ISSN 2089-8673 (Print) | ISSN 2548-4265 (Online)
Volume 13, Issue 2, July 2024

PNEUMONIA DIAGNOSIS THROUGH DEEP LEARNING: RESNET50V2 MODEL IMPLEMENTATION

Yufis Azhar[1], Wahyu Priyo Wicaksono[2], Zamah Sari[3]

[1,2,3] Informatika, Universitas Muhammadiyah Malang, Malang, Indonesia

email: yufis@umm.ac.id[1], wahyuwpw27@gmail.com[2], zamahsari@umm.ac.id[3]

Abstract

Pneumonia is a significant global health concern, particularly affecting young children and the elderly. It is a lung infection caused by bacteria, viruses, fungi, or parasites, leading to the alveoli filling with pus or fluid. This study addresses the challenge of accurately diagnosing pneumonia using chest X-ray images, a process traditionally dependent on the expertise of radiologists. The reliance on radiologists results in lengthy diagnosis times and high costs, particularly in regions with a shortage of medical professionals. This research presents a deep-learning approach to automate the classification of pneumonia using the ResNet50v2 model, which has been pre-trained on the ImageNet dataset. The dataset used in this study, obtained from the Guangzhou Women and Children's Medical Center, comprises 5,856 images, with 1,583 normal and 4,273 pneumonia cases. The images were preprocessed and augmented to enhance the model's robustness. The proposed model achieved an accuracy of 94%, demonstrating its potential in clinical settings to assist in the rapid and reliable diagnosis of pneumonia. This study contributes to the growing body of research in medical image analysis by employing a pre-trained ResNet50v2 model. It highlights the importance of leveraging advanced machine-learning techniques to improve diagnostic accuracy and efficiency.

Keywords: Pneumonia, Computer Vision, Deep Learning, ResNet50v2

Received: 16-12-2023 | **Revised:** 28-06-2024 | **Accepted:** 11-07-2024
DOI: https://doi.org/10.23887/v13i2.72068

INTRODUCTION

Pneumonia, a lung infection caused by bacteria, viruses, fungi, or parasites, leads to the lungs filling with fluid or pus, affecting individuals of all ages [1] [2]. According to the World Health Organization (WHO), pneumonia remains a leading cause of death among children under five years old, with 740,180 deaths in 2019 [3][4]. The diagnosis of pneumonia using X-ray images relies heavily on radiologists, leading to long diagnosis times and high costs [5][6][7]. The limited number of experienced radiologists in developing countries further exacerbates this issue [8].

Machine Learning, particularly Deep Learning, has shown promise in analyzing medical images, providing significant advancements in disease classification [9]. Convolutional Neural Networks (CNNs), a subset of Deep Learning, have proven effective in image-processing tasks [10]. It can run a computational process consisting of several layers of neural networks that are interconnected between one layer and another. It extracts features to allow the computer to learn information from the previous layer. Currently, Deep Learning plays an important role in the medical world. The medical image learning process in Deep Learning can extract information on medical images so that the analysis and classification process can be carried out effectively to help medical personnel diagnose diseases more accurately and quickly [11]. Convolution Neural Network is one of the approaches in deep learning widely used to solve various problems and produce good results [12]. Research on Deep Learning has been initiated to assist doctors in diagnosing diseases by utilizing X-ray images and overcoming cost problems in countries and hospitals with limited laboratory equipment [13].

Recent studies have utilized various pre-trained models, such as VGG16, DenseNet121, InceptionV3, and Xception, for pneumonia classification, achieving varying accuracy [14]. The total dataset used in that research is 5856 x-ray images from the Guangzhou Women and Children's Medical Center. The highest accuracy result obtained in this study was 90.5% in the proposed method.

Another study by Rachna Jain, Preeti Nagrath, Gaurav Kataria, V. Sirish Kaushik, and

D. Jude Hemanth entitled "Pneumonia Detection in Chest X-ray Images using Convolutional Neural Networks and Transfer Learning" proposed the use of four pre-trained models namely VGG16, VGG19, ResNet50, and Inception-V3 and two Convolution Neural Network models. The dataset used in this research is the same as the dataset in the previous research, where in this research, the dataset is obtained from the Kaggle site with the name "Chest X-Ray Images (Pneumonia)." The accuracy results obtained in the first model were 85.26%, while the accuracy results obtained in the second model were 92.31% [15].

