

Performance Comparison of VGG16, VGG19 and Alexnet Pre-Trained Transfer Learning Architecture Models in the Convolutional Neural Network Algorithm in Classification of Lung Disease

Fahri Aulia Alfarisi Harahap

Computer Science Department, Faculty of Mathematics and Natural Sciences,
Universitas Negeri Medan, Indonesia

Abstract.

Purpose: This study aims to comprehend the performance of transfer learning architectures VGG16, VGG19, and Alexnet in a Convolutional Neural Network for classifying lung diseases. Another objective is to determine the most superior transfer learning approach in this classification scenario.

Methods/Study design/approach: The dataset consists of 5 classes, normal lungs, pneumonia, bronchopneumonia, tuberculosis, and bronchitis. The data was sourced from Sinar Husni Deli Serdang Hospital through the radiology laboratory. The dataset was divided 80:20 for training and testing, with hyperparameters including a batch size of 32, 50 epochs, and optimization using Adaptive Momentum Optimization with a learning rate of 0.001.

Result/Findings: The research findings reveal that the VGG19 transfer learning architecture achieves the best performance with an accuracy of 59.17%, precision of 62%, recall of 59.2%, and an f-1 score of 58.8%. VGG16 ranks second with an accuracy of 55.83%, precision of 58%, recall of 55.8%, and an f-1 score of 55.2%. Alexnet has an accuracy of 49.17%, precision of 53.2%, recall of 49.2%, and an f-1 score of 50.6%. In an external test with 50 data points, VGG16 attains an accuracy of 54%, VGG19 scores 42%, and Alexnet records 46%. These models perform better in classifying normal lungs and tuberculosis compared to pneumonia, bronchopneumonia, and bronchitis.

Novelty/Originality/Value: The novelty of this research lies in analysis of lung image data demonstrates that homogeneity of RGB pixel values within a class supports transfer learning performance in classification. Conversely, heterogeneity in RGB pixel values can diminish the evaluation of that class.

Keywords: *Transfer Learning, VGG16, VGG19, Alexnet, Convolutional Neural Network, Lung Disease*

This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).



INTRODUCTION

Convolutional Neural Network (CNN) is one type of deep learning algorithm that can be utilized for disease classification and disease detection. This algorithm is employed to classify various disease types through medical image analysis. The potential of the CNN algorithm lies in its ability to handle complex and unstructured disease image data. This is evident from several studies, one of which is titled "An End-To-End Approach To Segmentation In Medical Images With CNN And Posterior-CRF"[1]. In this study, segmentation of Magnetic Resonance Imaging (MRI) image data was performed on brain stroke images, arteries within the body, and the detection of ischemic diseases in the brain. The research results indicate that employing the CNN algorithm can assist healthcare professionals in classifying humans affected by stroke or not, based on the results of brain MRI image and body arteries MRI image. Another study titled "Melanoma classification using Light-Fields with Morlet Scattering Transform and CNN: Surface Depth as a Valuable Tool to Increase Detection Rate" Discussing the classification of melanoma cancer using CNN, the research outcomes demonstrate that the CNN algorithm is capable of effectively classifying melanoma cancer. CNN excels at extracting crucial features from medical images through the implementation of convolution, pooling, and pattern recognition processes [2].

Based on information from the World Health Organization (WHO) in 2016, a significant portion of children under the age of 15 are exposed to highly polluted air, which poses risks to their health and development. According to WHO, approximately 600,000 children succumb to respiratory and urinary tract infections caused by the polluted air [3]. In the case of adults, according to the WHO, lung diseases fall under the category of Non-Communicable Diseases (NCDs), which are the leading cause of death worldwide. The WHO reports that approximately 3 million people die each year due to lung diseases, and about 64% of these deaths occur in developing countries. According to the American Cancer Society, lung cancer cases in the United States reached 236,740 in the year 2022, with 117,910 cases in males and 118,830 cases in females. Moreover, there were 130,180 deaths attributed to lung cancer [4]. In Indonesia, according to the WHO, in

the year 2019, there were 843,300 cases of Tuberculosis (TB) affecting individuals. The number of TB cases in Indonesia increased to around 845,000 in the year 2020, with the number of deaths due to TB surpassing 98,000 people. The lungs are crucial respiratory organs in humans.

