

## ***PNEUMONIA PREDICTION USING CONVOLUTIONAL NEURAL NETWORK***

**Vincentius Praskatama<sup>\*1</sup>, Christy Atika Sari<sup>2</sup>, Eko Hari Rachmawanto<sup>\*3</sup>, Noorayisahbe Mohd Yaacob<sup>4</sup>**

<sup>1,2,3</sup>Informatics Engineering, Faculty of Computer Science, University of Dian Nuswantoro, Indonesia

<sup>4</sup>Malaysia-Japan International Institute of Technology (MJIT), University of Technology Malaysia (UTM),  
Kuala Lumpur, Malaysia

Email: <sup>1</sup>[111202113456@mhs.dinus.ac.id](mailto:111202113456@mhs.dinus.ac.id), <sup>2</sup>[christy.atika.sari@dsn.dinus.ac.id](mailto:christy.atika.sari@dsn.dinus.ac.id), <sup>3</sup>[eko.hari@dsn.dinus.ac.id](mailto:eko.hari@dsn.dinus.ac.id),  
<sup>4</sup>[noorayisahbe@utm.my](mailto:noorayisahbe@utm.my)

(Article received: August 28, 2023; Revision: September 15, 2023; published: October 22, 2023)

### ***Abstract***

*Pneumonia is condition which our lungs become inflamed due to infection from viruses, bacteria, or fungi. Pneumonia can affect anyone, both adults and children. Because of this, prevention of pneumonia is important. Prevention can be done by the process of maintain our immunity and lungs. In this study, had been done classify pneumonia based on X-ray images. This study using X-ray images dataset with total data is 5840 images in .jpg extensions. With a total number of images from training data is 5216 images and number of images from the test data is 624 images. The dataset that used in this research has 2 main classes, namely class normal and pneumonia. Normal class indicates that the X-Ray results are not detected with pneumonia. While the pneumonia class indicates that the processed X-Ray results are diagnose affected by pneumonia. The purpose of this research is building model that can be used to classify pneumonia based on X-Ray images. The classification process carried out in this study uses the Convolutional Neural Network method. The purpose of using the CNN method in the classification process of this research is because, in the process, CNN can extract features automatically and independently, so that the data provided does not need to be preprocessing first, but the data still produces good extraction features and can provide accurate classification results. The results from the testing process is carried out to run or perform in the pneumonia classification process, the CNN model built obtained a classification test accuracy of 87.82051205635071%.*

**Keywords:** CNN, Deep Learning, Neural Network, Pneumonia, X-Ray.

## **1. INTRODUCTION**

Pneumonia is the infection that occurs in the human lungs [1]. Pneumonia usually occurs caused by the bacterial infection in the lungs [2]. For details, Pneumonia usually caused because the alveoli part in the lungs are infected by the fungi, bacteria or viruses [3]. It is estimated that around 1.4 million of young children die each year because of pneumonia, which 18% of children who die by the pneumonia infections are children that less than five years old [4]. Data from 2017 showed that 808,694 people who died due to this disease [5]. Shah et. al [6], argued that pneumonia is classified as a threatening disease by the WHO, which now has more than one million premature deaths caused by pneumonia. Pneumonia is also very life-threatening especially for people who have a compromised immune system, people who have respiratory diseases and people who smoke [7]. This can happen because people who have these diseases or habits, their lungs can be easily attacked by diseases, especially pneumonia. Jain et. al [8], suggested that pneumonia is the leading cause of child mortality in South Asia and Sub-Saharan Africa. Pneumonia is a dangerous disease, but can be prevented by administering drugs to maintain immunity [9]. Although it can be treated with

medication, the process of treating pneumonia as early as possible when it is diagnosed also needs to be done [10], considering how dangerous this disease is. Therefore, action from the authorities, especially in the health sector, is needed [11], to prevent and treat pneumonia.

Image classification is a process to predict data based on existing labels [12]. This classification process is carried out using different approaches such as using an algorithm to search for patterns from data, the process of using more complex and numerous features and so on [13]. Image classification is part of digital image processing. Image processing is part of artificial intelligence which performs the process of image improvement and analysis [14]. In the medical field, image processing can be used to analyze and diagnose patients according to their illness [15].

Deep learning is the part of Artificial Intelligence (AI) process or method that is now be the main thing in industry 4.0 (Internet of Things) [16]. This because in the process, deep learning has the better performance for processing big data when compared to the traditional machine learning methods [17]. Haque et. al [18], argued that when processed the classification using deep learning data does not need to be preprocessing, segmentation and feature

extraction first so that means the process of deep learning can be done directly using just raw data. This is can be done because in deep learning section, a process is carried out to build an artificial neural network that exactly similar to the human brain to learn and analyze the data [19]. Janiesch et al [20], explained that the implementation of deep learning method is very useful for used in processing high-dimensional data such as processing text, image, audio, speech and video data. With this existing advantages, deep learning can provide better performance for regression and classification tasks [21], especially in segmentation of images [22]. Ozbayoglu et. al [23], suggested that deep learning contains many ANN (Artificial Neural Network) layers, so the process of deep learning, produces high-levels data abstraction modeling. There are many layers that used when perform the deep learning process that is because in the process, deep iteration is carried out so that the model can perform a deeper analysis for the data [24]. Therefore, it is not surprising that deep learning is now be the common technique used for the image classification process or task [25].