In another study by Mohammad Rahimzadeh, Abolfazl Attar, "A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of Xception and ResNet50V2". The researcher proposed the use of pre-trained models Xception and ResNet50V2, with a dataset of 180 COVID-19 images and 42 Pneumonia images derived from the COVID chest x-ray dataset, coupled with 6012 Pneumonia images and 8851 Normal images obtained from the rsna pneumonia detection challenge. The test results of the method used gave good results, with an average accuracy of 91.40% [16]. Another study by Mahadar A, Mangukiya P, and Baraskar T titled "Comparison and Evaluation of CNN Architectures for Classification of Covid-19 and Pneumonia" proposed InceptionV3, ResNet50V2, MobilenetV2, and VGG16 models as base models and performed a fine-tuning process during model training. The data used in this study were 6939 x-ray images to classify COVID-19 and pneumonia. From the proposed tests, the fine-tuning process had a positive impact during model training, and the use of the ResNet50V2 pre-trained model got an accuracy result of 95%.[17].

Chest CT scans or X-ray images can be used to diagnose pneumonia. The application of CNN in healthcare has shown promising results, especially for diagnosing pneumonia based on X-ray images. Based on the results of previous studies, many variations of Convolution Neural Network models can be used. Previous studies have shown that applying the pre-trained ResNet50V2 model gave promising results, so this study proposes using the pre-trained ResNet50v2 model to improve the accuracy of pneumonia classification.

The ResNet50v2 architecture, an enhancement of ResNet50, offers improved performance in certain classification tasks due to adjustments in the relationship between residual blocks [18]. This study aims to demonstrate the efficacy of ResNet50v2 in classifying pneumonia from X-ray images, contributing to the existing body of research and addressing the challenges of pneumonia diagnosis in resource-constrained environments.

METHOD

This research was conducted through five stages. The first stage is dataset collection, the second is data separation, the third is data preprocessing and augmentation, the fourth is CNN modeling, and the last is model evaluation. The first stage is the collection of datasets containing chest X-ray images of pneumonia and normal. Next, the dataset that has been obtained is divided into training data and test data. The data that has been divided then goes through data preprocessing and data augmentation. After the process, the data is used as a source of information in the training process of the model built by utilizing the ResNet50V2 architecture. The last stage is the evaluation process, which contains information about the model's performance results during model training. Details about the research and its flow can be seen in Figure 1.

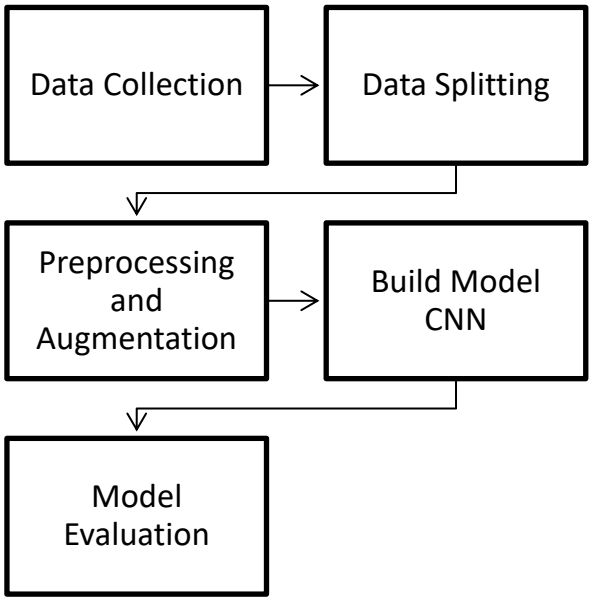

Figure 1. Research Flow

A. Dataset

In this study, the dataset is image data containing chest x-ray images. The dataset used in this study comes from the Kaggle site with the name Chest X-Ray Images (Pneumonia); the dataset in this study can be accessed on the Kaggle site at the following address: https://www.kaggle.com/datasets/paultimothymo

oney/chest-xray-pneumonia. The dataset's normal and pneumonia chest x-ray images were sourced from Guangzhou Women and Children's Medical Center, Guangzhou. In this study, the total images in the dataset amounted to 5856 images consisting of two categories, namely normal and pneumonia. The normal category comprises 1583 chest x-ray images, and the pneumonia category comprises 4273 chest x-ray images. Each normal and pneumonia category will be divided into train data and test data as a source of model training. Detailed examples of chest x-ray images used in this study can be seen in Figure 2 for the normal category and Figure 3 for the pneumonia category.

