JIMI Vol 1 No. 2 April 2024 P-ISSN : 3047-2113 E-ISSN : 3047-2121, Hal 19 - 29

DENASYA SMART PUBLISHER

JURNAL ILMIAH MULTIDISIPLIN ILMU

Halaman Jurnal: https://journal.smartpublisher.id/index.php/jimi
Halaman UTAMA Jurnal : https://journal.smartpublisher.id/

CNN INCEPTIONRESNET-V2 MACHINE LEARNING ARCHITECTURE FOR PNEUMONIA CHEST X-RAY CLUSTERING

Mohamad Firdaus [a*]
[a] mfirdausmumu@gmail.com, Universitas Indraprasta PGRI, Jakarta
[*]correspondence

ABSTRACT

In the last century, the use of machine learning, especially the Convolution Neural Network (CNN) has greatly helped the world of health (medicine). Through action research on image datasets, CNN succeeded and was able to show classification or grouping based on the same characteristics and properties on unlabeled images with higher accuracy and faster than other machine learning methods. This is very useful for the world of health, especially in the use of chest x-rays (chest x-rays) in the medical world. This study aims to optimize the CNN InceptionResNet-V2 architecture, for classifying Covid-19 disease, by training 4000 chest x-ray image datasets. The accuracy test results from InceptionResNet-V2 yielded 98%, with the precision of each CNN InceptionRestNet-V2 architecture class being Covid (99%), Lung_Opacity (97%), Normal (98%), Viral_Pneumonia (98%). The CNN InceptionRestNet architecture can help quickly and accurately produce chest x-rays.

Keywords: InceptionResNet, Convolution Neral Network, Chest X-Ray.

1. INTRODUCTION

The COVID-19 pandemic in Indonesia is part of the ongoing 2019 coronavirus disease pandemic worldwide. This disease is caused by a severe acute respiratory syndrome coronavirus 2. The first positive case of COVID-19 in Indonesia was detected on March 2, 2020 [1].

On June 23, 2021 the Government of Indonesia reported the number of confirmed cases of Covid-19 reaching 2,033,421 people [2]. This has an impact on the Hong Kong government's policy of banning all flights from Indonesia. The ban followed Indonesia's status which changed to become an A1 (extremely high risk) country [3].

Accurate diagnosis of COVID-19 can be done from an image chest x-ray, but requires medical personnel with special skills to be able to read the image data. Due to the limited number of these medical personnel, it makes it a long time (1-2 weeks) to get the diagnosis results.

Artificial intelligence technology or Artificial Intelligence (AI), specifically the Convolution Neural Network (CNN) method, is capable of predicting chest x-ray images of COVID-19 disease. This study aims to obtain the best results on the CNN InceptionRestNet architecture for detecting COVID-19 from chest x-ray images.

Pneumonia is one of the diseases that are often found in this world and to diagnose it requires certain skills. In this study (chest X-Ray pneumonia grouping), the authors used the Inception-ResNet-v2 deep learning architecture.

Received Maret 27, 2024; Revised April 2, 2024; Accepted April 25, 2024

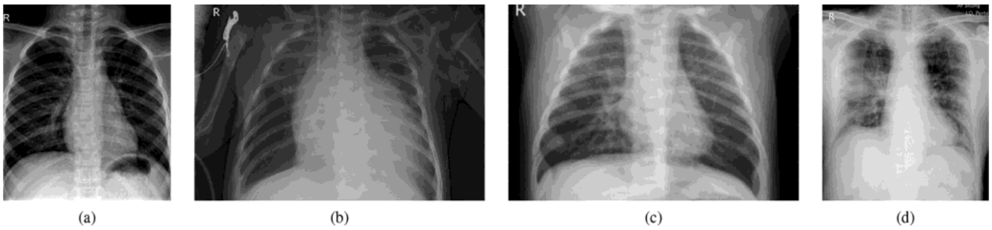


Fig. 1. Chest-x-ray samples showing a healthy patient (a), pneumonia from a bacteria infection (b), pneumonia from a virus infection (c), and pneumonia from a COVID-19 case (d).

For grouping or classify pneumonia and help doctors with their diagnosis, an effective solution is to analyze the chest X-ray image. This can be seen in figure 1, the spread of infection is clearly visible through x-ray technology, especially for severe conditions such as those associated with the SARS-CoV-2 virus. To fully exploit the information found in such images, deep learning approaches are being explored as they are already able to gain significant performance in several heterogeneous medical fields [18-20]. In detail, many solutions focus on binary classification between healthy people and COVID-19 patients [21-23].

