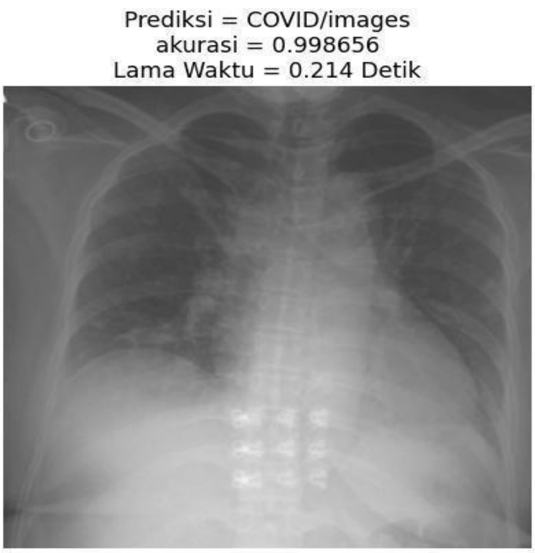
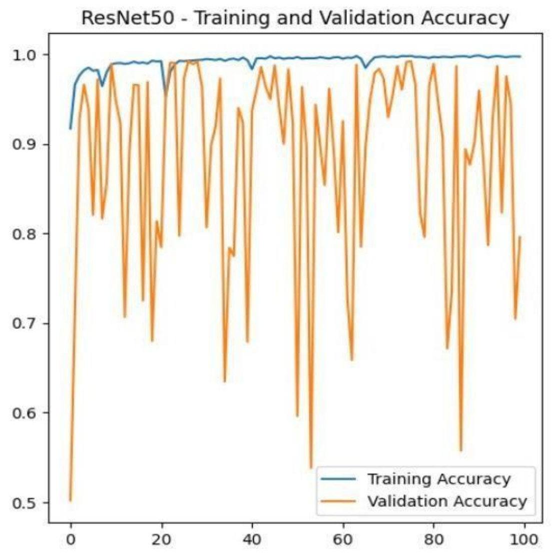


Figure 18 & 19 Covid and Normal Detection Result of the VGG - 19 Model

Jurnal Teknik Informatika (JUTIF) Vol. 6, No. 2, April 2025, Page. 483-500
P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4094

3.4. Training Results with the ResNet50 Model

The results from training the ResNet50 model are depicted in the graphs shown in Figures 20 and 21, which illustrate accuracy and loss over 100 epochs. The graphs indicated an accuracy of 0.99 and a loss value of 1.6.



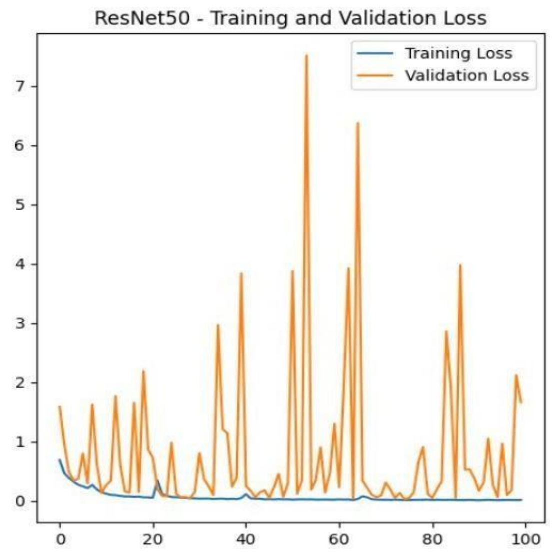


Figure 20 & 21 Accuracy and Loss Graph of the ResNet50 Model

After analyzing the graph plots, the model evaluation process continued with a confusion matrix, the results of which are displayed in Table 4. It shows that for the COVID class, 2049 data points were correctly predicted with no mispredictions. However, in the Normal class, 1137 data points were correctly predicted and 912 were mispredicted. Subsequently, the model underwent testing with test data to predict images. Figure 22 shows an image correctly predicted as the COVID class with an accuracy of 1.00 and a processing time of 1.683 seconds. Further testing involved predictions with the normal class. Figure 23 shows an image correctly predicted as the Normal class with an accuracy of 0.99 and a processing time of 0.070 seconds.