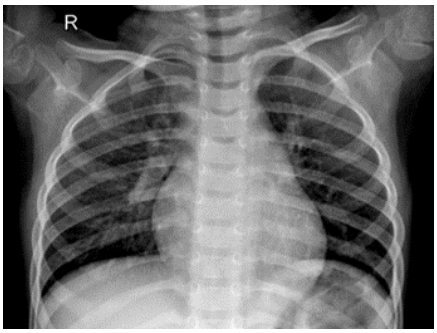

Figure 2. Normal X-Ray Image

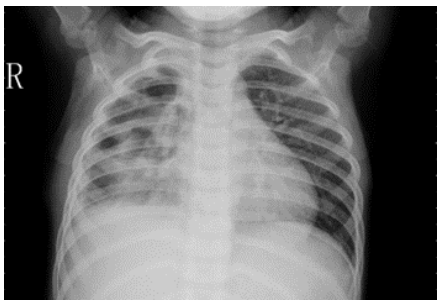

Figure 3. Pneumonia X-Ray Image

B. Split Dataset

In this study, the dataset that has been obtained consists of 5856 x-ray images. The next process is to divide the dataset into two parts: train and test data. Train data is a data collection used to train the model to classify pneumonia from x-ray images. Meanwhile, test data is used for the optimization process during model training. In this study, the train data contains x-ray image data with a total of 5232 data, while the test data contains x-ray images with 624 data. In the train and test data, there will be two categories of images: normal and pneumonia. Details about the number of datasets used in this study can be found in Table 1.

Table 1. Dataset Splitting

Label	Normal	Pneumonia	Total
Train	1349	3883	5232
Test	234	390	624

C. Data Preprocessing and Augmentation

In this research, the obtained data will undergo data preprocessing and augmentation. While no research specifically establishes one best preprocessing approach, experimentation and testing can increase the chances of finding the optimal configuration according to the dataset characteristics [19]. Data augmentation was performed on training data, applying transformations such as rotation, horizontal and vertical flips, and brightness adjustments to introduce variability and prevent overfitting [20].

In this research, there is a pre-processing process by changing the image size to 224 x 224 pixels. Using a data augmentation process that is useful for adding variations to the available image data. The data augmentation process uses the training data to modify the image to get variations. Augmentation can have a good effect on the model because it gets data with more diverse variations. Data augmentation also serves to reduce the occurrence of overfitting when training the model. Several parameters are used in the augmentation process; more details about the parameters and values can be seen in Table 2.

Table 2. Augmentation Parameters

Parameters	Value
rescale	1/255
zoom_range	0.3
horizontal_flip	True
vertical_flip	True
brightnes_range	[0.5, 2.0]
rotation_range	20

D. Build CNN Model

The main structures of CNN are convolution layers, pooling layers, and fully connected layers [21]. Residual Network (ResNet) is a CNN architecture developed to reduce the impact of vanishing gradient on Neural Networks and improve accuracy results [22]. There are several variations of the ResNet architecture, one of which is ResNet50V2. The ResNet50v2 model, pre-trained on the ImageNet dataset, was used

as the base model. Pre-trained models leverage previously learned features from large datasets, which can improve performance and reduce training time on smaller, task-specific datasets. Details of the ResNet50V2 architecture can be seen in Figure 4.

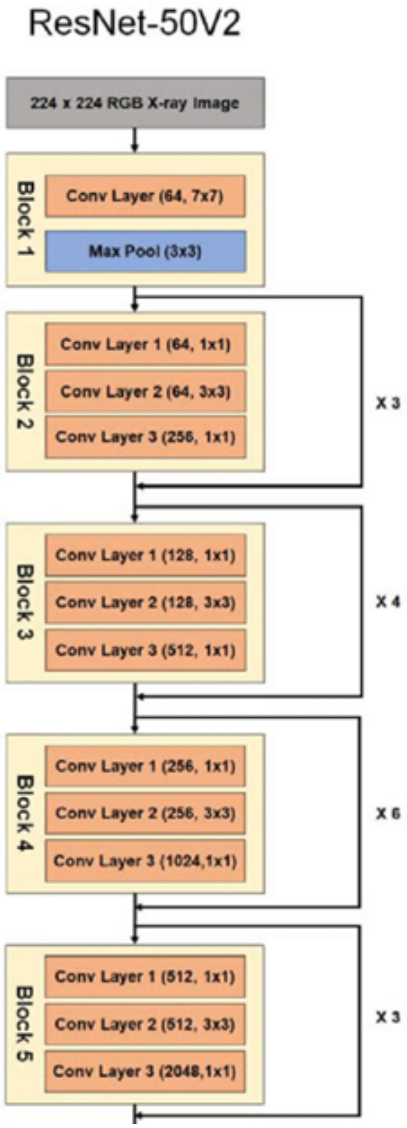

Figure 4. ResNet50V2 Architecture [23]

In this research, a CNN model was created by utilizing the pre-trained ResNet50V2 model. Modeling in this study is an input layer with a suggested size of 224x224 pixels, followed by a model architecture from ResNet50V2 with 'imagine' weights used as a base model. Subsequently, a Global Average Pooling layer was incorporated. Global Average Pooling aims to reduce overfitting during model training [24]. For the Fully Connected Layer, there is a dense layer with the number of neurons 64 and 128 and a Dropout Layer with a parameter value of 0.2. For the output layer in this study, one neuron with sigmoid activation was used in the model training process using the Adam optimizer with a learning rate of 0.001. Details of the entire model built in this study can be seen in Figure 5, and details regarding the value of each parameter in the model that has been built can be seen in Table 3.

