Classification of Pneumonia from Chest X-ray Images Using Keras Module TensorFlow

Franky Chandra Satria Arisgraha[1,a)], Ririees Rulaningtyas[1], Miranti Ayudya Kusumawardani[1]

[1] Biomedical Engineering, Faculty of Science and Technology, Universitas Airlangga, Surabaya, Indonesia

[a)] Corresponding author: franky-c-s-a@fst.unair.ac.id

Abstract. Pneumonia is a respiratory disease caused by bacteria and viruses that attack the alveoli, causing inflammation of the alveoli. This study aims to examine the ability of the Convolutional Neural Network (CNN) model to classify pneumonia and normal x-ray images. The method used in this research is to construct a CNN model from scratch by compiling layers one by one with the help of the Keras TensorFlow module, which consists of a Convolution layer, MaxPooling layer, Flatten layer, Dropout layer, and Dense layer. Data used in this research is from Guangzhou Women and Children Medical Center, Guangzhou, China. The total data used is 200 images divided into 160 test data, 20 training data, and 20 validation data. From the results of the research conducted, the model has the fastest processing speed of 9.6ms/epoch with a total of 20 epochs. The model has the highest accuracy value of 77% in the training process and an accuracy value of 80% in the testing process. The highest sensitivity value is 1.54 in training and 1.6 in testing. The highest specificity value is 0.77 in training and 0.8 in testing. It can be said that the model can do good classification. Pendahuluan / Introduction (first level heading).

INTRODUCTION

Pneumonia is an infection that causes inflammation of the alveoli and their filling with fluid or pus in one or both lungs (Franquet, 2018). Consequently, patients experience productive cough, fever, chills, chest pain, and difficulty breathing. Pneumonia can be caused by bacteria, viruses, or fungi, with bacterial infections being most common in adults. Pneumonia is a leading cause of death among children worldwide. The World Health Organization (WHO) estimates that this disease accounts for 15% of deaths in children under 5 years old (WHO, 2020). Annually, pneumonia affects approximately 450 million people, representing seven percent of the global population, and leads to around 4 million deaths. In Indonesia, based on the Basic Health Research (Riskesdas) conducted by the Ministry of Health in 2018, the estimated number of pneumonia cases accounts for 2% of the total population. This indicates a 0.2% increase from 2013 when it was 1.8% (Kemenkes, 2018).

Early detection of pneumonia is crucial for providing appropriate medical care to patients. One common technique for early detection involves the examination of X-ray thorax images by radiologists or physicians. However, this method has inherent weaknesses due to subjectivity, influenced by the reader's expertise, fatigue, and visual acuity. To address the limitations of traditional reading techniques, researchers have experimented with building a pneumonia X-ray image classification model using machine learning. Stephen et al. employed a deep learning model, Convolutional Neural Network, comprising 4 Convolutional layers, 4 MaxPooling layers, 1 Flatten layer, and 7 Dense layers, with additional input variations such as scaling, rotation, width shift, height shift, shear range, zoom range, and horizontal flip (Stephen et al., 2019). The results demonstrated an accuracy rate of 93%. Jain et al. tested several Convolutional Neural Network (CNN) models, including VGG16, VGG19, ResNet50, and Inception-v3, for pneumonia detection (Jain et al., 2020). All four CNN methods used showed effectiveness in feature extraction and automatic classification. The best performance was observed with RestNet50. Xiang Xu conducted experiments to obtain optimal feature extraction using the Graph-Knowledge Embedded CNN method (Yu et al., 2020). Combining features extracted using two different methods with graph-based feature reconstruction, Xiang Xu et al. achieved an accuracy of 98%.

