JURNAL ILMU PENDIDIKAN INDONESIA
Website:http://ejournal.uncen.ac.id/index.php/JIPI
Jurnal Terakreditasi Nasional, Keputusan, No 225/E/KPT/2022

Jurnal Ilmu Pendidikan Indonesia

Vol 14, No 2, Page 97 – 104, June 2026
P – ISSN 2338-3402, E – ISSN 2623-226X

CLASSIFICATION OF PNEUMONIA DISEASE USING THE MINI XCEPTION MODEL ON X-RAY DATA

Anggrainy Togi Marito Siregar[1]; Dina Jumiatul Fitri[2]; Happy Alyzhya Haay[3]; Rasi Kasim Samosir[4]

[1] Mathematics, Cenderawasih University, Indonesia; anggrainysiregar369@gmail.com
[2] Statistics, Cenderawasih University, Indonesia; dinajumiatulfitri@gmail.com
[3] Mathematics, Cenderawasih University, Indonesia; happyalyzhya98@gmail.com
[4] Planologi, Cenderawasih University, Indonesia; rk.samosir.93@gmail.com

Abstract: Pneumonia is a disease that often causes death in Indonesia. In general, many that cause a person to develop pneumonia include pneumonia due to bacterial, viral, mycoplasma pneumonia, fungal pneumonia. There are many ways to detect a patient grouped into one type of pneumonia. One way is to use an X-ray machine. X-ray is technology that can send waves of electromagnetic radiation briefly to scan the condition of the inside of the body. In this study, we tried to classify patients affected by bacterial pneumonia and viral pneumonia as well as normal people. The data we use is a picture of the lungs taken from the X-ray results. This research was conducted by applying the mini Xception model using the python program. The model can predict the results of X-ray scans that belong to the class of bacterial pneumonia very well, as seen from the value of precision and sensitivity of 80 and 97 percent, respectively. Viral pneumonia class can not be predicted as good as the two previous classes, but the results obtained are quite good as seen from the value of precision and sensitivity of 85 and 67 percent, respectively. The overall accuracy of the model obtained is 0.86.
Keywords: Pneumonia; X-ray Data; Mini Xception Model

1. INTRODUCTION

Pneumonia is a dangerous disease caused by the lack of oxygen in the lungs to circulate throughout the body. According to the 2013 Basic Health Research (Riskesdas), pneumonia is a disease that has received government attention, ranking 9th out of 10 leading causes of death in Indonesia, with a rate of 2.1%.

The bacteria that cause pneumonia are Streptococcus pneumoniae, Haemophilus influenzae, Staphylococcus aureus, group B streptococcus, as well as atypical chlamydia and mycoplasma. Other common causes are viruses, including respiratory syncytial virus (RSV), parainfluenza virus, adenovirus, and influenza virus (Leung et al., in Seyawati and Marwiati 2018).

According to the Indonesian Ministry of Health (2010), nearly 70% of pneumonia cases in the community are caused by bacteria, primarily Streptococcus pneumoniae

(50%) and Haemophilus influenzae (20%). In fact, pneumonia caused by viruses is less serious than pneumonia caused by bacteria. This, of course, presents a challenge in diagnosing pneumonia patients.

Pneumonia can be diagnosed in various ways, one of which is through a physical examination, which is usually characterized by symptoms such as a persistent cough, fever, sweating, chest pain, decreased appetite, and a rapid heartbeat. However, this method is very difficult to diagnose in individuals, as many diseases exhibit similar symptoms. A common alternative is an X-ray, a medical imaging technique that uses electromagnetic radiation waves to capture images or photographs of the inside of the body (Ghayoumi, 2017).

In this study, we used these X-ray results as data to distinguish the type of pneumonia a patient is suffering from. This data will then be processed using a technology known as machine learning to accurately predict pneumonia. The model we used in this study is Mini Xception.