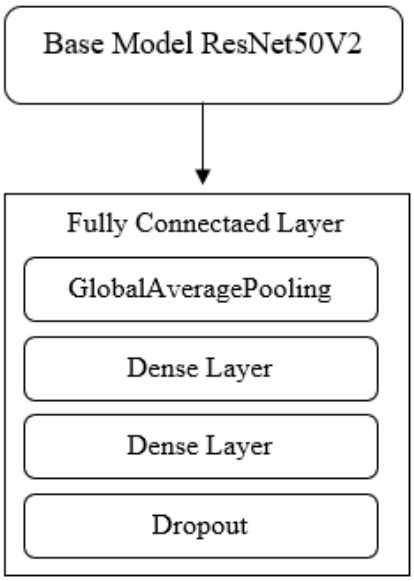

Figure 5. Model Architecture

Table 3. Values at Fully Connected Layer

Description	Value
Input Layer	224x224
Dense Layer	64, 128
Dropout Layer	0.2
Output Layer	Sigmoid

E. Model Evaluation

Model evaluation is a crucial stage in assessing the model's performance during the training process on the provided dataset. This evaluation is derived from the Classification Report and Confusion Matrix results obtained in this study. The Classification Report offers a comprehensive summary of the model's performance throughout training, highlighting several key metrics: accuracy, precision, recall, and F1 score.

Accuracy reflects the model's overall ability to correctly classify the instances within the dataset. It measures the proportion of true results (both true positives and true negatives) among the total number of cases examined. On the other hand, precision measures the model's accuracy in predicting positive instances, indicating how

many of the positive predictions made by the model are actually correct. This metric is particularly important in scenarios with a high cost of false positives.

Recall, also known as sensitivity, assesses the model's effectiveness in identifying all relevant instances within the dataset. It shows how well the model captures all true positive cases, highlighting the ability to detect the actual positives among the dataset. The F1 Score, a harmonic mean of precision and recall, provides a single metric that balances the two, giving a more comprehensive measure of the model's performance when precision and recall are important.

Equations detailing these metrics are provided: Equation 1 for Accuracy, Equation 2 for Precision, Equation 3 for Recall, and Equation 4 for F1 Score. These equations form the basis for the performance summary presented in the Classification Report, offering a quantitative analysis of the model's capabilities during the training phase.

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

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

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

$$F1\ Score = \frac{Precision*Recall}{Precision+Recall} \quad (4)$$

The results in the equation contain a summary of the model training process that can be used as information to determine whether the model used can provide good performance on the data provided. The next stage in the model evaluation process is using the confusion matrix. The next stage in the model evaluation process is using a confusion matrix. Confusion Matrix is one method that can be used to determine the model evaluation results after the training process. This evaluation process uses data on test data as a source of testing. Each image in the test data will go through a classification process to determine how much data can be classified correctly and incorrectly. The results obtained from this process are variables used to calculate the accuracy, precision, recall, and F1-Score values. There are four indicators in the Confusion Matrix. True Positive, True Negative, False Positive, and False Negative show the data classified correctly and incorrectly according to the category [25]. A detailed visualization of the confusion matrix can be seen in Figure 6.

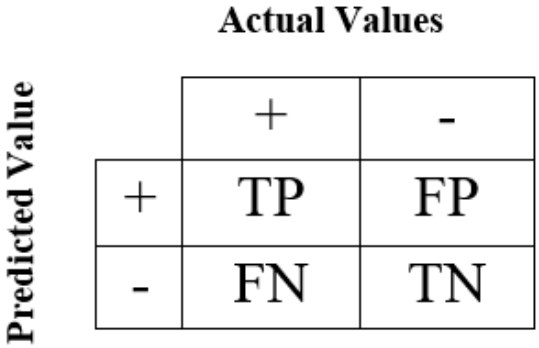

Figure 6. Confusion Matrix

Figure 6 is a visualization of the Confusion matrix. In the classification report and confusion matrix, there are several variables, such as:

TP (True Positive): Shows the number of normal images that can be classified correctly.
TN (True Negative): Shows the number of normal images classified as signs of pneumonia.
FP (False Positive): Shows the number of pneumonia images that can be classified correctly.
FN (False Negative): This shows the number of pneumonia images classified as normal.

RESULT AND DISCUSSION

The model in Figure 5 will be trained on the dataset provided, with 100 epochs divided into two stages. In the first stage, the model training is carried out on the freeze layer condition on the base model, ResNet50V2, with 50 epochs. In the second stage, a fine-tuning process is carried out by changing the setting of some of the base model layers to unfreeze and retrain for 50 epochs.