In order to classify the ability to classify pneumonia from chest X-ray images, with special attention to SARS-CoV-2 infection, experiments were carried out on a specially designed data set to contain data from various public collections that tackle the task of pneumonia classification. Note that the pooled collections have heterogeneous chest X-ray dimensions because the data are from different study groups. This can be seen in Figure 3. For the first three classes, they were taken from the well-known Kaggle pneumonia, and divided into training sets, development, and test sets as appropriate [24].

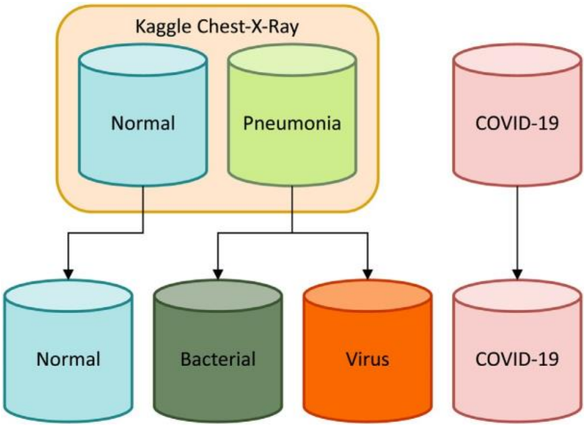


Figure 2. Dataset subdivision scheme.

This study aims to detect Viral Pneumonia from chest x-ray images with the InceptionResNet-V2 Architecture. Where the research results can be seen in table. 3 Precision Testing Results. Meanwhile, the input dataset is obtained from https://www.kaggle.com. We know that pneumonia is an acute respiratory infection disease caused by pathogenic bacteria or fungi and viruses. This infection affects one or both of our lungs and can cause mild to life-threatening effects of all ages. This infectious disease causes the world's largest number of deaths in children; where 15% of child deaths in 2017, data from the World Health Organization (WHO) statistics and among the possible sources are bacteria and viruses (infectious causes of pneumonia)[17].

2. RESEARCH METHODS

The current technological era almost all involves artificial intelligence (AI) and deep learning (DL) are the techniques most widely used in computational approaches in the field of Machine Learning (ML). Many applications are formed by involving ML [4], [5]. The most widely used type of DL in human life is the Convolutional Neural Network (CNN) [6]-[8],[9].

The deep layer compared to other ML makes this method fall into the DL category. The ability to be able to distinguish image objects, then this technique is better known as CNN (Figure 3). Several processes and terms in CNN will be explained briefly in this sub-section.

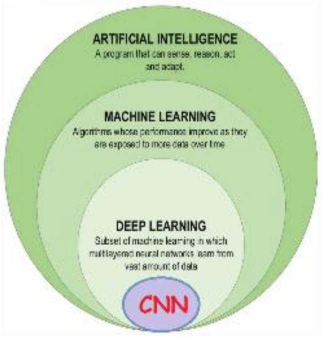
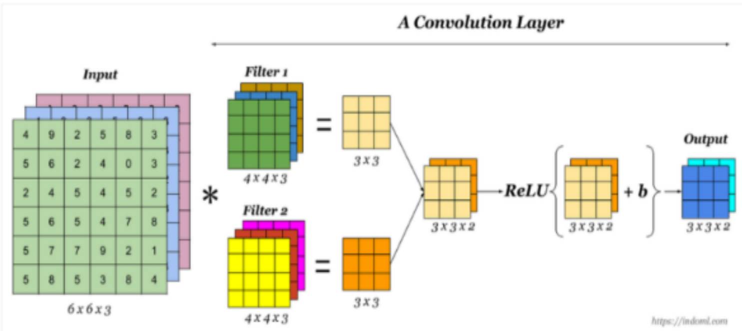


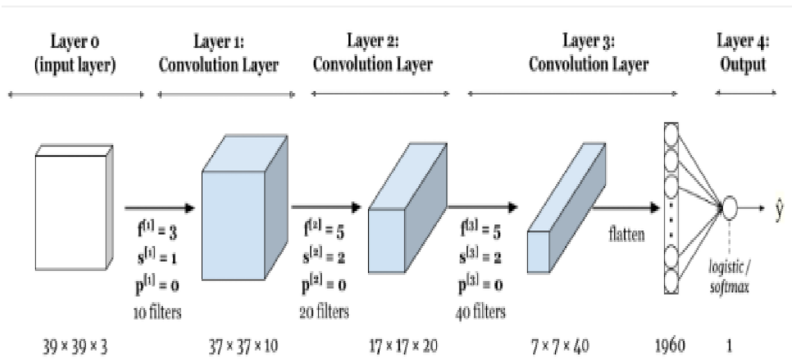
Figure 3. Classification of CNN in AI

2.1. Convolutional Neural Networks

CNN's initial thought was neural networks (neurons) in humans and animals [10]. CNN consists of a stack of convolutional layers with different parameters (filters, padding, stride and others). Each input is arranged in the form of a three-dimensional matrix, namely height (height), width (width) and depth (depth), generally the height value is the same as the width value. For the RGB matrix input, the channel value is 3, the channel represents the depth value in the matrix. The end of the CNN layer series produces a kernel model, then proceeds with the fully connected process. Fully Connected aims to classify images from a particular label. (Figure 4).