Based on this information, the authors aimed to build a CNN model capable of lightweight X-ray pneumonia image classification with high accuracy and efficiency using the Keras API based on the TensorFlow framework. The Keras API is an interface library designed to simplify the implementation of deep learning algorithms, reducing the required number of steps for code implementation running on TensorFlow (Chollet, F., 2015). TensorFlow itself is an open-source library created by the Google Brain team, utilized for large-scale numerical computations and machine learning projects, including deep learning (neural networks) (TensorFlow, 2018). Utilizing TensorFlow and Keras API aims to maximize research efficiency. The constructed Convolutional Neural Network (CNN) structure in this study comprises 6 convolutional layers, with 1 initial layer serving as input, 5 MaxPooling layers, 1 Flatten layer, 1 Dropout layer, and 2 Dense layers utilizing ReLu activation for convolution layers and softmax for classification (Dense) layer. The dataset used consists of X-ray images of pneumonia and normal cases obtained from female patients aged one to five years from the Guangzhou Women and Children's Medical Center, Guangzhou, China (Mooney, 2019)..

RESEARCH METHODOLOGY

Data

Chest X-Ray (CXR) or chest radiography is one of the most commonly used medical imaging applications for disease detection, including pneumonia. X-ray images can reveal areas of opacity or obstruction with brighter shades on the image. During the data collection phase, the best dataset with a high quantity and diversity of data was sought and compared. After selecting the suitable dataset, it was downloaded from the link provided: https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia. This dataset comprises X-ray images of pediatric patients aged one to five years from the Guangzhou Women and Children's Medical Center, Guangzhou, China (Mooney, 2019). Next, the data was examined and organized according to the defined scope, which consists of a total of 200 images distributed in an 80:10:10 ratio. Specifically, there are 160 training data, comprising 80 pneumonia and 80 normal images. Additionally, there are 20 test data, with 10 pneumonia and 10 normal images, and 20 validation data, consisting of 10 pneumonia and 10 normal images.

Design and Construction of the Model Structure

This research was conducted using Google Colaboratory, a cloud-based text editor. This application runs within a web browser, eliminating the need for installation on the user's device. The programming language used in this study is Python, with the TensorFlow library and the Keras API module. First, we conduct a preprocessing data. In the data preprocessing stage, data augmentation is performed on the image size and various image transformations. Image augmentation is a technique that applies different transformations to an image, resulting in multiple transformed copies of the same image, thereby expanding the dataset (Bharati, S., 2020).

The method used involves extracting data arrays from an image using the Image Data Generator function provided by the Keras library, allowing the data to be read by the CNN model that has been prepared. In this step, the image data is standardized to a uniform size by scaling the data to 1/255 and applying horizontal flip to introduce data variation. This research utilizes the Keras API, making use of several available layer classes, namely Sequential, Convolution 2D, MaxPool2D, Flatten, Dropout, and Dense, as the foundation for constructing the CNN structure. These classes play a crucial role in defining the architecture of the Convolutional Neural Network (CNN) used in the study. In each convolutional layer, the filter sizes used are 3, 32, 64, 128, and 256, and the Rectified Linear Unit (ReLU) activation function is applied. The ReLU activation function is chosen for its ability to introduce non-linearity to the model and avoid the vanishing gradient problem. As for the final Dense layer, the Softmax activation function is used. Softmax is commonly employed for multiclass classification tasks, where the output layer contains more than one neuron. It allows for the computation of class probabilities, making it suitable for determining the most likely class among multiple classes (Krizhevsky, 2012).

RESULT AND DISCUSSION

Data Augmentation

In the initial stage of the research, data preprocessing was performed, aiming to standardize the format of the input data and expand the data variations. In this study, scaling with a factor of 1/225 and horizontal flipping were utilized for data standardization. An example of the augmented image resulting from this process is shown in Fig. 1.

