COMPARATIVE CLASSIFICATION OF LUNG X-RAY IMAGES WITH CONVOLUTIONAL NEURAL NETWORK, VGG16, DENSENET121

Muhammad Ilham Prasetya [1*)] ; Yuris Alkhalifi [2] ; Rifki Sadikin [3] ; Yan Rianto [4]

[1,3,4] Computer Science
[1,3,4] Nusa Mandiri University
[1,3,4] www.nusamandiri.ac.id
[1] 14002329@nusamandiri.ac.id*), [3] rifki.rdq@nusamandiri.ac.id, [4] yan.yrt@nusamandiri.ac.id

[2] Computer Technology
[2] Bina Sarana Informatika University
[2] www.bsi.ac.id
[2] yuris.yak@bsi.ac.id

Abstract — *Lungs are one of the organs of the human body, and lung tissue will ultimately affect human abilities. The respiratory system exchanges oxygen and carbon dioxide in the blood. Problems that often occur are polluted air quality, many bacteria that attack the lungs, and lung disease that can cause shortness of breath, mobility difficulties, and hypoxia so that if not detected immediately, it can cause death. In this regard, this study aims to compare the classification of normal lungs with those of those suffering from Cardiomegaly. The preparation of this dataset contributes to improving the quality of the disease classification system on X-ray images. CNN, VGG 16, and DenseNet methods were chosen as classification methods to ensure performance and determine which method is the best for classifying Lung Diseases. It can be concluded that by using the DenseNet121 model, X-Ray images in this research dataset get an accuracy of 67.06%. For the VGG16 model, it gets an accuracy of 68.94%, and for the CNN model, it gets the highest accuracy of 80.54%.*

Keywords: *Chest X-Ray, CNN, DenseNet121, VGG16*

Abstract — Paru-paru merupakan salah satu organ tubuh manusia, dan jaringan paru-paru pada akhirnya akan mempengaruhi kemampuan manusia. Sistem pernapasan menukar oksigen dan karbon dioksida di dalam darah. Masalah yang sering terjadi adalah kualitas udara yang tercemar, banyak bakteri yang menyerang paru-paru, dan penyakit paru-paru dapat menyebabkan sesak napas, kesulitan mobilitas, dan hipoksia, sehingga jika tidak segera terdeteksi dapat menyebabkan kematian. Sehubungan dengan hal tersebut, tujuan penelitian yaitu komparasi klasifikasi paru-paru normal dengan paru-paru yang menderita Cardiomegaly. Penyusunan dataset ini sebagai bentuk kontribusi dalam meningkatkan kualitas sistem klasifikasi penyakit pada citra X-ray. Metode CNN, VGG 16 dan DenseNet dipilih sebagai metode klasifikasi guna memastikan kinerja dan metode tersebut mana yang paling terbaik untuk melakukan klasifikasi Penyakit Paru – Paru. Dapat disimpulkan bahwa dengan menggunakan model DenseNet121, citra X-Ray pada dataset penelitian ini mendapatkan akurasi sebesar 67,06%, untuk model VGG16 mendapatkan akurasi sebesar 68,94% dan untuk model CNN mendapatkan akurasi tertinggi yakni sebesar 80,54%.

Keywords*: Lung X-Ray, CNN, DenseNet121, VGG16*

PRELIMINARY

Technological developments in image processing are currently beneficial in the health or medical world, especially for detecting a human diseases. (Cholik, 2021)(Hadianti & Riana, 2018). Currently, there are many studies that explain image processing in the world of Health to detect a disease, such as skin disease (Triyono et al., 2021) breast cancer (Hilaliyah, 2021), cervical cancer (Agustyawati et al., 2021). , eye disease (Indraswari et al., 2022), Alzheimer's disease (Phiadelvira, 2021), Dental Abscess Disease (Setiaji et al., 2018) and many more. From several studies that have been mentioned, it is stated that current technological developments are very clearly useful and help experts to process disease detection through images. Image processing plays a role as image processing using a computer, becoming a better quality image. One of the image processing operations is object recognition from a digital image. An essential process in recognizing objects that are presented visually or in the form of prints is segmentation (Gao et al., 2018)

The lungs are one of the main organs in the human body that function for the respiratory system or the respiratory process (Idatin Nikmah et al., 2019). Problems that often occur in the respiratory system are polluted air quality, so many bacteria attack the lungs, lung disorders can cause sufferers to have difficulty breathing, have a problem doing activities, and lack of oxygen so that if it is not detected quickly, it can cause death (Jaisakthi et al., 2019).