For feature fusion, several approaches Benitez, et al (2016) and Ghimire, et al (2017) combine both geometric and visual features to address the weaknesses of both previous approaches. The approach used can produce better results in certain cases. The next approach is an advanced approach. This approach uses deep learning algorithms as the basis of its approach, such as convolutional neural networks (CNN). The main advantage of using CNN is its ability to directly learn from input images (Walecki, et al., 2017). Therefore, CNN is quite commonly used in image processing. In general, the use of CNN for feature extraction involves a set of fully connected layers in the last layer (Ghimire and Lee., 2013). Fully connected layers contain most of the parameters in CNN. To reduce the parameters in the last layer, for efficiency purposes, several CNN architectures have been developed, such as Inception V3 (Szegedy, et al., 2016; Chollet, 2016).

2. RESEARCH METHODS

Data

The data in this study used chest x-ray images taken from Kaggle. The downloaded data consists of two parts: training data and test data. The training data consists of 1,342 chest x-ray images in the normal category, 2,530 images in the bacterial

pneumonia category, and 1,345 images in the viral pneumonia category. The test data consists of 242 chest x-ray images in the normal category, 234 images in the bacterial pneumonia category, and 148 images in the viral pneumonia category.

Pre-Processing

The pre-processing stage involves exploring the acquired data. Before data preparation, the data is separated from the labels so that model development can begin. After separating the image data and labels, the images are resized by taking the value of each pixel and converting it into a 150 x 150 x 3 tensor.

Mini-Xception Model

The model used in this study is referred to as mini-Xception (Arriaga, et al., 2017). This model was optimized using the Adam optimizer (Diederik and Jimmy., 2014). This model combines the removal of fully connected layers with a combination of depth-wise separable convolutions and residual modules. Residual modules modify the mapping between the two layers, resulting in two distinct features (Suk and Prabhakaran., 2014).

The number of parameters in convolutional layers can be reduced by using depth-wise separable convolution. Depth-wise separable convolutions combine two distinct layers: depth-wise convolution and point-wise convolution. This layer separates spatial cross-correlations from channel cross-correlations (Chollet, 2016). This separation is achieved by first applying a D×D filter to each of the M input channels and then applying N 1×1×M convolution filters to combine the M input channels into N output channels. Applying 1×1×M convolutions combines each value in the feature mapping, taking into account the spatial relationships between the channels (Happy, et al., 2012).

The next architecture is a fully-convolutional neural network, consisting of four residual depth-wise separable convolutions. Each convolution is followed by a batch normalization operation and a ReLU activation function. The final layer applies global average pooling and a soft-max activation function to generate class predictions. The model architecture is shown in Figure 1 below.