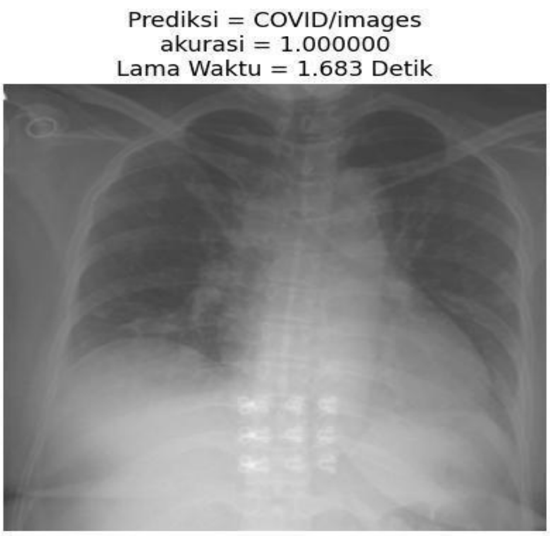


Figure 22 & 23 Covid and Normal Detection Result of the ResNet50 Model

P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4094

3.5. Training Results with the InceptionV3 Model

The results from training the InceptionV3 model are depicted in the graphs shown in Figures 24 and 25, which illustrate accuracy and loss over 100 epochs. The graphs indicated an accuracy of 0.99 and a loss value of 0.02.

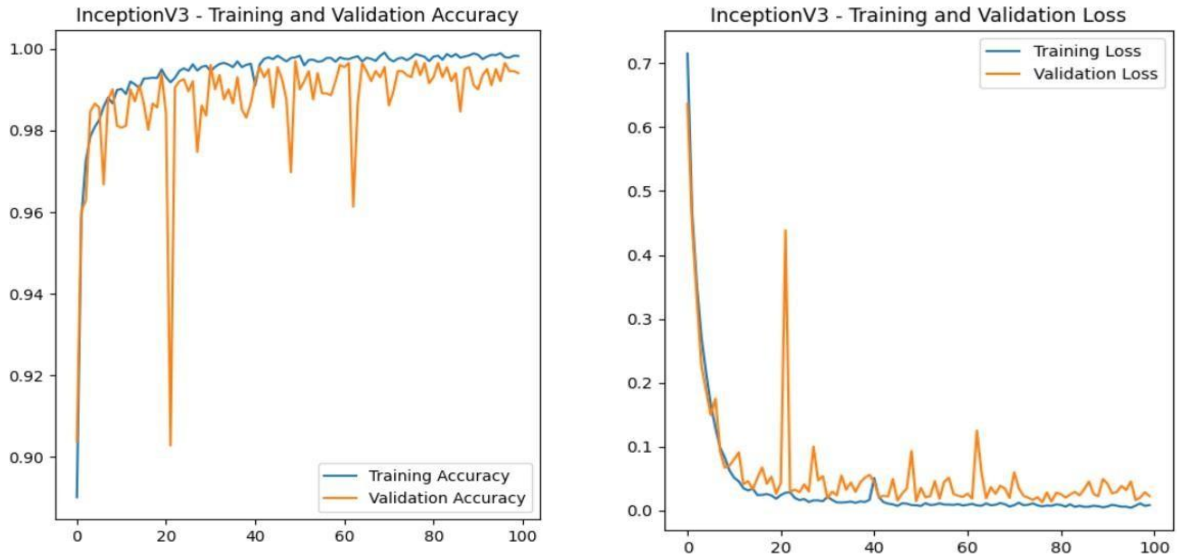


Figure 24 & 25 Accuracy and Loss Graph of the InceptionV3 Model

After analyzing the graph plots, the model evaluation process continued with a confusion matrix, the results of which are displayed in Table 4. It shows that for the COVID class, 2045 data points were correctly predicted while 4 were mispredicted. Similarly, in the Normal class, 2034 data points were correctly predicted and 15 were mispredicted. Subsequently, the model underwent testing with test data to predict images. Figure 26 shows an image correctly predicted as the COVID class with an accuracy of 0.99 and a processing time of 2.534 seconds. Further testing involved predictions with the normal class. Figure 27 shows an image correctly predicted as the Normal class with an accuracy of 0.99 and a processing time of 0.073 seconds.