Most lung issues arise from inflammation due to contaminated air carrying viruses or bacteria. This can result in various respiratory illnesses such as Pneumonia, Covid-19, Tuberculosis, Bronchitis, and others. Lung disorders can affect human airways and pose life-threatening risks if not addressed seriously. Consequently, unwanted consequences can arise, including breathing difficulties, impaired mobility, and oxygen deficiency. If not promptly treated, these conditions can lead to death [5]. According to the Head of the Radiology Laboratory at Sinar Husni Hospital, on average, there are about 15-16 patients daily who undergo chest X-ray examinations. The chest X-ray process in this laboratory takes approximately 13-15 minutes. Once the X-ray results are available, they are sent to a specialist doctor for the disease diagnosis process. The average waiting time for patients to receive their chest X-ray diagnosis results is 60 minutes. The reason for the relatively long time required for diagnosing lung diseases is due to the fact that doctors need to manually inspect and verify nodules in the lungs.

Recently, several studies have been conducted to address the issue of lung disease classification. a research titled "Pneumonia Disease Classification Using Convolutional Neural Network with Adaptive Momentum Optimization." In this study, two classes of data were used: normal lungs and pneumonia lung disease. The research findings indicate that the precision and recall values for pneumonia lung disease are 97% and 58%, respectively. For normal lungs, the precision and recall values are 70% and 98%, respectively. The overall accuracy achieved in this study is 78% [6]. Furthermore, a study titled "Pneumonia Disease Identification Based on Chest X-Ray Images Using Convolutional Neural Network [7]" investigated the identification of pneumonia disease. In this study, two classes were utilized: the normal lung class and the pneumonia lung class. The research outcomes reveal that for the normal lung class, the precision value is 93%, recall value is 75%, and f-1 score is 83%. As for the pneumonia class, the precision value is 87%, recall value is 96%, and f-1 score is 91%. The overall accuracy achieved in this study is 88%.

Therefore, based on the explanations provided above and several existing challenges, I am interested in conducting research using the convolutional neural network method to address the mentioned issues or shortcomings. This research will focus on the theme of comparing the performance of pre-trained transfer learning architecture models, namely VGG16, VGG19, and AlexNet, within the Convolutional Neural Network algorithm for lung disease classification.

METHODS

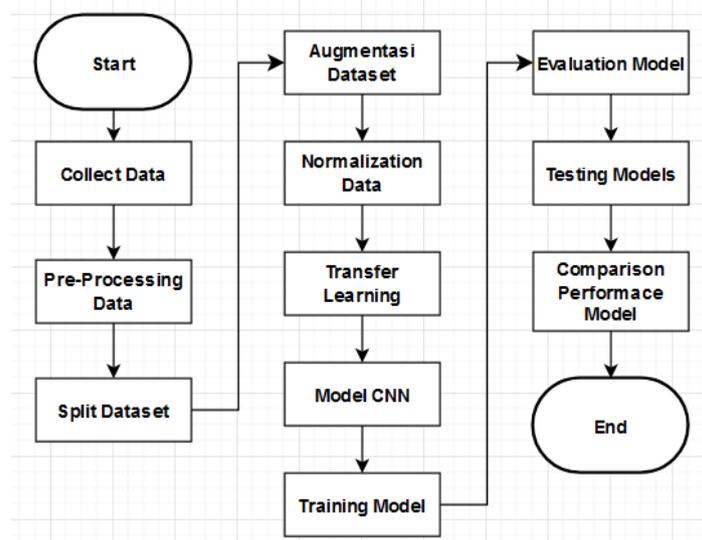


Figure 1. Research flow

Figure 1 presents, data is collected and gathered according to the existing categories or classes of data. The disease classes are determined based on the names of diseases recorded in the laboratory and radiology records of Sinar Husni Hospital, adjusted to the patients' names. Data collection is carried out manually. The collected data is then resized through image cropping. The cropping process is done with a 1:1 aspect ratio. The processed images facilitate the learning of Convolutional Neural Network (CNN) models and also

effectively eliminate noise present in the images [8]. Subsequently, the data undergoes a transformation into binary thresholding. The thresholding method is used to convert the images into binary form. Referring to the study titled "Gabor Filter Evaluation For Binary Thresholding Image Processing In Early Detection Application," the results indicate that the binary thresholding transformation process yields an average improvement of 5% compared to grayscale transformation [9]. This process enhances the differentiation of features in the images and separates objects from the background, thereby facilitating pattern recognition in the images [10]. Furthermore, all image data is resized once again to dimensions of 224x224. This resizing is done to match the pre-trained architectures of VGG16, VGG19, and AlexNet, which were trained using this size [11]. Finally, the image data is converted into pixel matrices as input for model creation.