Testing the model that has been built by utilizing ResNet50V2 as a base model, GlobalAveragePooling in the Pooling layer, two dense layers with several neurons of 64, 128, and a dropout layer with a weight of 0.2 in the Fully Connected Layer. This research used Adam Optimizer with a Learning Rate of 0.001. The performance results obtained by the model during the training process can be seen in Figure 7 and Figure 8. Figure 7 contains information about the graph of the accuracy of the model that has been built, and Figure 8 contains information about the loss graph of the loss graph on the model that has been built.

Figures 7 and 8 are the results of graphical representations that illustrate how the model's performance can be used as a source of information to find out whether the model can produce good performance during the training process. Figure 7 specifically illustrates information about the accuracy results presented in a graphical illustration, while Figure 8 illustrates information about the loss graph results presented in a graphical illustration. Both graphs have the same axis structure, with the x-axis indicating the number of epochs during model training and the y-axis depicting the accuracy and loss values corresponding to the results of each epoch during the model training process.

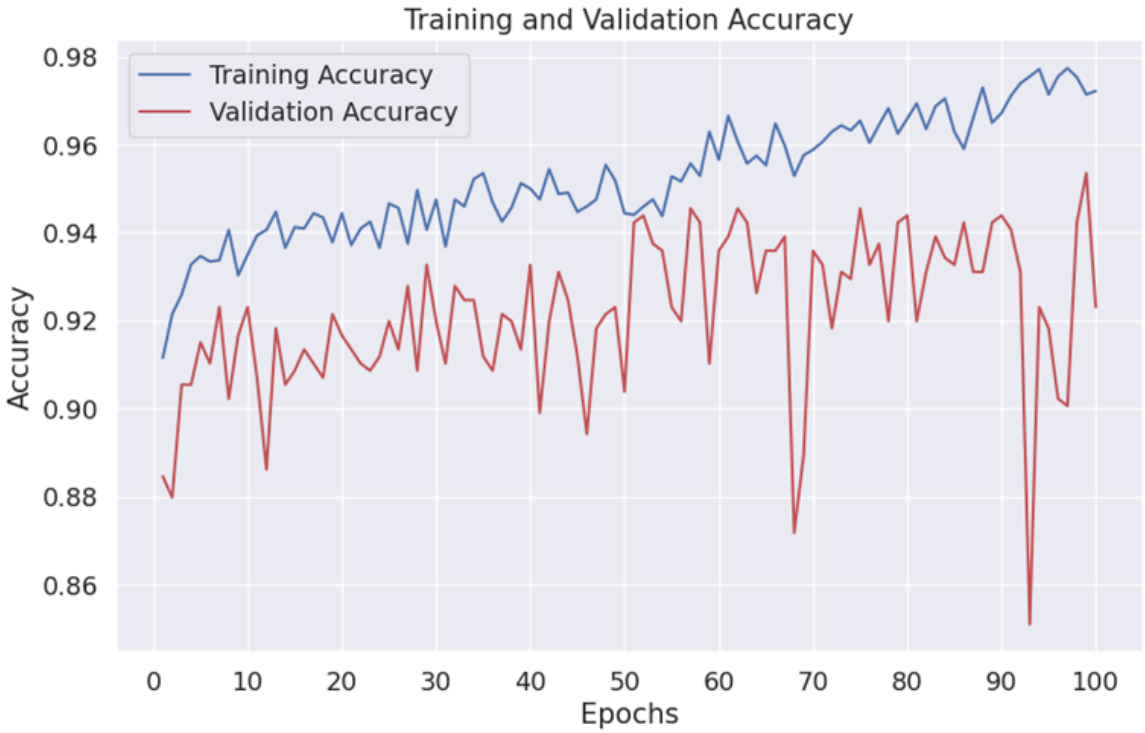

Figure 7. Accuracy Graph

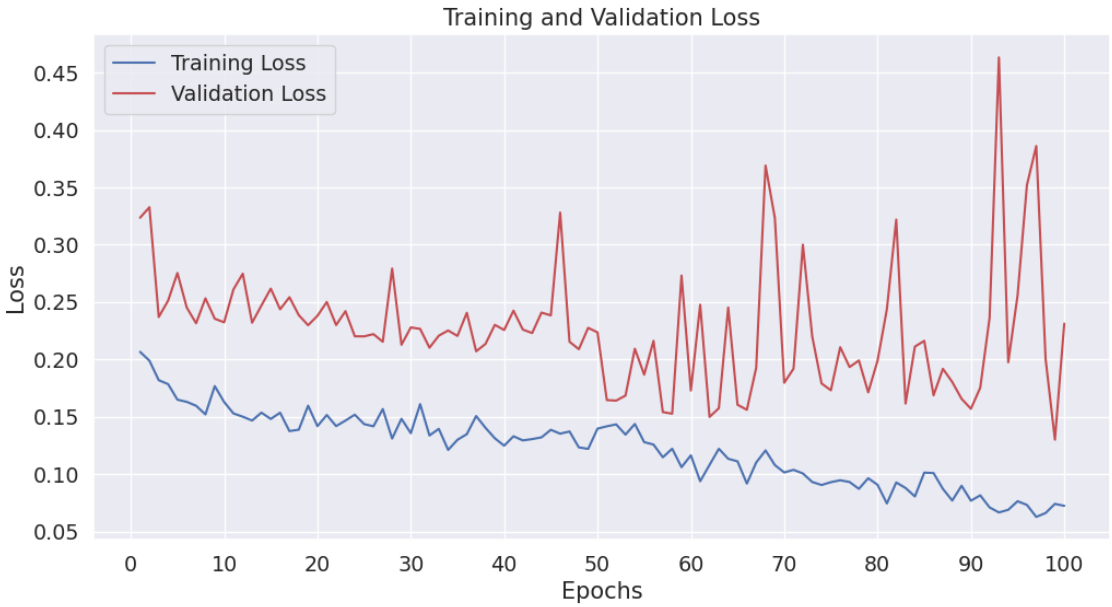

Figure 8. Loss Graph

The accuracy results depicted in Figure 7 indicate that the model achieved high accuracy throughout the training process. This consistent performance demonstrates the model's capacity to learn effectively. Notably, there is a marked increase in accuracy from the beginning to the end of the training iterations, underscoring the model's ability to improve over time.

In parallel, the loss graph results shown in Figure 8 reveal that the model attained a favorable loss value, indicative of its proficiency in minimizing error. The visual representation highlights that the model maintained a low loss value, which progressively decreased from the start to the conclusion of the training iterations.

The combined visualization of accuracy and loss graphs leads to the conclusion that the model was constructed with a strong learning capability. This conclusion is supported by the model's ability to achieve high accuracy values and demonstrate continuous improvement in accuracy throughout the training process. Additionally, the model's ability to maintain a relatively low loss value and exhibit a decline in loss over time further confirms its effectiveness.