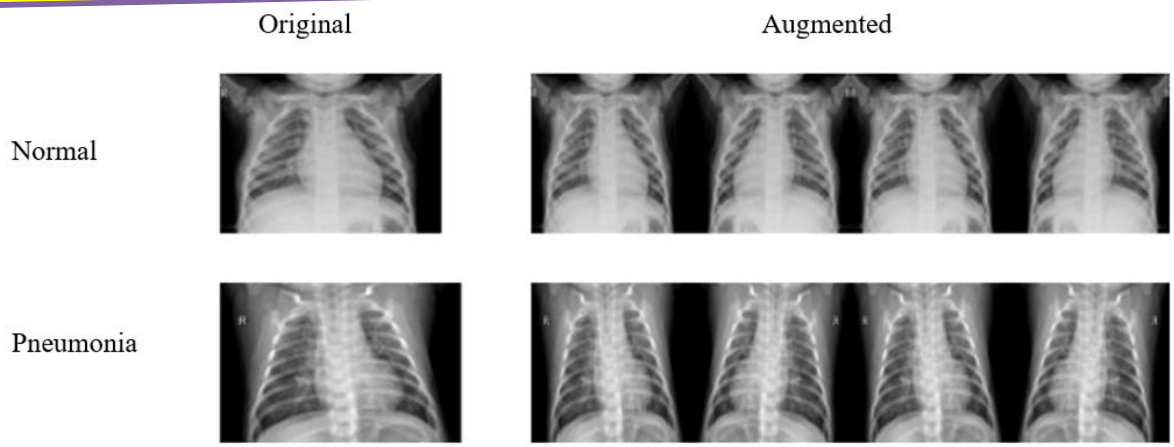

FIGURE 1. Samples of normal and pneumonia chest x-ray images during the data augmentation process are shown in Fig. 4.

The Convolutional Neural Network (CNN) Model Architecture

In this research, a Convolutional Neural Network (CNN) machine learning model was constructed with the following structure: 6 convolutional layers, 5 MaxPool layers, 1 Flatten layer, 1 Dropout layer, and 2 Dense layers, employing filter sizes of 3, 32, 64, 128, and 256. Unlike previous studies that utilized transfer learning approaches with pre-existing models that could be modified as needed, this research manually designed the model. Each layer was meticulously considered for its function and efficiency, utilizing modules from the Keras API to facilitate the process for each layer.

Results of the Model Classification

The research was conducted using 20 epochs for each input variation, namely input sizes of 150x150 pixels, 180x180 pixels, and 200x200 pixels. The accuracy values obtained from the training and testing processes by the CNN model are presented in TABLE 1.

TABLE 1. Comparison of Training and Testing Accuracy Result

Input Size	Training Accuracy	Testing Accuracy
150x150	72%	80%
180x180	74%	80%
200x200	77%	80%

TABLE 1 illustrates the accuracy performance in both the training and testing processes of the CNN model for each testing scenario with different input image sizes. Each scenario was run for 20 epochs. The values in TABLE 1 were obtained by executing the eval command provided by Keras TensorFlow.

In the first scenario, the image size was set to 150x150 pixels as input to the model. The accuracy achieved was 72% for training and 80% for testing. The model's performance with an input size of 150x150 pixels during the training process can be observed in FIGURE 2.