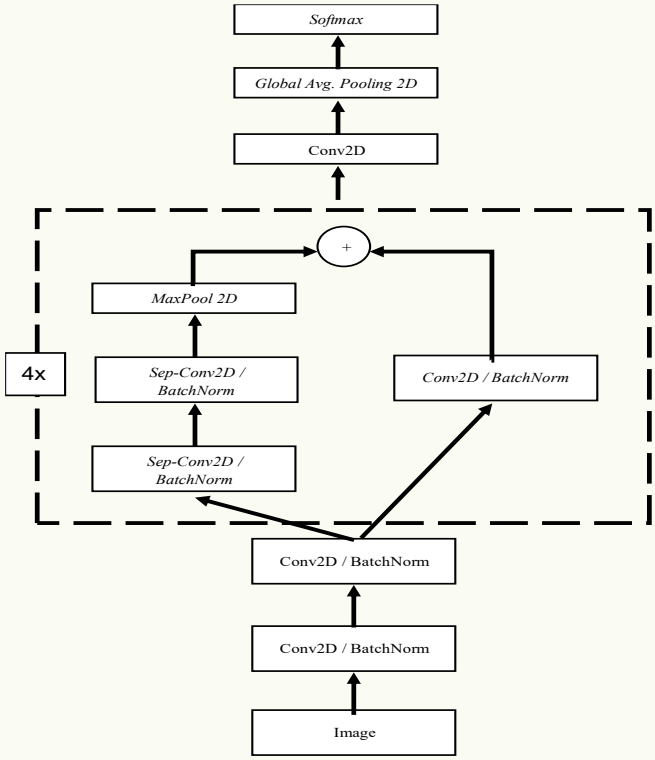

Figure 1. Mini-Xception architecture

Model Evaluation

In general, when classifying data, we can evaluate the classification model using a confusion matrix. The key points to consider in a confusion matrix are the percentage values of accuracy, sensitivity, and precision (Khan, et.al., 2013 ; Siddiqi, et al., 2015). The following illustrates how to calculate a confusion matrix using True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN).

True Positive (TP) is the probability that the model correctly predicts that the patient is ill. True Negative (TN) is the probability that the model correctly predicts that the individual is not suffering from the disease. A false positive (FN) is the probability that the model correctly predicts that the individual is ill, but in fact, they are not. Based on Kaulard, et al. (2012), describe that false positive (FP) is the probability that the model correctly predicts that the individual is not ill, but in fact, they are. The values for determining accuracy, precision, and sensitivity are as follows:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$precision = \frac{TP}{TP+FP} \quad (2)$$

$$Sensitivity = \frac{TP}{TP+FN} \quad (3)$$

3. RESULTS AND DISCUSSIONS

The purpose of this study was to classify whether a person had pneumonia based on chest X-ray data. If a person was classified as having pneumonia, a distinction was also made between bacterial and viral causes. The following data was used in this study.

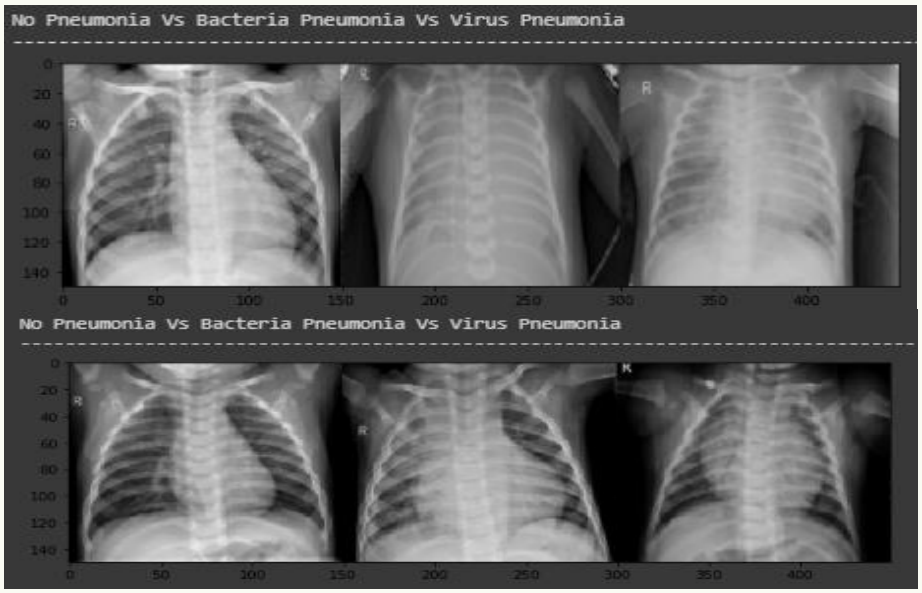

Figure 2. Normal X-Ray Results, Bacterial Pneumonia, And Viral Pneumonia

The developed model was run using Google Collaboratory. The model training process is illustrated in Figure 3 below.