cnn_model = models.Sequential()
cnn_model.add(layers.Conv2D(32,(3,3), activation='relu', input_shape=(6,6,3))



cnn_model = models.Sequential()
cnn_model.add(layers.Conv2D(32,(3,3), activation='relu', input_shape=(39,39,3)))
cnn_model.add(layers.Conv2D(64,(5,5), activation='relu'))
cnn_model.add(layers.Conv2D(64,(5,5), activation='relu'))
cnn_model.add(layers.Flatten()).

Figure 4. The convolution process (top) which is part of the overall CNN layer (bottom) [11].

2.1.1. Convolution process

The convolution process is the calculation of the dot product between the input matrix and the filter matrix. The filter matrix is a rectangular matrix (height=width) which is a collection of weight values, sometimes also referred to as a kernel matrix. The height and width sizes are always smaller than the input, but the depth (channel) value is the same as the input depth. The convolution process generally uses more than one type of kernel matrix (filter 1, filter 2, and so on). The size of the output matrix for rows and columns can be used with the conditions O = I – F+1. (Figure 5).

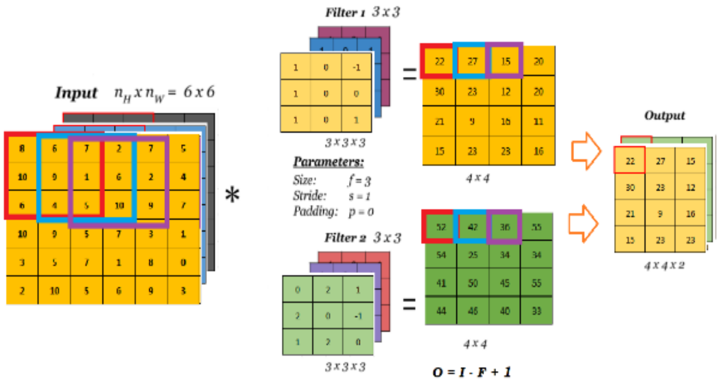


Figure 5. Details of the convolution process.

2.1.2. Stride process

The stride process is actually part of the convolution process itself. The filter matrix will scan the input matrix by convolution of the two matrices according to the size of the filter matrix. After the first process is complete, proceed with the second convolution process by moving the input matrix (second input matrix) which will be convolved with the filter matrix. The shift of the first and second input matrices corresponds to the stride value specified as a parameter in each layer. This process is repeated for the next convolution, resulting in a matrix with a smaller size. (Figure 6).

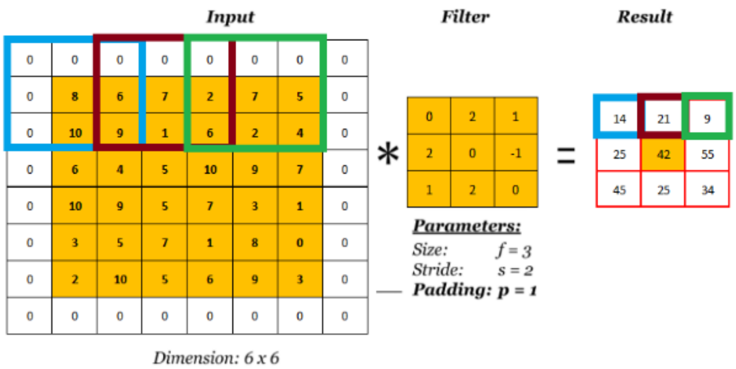


Figure 6. Stride and padding

2.1.3. padding process

Padding is the process of adding the same large value on each side of the input image (usually a value of 0 or one). Even though the input matrix is larger than before, in the convolution process it will still be smaller than the original input matrix. The purpose of padding is to provide a clearer boundary of an image, so that the process can be more focused. Padding is sometimes also used to adjust the size of the output matrix by adding values on each side of the image so that it matches the input matrix that will be used in the next convulsive layer. (Figure 6).