Research by (Baltruschat et al., 2018) conducted comparisons of deep learning on approaches to X-ray classification. This study used the extended model of the ResNet-50 architecture. In a systematic evaluation, using 5-fold re-sampling and multi-label loss functions, comparing the performance of different approaches to pathology classification with ROC statistics, and analyzing differences between classifiers using rank correlation. Overall, the researchers observed a fair amount of spread in the performance achieved and concluded that ResNet-38 X-ray-only, integrating non-image data yielded the best overall results. Next, a class activation map is used to understand the classification process, and a detailed analysis of the impact of non-image features is provided.

Another research by (Bharati et al., 2020) introduces that Detecting Lung Disease in a Timely is very important. In a lot of image processing and image modeling, in this study, Various forms of deep learning existing techniques including *convolutional neural network* (CNN), *vanilla neural network*, *visual geometry group based* neural network (VGG), and capsule network, are applied for the prediction of lung disease. Therefore, the researcher proposes a new deep *hybrid learning framework* combining VGG, data augmentation, and spatial transformer network (STN) with CNN. This new *hybrid* method is referred to here as VGG Data STN with CNN (VDSNet). As an implementation tool, with a complete and sampled data set, VDSNet outperforms existing methods in terms of several metrics including precision, gain, F0.5 score, and validation accuracy. For the entire dataset case, VDSNet showed a validation accuracy of 73%. At the same time, vanilla grey, vanilla RGB, hybrid CNN, and VGG, as well as the modified capsule network, had accuracy values of 67.8%, 69%, 69.5%, and 63.8%, respectively. Full data and samples, VDSNet outperforms existing methods in terms of several metrics including precision, gain, F0.5 score, and validation accuracy. For the entire dataset case, VDSNet shows a validation accuracy of 73%. At the same time, vanilla Grey, vanilla RGB, hybrid CNN and VGG, and the modified capsule network have accuracy values of 67.8%, 69%, 69.5%, and 63.8%, respectively. Each. When a sample data set is used instead of a complete data set, VDSNet requires substantially less training time at the expense of slightly lower validation accuracy.

The study was (Chamveha et al., 2020)conducted to calculate *the cardiothoracic ratio* (CTR) of *Chest X-Ray* by applying a U-Net-based *deep learning model* with VGG16 to extract lung and heart masks from chest X-rays, which segmented the U-Net model image based on pixels individually. High-speed accuracy across a wide range of segmentation using an end-to-end network of *encoder-decoders containing encoders* that perform feature extraction into output. As well as the area of the heart mask obtained. The results were obtained with a success rate of 76.5% accuracy.

As for the research that will be carried out in this study, a comparison of the classification of normal lungs with lungs suffering from *Cardiomegaly will be carried out*. The preparation of this dataset contributes to improving the quality of the *Cardiomegaly* classification system on X-ray images. The model that will be used in this study is to use the CNN, VGG16, and DenseNet121 models as a classification method to ensure performance and determine which way is the best for classifying lung disease.

MATERIALS AND METHODS

This stage is the stage that describes the method. The following are the methods or stages that will be carried out in this research.