The evaluation stage is after knowing the accuracy and loss graph results by testing the test data using the Confusion matrix. The confusion matrix results show the performance of the model built to classify. The results obtained in this confusion matrix contain the number of image data in the test data that can be classified correctly and incorrectly on each label. The results of the confusion matrix model built in this study can be seen in Figure 9.

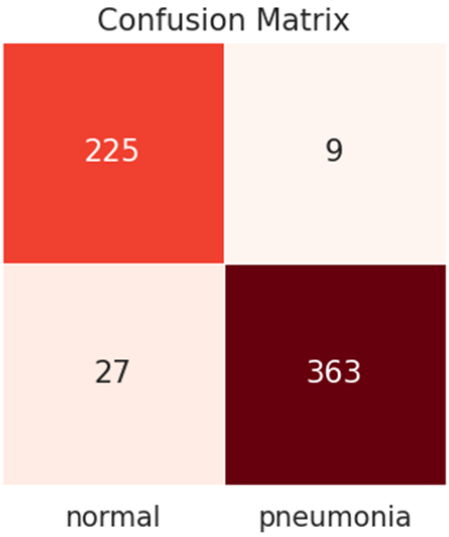

Figure 9. Confusion Matrix

Figure 9 presents a visualization of the confusion matrix, which illustrates the performance of the constructed model. The results displayed in the confusion matrix are derived from the classification process applied to the testing dataset, which comprises 624 data points. This matrix clearly indicates the model's classification capabilities by displaying the number of correctly and incorrectly classified instances.

The results reveal that the model demonstrates strong performance in classification tasks, as evidenced by the high number of images correctly identified according to their actual labels. Specifically, the confusion matrix shows that the model accurately classified 588 out of the 624 data points. This includes 225 images correctly identified as normal and 363 images correctly identified as pneumonia.

Conversely, the model misclassified 36 data points, with nine normal images incorrectly labeled and 27 pneumonia images inaccurately classified. These results underscore the model's effectiveness in distinguishing between normal and pneumonia cases while also highlighting areas for potential improvement in reducing misclassifications.

The classification report can be used to obtain accuracy, precision, recall, and F1 Score results. It is used to determine the results during the model training process. The performance results obtained through the classification report can be seen in Table 4, and the results can be compared with previous and proposed research.

Table 4. Evaluation Result

Model	Accuracy	Precision	Recall	F1-Score
Gaobo Liang, Lixin Zheng (2020) [14]	90%	89%	96%	92%
Proposed Model	**94%**	**93%**	**95%**	**94%**

The information presented in Table 4 provides a comprehensive comparison of the performance metrics between previous research and the new approach proposed by the authors. The performance results of the proposed methodology demonstrate a significant

improvement over earlier studies. Specifically, the accuracy rate has increased to 94%, while precision has reached 93%. The recall metric shows an impressive 95%, and the F1-Score stands at 94%. These enhanced performance metrics underscore the effectiveness of the research methodology introduced in this study.