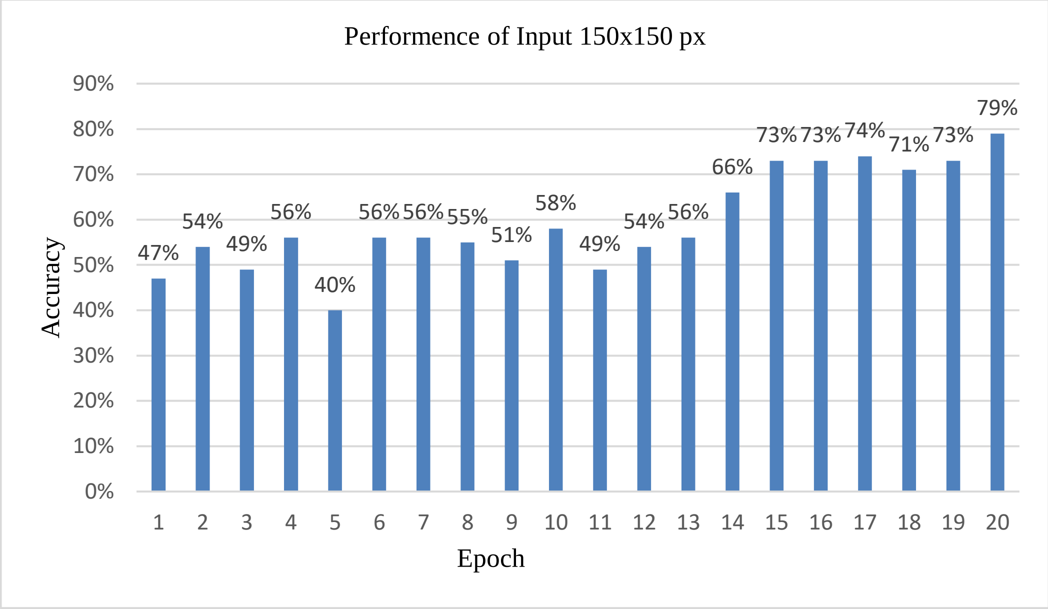

FIGURE 2. Accuracy Graph against Epochs for Input 150x150 Pixels

In the second scenario, the image size was set to 180x180 pixels as input to the model. The accuracy achieved was 74% for training and 80% for testing. The model's performance with an input size of 180x180 during the training process can be observed in FIGURE 3.

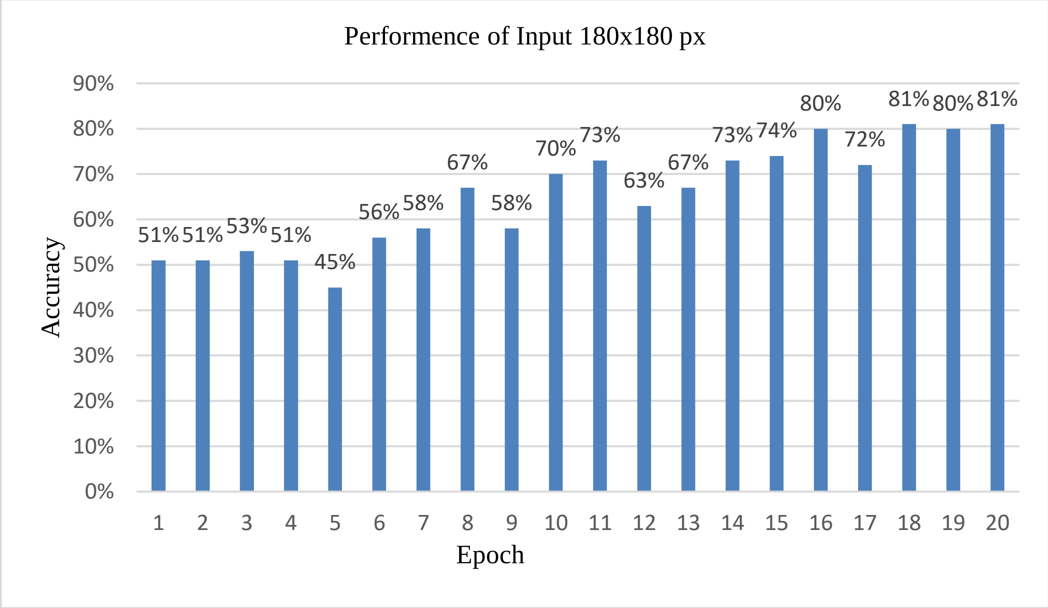

FIGURE 3. Accuracy Graph against Epochs for Input 180x180 Pixels

In the third scenario, the image size was set to 200x200 pixels as input to the model. The accuracy achieved was 77% for training and 80% for testing. The model's performance with an input size of 200x200 pixels during the training process can be observed in FIGURE 4.