2.1.4. Pooling

Actually the pooling process is identical to data sampling in statistics, the difference is that in pooling there is no lost information. Pooling aims to reduce the size of the matrix so that the calculations performed by the machine become faster. There are many types of pooling, but the most commonly used are max pooling and average pooling. Max pooling is taking the largest value in a certain matrix size, while average pooling is taking the average value in a certain matrix size. (Figure 6a).

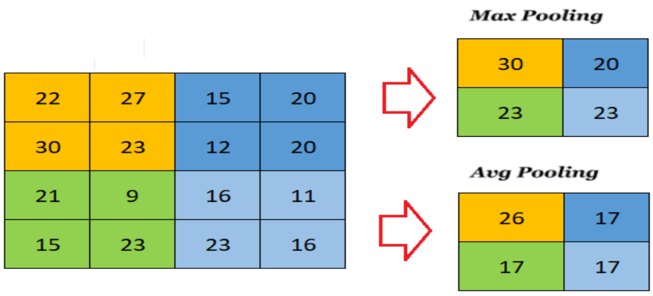


Figure 6a. Pooling process.

2.1.5. Activation Function

The activation function is given after the pooling process, the goal is that the resulting value is within a certain range of values (usually ranging from 0 to 1) and also functions to make decisions to take appropriate results and eliminate unwanted (bad) results, map input to output. non-linear, providing more complex learning abilities. The activation function also has the ability to distinguish very important features [8].

Activation function	Equation	Example	1D Graph
Unit step (Heaviside)	$\phi(z) = \begin{cases} 0, & z < 0, \\ 0.5, & z = 0, \\ 1, & z > 0, \end{cases}$	Perceptron variant	
Sign (Signum)	$\phi(z) = \begin{cases} -1, & z < 0, \\ 0, & z = 0, \\ 1, & z > 0, \end{cases}$	Perceptron variant	
Linear	$\phi(z) = z$	Adaline, linear regression	
Piece-wise linear	$\phi(z) = \begin{cases} 1, & z \geq \frac{1}{2}, \\ z + \frac{1}{2}, & -\frac{1}{2} < z < \frac{1}{2}, \\ 0, & z \leq -\frac{1}{2}, \end{cases}$	Support vector machine	
Logistic (sigmoid)	$\phi(z) = \frac{1}{1 + e^{-z}}$	Logistic regression, Multi-layer NN	
Hyperbolic tangent	$\phi(z) = \frac{e^{z} - e^{-z}}{e^{z} + e^{-z}}$	Multi-layer Neural Networks	
Rectifier, ReLU (Rectified Linear Unit)	$\phi(z) = max(0, z)$	Multi-layer Neural Networks	
Rectifier, softplus	$\phi(z) = \ln(1 + e^{z})$	Multi-layer Neural Networks	

Figure 7. Types of activation functions [13]

2.1.6. Fully Connected Layer (FC)

This process is at the end of each CNN layer, called fully connected because all neurons (nodes) are connected to neurons in the layers before and after. FC functions as a classification, technically this process is none other than the conventional neural network layer method (feed forward neural network) [10], for details can be seen in Figure 8.

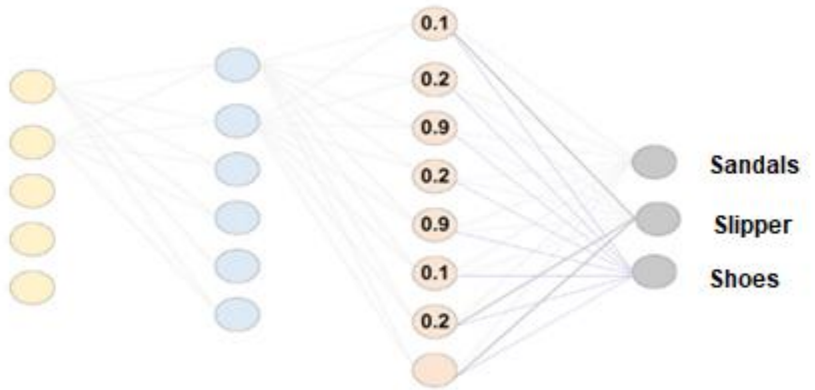


Figure 8. Fully connected (FC) network

2.2. CNN Architecture

From the competition conducted by the ImageNet Large Scale Visual Recognition Challenge (ILSVRC), there are many CNN architectures that have been created as an effort to improve accuracy by testing certain input datasets that are used as standards (MNIST [7], imagenet, coconut etc.) (Figure 8). However, many architectures prefer to add layers to great depths rather than devise a structural design with a new method. This results in the need for high computer performance to be able to run the architecture [12].