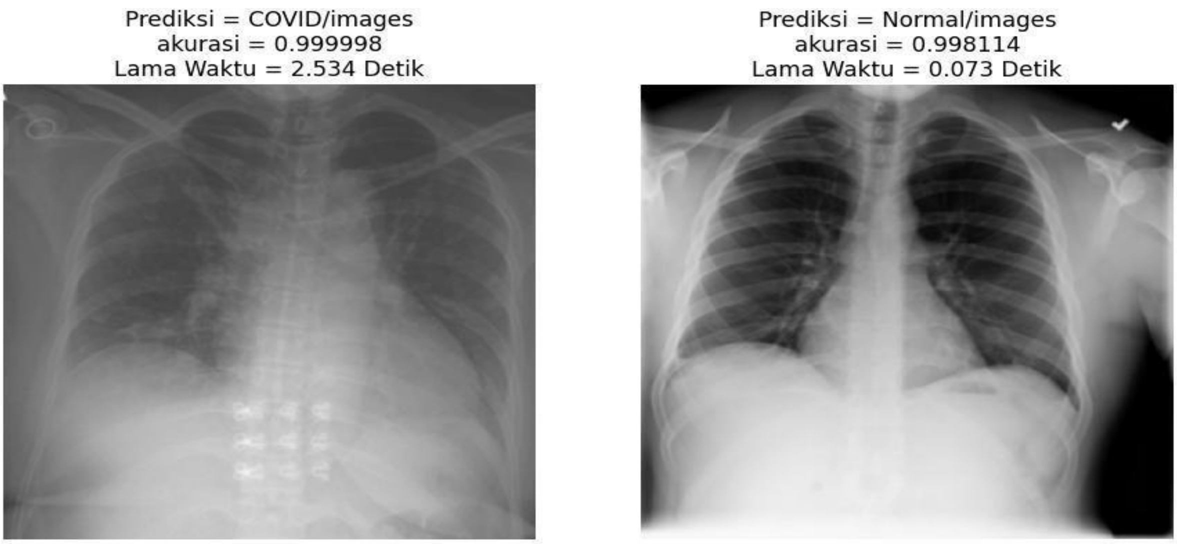


Figure 26 & 27 Covid and Normal Detection Result of the InceptionV3 Model

P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 6, No. 2, April 2025, Page. 483-500
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4094

3.6. Performance Comparison of Models

After training the five proposed models, their performance was tested using classification reports. The results are displayed in Table 3.

Table 3. Performance Result from Each Model using Classification Report

MODEL	CLASS	PRECISION	RECALL	ACCURACY
CNN	Covid	96%	98%	97%
	Normal	98%	95%	
VGG – 16	Covid	98%	100%	99%
	Normal	100%	98%	
VGG – 19	Covid	100%	99%	100%
	Normal	99%	100%	
ResNet50	Covid	69%	100%	78%
	Normal	100%	55%	
InceptionV3	Covid	99%	100%	100%
	Normal	100%	99%	

The confusion matrix result are displayed in Table 4.