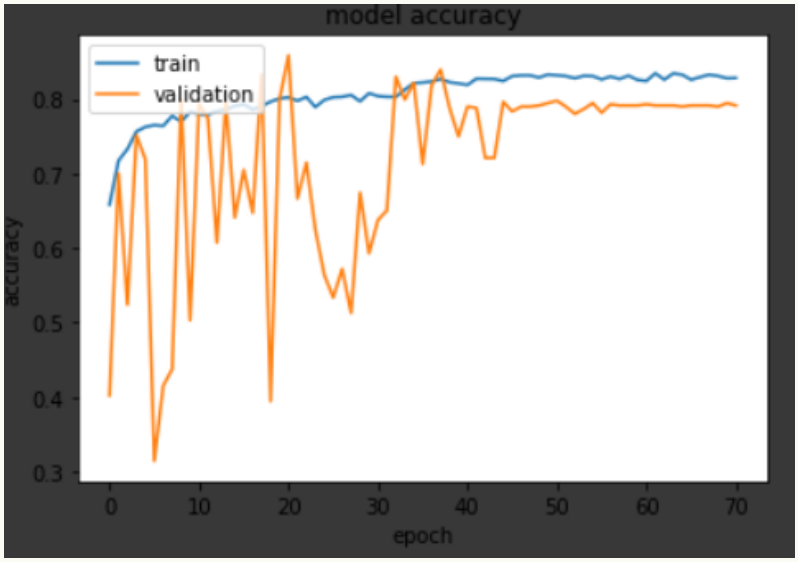

Figure 3. Mini-Xception Model Evaluation Graph

Figure 3 shows that the validation score of the Mini-Xception model fluctuates at the beginning of the epoch and becomes more stable as the epoch progresses. The model selected was the one with the highest accuracy on the

validation data.

The results of the precision and sensitivity calculations for the model validation results are presented in Table 1 below.

Table 1. Model Evaluation Scores

Class	Criteria	Scoring
Normal	Presisi	0.94
	Sensitivitas	0.86
Bacterial Pneumonia	Presisi	0.80
	Sensitivitas	0.97
Viral Pneumonia	Presisi	0.85
	Sensitivitas	0.67

Table 1 shows that the model can distinguish between normal and pneumonia classes very well, as evidenced by a precision value of 94 percent and a sensitivity of 86 percent. The model also performed very well in predicting X-ray scans that included bacterial pneumonia, as evidenced by precision and sensitivity values of 80 and 97 percent, respectively. While viral pneumonia was not as well predicted as the previous two classes, the results were quite good, as evidenced by precision and sensitivity values of 85 and 67 percent, respectively. The overall accuracy of the model was 0.86.

The following figure illustrates the distribution of labels between the model's predictions and the actual results.

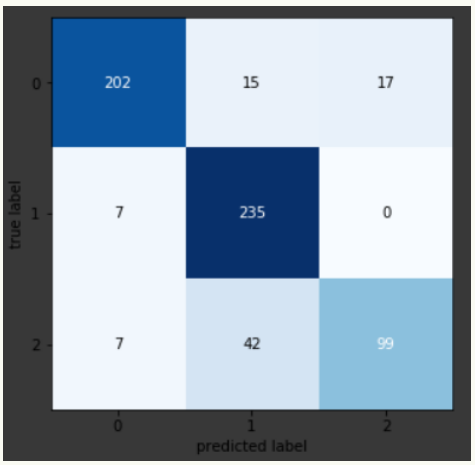

Figure 4. Confusion Matrix of the Model

Figure 2 shows that the model can distinguish well between classes labeled 0, or the normal class, as seen by the 14 prediction errors out of 216 data points belonging to that class. Class labeled 1, or bacterial pneumonia, contains information that the model can correctly predict 235 of the 292 data points belonging

to that class. Class labeled 2, or viral pneumonia, can correctly predict 99 of the 116 data points belonging to that class.