The next step involves data normalization, achieved by dividing all pixel values of the images by 255. According to the study "Image Recognition Performance Enhancements Using Image Normalization," data normalization minimizes model complexity and aids in faster convergence during training, stabilizing activations and facilitating model learning [12]. Data normalization is performed after splitting the dataset to prevent leakage of information from the testing data into the training process, which would occur if normalization were done before dataset splitting. The subsequent phase defines the transfer learning architectures to be used: VGG16, VGG19, and AlexNet architectures. These transfer learning architectures are connected to the fully connected layers of the CNN model. The next step involves creating the CNN model, which entails designing the fully connected layers that connect to the pre-defined transfer learning architectures. The hyperparameters used for creating the CNN model are outlined in the table 1.

Table 1. Hyperparameter Model

Optimization	Adam
Learning Rate	1×10^{-3}
Epoch	50
Batch	32

Table 1 presents the hyperparameters employed in this research. The optimization technique utilized is Adaptive Momentum Optimization (Adam). Adam optimization is chosen based on the study titled "A Novel Convolutional Neural Networks Based Spinach Classification and Recognition System" [12]. This study compared Adam optimization with Rmsprop, Adagrad, Adamax, and Nadam optimizations, concluding that Adam optimization achieves the highest validation accuracy and lowest validation loss with a learning rate of 1×10^{-3} .

In this phase, the model training process takes place. The CNN model, as defined along with the hyperparameters in Table 1, is trained with a batch size of 32 and for 50 epochs. Referring to the study "Optimasi Hyperparameter CNN Untuk Klasifikasi Penyakit Padi" [13], it is concluded that a batch size of 32 yields the highest validation and testing accuracy compared to batch sizes of 8 and 16. Once the model is trained, the training and loss values are plotted. Finally, the model is saved in the h5 format. The subsequent phase involves model evaluation. In this step, the transfer learning CNN architectures (VGG16, VGG19, and AlexNet) are evaluated based on accuracy, precision, recall, and F-1 score values derived from the confusion matrices of each model architecture. Subsequently, testing is conducted using data beyond the training and testing datasets. 50 external data points are used as test data, with each class containing 10 data points. The accuracy of the model's external data testing is calculated by dividing the number of correct predictions by the total number of test data points. The final phase involves comparing the performance of the transfer learning architectures (VGG16, VGG19, and AlexNet) based on accuracy, precision, recall, and F-1 score values. This comparison is conducted to determine which transfer learning architecture yields the best performance.

RESULT AND DISCUSSION

The data used in this study are primary data obtained from human lung X-ray images acquired from the Radiology Laboratory of Sinar Husni Deli Serdang Hospital. The total number of acquired data is 1250 images. Lung X-ray images were randomly collected from various years, starting from 2010 to 2022, based on their respective classes. The data consist of 5 classes, namely 250 images of Normal, 250 images of Bronchopneumonia, 250 images of Pneumonia, 250 images of Bronchitis, and 250 images of Tuberculosis. These images have varying pixel intensity values and are in the .jpg file format. Regions in the X-ray lung images that appear white indicate the presence of diseases.

Table 2. Split dataset

Classes	Training	Testing	Validation
<i>Bronkitis</i>	192	48	10
<i>BronkoPneumonia</i>	192	48	10
<i>Normal</i>	192	48	10
<i>Pneumonia</i>	192	48	10
<i>Tuberculosis</i>	192	48	10
Total Number	960	240	50

In Table 2 the total dataset used is 1250, and the datasets used for training, testing, and validation outside of training and testing are 960, 240, and 50 respectively. The value 960 is derived from 80% of 1200, 240 is derived from 20% of 1200, and 50 is obtained by subtracting 50 from 1250. For each class in the training dataset, there are 192 image data, whereas in the testing dataset for each class, there are 48 image data. In the dataset for validation outside of training and testing for each class, there are 10 image data.