Table 4. Performance Result from Each Model using Classification Report

MODEL	PREDICTION			
	COVID		NORMAL	
	CORRECT	INCORRECT	CORRECT	INCORRECT
CNN	2006	43	1955	94
VGG – 16	2047	2	2013	36
VGG – 19	2034	15	2047	2
ResNet50	2049	0	1137	912
InceptionV3	2045	4	2034	16

4. DISCUSSIONS

The results of this study demonstrate that the more complex Convolutional Neural Network (CNN) architectures, specifically VGG-19 and InceptionV3, are highly effective in classifying COVID-19 and normal lung X-ray images. Both models achieved an impressive accuracy of 100%, significantly outperforming the other models tested, including the CNN handcraft, VGG-16, and ResNet50. VGG-16 also exhibited strong performance with a 99% accuracy, suggesting that deep CNN architectures are well-suited for medical image classification tasks, particularly in the detection of COVID-19. The CNN handcrafted model, while simpler, still performed reasonably well with a 97% accuracy, although it lagged behind the more advanced architectures. In contrast, ResNet50 exhibited the weakest performance, with an accuracy of 78%, largely due to significant misclassification in the normal class. This indicates that while ResNet50 is typically effective in other

Vol. 6, No. 2, April 2025, Page. 483-500
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4094

image classification tasks, it may not be the most suitable model for COVID-19 detection without further tuning or dataset adjustments.

One key factor contributing to the success of VGG-19 and InceptionV3 could be their ability to capture more complex and fine-grained features from the X-ray images due to their deeper architectures and the inclusion of more convolutional layers. These models likely benefited from the use of data augmentation and random oversampling, which helped mitigate the issue of class imbalance and reduce overfitting, particularly in the COVID-19 class. The underperformance of ResNet50 may suggest that its shortcut connections, while useful in preventing vanishing gradients, might not be as beneficial in handling the specific patterns present in medical images of lung infections. Future research could focus on refining the ResNet50 model or exploring other CNN architectures.

5. CONCLUSION

Based on the testing of the five models proposed in this study, it can be concluded that the VGG-19 and InceptionV3 models exhibited the best performance, achieving an overall accuracy of 100%. These models were able to predict nearly all data with high precision, with minimal errors for both COVID and normal classes, as evidenced by the confusion matrix results.

The VGG-16 model also demonstrated good performance with an accuracy of 99%, though it allowed for some predictive errors. Conversely, the CNN model achieved an accuracy of 97% with a few prediction errors in both classes, while the ResNet50 model achieved an accuracy of 78% with a significant number of predictive errors in the Normal class, marking the lowest accuracy. From these findings, it can be concluded that VGG-19 and InceptionV3 are the most suitable models for classifying COVID and normal images in this research.

For future research, it is recommended to explore other CNN models such as MobileNet, and to use oversampling techniques such as SMOTE, along with different tuning techniques to prevent stagnation and Overfitting during the learning process.

REFERENCES

[1] "Mengulas Riwayat Pandemi Dunia - Update & Tren." Accessed: Jul. 02, 2024. [Online]. Available: https://beranisehat.com/mengulas-riwayat-pandemi-dunia/

[2] R. Filip, R. Gheorghita Puscaselu, L. Anchidin-Norocel, M. Dimian, and W. K. Savage, "Global Challenges to Public Health Care Systems during the COVID-19 Pandemic: A Review of Pandemic Measures and Problems," Aug. 01, 2022, *MDPI*. doi: 10.3390/jpm12081295.

[3] M. Chandra, K. Kumar, P. Thakur, S. Chattopadhyaya, F. Alam, and S. Kumar, "Digital technologies, healthcare and Covid-19: insights from developing and emerging nations," Mar. 01, 2022, *Springer Science and Business Media Deutschland GmbH*. doi: 10.1007/s12553-022-00650-1.

[4] M. Ciotti, M. Ciccozzi, A. Terrinoni, W.-C. Jiang, C.-B. Wang, and S. Bernardini, "The COVID-19 pandemic," *Crit Rev Clin Lab Sci*, vol. 57, no. 6, pp. 365–388, Aug. 2020, doi: 10.1080/10408363.2020.1783198.

[5] K. D. Vihta *et al.*, "Omicron-Associated Changes in Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Symptoms in the United Kingdom," *Clinical Infectious Diseases*, vol. 76, no. 3, pp. E133–E141, Feb. 2023, doi: 10.1093/cid/ciac613.