CNN or Convolutional Neural Network is the method used to perform deep learning process in image processing area. Desai et al [26], suggested that CNN is the form of ANN or artificial neural network method that widely using to examine, identify or classify images. In CNN architecture, neurons are implemented in the form of layers, where the layer consists of the input layer, the hidden layer and the output layer [27]. In its implementation, CNN can be used to classify or segment images [28]. CNN in concept works like an ordinary artificial neural network, but has the characteristic of having neurons that contain weights and biases that can be learned [29]. These weights and biases are used for neurons to connect with other layers [30]. Because of this, CNN has the advantage of being able to automatically perform detection or classification without having to be instructed by humans first [31]. This can be done because in the process, CNN performs feature extraction automatically using the convolution layer that is built [32].

This study that conducted will used Convolutional Neural Network (CNN) to classify and predict pneumonia that will be based on X-Ray result images. Purpose from conducting this study that is to build the model where that model can be classify pneumonia and also aims to see and make predictions from the X-Ray results of patients who are given diagnosed by pneumonia or not. So that will be hoped that earlier action can be taken when it is diagnosed

as having pneumonia. The purpose from using CNN in this research is because CNN method works with an artificial neural network (ANN) system, so that in the implementation process, can be well done as humans learn something. CNN is also used because the advantage from the method so can being able to perform automatic feature extraction with layers that we built so that means can easily used for classification and prediction on X-Ray result images.

In a study that conducted by Mujahid et. al [33], discussed about the classification of pneumonia using X-Ray result images and pretrained CNN models, namely InceptionV3, VGG and ResNet50. The purpose from this research that is building model that can perform classification based on X-Ray images and compare which method is better and more efficient to be able to perform the image classification process. The results obtained after research and testing using the pretrained VGG model get an accuracy of 98.06%, with InceptionV3 getting an accuracy of 99.29% and using ResNet50 getting an accuracy 99.8%.

Research that had been conducted by Cakraborty et. al [34] in 2021 discusses about classification process for Covid-19 and pneumonia based on X-Ray result images using the CNN transfer learning process. The purpose from conducting this research is building model that can classify pneumonia and covid-19 using X-Ray result images and to see the performance of VGG16 to classify X-Ray result images. Results that obtained from doing this study is by using VGG16 transfer learning, getting a test accuracy of 97.11%.

## **2. RESEARCH METHODS**

### **2.1. Workflow for classification process**

In the classification process, this research will use Jupyter notebook IDE and python programming language to implement the system. First, data reading and deep learning modeling using CNN are performed. The classification process that will be perform in this research is divided into 3 stages that is training stages, validation stages and testing stages. The data used is training data which is divided again by 20% for model validation and testing data. The purpose of taking 20% as validation data is so that the model built does not occur overfitting. In order to be able to correctly identify patterns from the data, training and testing methodologies shall be applied so as to demonstrate that a model is reliable in predicting new data. Figure 1 gives a view of the classification process.

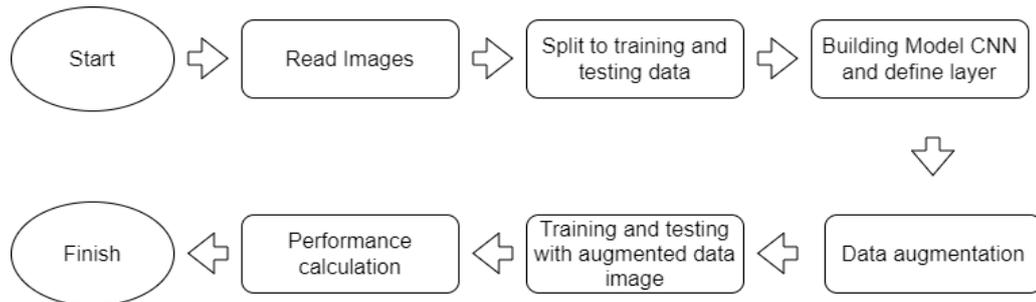


Figure 1. Workflow for classification process

Figure 4 shows the stages of the process taken to be able to perform the pneumonia classification process based on X-Ray images. An explanation of the process is given below.

1. The first thing to do to perform the image classification process is to read the image data that will be used either to carry out the training process or to carry out the testing process on the testing model on the model.
2. After the data reading process has been carried out, the performing split the data into training data and testing data. For the train data, it is further divided into 20% validation data so that later validation can be carried out so that the model built does not experience overfitting.
3. Then after dividing the data, the process of building a model with the layers to be used will be carried out. For the model to be used, namely the Convolutional Neural Network model. In this study, layers that had been built are layers ranging from convolution to fully-connected layers to perform the pneumonia classification process based on X-Ray images. The layers used and their parameters are given in Figure 2.
4. After model building and data reading, the data augmentation process is carried out. This process aims to perform further processing on the image so that the processed image can be the same and there is no overfitting on the model [33]. In this study, the data augmentation performed is image rotation by 30, image length shifting by 0.2, image width shifting by 0.2, shear by 0.2, zoom in image by 0.2. Then the augmentation performed on the image is a horizontal flip process and uses the fill mode, namely nearest. By doing this augmentation, it is hoped that the images used in both the training and testing processes can be equal in value so that there is no gap and cause overfitting when the model performs the training and testing process.
5. Then after data augmentation, the model is trained to recognize the pattern. While the model is training, the validation process will also be carried out so that the model built does not occur overfitting. After the model is completed for training and validation, a testing process can be carried out on the model to be

able to see how the robustness or resilience of the model in performing classification.