The superior classification performance highlights the robustness and reliability of the proposed approach, indicating its potential for practical applications in the relevant field. The accuracy improvement suggests a more precise identification of true positive cases, while the high precision rate reflects the model's ability to correctly classify positive instances with minimal false positives. The elevated recall rate ensures that the model successfully identifies most positive cases, thus minimizing false negatives. Furthermore, the balanced F1-Score consolidates the precision and recall improvements, showcasing a well-rounded enhancement in classification performance.

These advancements in performance metrics validate the proposed research methodology and demonstrate its significant contribution to the field. The improved results pave the way for further research and development, emphasizing the potential for practical implementation and broader adoption. The findings reflect positively on the progress and impact of the authors' research, establishing a new benchmark for future studies in this domain.

CONCLUSION

The conclusions obtained from this study were obtained from the results of model testing using the ResNet50V2 architecture. The model built using the ResNet50V2 architecture during training can provide good results for the pneumonia disease classification process on the given dataset. This is indicated by an increase in accuracy during the model training process and a decrease in the loss value during the model training process. The model built produced higher accuracy than previous studies; the results obtained in this study produced an accuracy of 94%, precision reaching 93%, recall showing 95%, and F1-Score results of 94. The results obtained in the confusion matrix show good performance, and the model built can classify normal and pneumonia images in the test data with satisfactory results. The results obtained from testing the model using the confusion matrix were able to classify normal and pneumonia images correctly, with as much as 588 data. They only misclassified normal and pneumonia images, with as much as 36 data from 624 test data images.

Future research could explore using other pre-trained models and fine-tuning techniques to enhance classification performance. By experimenting with different pre-trained architectures, researchers can identify models that offer improved accuracy and efficiency for specific tasks. Additionally, fine-tuning these models on domain-specific datasets can help tailor their performance to better suit particular classification challenges.

Moreover, investigating the impact of different data augmentation strategies on model performance could provide valuable insights into optimizing training processes. Data augmentation techniques, such as rotation, scaling, cropping, and color adjustments, can significantly affect the model's robustness and generalizability. Understanding which augmentation strategies yield the best results for specific data types can help develop more effective training pipelines.

Exploring these avenues can contribute to a deeper understanding of how various techniques and methodologies can be leveraged to improve the performance of classification models, ultimately leading to more accurate and reliable outcomes in practical applications.

REFERENCES

[1] I. M. Dendi Maysanjaya, “Klasifikasi Pneumonia pada Citra X-rays Paru-paru dengan Convolutional Neural Network (Classification of Pneumonia Based on Lung X-rays Images using Convolutional Neural Network),” 2020.

[2] I. Rudan, L. Tomaskovic, C. Boschi-Pinto, and H. Campbell, “Global estimate of the incidence of clinical pneumonia among children under five years of age,” *Bull World Health Organ*, vol. 82, no. 12, pp. 895–903, 2004, doi: /S0042-96862004001200005.

[3] D. Varshni, K. Thakral, L. Agarwal, R. Nijhawan, and A. Mittal, “Pneumonia Detection Using CNN based Feature Extraction,” *Proceedings of 2019 3rd IEEE International Conference on Electrical, Computer and Communication Technologies, ICECCT 2019*, pp. 1–7, 2019, doi: 10.1109/ICECCT.2019.8869364.

[4] who.int, “Pneumonia,” https://www.who.int/. Accessed: Jan. 21, 2022. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/pneumonia

[5] R. Rahmadewi and R. Kurnia, “Klasifikasi Penyakit Paru Berdasarkan Citra Rontgen dengan Metoda Segmentasi

Sobel," *JURNAL NASIONAL TEKNIK ELEKTRO*, vol. 5, no. 1, Jan. 2016, doi: 10.20449/jnte.v5i1.174.

[6] H. Müller, N. Michoux, D. Bandon, and A. Geissbuhler, "A review of content-based image retrieval systems in medical applications - Clinical benefits and future directions," *Int J Med Inform*, vol. 73, no. 1, pp. 1–23, 2004, doi: 10.1016/j.ijmedinf.2003.11.024.

[7] I. D. Apostolopoulos and T. A. Mpesiana, "Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks," *Phys Eng Sci Med*, vol. 43, no. 2, pp. 635–640, Jun. 2020, doi: 10.1007/s13246-020-00865-4.