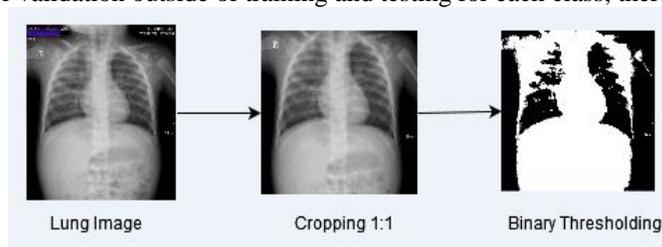


Figure 2. Processing image dataset

Figure 2 shows the pre-processing data stage was conducted on all X-ray image data by cropping the original image size to a 1:1 aspect ratio. This process was carried out to eliminate noise readings such as names, dates, hospital names, and other readings that could hinder the optimal learning process of the model. This process does not eliminate the main features of the lung images. Subsequently, the pre-processing stage was performed on each image by transforming the image into a binary thresholding form. This process can be observed in Figure 2. Image data analysis will be conducted based on the RGB values for each class. To perform the data analysis, the average calculation of the standard deviation of each RGB value from all image data and the average standard deviation of RGB within each class will be used.

Table 3. Mean standard deviation red

Class	Average Std Red
Bronchitis	65.87056910569106
Bronchopneumonia	64.943
Normal	62.07584
Pneumonia	63.609397590361446
Tuberculosis	62.990719999999996

Table 4. Mean standard deviation green

Class	Average Std Green
Bronchitis	68.23126016260163
Bronchopneumonia	65.9644
Normal	64.1456
Pneumonia	65.27502008032128
Tuberculosis	63.697599999999994

Table 5. Mean standard deviation blue

Class	Average Std Blue
Bronchitis	69.50028455284553
Bronchopneumonia	66.46992
Normal	64.75211999999999
Pneumonia	66.24148594377509
Tuberculosis	64.22888

Table 6. Mean standard deviation RGB

Class	Average Std RGB
Bronchitis	68.98138211382113
Bronchopneumonia	66.22772
Normal	64.791
Pneumonia	66.04666666666667
Tuberculosis	64.15164