6. After the model has completed the testing process, the model performance can be calculated when perform the process of classification. In this process, the performance from the model after testing will be calculated using the confusion matrix and also the classification report so that it can be seen how effective and good the model is in classifying data.

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 150, 150, 32)	320
max_pooling2d (MaxPooling2D)	(None, 75, 75, 32)	0
conv2d_1 (Conv2D)	(None, 75, 75, 64)	18496
max_pooling2d_1 (MaxPooling2D)	(None, 37, 37, 64)	0
conv2d_2 (Conv2D)	(None, 37, 37, 96)	55392
max_pooling2d_2 (MaxPooling2D)	(None, 18, 18, 96)	0
conv2d_3 (Conv2D)	(None, 18, 18, 96)	83040
max_pooling2d_3 (MaxPooling2D)	(None, 9, 9, 96)	0
flatten (Flatten)	(None, 7776)	0
dense (Dense)	(None, 512)	3981824
dropout (Dropout)	(None, 512)	0
dense_1 (Dense)	(None, 2)	1026

-----  
 Total params: 4140098 (15.79 MB)  
 Trainable params: 4140098 (15.79 MB)  
 Non-trainable params: 0 (0.00 Byte)

Figure 2. Layer for CNN

## 2.2. Dataset

Pneumonia is a disease that occurs usually because it is caused by a bacterial infection in the lungs [2]. Pneumonia is caused because the alveoli in the lungs are infected by fungi, bacteria or viruses [3]. Pneumonia is classified as a threatening disease by the WHO which now has more than one million premature deaths caused by pneumonia [6]. Purpose of doing this research is to perform a classification process based on the patient's X-Ray results whether the patient is infected with pneumonia or not. So that after the diagnosis stage, further medical action and treatment can be given as early as possible so that the disease does not claim lives.



Figure 3. X-Ray normal lungs



Figure 4. X-Ray pneumonia lungs

Figures 3 and 4 show image that contains X-Ray results from patients with pneumonia (Figure 4) and patients without pneumonia (Figure 3). This research will use the X-Ray result image dataset. That the dataset is obtained from the kaggle.com website. Dataset that used divided into two main classes, namely class normal and pneumonia. The total images used in the dataset in this study were 5840 images. With a distribution of 1575 data in class of normal and 4265 data in class of pneumonia. Of the total 5840 data, it is further divided into 5216 training data with a distribution of 1341 image data in the normal class and 3875 image data in the pneumonia class. And testing data as much as 624 image data with a distribution of 234 images that include in class normal and 390 images that include in class of pneumonia. The purpose of dividing the data is to later carry out the model training and model testing processes. In the testing data, it will be divided again by 20% which aims to carry out the validation process on the model. So the model that build does not occur overfitting. When the data are split, a training process will then be performed so that model that build can find patterns from the image data used. So that after the training process and understanding the pattern from the data, the model can perform a new data prediction process using test data.

### 2.3. Convolutional Neural Network (CNN)

The method used to perform the deep learning in image processing is CNN or Convolutional Neural Network. The neurons are created in the CNN architecture as the layers, consisting of an input layer, a hidden layer and an output layer [27]. In the concept of CNN that is works like an ordinary ANN or artificial neural network architecture, but had the characteristic of having neurons that contain weights and biases that can be learned [29]. These weights and biases are used for neurons to connect with other layers [30]. Because of this, CNN has the advantage of being able to automatically perform detection or classification without having to be instructed by humans first [31]. In the process, there are layers that are usually used in the process of building CNN models. The layer is a convolutional layer that is used to perform feature extraction from given data. For the formula of the convolution layer mathematically seen in point 1. Where: P is the input, O is the kernel that

used, and z and j are the indices of P and O, respectively.

$$(P * O)(z, j) = \sum_m \sum_n P(z + m, j + n) * O(z, j) \quad (1)$$

Then there is a pool layer that used for reduce dimensions of the features so that the computation process in the next layer can be faster. In this study, the pooling layer used will use the MaxPooling process which the process from this MaxPooling layer is taking the maximum of value from the features that used. For the calculation of the MaxPooling layer seen in point 2.

$$MaxPool(X, Y)(i, j) = Max_{o,p} X_{(i*s+o), (j*s+p)} \quad (2)$$

Where: X is the input, Y is the kernel used, i and j are matrices for the index of the MaxPooling process results. Whereas o and p are the position of the kernel (Y) and S is how much the kernel is shifted according to the input.