[6] F. A. AlMalki, S. Albukhaty, A. A. Alyamani, M. N. Khalaf, and S. Thomas, "The relevant information about the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) using the five-question approach (when, where, what, why, and how) and its impact on the environment," May 01, 2023, *Springer Science and Business Media Deutschland GmbH*. doi: 10.1007/s11356-022-18868-x.

[7] J. Kesehatan Manarang, A. Suswanti Agustina, and R. Fajrunni, “Volume 6, Nomor Khusus, Oktober 2020 Perbandingan Metode RT-PCR dan Tes Rapid … 47.” [Online]. Available: http://jurnal.poltekkesmamuju.ac.id/index.php/m

[8] C. Parellada *et al.*, “Comparative study between an antigen test and the reverse transcription-polymerase chain reaction (RT-PCR) test for the diagnosis of COVID-19 in pediatrics,” *Arch Argent Pediatr*, vol. 121, no. 5, Oct. 2023, doi: 10.5546/aap.2022-02908.eng.

[9] S. Ilahiyah and A. Nilogiri, “Implementasi Deep Learning Pada Identifikasi Jenis Tumbuhan Berdasarkan Citra Daun Menggunakan Convolutional Neural Network”.

[10] S. M. Kasongo, “A deep learning technique for intrusion detection system using a Recurrent Neural Networks based framework,” *Comput Commun*, vol. 199, pp. 113–125, Feb. 2023, doi: 10.1016/j.comcom.2022.12.010.

[11] L. Chen, S. Li, Q. Bai, J. Yang, S. Jiang, and Y. Miao, “Review of Image Classification Algorithms Based on Convolutional Neural Networks,” *Remote Sens (Basel)*, vol. 13, no. 22, 2021, doi: 10.3390/rs13224712.

[12] N. Awalia and A. Primajaya, “IDENTIFIKASI PENYAKIT LEAF MOLD DAUN TOMAT MENGGUNAKAN MODEL DENSENET121 BERBASIS TRANSFER LEARNING,” vol. 8, no. 1, 2022, [Online]. Available: http://ejournal.fikom-unasman.ac.id

[13] M. Krichen, “Convolutional Neural Networks: A Survey,” *Computers*, vol. 12, no. 8, Aug. 2023, doi: 10.3390/computers12080151.

[14] G. Imam Andaru and D. Hatta Fudholi, “Pengembangan Sistem Deteksi On-Shelf Availability Produk Menggunakan Algoritma YOLOV8 pada Aplikasi Beregerak,” 2024. [Online]. Available: https://journal.stmiki.ac.id

[15] E. N. Arrofiqoh and H. Harintaka, “IMPLEMENTASI METODE CONVOLUTIONAL NEURAL NETWORK UNTUK KLASIFIKASI TANAMAN PADA CITRA RESOLUSI TINGGI,” *GEOMATIKA*, vol. 24, no. 2, p. 61, Nov. 2018, doi: 10.24895/jig.2018.24-2.810.

[16] M. F. Naufal *et al.*, “Analisis Perbandingan Algoritma Klasifikasi Citra Chest X-ray Untuk Deteksi Covid-19,” *Teknika*, vol. 10, no. 2, pp. 96–103, Jun. 2021, doi: 10.34148/teknika.v10i2.331.

[17] L. Anggraini Susanti, A. M. Soleh, B. Sartono, I. Pertanian Bogor, and P. Korespondensi, “DEEP LEARNING IMAGE CLASSIFICATION RONTGEN DADA PADA KASUS COVID-19 MENGGUNAKAN ALGORITMA CONVOLUTIONAL NEURAL NETWORK DEEP LEARNING IMAGE CLASSIFICATION CHEST X-RAY IN COVID-19 CASES USING THE CONVOLUTIONAL NEURAL NETWORK ALGORITHM”, doi: 10.25126/jtiik.2023107142.

[18] M. Rahimzadeh and A. Attar, “A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of Xception and ResNet50V2,” *Inform Med Unlocked*, vol. 19, Jan. 2020, doi: 10.1016/j.imu.2020.100360.