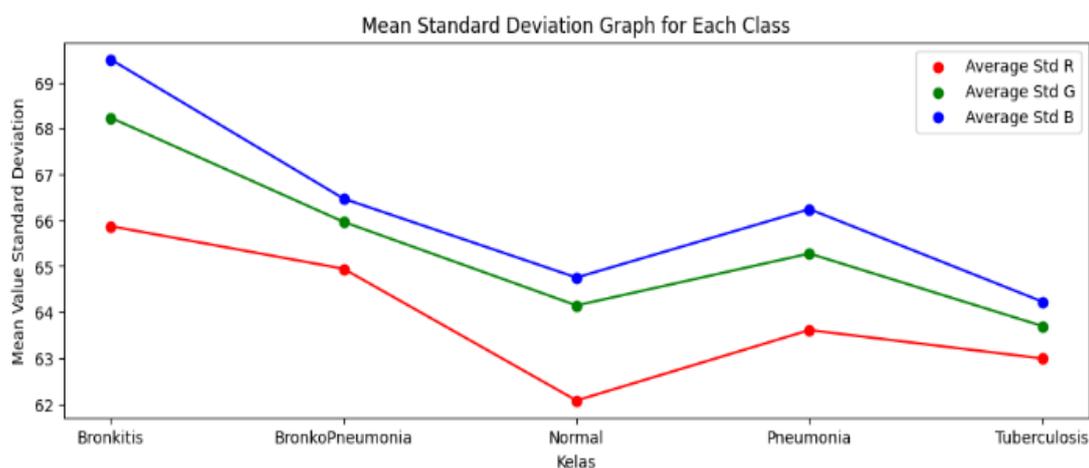


Figure 3. Line chart of mean standard deviation of rgb values each class

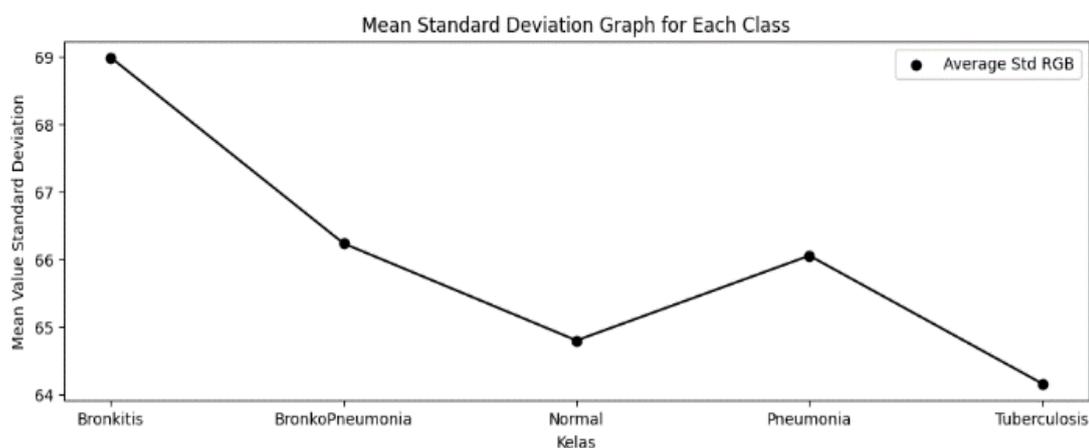


Figure 4. Line chart of mean standard deviation of RGB

Table 3, Table 4, Table 5 and Table 6 show the RGB values for each class of lung disease. Table 3 displays the standard deviation of red values, Table 4 displays the standard deviation of green values, Table

Neural Network (CNN) for the classification of lung diseases across 5 classes. The outcomes of the model training will be presented as follows:

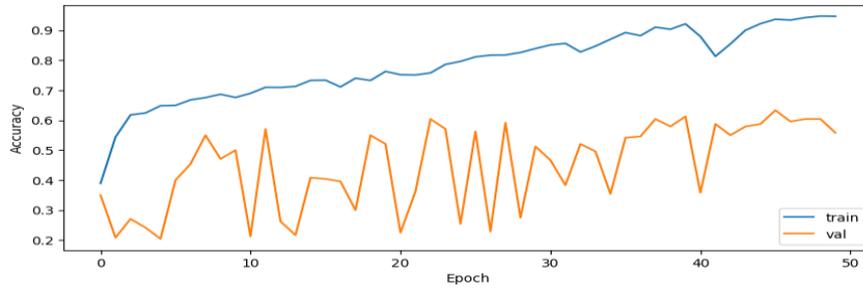


Figure 6. VGG16 accuracy

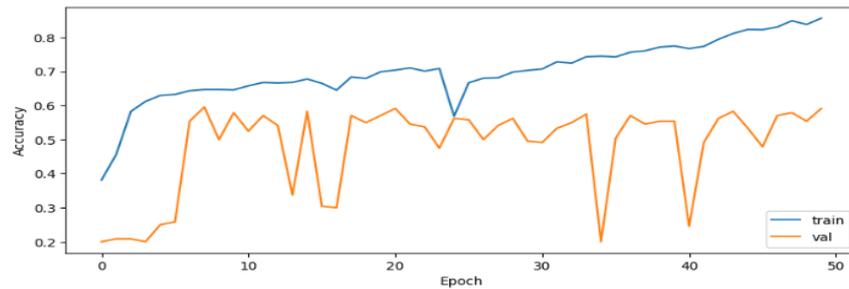


Figure 7. VGG19 accuracy

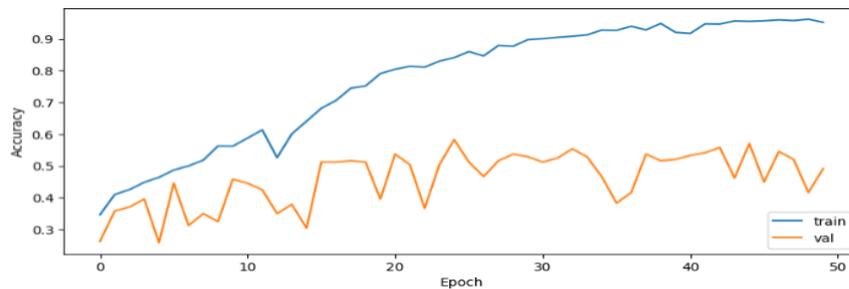


Figure 8. AlexNet accuracy

Table 7 Summary of training results for each architecture

Architecture	Accuracy	Loss	Val Accuracy	Val Loss
VGG16	94.70%	0.1480	55.83%	2.4488
VGG19	85.70%	0.3490	59.17%	1.2439
ALEXNET	95.26%	0.1622	49.17%	3.7588