Furthermore, there is a fully-connected layer that had been used for connect results that obtained from the process in the convolution layer and the pooling layer so that the classification process can be carried out later. In each layer, an activation function is used to provide a non-linearity process in the model. When build model, we also need the activation function. Activation function that used in the convolutional layer and pooling layer is RELU (Rectified Linear Activation) which can be seen mathematically in point 3 and fully connected layer use activation function Softmax to perform the multiclass classification process which seen mathematically in point 4.

$$RELU(i) = \max(0, i) \quad (3)$$

$$Softmax(X_m) = \frac{e^{X_m}}{\sum_{n=1}^k e^{X_n}} \quad (4)$$

With: X(m) is the value for class m, while k is total classes and then e is euler number = 2.71828.

### 2.4. Confusion matrix

A Confusion Matrix is the measurement metric used to determine how well this constructed model performs. In this research, the confusion matrix is calculated after testing the model. The goal is to find out how the performance of the model that has been trained when testing. The values that include in the confusion matrix are true positive, true negative, false positive and false negative. These values also can be used to create a report from the classification results when using the model. The values in the report are recall, f1-score, precision and support. Recall is the value that used for see the positive prediction results compared to the actual positive value, so that the

model's performance is obtained to be able to carry out the positive class prediction process correctly. Recall calculation is given in point 5. Precision is a value used to see the performance of the model in carrying out the positive prediction process or can be referred to as the accuracy of the model in making predictions. To see the balance between Recall and Precision values, the f1 score is used. For the calculation of the f1 score mathematically can be seen in point 7.

$$Rec = \frac{True\ Pos}{True\ pos+False\ neg} \tag{5}$$

$$Prec = \frac{True\ pos}{True\ pos+False\ pos} \tag{6}$$

$$Skor\ F1 = 2 * \frac{Prec*Rec}{Prec+Rec} \tag{7}$$

Example of visualization from the confusion matrix that used in this research is given in Figure 5. With the number that used is from randomize because purpose just for visualize and analyze the matrix.

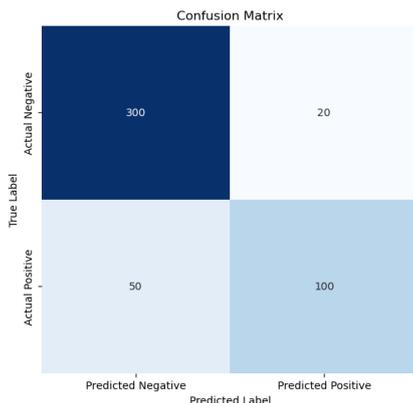


Figure 5. Confusion Matrics Visualization

Figure 5 had been seen the visualization from the confusion matrix. Had been seen in the visualization, the TP value is 300, the FP value is 100, the TN value is 300 and the FN value is 50. These values are random values used to visualize the confusion matrix. By using this data, we can calculate the recall values, precision values and f1-score values based from formulas that given in points 5, 6 and 7. After calculating the recall, precision and f1-score based on random data that already given, the results obtained are a recall value of 66.66%, a precision value of 83.33% and for the f1 score value obtained which is 74.07%. This value will be useful to be able to see the performance and evaluate the Convolutional Neural Network (CNN) model that has been built and trained.

### 3. RESULTS AND DISCUSSION

In this research, we will use the jupyter notebook IDE as a tool or tool used to write program

code to build the system. And as a programming language used, namely using Python and several python libraries used for the deep learning process. The deep learning technic that had been used from this research is Convolutional Neural Network or CNN algorithm. The purpose from using this algorithm is to see how effective the Convolutional Neural Network algorithm in order to classifying pneumonia based on patient X-Ray result images. The data that used to conduct the training and testing process is obtained from kaggle.com. After the data is inputted, the data reading and data normalization process will be carried out so that the inputted image becomes in double form or the pixel value ranges from 0 to 255. After the normalization process is done, another process is perform for the division of training and testing data. For training data, that divided again into 2 parts, namely train data and validation data whose data is taken from 20% of the test data. The purpose of dividing the validation data is so that when the model is training according to epoch or iteration, the model is simultaneously validated so that it can see whether overfitting occurs or not in the CNN model built. The division of data into testing data is intended so that previously trained models can be tested so that can meause the performance of the model in order to predicting new data in addition to training and validation data. The calculation of performance in this study is carried out using the confusion matrix and also the classification report so that it can be seen which model guesses are correct or wrong and the final performance of the model after the classification process. For visualization of the accuracy graph and loss graph results that happen during train and validate model are given in Figures 6 and 7.