In this study, we will try the CNN architecture that can be done by computers, namely the CNN InceptionRestNet architecture, with a few adjustments to get the expected results. This CNN architecture will carry out the training and testing process using a chest x-ray image data set to detect the presence of the type of covid-19 disease. The image data set for this study was obtained or obtained from https://www.kaggle.com/code/ernestomaisuls/covid-19-x-ray-modified-xception/data.

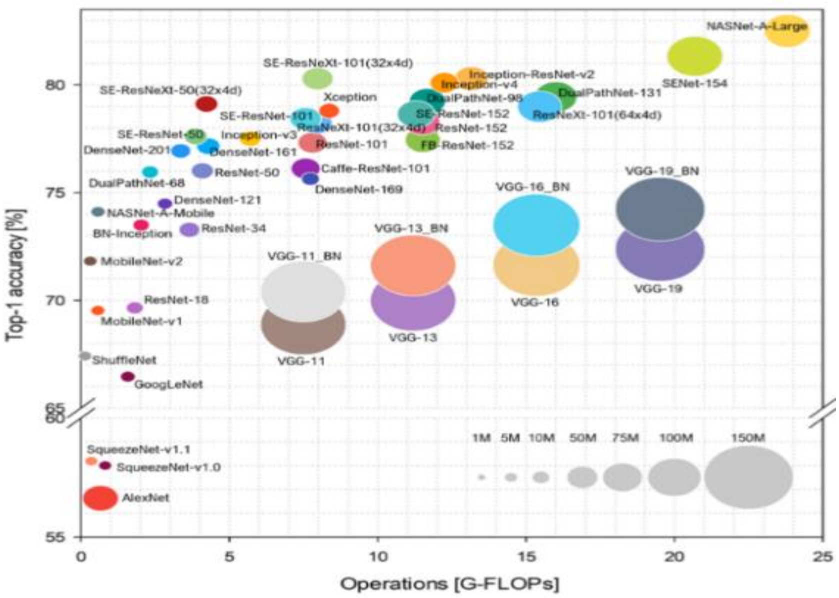


Figure 9. Comparison of accuracy of CNN architectures based on ILSVRC [12]

2.2.1. InceptionResNet-V2

As an effort to develop CNN, Szegedy (2014) introduced Inception in his paper entitled "Going Deeper with Convolution" [14]. Where Convolution is an image extraction action to get a model in the form of a kernel matrix. In this process, filtering is carried out which shifts with a certain "stride value" in an input image. Or a process to see the value of a parameter that determines how much the number of filter shifts is in an input image. Furthermore, the results of the convolution become the input for the fully connected part for the classification process [15]. This aims to get a more reliable network and increase the level of classification accuracy produced by using this InceptionRestNet architecture.

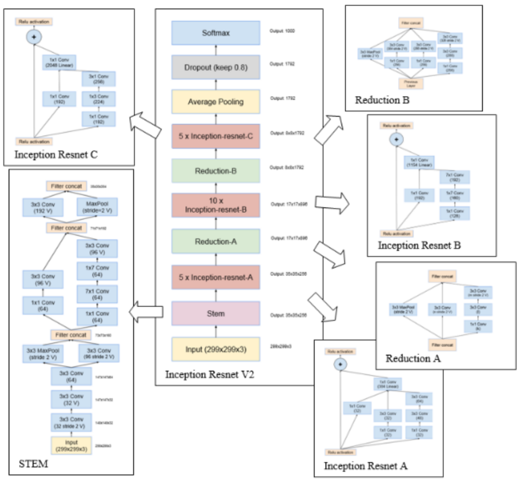


Figure 10. Inception Resnet-V2 module block

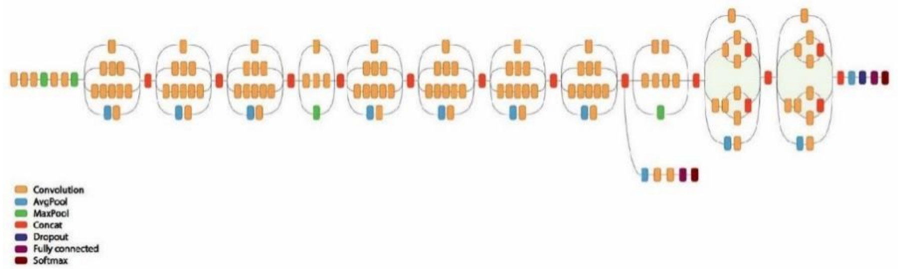


Figure 11. Overall diagram of Inception Resnet V2 [16]

2.3. Data preparation and research stage

This study uses an input dataset of 4000 images, divided into 4 classes, namely:

0 – Normal 1,000 images (25%)
1 – Viral Pneumonia 1,000 images (25%)
2 – Lung_Opacity 1,000 images (25%)
3 – Covid 1,000 images (25%)

Of the 1000 image data for each class, 80% (800 images) are divided for training data and 20% (200 images) for data validation. (Figure 12).