4. CONLUSIONS AND SUGGESTIONS

CONLUSIONS

The mini-Xception model used can predict pneumonia caused by bacteria and viruses quite well. The essence of using the mini-Xception model is to reduce the number of model parameters. This reduction in model parameters can affect the model's training speed and thus its predictions. Applications using the results of this model training will be more efficient than those using models that use more parameters in their development. This model can be further developed by training it with more data. Furthermore, further exploration of the data used can be carried out, such as determining observable positions to determine whether a person has pneumonia or not. This determination certainly requires experts such as surgeons.

SUGGESTIONS

Further research is recommended using a larger amount of X-ray data to ensure more accurate and stable model classification results. Furthermore, comparisons with other deep learning models are needed to determine which method performs best in detecting pneumonia.

5. ACKNOWLEDGMENTS

The authors would like to thank Cenderawasih University for its support in conducting this research. They also thanks all those who assisted in data collection, research, and the preparation of this journal.

BIBLIOGRAPHY

Arriaga O., Ploger P. G., and Valdenegro M (2017). *Real-time Convolutional Neural Networks for Emotion and Gender Classification*. arXiv:1710.07557v1 [cs.CV] 20 Oct 2017.

Benitez-Quiroz C.F., Srinivasan R., Martinez A.M. (2016) EmotioNet: *An accurate, real-time algorithm for the automatic annotation of a million facial expressions in the wild; Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; Las Vegas, NV, USA. 26 June–1 July 2016; pp. 5562–5570.

Chollet F. (2016) Xception: *Deep learning with depthwise separable convolutions*. 2016. CoRR, abs/1610.02357.

Diederik Kingma and Jimmy Ba. (2014). *A method for stochastic optimization. arXiv preprint arXiv*:1412.6980, 2014.

Ghayoumi M. (2017). *A quick review of deep learning in facial expression. J. Commun. Comput*. 2017; 14:34–38.

Ghimire D., Lee J. (2013). *Geometric feature-based facial expression recognition in image sequences using multi-class AdaBoost and support vector machines.* Sensors. 2013; 13:7714–7734. doi: 10.3390/s130607714.

Ghimire D., Jeong S., Lee J., Park S.H. (2017) *Facial expression recognition based on local region specific features and support vector machines. Multimed.* Tools Appl. 2017; 76:7803–7821. doi: 10.1007/s11042-016-3418-y.

Happy S.L., George A., Routray A. (2012). *A real time facial expression classification system using local binary patterns; Proceedings of the 4th International Conference on Intelligent Human Computer Interaction*; Kharagpur, India. 27–29 December 2012; pp. 1–5.

Khan R.A., Meyer A., Konik H., Bouakaz S. (2013). *Framework for reliable, real-time facial expression recognition for low resolution images. Pattern Recognit*. Lett. 2013;34: 1159–1168. doi: 10.1016/j.patrec.2013.03.022.

Kaulard K., Cunningham D.W., Bülthoff H.H., Wallraven C. (2012). *The MPI facial expression database—A validated database of emotional and conversational facial expressions.* PLoS ONE. 2012;7:e32321. doi: 10.1371/journal.pone.0032321.

Suk M. and Prabhakaran B. (2014). *Real-time mobile facial expression recognition system—A case study; Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops;* Columbus, OH, USA. 24–27 June 2014; pp. 132–137.

Siddiqi M.H., Ali R., Khan A.M., Park Y.T., Lee S. (2015). *Human facial expression recognition using stepwise linear discriminant analysis and hidden conditional random fields. IEEE Trans. Image* Proc. 2015; 24:1386–1398. doi: 10.1109/TIP.2015.2405346.

Szegedy C., Vanhoucke V., Ioffe Sergey., Shlens Jon., and Wojna Z. (2016). *Rethinking the inception architecture for computer vision. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016; pp. 2818–2826.

Walecki R., Rudovic O., Pavlovic V., Schuller B., Pantic M. (2017). *Deep structured learning for facial action unit intensity estimation; Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; Honolulu, HI, USA*. 21–26 July 2017; pp. 3405–3414.