Figures 6 and 7 show the accuracy and loss visualization graphs experienced by builded model during the train and validation process. In the graph, that can been seen performance that in the Convolutional Neural Network model built does not experience overfitting because in the process that can be seen there is no significant difference from the line between training and validation. In the training process, the accuracy is 95.73% and in the validation process, the model gets an accuracy of 95.69%. For the loss value obtained in the training process is 11.66% and during the validation process, the loss value obtained is 11.98%. From the values that obtained, that can been seen that the Convolutional Neural Network model does not experience overfitting. In the training process, using an epoch with a value of 50 with an early stopping process based on the validation loss value. So that if the validation loss value does not see an increase, it will be stopped. Therefore, in the train and validation process, the epoch stops at a value of 33. Which means that after iterating at epoch 33, the validation loss value will remain stagnant or show no change, so the iteration stops at epoch 33. After had been train

and validate, then the model testing process carried out again. The purpose of this process is to see the performance of the model when faced with new data other than the data used to conduct training and validation. After the testing process, the Convolutional model that has been built gets a testing accuracy of 87.82051205635071% or about 88% accuracy of the model's accuracy in classification. This accuracy is considered good and shows that the Convolutional Neural Network that built can accurately classify pneumonia diagnose by the X-Ray result image that given. After the test process is carried out and the accuracy is obtained, so the detail performance from the model testing process can be calculated. The calculation of performance will use the classification report. The results from the confusion matrix that had been obtained from the testing process given in Figure 8.

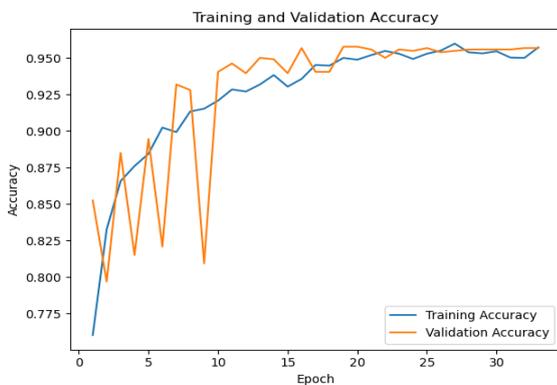


Figure 1. Accuracy Curve

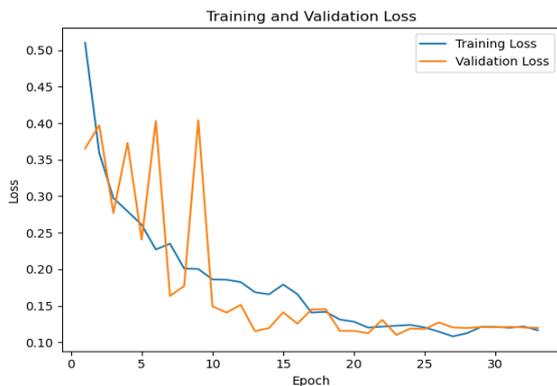


Figure 7. Loss Curve

Figure 8 shows detail result from model testing process for pneumonia classification based on the given X-Ray image data. Based on confusion matrix that obtain, the testing accuracy is 87.82051205635071% or about 88%. This can be explained because the model still makes some wrong classifications, namely there are 70 data that should be normal but are predicted to be pneumonia and there are 6 data that should be pneumonia but are predicted to be normal. But behind this, the model can classify data well as seen from 164 normal data that was successfully guessed by the model and 384 pneumonia data that was successfully guessed by the

model. With an accuracy of around 88%, it also shows that the model's performance for classification is very good and efficient.

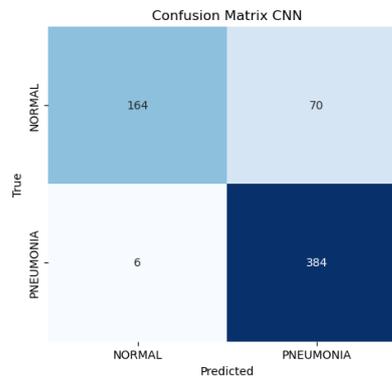


Figure 2. Confusion matrix for classification with CNN

Then after calculating the confusion matrix, it will be carried out again for the classification report calculation process. This calculation aims to measure the performance and accuracy that obtain from model process and see the harmonic value that include in the model after testing process. For the classification report value obtained, namely the Precision values, Recall values, F1-Score values and Support values. For the testing classification report can be seen in Figure 9.

Figure 9 shows the classification report results from the testing process of the Convolutional Neural Network (CNN) model built to classify and predict pneumonia based on the X-Ray images provided. That can be seen in Figure 9, the precision value that obtain from the class of normal is 96% and the precision value from the class of pneumonia is 85%. This shows that the Convolutional Neural Network Model built can well and accurately detect the normal class and is less accurate in detecting the pneumonia class. Then the recall value that obtain from the class of normal is 70% and the recall value from the class of pneumonia is 98%. This shows that the Convolutional Neural Network model built can better classify the pneumonia class than the normal class. As for the F1 score value in the normal class, it is around 81% and the pneumonia class gets a value of 91%. This shows that the pneumonia class gets a harmonic value or a balance value between precision and recall that is very good when compared to the normal class. This can happen because there is a gap in the value between precision and recall in the class of normal, which affects the harmonic value or balance value for the class of normal. For the support value, it is obtained from the amount of data carried out to carry out the testing process. The support value shows that the class of normal data that used during the test process is 234 data and the class of pneumonia data that used during the test process is 390 data.