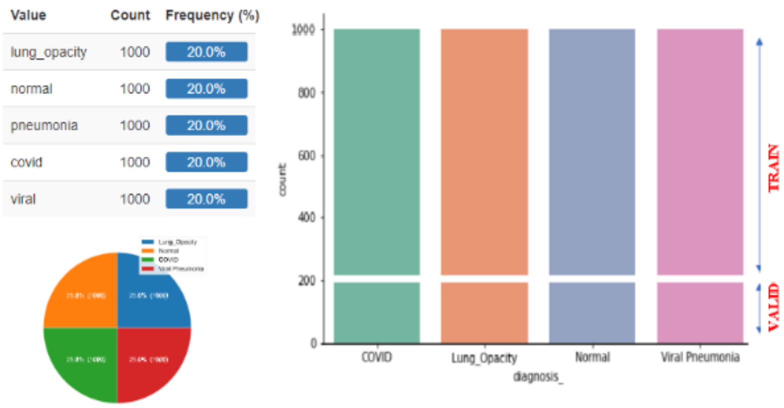


Figure 12. Pie and histogram dataset

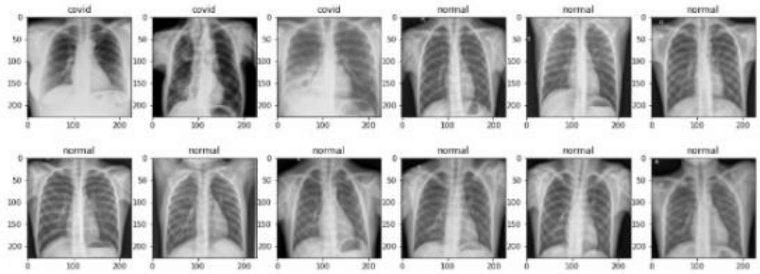


Figure 13. Chest X-Ray Image Dataset

The input dataset (train and validation data) is used as training data for the CNN InceptionRestNet architecture, the results of the training are in the form of a model that can be used to classify images. To obtain a large measurement of accuracy, testing is carried out using data testing as many as 4000 images (1000 images per class). (Figure 14)

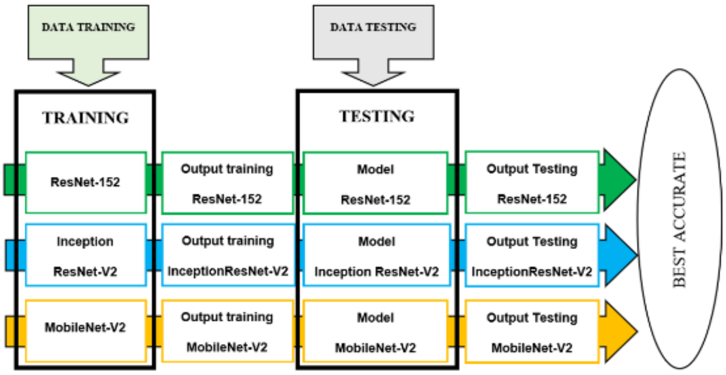
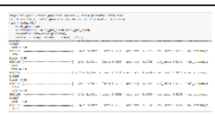
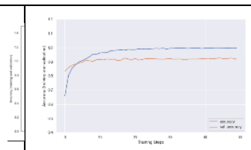
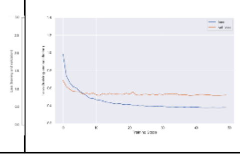


Figure 14. Research stage

3. RESULT AND DISCUSSION

The results of the training process with input data obtained values of train accuracy, train loss, val accuracy and val loss from the InceptionRestNet architecture, as follows: (Table 1)

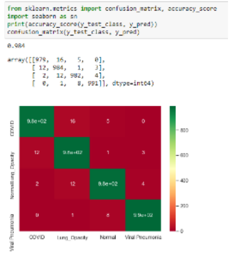
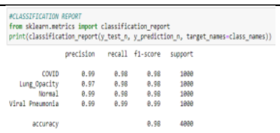
Table 1. Training Results

		InceptionResNet
TRAIN	Loss	0.3769
	Acc	0.9958
VALID	Loss	0.52
	Acc	0.9175
History	Train Acc, Val Acc, Train Loss, Val Loss	
Grafik	Train Acc Valid Acc	
	Train Loss Valid Loss	

From table 1 it can be seen that the percentage for train and valid accurate on the CNN InceptionResNet-V2 architecture is very good (> 90%), but for valid loss it still looks quite large (> 40%) and needs to be increased so that the value becomes smaller. As for the train loss is <40%.