Figure 6, Figure 7 and Figure 8 show graphs of the training results of models that have been trained previously, and the training results are displayed in Table 7. In Table 7 the highest training accuracy is achieved by the AlexNet architecture with a value of 95.26%, but it has the lowest validation accuracy at 49.17%. On the other hand, the highest validation accuracy is attained by the VGG19 architecture at 59.17%, while it obtains the lowest training accuracy of 85.70%. The lowest loss value is observed in the AlexNet architecture at 0.1480, whereas the lowest validation loss value is obtained by the VGG19 architecture at 1.2439. The higher training accuracy of the AlexNet architecture compared to the other two architectures is attributed to its simplicity when contrasted with the complex structures of VGG16 and VGG19. This allows the AlexNet model to undergo training faster and potentially prevents overfitting. This is evident from the validation loss values of AlexNet, which are higher than those of VGG16 and VGG19. The VGG19 architecture achieves the lowest training accuracy among VGG16, AlexNet, and itself. However, it obtains the highest validation accuracy and the lowest validation loss, indicating that the VGG19 architecture might

provide better generalization. To determine the best-performing transfer learning architecture, further evaluation using confusion matrices and classification reports is necessary for each model.

Table 8. Classification report VGG16

Classes	Precision	Recall	F-1 Score
Tuberculosis	85%	92%	88%
Bronchitis	39%	62%	48%
Normal	93%	79%	85%
Bronchopneumonia	21%	21%	21%
Pneumonia	52%	25%	34%
Accuracy		55.83%	

Table 8 it can be observed that the performance of the tested VGG16 model varies considerably across each class of data. In the Tuberculosis class, the model achieves a precision of 85%, recall of 92%, and an F1 score of 88%. Moving on to the Bronchitis class, the model's precision is 39%, recall is 62%, and the F1 score is 48%. The Normal class attains a precision of 93%, recall of 79%, and an F1 score of 85%. For the BronchoPneumonia class, the precision is 21%, recall is 21%, and the F1 score is 21%. Lastly, in the Pneumonia class, the precision is 52%, recall is 25%, and the F1 score is 34%. In this model, an accuracy value of 55.83% is obtained. This accuracy represents the testing data's (20%) accuracy against the training data (80%) during the model's training.

Table 9. Classification report VGG19

Classes	Precision	Recall	F-1 Score
Tuberculosis	75%	94%	83%
Bronchitis	48%	65%	55%
Normal	92%	69%	79%
Bronchopneumonia	31%	35%	33%
Pneumonia	64%	33%	44%
Accuracy		59.17%	

Table 9 it is evident that the performance of the tested VGG19 model exhibits notable variation across each class of data. In the Tuberculosis class, the model achieves a precision of 75%, recall of 94%, and an F1 score of 83%. Transitioning to the Bronchitis class, the model's precision is 48%, recall is 65%, and the F1 score is 55%. The Normal class attains a precision of 92%, recall of 69%, and an F1 score of 79%. For the BronchoPneumonia class, the precision is 31%, recall is 31%, and the F1 score is 31%. Lastly, in the Pneumonia class, the precision is 64%, recall is 33%, and the F1 score is 44%. In this model, an accuracy value of 59.71% is obtained. This accuracy represents the testing data's (20%) accuracy against the training data (80%) during the model's training.

Table 10. Classification report Alexnet

Classes	Precision	Recall	F-1 Score
Tuberculosis	89%	65%	75%
Bronchitis	34%	25%	29%
Normal	83%	81%	82%
Bronchopneumonia	22%	31%	26%
Pneumonia	38%	44%	41%
Accuracy		49.17%	

Table 10 it is evident that the performance of the tested AlexNet model demonstrates significant

variation across each class of data. In the Tuberculosis class, the model achieves a precision of 89%, recall of 65%, and an F1 score of 75%. Moving on to the Bronchitis class, the model's precision is 34%, recall is 25%, and the F1 score is 29%. The Normal class attains a precision of 83%, recall of 81%, and an F1 score of 82%. For the BronkoPneumonia class, the precision is 22%, recall is 31%, and the F1 score is 26%. Lastly, in the Pneumonia class, the precision is 38%, recall is 44%, and the F1 score is 41%. In this model, an accuracy value of 49.17% is obtained. This accuracy represents the testing data's (20%) accuracy against the training data (80%) during the model's training.