[8] M. Haloi, K. R. Rajalakshmi, and P. Walia, "Towards Radiologist-Level Accurate Deep Learning System for Pulmonary Screening," Jun. 2018, [Online]. Available: http://arxiv.org/abs/1807.03120

[9] H. Liu, L. Wang, Y. Nan, F. Jin, Q. Wang, and J. Pu, "SDFN: Segmentation-based deep fusion network for thoracic disease classification in chest X-ray images," *Computerized Medical Imaging and Graphics*, vol. 75, pp. 66–73, Jul. 2019, doi: 10.1016/j.compmedimag.2019.05.005.

[10] M. Z. Alom *et al.*, "A state-of-the-art survey on deep learning theory and architectures," *Electronics (Switzerland)*, vol. 8, no. 3. MDPI AG, Mar. 01, 2019. doi: 10.3390/electronics8030292.

[11] M. I. Razzak, S. Naz, and A. Zaib, "Deep learning for medical image processing: Overview, challenges and the future," in *Lecture Notes in Computational Vision and Biomechanics*, vol. 26, Springer Netherlands, 2018, pp. 323–350. doi: 10.1007/978-3-319-65981-7_12.

[12] S. Indolia, A. K. Goswami, S. P. Mishra, and P. Asopa, "Conceptual Understanding of Convolutional Neural Network- A Deep Learning Approach," in *Procedia Computer Science*, Elsevier B.V., 2018, pp. 679–688. doi: 10.1016/j.procs.2018.05.069.

[13] B. Sekeroglu and I. Ozsahin, "Detection of COVID-19 from Chest X-Ray Images Using Convolutional Neural Networks," *SLAS Technol*, vol. 25, no. 6, pp. 553–565, Dec. 2020, doi: 10.1177/2472630320958376.

[14] G. Liang and L. Zheng, "A transfer learning method with deep residual network for pediatric pneumonia diagnosis," *Comput Methods Programs Biomed*, vol. 187, Apr. 2020, doi: 10.1016/j.cmpb.2019.06.023.

[15] R. Jain, P. Nagrath, G. Kataria, V. Sirish Kaushik, and D. Jude Hemanth, "Pneumonia detection in chest X-ray images using convolutional neural networks and transfer learning," *Measurement (Lond)*, vol. 165, Dec. 2020, doi: 10.1016/j.measurement.2020.108046.

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

[17] A. Mahadar, P. Mangukiya, and T. Baraskar, "Comparison and Evaluation of CNN Architectures for Classification of Covid-19 and Pneumonia," in *Proceedings of the IEEE International Conference Image Information Processing*, Institute of Electrical and Electronics Engineers Inc., 2021, pp. 110–115. doi: 10.1109/ICIIP53038.2021.9702676.

[18] K. U. Ahamed *et al.*, "A deep learning approach using effective preprocessing techniques to detect COVID-19 from chest CT-scan and X-ray images," *Comput Biol Med*, vol. 139, Dec. 2021, doi: 10.1016/j.compbiomed.2021.105014.

[19] X. Pei *et al.*, "Robustness of machine learning to color, size change, normalization, and image enhancement on micrograph datasets with large sample differences," *Mater Des*, vol. 232, Aug. 2023, doi: 10.1016/j.matdes.2023.112086.

[20] C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," *J Big Data*, vol. 6, no. 1, Dec. 2019, doi: 10.1186/s40537-019-0197-0.

[21] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in *Proceedings of 2017 International Conference on Engineering and Technology, ICET 2017*, Institute of Electrical and Electronics Engineers Inc., Mar. 2018, pp. 1–6. doi: 10.1109/ICEngTechnol.2017.8308186.

[22] P. Kaviya, P. Chitra, and B. Selvakumar, "A Unified Framework for Monitoring Social Distancing and Face Mask Wearing Using Deep Learning: An

Approach to Reduce COVID-19 Risk," *Procedia Comput Sci*, vol. 218, pp. 1561–1570, 2023, doi: 10.1016/j.procs.2023.01.134.

[23] M. E. Karar, E. E. D. Hemdan, and M. A. Shouman, "Cascaded deep learning classifiers for computer-aided diagnosis of COVID-19 and pneumonia diseases in X-ray scans," *Complex and Intelligent Systems*, vol. 7, no. 1, pp. 235–247, Feb. 2021, doi: 10.1007/s40747-020-00199-4.

[24] R. L. Kumar, J. Kakarla, B. V. Isunuri, and M. Singh, "Multi-class brain tumor classification using residual network and global average pooling," *Multimed Tools Appl*, vol. 80, no. 9, pp. 13429–13438, Apr. 2021, doi: 10.1007/s11042-020-10335-4.

[25] M. L. Huang and Y. C. Liao, "A lightweight CNN-based network on COVID-19 detection using X-ray and CT images," *Comput Biol Med*, vol. 146, Jul. 2022, doi: 10.1016/j.compbiomed.2022.105604.