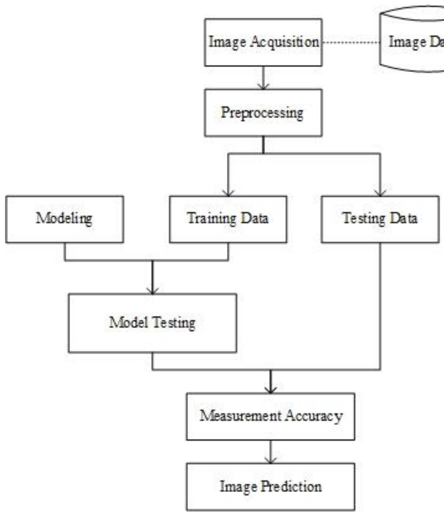

Figure 1. Research Method

Dataset

The dataset in this study was taken from the Kaggle dataset (National Institutes of Health Chest X-Ray Dataset, 2018) by taking 5826 image data by experimenting with 2 classes, namely the

Cardiomegaly class and the No Finding class. The number of images from each category can be seen in Table 1.

Table 1
Number of Images from Each Class

No	Class	Number of Images
1	Cardiomegaly	3050 Image
2	No Finding	2776 Image

As for the example image of each of these classes can be seen in Figure 2.

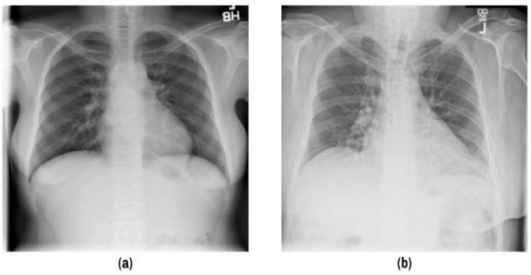

Figure 2. (a) *No finding class image* (b) *Cardiomegaly class image*

Preprocessing

At this stage, *preprocessing is carried out* to correct the data used by suppressing unwanted distortion and some important unwanted features. With a dataset of 5826 For image scaling, this research uses an image with a size of 32x32 Pixels by maintaining three-channel colors. As for the distribution of image data, the dataset will be divided into three parts, namely training data, validation data, and testing data, with details as shown in Table 2.

Table 2
Dataset

20% Testing Data	80%	
	Training Data	Data Validation
586 Image	5192 Image	50 Images

The dataset that has been preprocessed will then be randomized with *random_state* = 101, which can be seen in Figure 3.

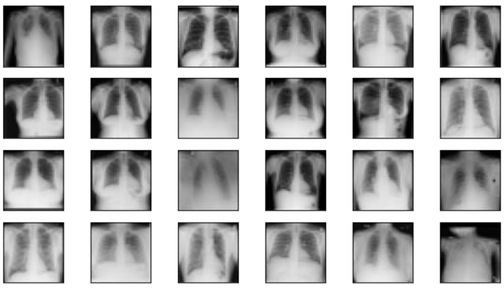

Figure 3. Image of Dataset after *Preprocessing* with *random_state*

Modeling

1. CNN models

In the CNN model, the stages of the model are made, as shown in Figure 4.

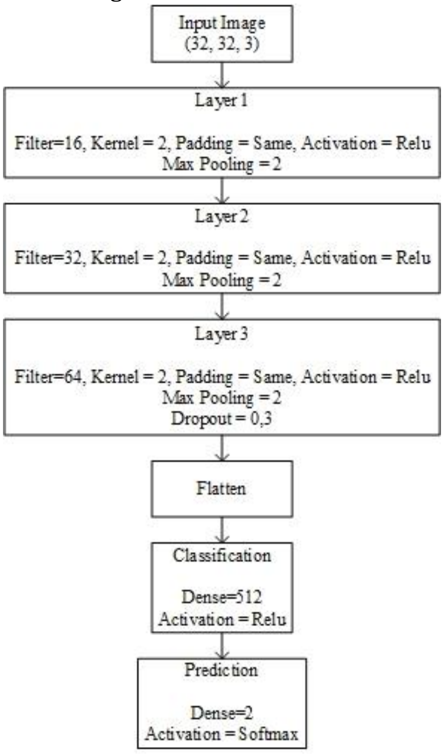

Figure 4. CNN Model

2. VGG16 and DenseNet121 . models

In this model, the stages of the model are made, as shown in Figure 5.