According to the research "A REVIEW ON EVALUATION METRICS FOR DATA CLASSIFICATION EVALUATIONS," to obtain a performance evaluation of a classification model, one can calculate the average precision, recall, and F1 score across all classes present [15].

$$Average_{Precision} = \frac{\sum_{i=1}^l \frac{tp_i}{tp_i + fp_i}}{l} \quad (1)$$

$$Average_{Recall} = \frac{\sum_{i=1}^l \frac{tp_i}{tp_i + fn_i}}{l} \quad (2)$$

$$Average_{F1-Score} = \frac{\sum_{i=1}^l 2 \times \frac{tp_i}{tp_i + fp_i} \times \frac{tp_i}{tp_i + fn_i}}{\frac{tp_i}{tp_i + fp_i} + \frac{tp_i}{tp_i + fn_i}} \quad (3)$$

Table 11 Summary of transfer learning architecture evaluation

Architecture	Accuracy	Precision	Recall	F-1 Score
VGG16	55.83%	58%	55.8%	55.2%
VGG19	59.17%	62%	59.2%	58.8%
ALEXNET	49.17%	53.2%	49.2%	50.6%

Table 11 shows it can be observed that the VGG19 transfer learning architecture yields superior performance compared to the other two transfer learning architectures. VGG19 achieves the highest accuracy of 59.17%, surpassing VGG16's accuracy by 3.34% and AlexNet's accuracy by 10%. Furthermore, the VGG19 transfer learning architecture achieves the highest precision at 62%, which is 4% higher than VGG16's precision and 8.8% higher than AlexNet's precision. Similarly, in terms of recall, the VGG19 transfer learning architecture achieves the highest recall value of 59.2%, outperforming VGG16's recall by 3.4% and AlexNet's recall by 10%. Lastly, the highest F1 score is attained by the VGG19 transfer learning architecture, measuring 58.8%. This F1 score is 3.6% higher than VGG16's F1 score and 9.6% higher than AlexNet's F1 score.

The accuracy results of the three architectural models, VGG16, VGG19, and AlexNet, can be deemed as low due to their inability to accurately classify Bronchitis, Pneumonia, and Bronkopneumonia diseases. These models struggle to extract relevant features for the classification of these diseases, possibly due to the challenges in identifying distinct and accurate patterns within the images. Additionally, the lack of detailed information in lung class images, such as low contrast and poor resolution, makes it challenging for the models to recognize patterns clearly and precisely. Based on Tables 3 4, 5 and 6 these three models demonstrate proficient classification only for Tuberculosis and normal lung conditions.

CONCLUSION

The evaluation results of the Convolutional Neural Network models utilizing Transfer Learning with VGG16, VGG19, and AlexNet architectures yielded varying outcomes. In the Transfer Learning VGG16 architecture, an accuracy of 55.38% was achieved, accompanied by a precision of 58%, recall of 55.8%,

and an F-1 score of 55.2%. On the other hand, the Transfer Learning VGG19 architecture yielded an accuracy of 59.17%, a precision of 62%, a recall of 59.2%, and an F-1 score of 58.8%. Lastly, employing the Transfer Learning AlexNet architecture resulted in an accuracy of 49.17%, a precision of 53.2%, a recall of 49.2%, and an F-1 score of 50.6%. Notably, this research determined that the best-performing model was the Transfer Learning VGG19 architecture. Based on the analysis of lung image data, which involved calculating the standard deviation of RGB values for each class, a conclusion was drawn. It was observed that the more homogeneous the RGB pixel values within a class of lung image data, the better the performance of the transfer learning model in performing classification. Conversely, when the RGB pixel values within a class of data are more heterogeneous, the evaluation outcomes for that class tend to be less favorable.