[19] Jalu Nusantoro, Faldo Fajri Afrinanto, Wana Salam Labibah, Zamah Sari, and Yufis Azhar, “Detection of Covid-19 on X-Ray Image of Human Chest Using CNN and Transfer Learning,” *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 6, no. 3, pp. 430–441, Jun. 2022, doi: 10.29207/resti.v6i3.4118.

[20] D. Iskandar Mulyana, M. Ainur Rofik, and M. Ohan Zoharuddin Zakaria, “Klasifikasi Kendaraan pada Jalan Raya menggunakan Algoritma Convolutional Neural Network (CNN)”.

[21] J. A. Guzmán-Torres, L. A. Morales-Rosales, I. Algredo-Badillo, G. Tinoco-Guerrero, M. Lobato-Báez, and J. O. Melchor-Barriga, “Deep learning techniques for multi-class classification of asphalt damage based on hamburg-wheel tracking test results,” *Case Studies in Construction Materials*, vol. 19, Dec. 2023, doi: 10.1016/j.cscm.2023.e02378.

[22] Faiz Nashrullah, Suryo Adhi Wibowo, and Gelar Budiman, “The Investigation of Epoch Parameters in ResNet-50 Architecture for Pornographic Classification,” *Journal of Computer, Electronic, and Telecommunication*, vol. 1, no. 1, Jul. 2020, doi: 10.52435/complete.v1i1.51.

Vol. 6, No. 2, April 2025, Page. 483-500
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4094

[23] U. UNGKAWA and G. AL HAKIM, "Klasifikasi Warna pada Kematangan Buah Kopi Kuning menggunakan Metode CNN Inception V3," *ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika*, vol. 11, no. 3, p. 731, Jul. 2023, doi: 10.26760/elkomika.v11i3.731.

[24] R. Z. Fadillah, A. Irawan, M. Susanty, and I. Artikel, "Data Augmentasi Untuk Mengatasi Keterbatasan Data Pada Model Penerjemah Bahasa Isyarat Indonesia (BISINDO)," *JURNAL INFORMATIKA*, vol. 8, no. 2, 2021, [Online]. Available: http://ejournal.bsi.ac.id/ejurnal/index.php/ji

[25] K. Alomar, H. I. Aysel, and X. Cai, "Data Augmentation in Classification and Segmentation: A Survey and New Strategies," *J Imaging*, vol. 9, no. 2, Feb. 2023, doi: 10.3390/jimaging9020046.

[26] V. Werner de Vargas, J. A. Schneider Aranda, R. dos Santos Costa, P. R. da Silva Pereira, and J. L. Victória Barbosa, "Imbalanced data preprocessing techniques for machine learning: a systematic mapping study," *Knowl Inf Syst*, vol. 65, no. 1, pp. 31–57, Jan. 2023, doi: 10.1007/s10115-022-01772-8.

[27] T. Wongvorachan, S. He, and O. Bulut, "A Comparison of Undersampling, Oversampling, and SMOTE Methods for Dealing with Imbalanced Classification in Educational Data Mining," *Information (Switzerland)*, vol. 14, no. 1, Jan. 2023, doi: 10.3390/info14010054.

[28] A. Y. Triyanto and R. Kusumaningrum, "Implementasi Teknik Sampling untuk Mengatasi Imbalanced Data pada Penentuan Status Gizi Balita dengan Menggunakan Learning Vector Quantization Implementation of Sampling Techniques for Solving Imbalanced Data Problem in Determination of Toddler Nutritional Status using Learning Vector Quantization," vol. 19, pp. 39–50, 2017.

[29] S. Park, Y. Hong, B. Heo, S. Yun, and J. Y. Choi, "The Majority Can Help the Minority: Context-rich Minority Oversampling for Long-tailed Classification," 2022.

[30] S. Diantika, "PENERAPAN TEKNIK RANDOM OVERSAMPLING UNTUK MENGATASI IMBALANCE CLASS DALAM KLASIFIKASI WEBSITE PHISHING MENGGUNAKAN ALGORITMA LIGHTGBM," 2023.

Vol. 6, No. 2, April 2025, Page. 483-500
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.2.4094