The result of the training process is a model that can be used for unlabeled image classification. To find out more complete measurement results, a test was carried out using the same amount of test data as the input data, namely 4000 images (1000 images per class). The testing program that is run produces output in the form of a confusion matrix and a classification report. The results of this testing output are collected in the following table: (Table 2)

Table 2. Testing Results

<table>
<tr><th></th><th>InceptionRestNet</th></tr>
<tr><td>Confussion Matric</td><td>
</td></tr>
<tr><td>Classification Report</td><td>
</td></tr>
</table>

In order to be more detailed and clearly visible, the precision classification report values from table 2 are summarized in a table which can be seen in table 3 and for graphs can be seen in figure 18 below:

Table 3. Precision Testing Results

		InceptionResNet
PRECISSION	Covid	0.98
	Lung Opacity	0.97
	Normal	0.99
	Viral Pneumonia	0.98
	Accurate	**0.98**

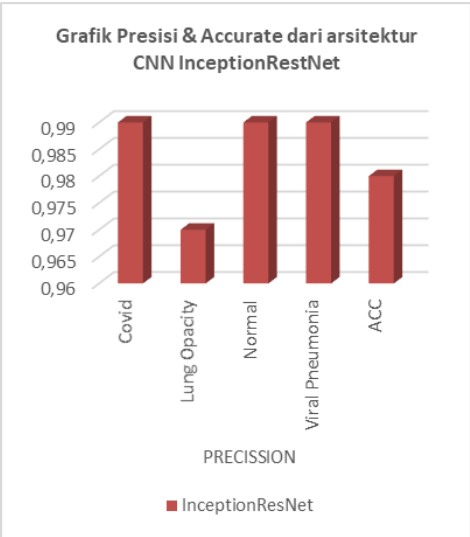


Figure 15. Precision & Accurate Graph of the CNN InceptionRestNet architecture

4. KESIMPULAN DAN SARAN

The results of this research optimization are:

a. The results of research to predict the type of covid disease from Chest X Ray image data using the CNN InceptionRestNet architecture are very good with an accuracy value of more than 95%. And for the results of the CNN InceptionResNet-V2 architecture, the accuracy is 98% and the precision for each class is more than 95%, namely: Covid (99%), Lung_Opacity (97%), Normal (99%), Viral_Pneumonia (99%).
b. During a time when the COVID-19 pandemic is getting worse and there are new variants with a higher transmission rate, the application of artificial intelligence (AI) technology in the world of medicine (medical) is urgently needed to help ease the work of doctors.
c. To detect Viral Pneumonia from chest x-ray images on CT Scan equipment available in almost all hospitals in every city in Indonesia, it would be better analyzed using the CNN InceptionRestNet-V2 architecture.
d. The results of the model from this study can be used (embedding) in the form of a mobile internet or in the form of a web-based system, so that at any time and anywhere the patient is able to scan and get predictive results in real time at low cost, accurate and fast.