REFERENCES

- [1] S. Chen, Z. Sedghi, F. Dubost, G. Van Tulder, and M. De, "An end-to-end approach to segmentation in medical images with CNN and posterior-CRF," *Med. Image Anal.*, vol. 76, p. 102311, 2022, doi: 10.1016/j.media.2021.102311.
- [2] P. M. M. Pereira *et al.*, "Melanoma classification using light-Fields with morlet scattering transform and CNN : Surface depth as a valuable tool to increase detection rate," *Med. Image Anal.*, vol. 75, p. 102254, 2022, doi: 10.1016/j.media.2021.102254.
- [3] B. D. Prasetyo, "LEMBAR PENGESAHAN KLASIFIKASI CITRA X-RAY PARU-PARU ANAK PNEUMONIA DAN NON-PNEUMONIA MENGGUNAKAN METODE SEGMENTASI DAN DETEKSI TEPI TUGAS AKHIR Diajukan sebagai Salah Satu Syarat untuk Memperoleh Gelar Sarjana Teknik pada Program Studi Teknik Elektro Faku," 2020.
- [4] I. Buana and D. A. Harahap, "Asbestos, Radon Dan Polusi Udara Sebagai Faktor Resiko Kanker Paru Pada Perempuan Bukan Perokok," *AVERROUS J. Kedokt. dan Kesehat. Malikussaleh*, vol. 8, no. 1, p. 1, 2022, doi: 10.29103/averrous.v8i1.7088.
- [5] R. Moch Diar, R. Y. Fu'adah, and K. Usman, "Klasifikasi Penyakit Paru-Paru Berbasis Pengolahan Citra X Ray Menggunakan Convolutional Neural Network (Classification Of The Lung Diseases Based On X Ray Image Processing Using Convolutional Neural Network)," *e-Proceeding Eng.*, vol. 9, no. 2, pp. 476–484, 2022.
- [6] L. A. Andika, H. Pratiwi, and S. S. Handajani, "Klasifikasi Penyakit Pneumonia Menggunakan Metode Convolutional Neural Network Dengan Optimasi Adaptive Momentum," *Indones. J. Stat. Its Appl.*, vol. 3, no. 3, pp. 331–340, 2019, doi: 10.29244/ijjsa.v3i3.560.
- [7] N. P. Ekananda and D. Rimirasih, "Identifikasi Penyakit Pneumonia Berdasarkan Citra Chest X-Ray Menggunakan Convolutional Neural Network," *J. Ilm. Inform. Komput.*, vol. 27, no. 1, pp. 79–94, 2022, doi: 10.35760/ik.2022.v27i1.6487.
- [8] T. Y. Liu, H. Q. Zhang, H. X. Long, J. Shi, and Y. H. Yao, "Convolution neural network with batch normalization and inception-residual modules for Android malware classification," *Sci. Rep.*, vol. 12, no. 1, pp. 1–17, 2022, doi: 10.1038/s41598-022-18402-6.
- [9] B. Baba, U. Al, A. Mandar, U. Hasanuddin, I. S. Areni, and U. Hasanuddin, "Gabor Filter Evaluation for Binary Threshold Image Processing in Early Detection Application," no. May 2020, 2018.
- [10] S. Bhahri and Rachmat, "Transformasi Citra Biner Menggunakan," *J. Sist. Inf. dan Teknol. Inf.*, vol. 7, no. 2, pp. 195–203, 2018.
- [11] F. D. Adhinata, N. Annisa, F. Tanjung, W. Widayat, and G. R. Pasfica, "Comparative Study of VGG16 and MobileNetV2 for Masked Face Recognition," vol. 7, no. 2, pp. 230–237, 2021, doi: 10.26555/jiteki.v7i2.20758.
- [12] S. Sennan, D. Pandey, Y. Alotaibi, and S. Alghamdi, "A Novel Convolutional Neural Networks Based Spinach Classification and Recognition System," *Comput. Mater. Contin.*, vol. 73, no. 1, pp. 343–361, 2022, doi: 10.32604/cmc.2022.028334.
- [13] A. Julianto, A. Sunyoto, D. Ferry, and W. Wibowo, "Optimasi Hyperparameter Convolutional Neural Network Untuk Klasifikasi Penyakit Tanaman Padi (Optimization of Convolutional Neural Network Hyperparameters for Classification of Rice Plant Diseases)", [Online]. Available: <https://www.kaggle.com/tedisetiady/leaf-rice-dis>
- [14] R. Poojary, R. Raina, and A. K. Mondal, "Effect of data-augmentation on fine-tuned CNN model performance," vol. 10, no. 1, pp. 84–92, 2021, doi: 10.11591/ijai.v10.i1.pp84-92.
- [15] H. M and S. M.N, "A Review on Evaluation Metrics for Data Classification Evaluations," *Int. J. Data Min. Knowl. Manag. Process*, vol. 5, no. 2, pp. 01–11, 2015, doi: 10.5121/ijdkp.2015.5201.