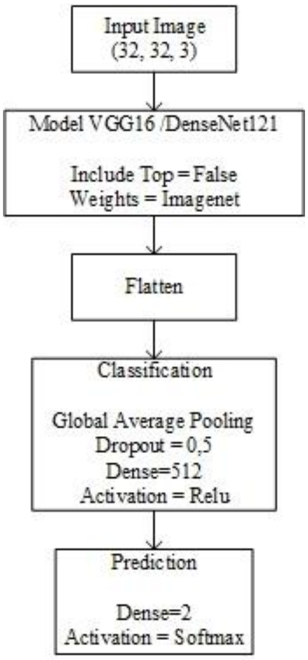

Figure 5. VGG16 and DenseNet121. Models

Techno Nusa Mandiri: Journal of Computing and Information Technology
Vol. 19, No. 1 March 2022 | DOI https://doi.org/10.33480/techno.v19i1.3010

Testing Stage

In this section, the data has been processed into the tested model. After being tested with the model, the accuracy value, *loss value, and the Confusion Matrix* truth table are obtained. After testing the DenseNet121 and VGG16 architectures to get accuracy in the classification of lung diseases after getting the results, a comparison is made with the test model. After comparing the accuracy of the three models, image prediction will also be made using the model with the highest accuracy.

RESULTS AND DISCUSSION

Convolutional Neural Network Model

After the data is entered into the CNN model, the accuracy and loss values in the training data and valid data can be seen in Figures 6 and 7.

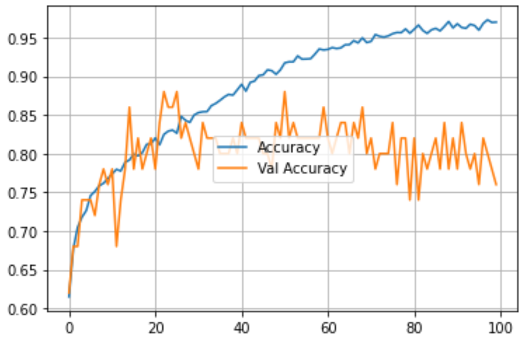

Figure 6. CNN Accuracy Value

The higher the value, the better the experimental results for the accuracy value. From Figure 6 above, it is known that the accuracy value of the training data (blue) tends to increase in each epoch. Still, the accuracy value of the validation data (yellow) increases to epoch 10; the accuracy tends to rise and fall until the last epoch. And for the accuracy value at the end of the epoch, the accuracy value is 96.99% for the training data, and the accuracy value is 76% for the validation data.

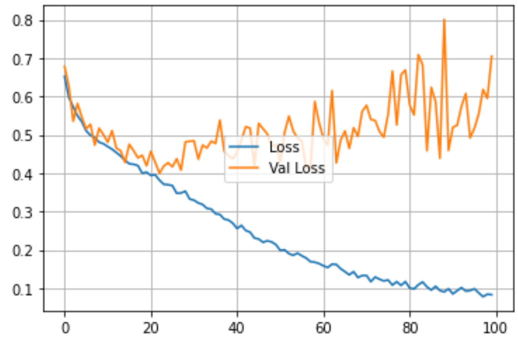

Figure 7. CNN *Loss Value*

For the loss value, the lower the value, the better the experimental results. From Figure 7 above, it is known that the *loss value* in the training data (blue) tends to decrease in each *epoch*, but the lost *weight* in the validation data (yellow) decreases until *epoch* 20. The *loss value* tends to rise and fall until the last *epoch*. And for the *loss value* at the end of the *epoch, the loss* value is 0.0830 for the training data, and the accuracy value is 0.7050 for the validation data.

Then from the CNN model on the *epoch value,* which has the lowest *loss value* that was successfully stored, namely at the 23rd *epoch with a value of 0.3997, the next step is to test the data on the testing data.* After the data testing process, the accuracy value was 80.54%.