REFERENCES

[1] Yandri Daniel Damaledo. "2 Maret 2020 Kasus Corona Pertama di Indonesia Diumumkan oleh Presiden Jokowi dan Menkes Terawan. ", https://tirto.id/gaKw. 2 Maret 2021. 2021
[2] WHO Indonesia, “Coronavirus Disease 2019 (COVID-19),” 2021.
[3] Indra Bakti, Mohamad Firdaus, “Arsitektur Convolutional Neural Network InceptionResNet-V2 Untuk Pengelompokan Pneumonia Chest X-Ray”, JUKOMTEK, 2023
[4] Indra Bakti, Mohamad Firdaus, “Classification of Image Files of Lung X-Ray Results with Architecture Convolution Neural Network (CNN)”, Journal of Information Technology, 2023
[5] Mohamad Firdaus, “Analisis Pemilihan Enterprise Resource Planning (ERP) Menggunakan Analytical Hierarchy Process (AHP) Studi Kasus Pada PT. Laba Usaha Mandiri”, Jurnal Teknologi Informasi (JUTECH), 2023.
[6] L. M. R. Rere, M. I. Fanany, and A. M. Arymurthy, “Simulated Annealing Algorithm for Deep Learning,” Procedia Comput. Sci., vol. 72, pp. 137–144, 2015.
[7] L. M. Rasdi Rere, M. I. Fanany, and A. M. Arymurthy, “Metaheuristic Algorithms for Convolution Neural Network,” Comput. Intell. Neurosci., vol. 2016, 2016.
[8] V. Ayumi, L. M. R. Rere, M. I. Fanany, and A. M. Arymurthy, “Optimization of convolutional neural network using microcanonical annealing algorithm,” 2016 Int. Conf. Adv. Comput. Sci. Inf. Syst. ICACSIS 2016, pp. 506–511, 2017.
[9] L. M. R. Rere, R. Dalam, and K. Baru, “Studi Pengenalan Ekspresi Wajah Berbasis Convolutional Neural Network,” vol. 3, 2019.
[10] L. Alzubaidi et al., Review of deep learning: concepts, CNN architectures, challenges, applications, future directions, vol. 8, no. 1. Springer International Publishing, 2021.
[11] B. Prijono, “Student Notes: Convolutional Neural Networks (CNN) Introduction,” Indoml.com. [Online]. Available: https://indoml.com/2018/03/07/student-notes-convolutional-neural-networks-cnn-introduction/.
[12] S. Bianco, R. Cadene, L. Celona, and P. Napoletano, “Benchmark analysis of representative deep neural network architectures,” IEEE Access, vol. 6, pp. 64270–64277, 2018.
[13] Sebastian Raschka, “Activation Functions for Artificial Neural Networks,” github.io, 2020. [Online]. Available: http://rasbt.github.io/mlxtend/user_guide/general_concepts/activation-functions/.
[14] G. Zeng, Y. He, Z. Yu, X. Yang, R. Yang, and L. Zhang, “Going Deeper with Convolutions,” J. Chem. Technol. Biotechnol., vol. 91, no. 8, pp. 2322–2330, 2016.
[15] M. Längkvist, L. Karlsson, and A. Loutfi, “Inception-v4, InceptionResNet and the Impact of Residual Connections on Learning,” Pattern Recognit. Lett., vol. 42, no. 1, pp. 11–24, 2014.

[16] N. Van Hieu and N. L. H. Hien, "Recognition of plant species using deep convolutional feature extraction," Int. J. Emerg. Technology, vol.,no. 3, pp. 904 – 910, 2020.

[17] S.B. Smith, G.W. Ruhnke, C.H. Weiss, G.W. Waterer, R.G. Wunderink, Trends in pathogens among patients hospitalized for pneumonia from 1993 to 2011, JAMA Intern. Med. 174 (11) (2014) 1837–1839, doi:10.1001/jamainternmed. 2014.4344.

[18] W. William, A. Ware, A.H. Basaza-Ejiri, J. Obungoloch, A review of image analysis and machine learning techniques for automated cervical cancer screening from papsmear images, Comput. Methods Programs Biomed. 164 (2018) 15– 22, doi:10.1016/j.cmpb.2018.05.034.

[19] R. Karthik, R. Menaka, A. Johnson, S. Anand, Neuroimaging and deep learn- ing for brain stroke detection-a review of recent advancements and future prospects, Comput. Methods Programs Biomed. 187 (2020) 105728, doi:10. 1016/j.cmpb.2020.105728.

[20] D. Avola, L. Cinque, A. Fagioli, G. Foresti, A. Mecca, Ultrasound medical imaging techniques: a survey, ACM Comput. Surv. (CSUR) 54 (3) (2021) 1–38, doi:10. 1145/3447243.

[21] S. Minaee, R. Kafieh, M. Sonka, S. Yazdani, G.J. Soufi, Deep-COVID: predicting COVID-19 from chest x-ray images using deep transfer learning, Med. Image Anal. 65 (2020) 101794, doi:10.1016/j.media.2020.101794.

[22] L. Brunese, F. Mercaldo, A. Reginelli, A. Santone, Explainable deep learning for pulmonary disease and coronavirus COVID-19 detection from x-rays, Com- put. Methods Programs Biomed. 196 (2020) 105608, doi:10.1016/j.cmpb.2020. 105608.

[23] W.M. Shaban, A.H. Rabie, A.I. Saleh, M. Abo Elsoud, Detecting COVID-19 patients based on fuzzy inference engine and deep neural network, Appl. Soft Comput. 99 (2021) 106906, doi:10.1016/j.asoc.2020.106906.

[24] D.S. Kermany, M. Goldbaum, W. Cai, C.C. Valentim, H. Liang, S.L. Baxter, A. McKeown, G. Yang, X. Wu, F. Yan, et al., Identifying medical diagnoses and treatable diseases by image-based deep learning, Cell 172 (5) (2018) 1122–1131, doi:10.1016/j.cell.2018.02.010.