After the process of calculating the precision, recall and f1 score values, from that classification report also can calculate the micro avg and wighted

avg values from the value of precision, value of recall and value of f1-score that already obtained. Micro avg is a value which is the average from the value is already obtained in each precision, recall and f1-score value. While the weighted avg value is used to calculate the value in each class, but the value will be multiplied by the weight value of each class so that it can provide an overview from the performance of model that built for used to classify each class by considering the imbalance in the sample. It can be seen in the classification report given in Figure 8, the avg macro value on the precision metric is 91%, on the recall metric is 84% and on the f1-score metric is 86%. These results indicate that the model built has an average value of precision or good accuracy. Meanwhile, the weighted avg value obtained in the precision metric is 89%, in the recall metric is 88% and in the f1-score metric is 87%. This value shows that the Convolutional Neural Network built has the good performance in performing the pneumonia classification process based on the test data.

	precision	recall	f1-score	support
NORMAL	0.96	0.70	0.81	234
PNEUMONIA	0.85	0.98	0.91	390
accuracy			0.88	624
macro avg	0.91	0.84	0.86	624
weighted avg	0.89	0.88	0.87	624

Figure 3. Classification report from CNN Model

#### 4. DISCUSSION

An accuracy of 87.82051205635071% was achieved during the classification process, following tests to classify and predict pneumonia using XRay result scans. This accuracy shows that the Convolutional Neural Network model built can accurately perform the pneumonia classification process. The model built can well distinguish

between images that have normal classes and pneumonia classes. However, because the test accuracy does not reach 100%, it can still be remembered that the model built can make a prediction error that the patient is diagnosed with pneumonia. From the testing process, the average values from precision, from recall and from f1-score are 91%, 84% and 86%. So that the Convolutional Neural Network that had been built in this research can accurately perform classification. In this study, the accuracy obtained is 87.82051205635071%, which is a good value, but still does not reach the perfect value for classification which is 100%. This is due to the large number of variations in the image in the form of variations in the position in performing the X-Ray process, the lighting used and the quality of each image contained in the dataset used and also because some of the images in the dataset have a fairly similar resemblance between the normal and pneumonia classes, so that because of the visual similarity, the model built can still be wrong in the prediction and classification process.

Many studies have also been conducted to classify and predict pneumonia using the CNN method. For the process of accuracy results carried out in other studies can be seen in Table 1. Based on Table 1 had been seen the comparison results from the accuracy of the Convolutional Neural Network model testing results that have been carried out in previous studies with the model created in this study. As had been seen in the table, the Convolutional Neural Network model built in this study obtained good accuracy when compared to the tests that have been carried out. It proves that a model developed in this study is capable of performing the pneumonia classification process more accurately. The results also show that, when it comes to pneumonia prediction using Xray images, the model which was used in this study is more efficient.

Table 1. Comparison on previous research

Research	Method	Testing Accuracy
[35]	CNN with Extreme Learning Machine	80.77%
[36]	CNN	77.56%
[37]	CNN	87.65%
[38]	CNN, pre trained CNN VGG16 and ResNet50	CNN is 75%, VGG16 is 80% and ResNet50 is 85%
Our	CNN	87.82051205635071%

#### 5. CONCLUSION

After this study has carried out the testing process and also the discussion process to build a Convolutional Neural Network model that can be useful for performing the classification process based on X-Ray result images, can concluded that model had been built can properly perform the pneumonia classification process based on the patient's X-Ray result image. This can be proven from the accuracy obtained in the model testing process, which obtained an accuracy of 87.82051205635071%, and obtained the best precision value of 96% obtained in the class of normal. The best recall value is 98% in the class of

pneumonia and the best f1-score value is 91% in the class of pneumonia. From these values, can be concluded that the Convolutional Neural Network model built in this study can accurately and well perform the classification process that patients have pneumonia or normal based on X-Ray images.

It is expected that it will be capable of using more sophisticated methods in future research. Such as using transfer learning Convolutional Neural Network or existing architectures such as VGG16, DenseNet, MobileNet, GoogleNet, and others. In future research, also expected to able adding more layers to the layers that will be used so that the Convolutional Neural Network model built can be

more complex and is expected to be more accurate to carry out the classification process. In the future also expected to able carry out the multi-class classification process and by using colored images.