Transfer Learning Model

DenseNet 121

After the data is entered into the DenseNet121 model, the accuracy and *loss values* in the training data and valid data are obtained, as shown in Figures 8 and 9.

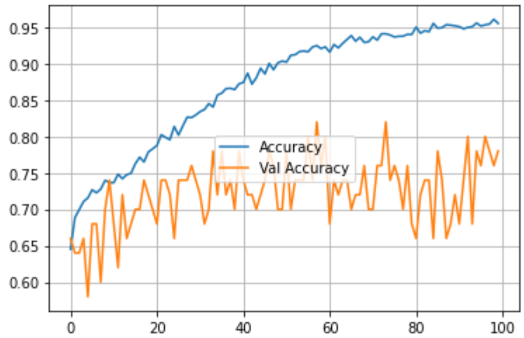

Figure 8. DenseNet121 Accuracy Value

From Figure 8 above, it is known that the accuracy value of the training data (blue color) tends to increase in each *epoch. Still, the* accuracy value of the validation data (yellow color) tends to rise and fall until the last *epoch*. And for the accuracy value at the end of the *epoch,* the accuracy value is 95.54% for the training data, and the accuracy value is 78% for the validation data.

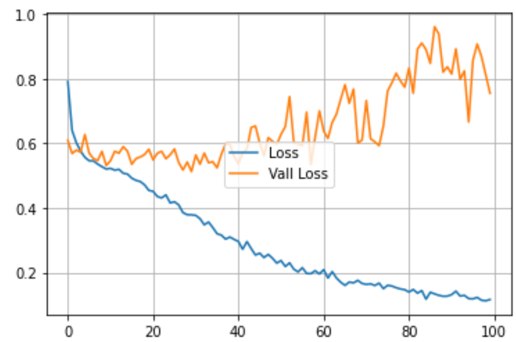

Figure 9. DenseNet121 *Loss Value*

From Figure 9 above, it is known that the *loss value* in the training data (blue) tends to decrease in each *epoch*, but the *loss value* in the validation data (yellow) tends to rise and fall until the last *epoch*. And for the *loss value* at the end of the *epoch, the loss* value is 0.1165 for the training data, and the accuracy value is 0.7555 for the validation data.

Then from the DenseNet121 model on the *epoch value with* the lowest *loss value* that was successfully stored, namely the 97th *epoch with a value of 0.0019, the next step is to test the data-on-data testing.* After the data testing process, the accuracy value was 67.06%.

VGG16

After the data is entered into the VGG16 model, the accuracy and *loss values* in the training data and valid data can be seen in Figures 10 and 11.

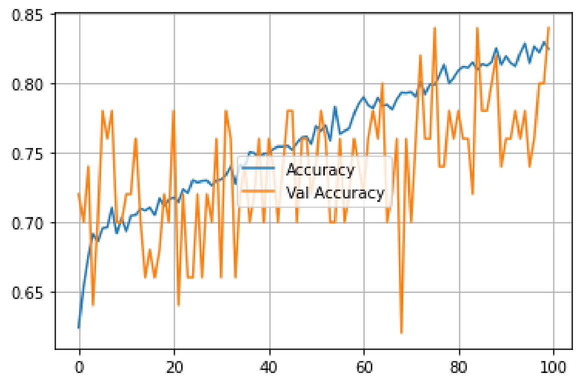

Figure 10. VGG16. Accuracy Value

From Figure 10 above, it is known that the accuracy value of the training data (blue color) tends to increase in each *epoch. Still, the* accuracy value of the validation data (yellow color) tends to rise and fall until the last *epoch*. And for the accuracy value at the end of the *epoch,* the accuracy value is 82.48% for the training data, and the accuracy value is 84% for the validation data.