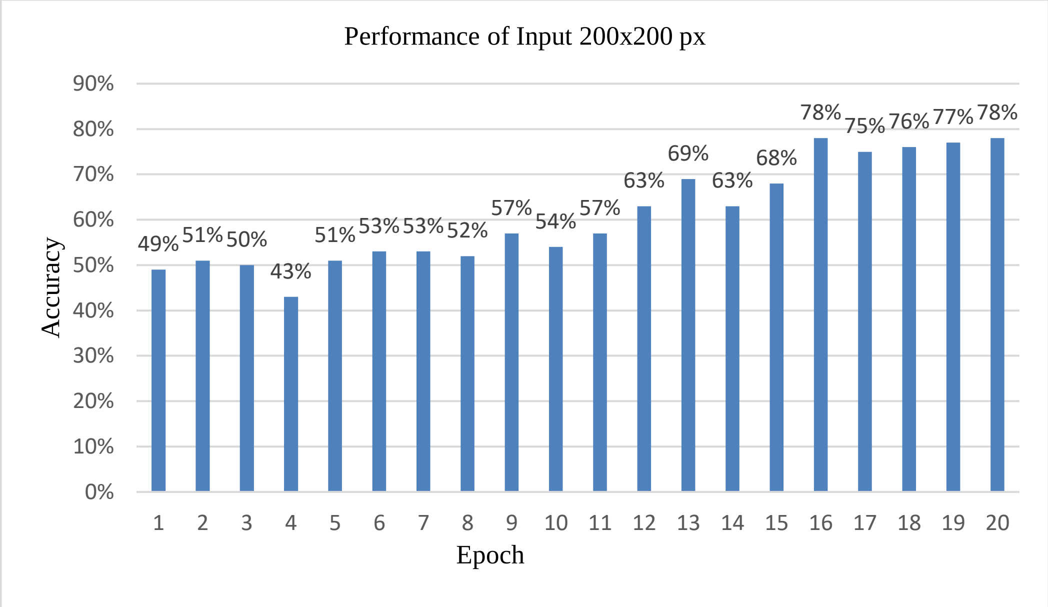

FIGURE 4. Accuracy Graph against Epochs for Input 200x200 Pixels

From the obtained accuracy data, it can be observed that the model produces stable accuracy values with the three different input sizes, yielding 80% accuracy during testing. The efficiency of the model's execution time can be seen in TABLE 2.

TABLE 2. Average Processing Time per 1 Epoch

Input Size (Pixels)	**Average Processing Time (ms/epoch)**
150x150	11.95
180x180	9.6
200x200	17.2

The fastest average time required for the model to complete one epoch was obtained with an input size of 180x180 pixels, which took 9.6 ms/epoch. This indicates that the model exhibits efficient processing speed. The prediction results for the test data, with a total of 160 samples consisting of 80 Pneumonia and 80 normal images, were compared to the confusion matrix table as show in TABLE 3.

TABLE 3. Confusion Matrix Table for Training - Input 150x150 Pixels

		Doctors Diagnose Result	
		Pneumonia	Normal
Programs Result	Pneumonia	57.6	22.4
	Normal	22.4	57.6

TABLE 4. Confusion Matrix Table for Training - Input 180x180 Pixels

		Doctors Diagnose Result	
		Pneumonia	Normal
Programs Result	Pneumonia	59.2	20.8
	Normal	20.8	59.2

TABLE 5. Confusion Matrix Table for Training - Input 200x200

		Doctors Diagnose Result	
		Pneumonia	Normal
Programs Result	Pneumonia	61,6	18,4
	Normal	18,4	61,6

TABLE 6. The Results of Accuracy, Sensitivity, and Specificity Calculation for Training

Input Size (Pixels)	**Accuracy (%)**	**Sensitivity (%)**	**Specify (%)**
150x150	0,72	0,72	0,72
180x180	0,74	0,74	0,74
200x200	0,77	0,77	0,77

Meanwhile, for the testing calculation, a single Confusion Matrix was utilized because the number of data and the accuracy generated by the model were the same for all three experimental scenarios. The dataset consisted of 20 samples, with an equal distribution of 10 pneumonia and 10 normal images, and input sizes of 150x150 pixels, 180x180 pixels, and 200x200 pixels.