## REFERENCES

- [1] A. F. Al Mubarak, J. A. M. Dominique and A. H. Thias, "Pneumonia Detection with Deep Convolutional Architecture," 2019 International Conference of Artificial Intelligence and Information Technology (ICAIIIT), Yogyakarta, Indonesia, 2019, pp. 486-489, doi: 10.1109/ICAIIIT.2019.8834476.
- [2] Murphy, C. N., Fowler, R., Balada-Llasat, J. M., Carroll, A., Stone, H., Akerele, O., Buchan, B., Windham, S., Hopp, A., Ronen, S., Relich, R. F., Buckner, R., Warren, D. A., Humphries, R., Campeau, S., Huse, H., Chandrasekaran, S., Leber, A., Everhart, K., ... Bourzac, K. M. (2020). Multicenter evaluation of the BioFire FilmArray Pneumonia/ Pneumonia plus panel for detection and quantification of agents of lower respiratory tract infection. *Journal of Clinical Microbiology*, 58(7). <https://doi.org/10.1128/JCM.00128-20>
- [3] Çınar, A., Yıldırım, M., & Eroğlu, Y. (2021). Classification of pneumonia cell images using improved ResNet50 model. *Traitement Du Signal*, 38(1), 165–173. <https://doi.org/10.18280/TS.380117>
- [4] Rahman, T., Chowdhury, M. E. H., Khandakar, A., Islam, K. R., Islam, K. F., Mahbub, Z. B., Kadir, M. A., & Kashem, S. (2020). Transfer learning with deep Convolutional Neural Network (CNN) for pneumonia detection using chest X-ray. *Applied Sciences (Switzerland)*, 10(9). <https://doi.org/10.3390/app10093233>
- [5] Alsharif, R., Al-Issa, Y., Alqudah, A. M., Qasmieh, I. A., Mustafa, W. A., & Alquran, H. (2021). Pneumoninet: Automated detection and classification of pediatric pneumonia using chest x-ray images and cnn approach. *Electronics (Switzerland)*, 10(23). <https://doi.org/10.3390/electronics10232949>
- [6] Shah, S., Mehta, H., & Sonawane, P. (2020). Pneumonia detection using convolutional neural networks. *Proceedings of the 3rd International Conference on Smart Systems and Inventive Technology, ICSSIT 2020*, 933–939. <https://doi.org/10.1109/ICSSIT48917.2020.9214289>
- [7] Chouhan, V., Singh, S. K., Khamparia, A., Gupta, D., Tiwari, P., Moreira, C., Damaševičius, R., & de Albuquerque, V. H. C. (2020). A novel transfer learning based approach for pneumonia detection in chest X-ray images. *Applied Sciences (Switzerland)*, 10(2). <https://doi.org/10.3390/app10020559>
- [8] Jain, R., Nagrath, P., Kataria, G., Sirish Kaushik, V., & Jude Hemanth, D. (2020). Pneumonia detection in chest X-ray images using convolutional neural networks and transfer learning. *Measurement: Journal of the International Measurement Confederation*, 165. <https://doi.org/10.1016/j.measurement.2020.108046>
- [9] E. Ayan and H. M. Ünver, "Diagnosis of Pneumonia from Chest X-Ray Images Using Deep Learning," 2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT), Istanbul, Turkey, 2019, pp. 1-5, doi: 10.1109/EBBT.2019.8741582.
- [10] Rahman, T., Chowdhury, M. E. H., Khandakar, A., Islam, K. R., Islam, K. F., Mahbub, Z. B., Kadir, M. A., & Kashem, S. (2020). Transfer learning with deep Convolutional Neural Network (CNN) for pneumonia detection using chest X-ray. *Applied Sciences (Switzerland)*, 10(9). <https://doi.org/10.3390/app10093233>
- [11] Wahyudi Setiawan and Fitri Damayanti, 2020, *J. Phys.: Conf. Ser.* 1477 052055 <https://doi.org/10.1088/1742-6596/1477/5/052055>
- [12] Moen, E., Bannon, D., Kudo, T., Graf, W., Covert, M., & van Valen, D. (2019). Deep learning for cellular image analysis. In *Nature Methods* (Vol. 16, Issue 12, pp. 1233–1246). Nature Research. <https://doi.org/10.1038/s41592-019-0403-1>
- [13] Bolón-Canedo, V., & Remeseiro, B. (2020). Feature selection in image analysis: a survey. *Artificial Intelligence Review*, 53(4), 2905–2931. <https://doi.org/10.1007/s10462-019-09750-3>
- [14] Sains dan Teknologi, J., Jumadi, J., & Sartika, D. (2021). *PENGOLAHAN CITRA DIGITAL UNTUK IDENTIFIKASI OBJEK MENGGUNAKAN METODE HIERARCHICAL AGGLOMERATIVE CLUSTERING*. <https://doi.org/10.23887/jstundiksha.v10i2.33636>
- [15] Maier, A., Syben, C., Lasser, T., & Riess, C. (2019). A gentle introduction to deep learning in medical image processing. In *Zeitschrift für Medizinische Physik* (Vol. 29, Issue 2, pp. 86–101). Elsevier GmbH. <https://doi.org/10.1016/j.zemedi.2018.12.003>