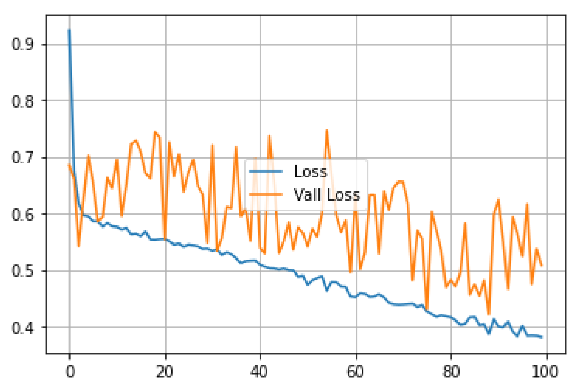

Figure 11. VGG16. *Loss Value*

From Figure 11 above, it is known that the *loss value* in the training data (blue) tends to decrease in each *epoch*, but the *loss value* in the validation data (yellow) tends to rise and fall until the last *epoch*. And for the *loss value* at the end of the *epoch, the loss* value is 0.3815 for the training data, and the accuracy value is 0.5082 for the validation data.

Then from the VGG16 model, the *epoch value with* the lowest *loss value* was successfully saved, namely at the 89th *epoch with a value of 0.4216. The next step is to test the data-on-data testing.* After the data testing process, the accuracy value is 68.94%.

Comparison Value Accuracy

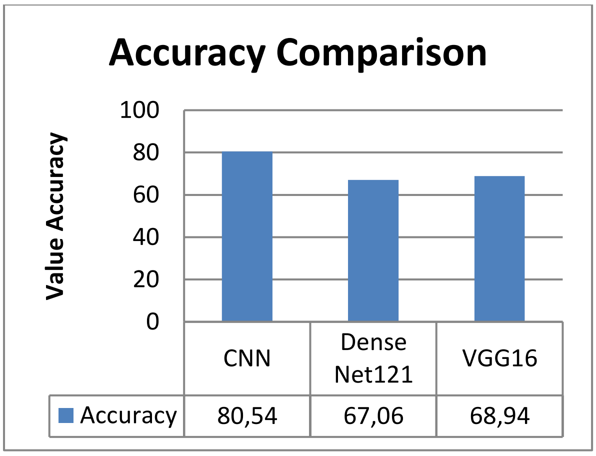

Figure 12. Comparison Graph of Accuracy Values

From Figure 12 above, it is known that the accuracy value of the testing data for the CNN model is 80.54%, the DenseNet121 model is 67.06%, and the VGG16 model is 68.94%.

Image Prediction

The X-Ray image will be predicted using the best model among the three proposed models, namely the CNN model. The results of image prediction can be seen in Figure 13.

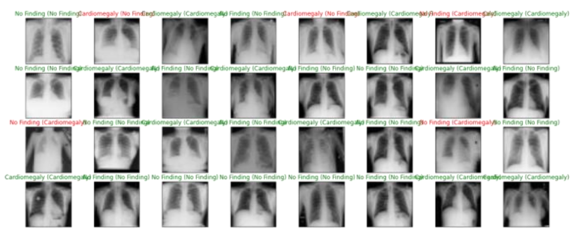

Figure 13. Prediction of X-Ray Image

In Figure 13, it is known that the results of image prediction using the CNN model, from a total of 32 images that were predicted, 27 images were predicted correctly, and five images that were expected to be wrong.

Vol. 19, No. 1 March 2022 | DOI https://doi.org/10.33480/techno.v19i1.3010

CONCLUSION

This study proposes Chest X-ray to classify lung diseases by conducting several research stages starting from image input, preprocessing stage, conducting data modeling using Conventional CNN, VGG16 and DenseNet121, then entering the data training stage, data testing, model testing, measurement accuracy and the last is the stage of predicting the image. so as to get the accuracy of each - each model test. It can be concluded that by using the DenseNet121 model, X-Ray images in this research dataset get an accuracy of 67.06%, for the VGG16 model it gets an accuracy of 68.94% and for the CNN model it gets the highest accuracy of 80.54%.