TABLE 7. Confusion Matrix Table – Testing Result

		Doctors Diagnose Result	
		Pneumonia	Normal
Programs Result	Pneumonia	8	2
	Normal	2	8

TABLE 8. Confusion Matrix Table – Testing Result

Input Size	**Accuracy (%)**	**Sensitivity (%)**	**Specify (%)**
150^2, 180^2, and 200^2	0,8	0,8	0,8

From the conducted experiments, it can be observed that the input size with the highest accuracy during the training process is the input size of 200x200 pixels, achieving an accuracy rate of 77%. There is a significant difference of 4% between the input sizes of 150x150 pixels and 200x200 pixels, and 3% between 180x180 pixels and 200x200 pixels. Meanwhile, during the testing process, all three scenarios achieved the same accuracy rate of 80%. Subsequently, an experiment beyond the scope of the research was conducted by running the training command for a second time, i.e., using 40 epochs. The results showed that with an input size of 150x150 pixels, the model achieved an accuracy of 95%, while for 180x180 and 200x200 px, the model experienced overfitting. In the case of overfitting, the model obtained perfect accuracy during training but significantly lower accuracy during testing, indicating the model's inability to recognize features beyond the training data. The prediction results of the model in this study were satisfactory, even though the model was manually constructed by assembling the CNN layers one by one according to their respective functions, rather than using transfer learning approaches. The model provided accurate predictions for data outside the training, testing, and validation datasets, correctly classifying pneumonia as [0,1] and normal as [0,0]. The code and results are available in the appendix. It can be concluded that the model in this study can classify pneumonia and normal chest X-ray images similar to transfer learning methods in previous research. However, further optimization is required to improve the accuracy and enhance stability.

CONCLUSION

The CNN model achieved an accuracy of 80% on the test data for all three input sizes. However, on the training data, it achieved accuracy of 72% for 150x150 pixels, 74% for 180x180 pixels, and 77% for 200x200 pixels. For the 150x150 pixels input, the model reached a maximum accuracy of 95% when trained for 40 epochs, but it experienced overfitting for the 180x180 pixels and 200x200 pixels inputs. Regarding the sensitivity and specificity values, the training results in this study showed sensitivity and specificity of 72% for the 150x150 pixels input, 74% for 180x180 pixels, and 77% for 200x200 pixels. However, during testing, all three input scenarios had the same sensitivity and specificity values, which were both 80%.

ACKNOWLEDGMENTS

I express my gratitude to Lecturers in Physics Department Faculty of Science and Technology Universitas Airlangga for the opportunity to submit our paper.

REFERENCES

1. Allaouzi, I., & Ahmed, M. B. (2019). A novel approach for multi-label chest X-ray classification of common thorax diseases. *IEEE Access, 7*, 64279-64288.
2. Bharati, S., Podder, P., & Hossain, R. (2020). Hybrid deep learning for detecting lung diseases from X-ray images. *Journal of Informatics in Medicine Unlocked, 20*(2020). Publisher: Elsevier.
3. Chollet, F. (2015). Keras. Retrieved December 27, 2022, from https://keras.io.
4. Firmansyah, I., Hayadi, B. H. (2022). Komparasi Fungsi Aktivasi Relu Dan Tanh Pada Multilayer Perceptron. *JIKO: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika, 8*(2), 200-206. September 2022.
5. Franquet, T. (2018). Imaging of community-acquired pneumonia. *Journal of Thoracic Imaging*.
6. Geitgey, A. (n.d.). Machine Learning is Fun Part 3: Deep Learning and Convolutional Neural Networks. Retrieved December 27, 2022, from https://medium.com/@ageitgey/machine-learning-is-fun-part-3-deep-learning-and-convolutional-neural-networks-f40359318721.
7. Hariyani, Y. S., et al. (2020). Deteksi Penyakit Covid-19 Berdasarkan Citra X-Ray Menggunakan Deep Residual Network. *ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika, 8*(2), 443-453. May 2020.
8. Heidari, M., Mirniaharikandeheia, S., Khuzanib, A.Z., Danalaa, G., Qiu, Y., & Zheng, B. (2020). Improving the performance of CNN to predict the likelihood of COVID-19 using chest X-ray images with preprocessing algorithms. *International Journal of Medical Informatics, 144*, 104284. December 2020. Publisher: Elsevier.
9. Howard, A., Sandler, M., Chu, G., Chen, L. C., Chen, B., Tan, M., ... & Le, Q. V. (2019). Searching for mobilenetv3. *Proceedings of the IEEE International Conference on Computer Vision* (pp. 1314-1324).
10. Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., … Cao, B. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet, 395*(10223), 497–506