- [16] Sarker, I. H. (2021). Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. In *SN Computer Science* (Vol. 2, Issue 6). Springer. <https://doi.org/10.1007/s42979-021-00815-1>
- [17] Liu, H., & Lang, B. (2019). Machine learning and deep learning methods for intrusion detection systems: A survey. In *Applied Sciences (Switzerland)* (Vol. 9, Issue 20). MDPI AG. <https://doi.org/10.3390/app9204396>
- [18] Rizwan I Haque, I., & Neubert, J. (2020). Deep learning approaches to biomedical image segmentation. In *Informatics in Medicine Unlocked* (Vol. 18). Elsevier Ltd. <https://doi.org/10.1016/j.imu.2020.100297>
- [19] Wang, P., Fan, E., & Wang, P. (2021). Comparative analysis of image classification algorithms based on traditional machine learning and deep learning. *Pattern Recognition Letters*, 141, 61–67. <https://doi.org/10.1016/j.patrec.2020.07.042>
- [20] Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets*, 31(3), 685–695. <https://doi.org/10.1007/s12525-021-00475-2>
- [21] Huang, H., Peng, Y., Yang, J., Xia, W., & Gui, G. (2020). Fast Beamforming Design via Deep Learning. *IEEE Transactions on Vehicular Technology*, 69(1), 1065–1069. <https://doi.org/10.1109/TVT.2019.2949122>
- [22] Liu, X., Song, L., Liu, S., & Zhang, Y. (2021). A review of deep-learning-based medical image segmentation methods. *Sustainability (Switzerland)*, 13(3), 1–29. <https://doi.org/10.3390/su13031224>
- [23] Ozbayoglu, A. M., Gudelek, M. U., & Sezer, O. B. (2020). Deep learning for financial applications: A survey. In *Applied Soft Computing Journal* (Vol. 93). Elsevier Ltd. <https://doi.org/10.1016/j.asoc.2020.106384>
- [24] Nikou, M., Mansourfar, G., & Bagherzadeh, J. (2019). Stock price prediction using DEEP learning algorithm and its comparison with machine learning algorithms. *Intelligent Systems in Accounting, Finance and Management*, 26(4), 164–174. <https://doi.org/10.1002/isaf.1459>
- [25] Arnal Barbedo, J. G. (2019). Plant disease identification from individual lesions and spots using deep learning. *Biosystems Engineering*, 180, 96–107. <https://doi.org/10.1016/j.biosystemseng.2019.02.002>
- [26] Desai, M., & Shah, M. (2021). An anatomization on breast cancer detection and diagnosis employing multi-layer perceptron neural network (MLP) and Convolutional neural network (CNN). In *Clinical eHealth* (Vol. 4, pp. 1–11). KeAi Communications Co. <https://doi.org/10.1016/j.ceh.2020.11.002>
- [27] Bora, M. B., Daimary, D., Amitab, K., & Kandar, D. (2020). Handwritten Character Recognition from Images using CNN-ECOC. *Procedia Computer Science*, 167, 2403–2409. <https://doi.org/10.1016/j.procs.2020.03.293>
- [28] Polsinelli, M., Cinque, L., & Placidi, G. (2020). A light CNN for detecting COVID-19 from CT scans of the chest. *Pattern Recognition Letters*, 140, 95–100. <https://doi.org/10.1016/j.patrec.2020.10.001>
- [29] Bhatt, D., Patel, C., Talsania, H., Patel, J., Vaghela, R., Pandya, S., Modi, K., & Ghayvat, H. (2021). Cnn variants for computer vision: History, architecture, application, challenges and future scope. In *Electronics (Switzerland)* (Vol. 10, Issue 20). MDPI. <https://doi.org/10.3390/electronics10202470>
- [30] Kattenborn, T., Leitloff, J., Schiefer, F., & Hinz, S. (2021). Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. In *ISPRS Journal of Photogrammetry and Remote Sensing* (Vol. 173, pp. 24–49). Elsevier B.V. <https://doi.org/10.1016/j.isprsjprs.2020.12.010>
- [31] Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, 8(1). <https://doi.org/10.1186/s40537-021-00444-8>
- [32] Agarwal, M., Gupta, S. K., & Biswas, K. K. (2020). Development of Efficient CNN model for Tomato crop disease identification. *Sustainable Computing: Informatics and Systems*, 28. <https://doi.org/10.1016/j.suscom.2020.100407>
- [33] Mujahid, M., Rustam, F., Álvarez, R., Luis Vidal Mazón, J., Díez, I. de la T., & Ashraf, I. (2022). Pneumonia Classification from X-ray Images with Inception-V3 and Convolutional Neural Network. *Diagnostics*, 12(5). <https://doi.org/10.3390/diagnostics12051280>
- [34] Chakraborty, S., Paul, S., & Hasan, K. M. A. (2022). A Transfer Learning-Based Approach with Deep CNN for COVID-19 and Pneumonia-Affected Chest X-ray Image

Classification. *SN Computer Science*, 3(1).

<https://doi.org/10.1007/s42979-021-00881-5>

- [35] Nugroho, B., & Yulia, E. (2021). *KINERJA METODE CNN UNTUK KLASIFIKASI PNEUMONIA DENGAN VARIASI UKURAN CITRA INPUT*. 8(3), 533–538. <https://doi.org/10.25126/jtiik.202184515>
- [36] Perdananto, A. (2019). Penerapan deep learning pada Aplikasi prediksi penyakit Pneumonia berbasis Convolutional Neural networks. *Journal of Informatics and Communication Technology (JICT)*, 1(2), 1-10. [https://doi.org/10.52661/j\\_ict.v1i2.34](https://doi.org/10.52661/j_ict.v1i2.34)
- [37] Omar, H., & Babalık, A. (2019). Detection of pneumonia from X-ray images using convolutional neural network. *Proceedings Book*, 183.
- [38] Wei, X., Chen, Y., & Zhang, Z. (2020). Comparative Experiment of Convolutional Neural Network (CNN) Models Based on Pneumonia X-ray Images Detection. *Proceedings - 2020 2nd International Conference on Machine Learning, Big Data and Business Intelligence, MLBDBI 2020*, 449–454. <https://doi.org/10.1109/MLBDBI51377.2020.000095>.