REFERENCE

Agustyawati, D. N., Fauzi, H., & Pratondo, A. (2021). PERANCANGAN APLIKASI DETEKSI KANKER SERVIKS MENGGUNAKAN METODE CONVOLUTIONAL NEURAL NETWORK APPLICATION DESIGN OF SERVIKS CANCER DETECTOR BASED USING CONVOLUTIONAL NETWORK. *E-Proceeding of Engineering*, *08*(04).

Setiaji, R. A., Hidayat, B., & Suhardjo. (2018). SINTESIS PENELITIAN DETEKSI PENYAKIT ABSES PADA GIGI MANUSIA MELALUI CITRA PERIAPIKAL RADIOGRAF DOMAIN SPASIAL. *EProceedings of Engineering*, *05*(03).

Baltruschat, I. M., Nickisch, H., Grass, M., Knopp, T., & Saalbach, A. (2018). Comparison of deep learning approaches for multi-label chest X-Ray classification. *ArXiv*, 1–10.

Bharati, S., Podder, P., & Mondal, M. R. H. (2020). Hybrid deep learning for detecting lung diseases from X-ray images. *Informatics in Medicine Unlocked*, *20*. https://doi.org/10.1016/j.imu.2020.100391

Chamveha, I., Promwiset, T., Tongdee, T., Saiviroonporn, P., & Chaisangmongkon, W. (2020). Automated cardiothoracic ratio calculation and cardiomegaly detection using deep learning approach. *ArXiv*, 1–11. https://arxiv.org/abs/2002.07468

Cholik, C. A. (2021). *Perkembangan Teknologi Informasi Komunikasi ICT dalam Berbagai Bidang*. *2*(2), 2746–1209.

Gao, M., Bagci, U., Lu, L., Wu, A., Buty, M., Shin, H. C., Roth, H., Papadakis, G. Z., Depeursinge, A., Summers, R. M., Xu, Z., & Mollura, D. J. (2018). Holistic classification of CT attenuation patterns for interstitial lung diseases via deep convolutional neural networks. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging and Visualization*, *6*(1), 1–6. https://doi.org/10.1080/21681163.2015.1124249

Hadianti, S., & Riana, D. (2018). Segmentasi Citra Bemisia Tabaci Menggunakan Metode K-Means. *Seminar Nasional Inovasi Dan Tren (SNIT)*.

Hilaliyah, P. K. (2021). *Deteksi Dini Kanker Payudara Pada Citra Histopatologi Menggunakan Metode Convolution Neural Network (CNN)*.

Idatin Nikmah, Z., Kurnia Aditya, S., & Gustri Wahyuni, E. (2019). Aplikasi Web Untuk Pendeteksi Penyakit Paru-Paru Menggunakan Metode Certainty Factor. *Seminar Nasional Informatika Medis*.

Indraswari, R., Herulambang, W., & Rokhana, R. (2022). Deteksi Penyakit Mata Pada Citra Fundus Menggunakan Convolutional Neural Network (CNN) Ocular Disease Detection on Fundus Images Using Convolutional Neural Network (CNN). *TechnoCom*, *21*(2), 378–389. https://www.kaggle.com/datasets/jr2ngb/cataractdataset

Jaisakthi, S. M., Mirunalini, P., Thenmozhi, D., & Vatsala. (2019). Grape leaf disease identification using machine learning techniques. *ICCIDS 2019 - 2nd International Conference on Computational Intelligence in Data Science, Proceedings*. https://doi.org/10.1109/ICCIDS.2019.8862084

National Institutes of Health Chest X-Ray Dataset. (2018). *NIH Chest X-rays*. Kaggle. https://www.kaggle.com/nih-chest-xrays/data

Phiadelvira, B. Y. (2021). *Klasifikasi Kanker Serviks berdasarkan Citra Kolposkopi menggunakan Convolutional Neural Network (CNN) Model Alexnet*.

Triyono, L., Nur, A., Thohari, A., Hestiningsih, I., & Yobioktabera, A. (2021). KLASIFIKASI PENYAKIT KULIT MENGGUNAKAN METODE CONVOLUTIONAL NEURAL NETWORK. *Sentrikom*, *04*(01), 37–44.