

Classification of Breast Cancer with Transfer Learning on Convolutional Neural Network Models

Mesrawati Hulu¹⁾, Bayu Angga Wijaya²⁾, Resel³⁾, Nestina Halawa⁴⁾, Angki Angkota Tarigan⁵⁾

^{1,2,3,4,5)} Universitas Prima Indonesia

¹⁾mesrahulu8@gmail.com, ²⁾bayuanggawijaya@unprimdn.ac.id, ³⁾shabirresel@gmail.com,

⁴⁾nestinahalawa08@gmail.com, ⁵⁾angkiangkotatrg@gmail.com

Submitted : Jun 26, 2024 | **Accepted :** Jul 18, 2024 | **Published :** July 19, 2024

Abstract: Breast cancer is a serious medical condition and a leading cause of death among women. Early and accurate diagnosis is crucial for improving patient outcomes. This study explores the use of Convolutional Neural Networks (CNNs) with Transfer Learning using DenseNet121 and ResNet50 models to enhance breast cancer classification via mammography. Transfer Learning enables CNN models to leverage knowledge learned from larger datasets such as ImageNet to improve performance on specific breast cancer datasets. The dataset comprised medical images with three breast variations: benign, malignant, and normal, totaling 531 data points. Data was split with a 70% training and 30% validation ratio. Two CNN models, AlexNet and ResNet50, were evaluated to compare their performance in classifying these breast cancer types. The experimental results show that AlexNet achieved a training accuracy of 98.01%, while ResNet50 achieved 64.07%. AlexNet demonstrated superior performance in identifying complex patterns in mammography images, resulting in more accurate classification of different breast cancer types. These findings highlight the potential of deep learning applications to support more precise and effective medical diagnostics for breast cancer. This research contributes significantly to the development of AI technologies in healthcare aimed at improving early detection of breast cancer. The implications of this study could expand our understanding of Transfer Learning applications in medical contexts, driving further advancements in this field to enhance patient care and prognosis.

Keywords: Transfer Learning; Convolutional Neural Network; AlexNet; ResNet; Classification

INTRODUCTION

Cancer is a term referring to a group of diseases characterized by uncontrolled cell proliferation that can invade surrounding tissues (Khandezamin et al., 2020). The growth of these abnormal cells can form a mass or lump, called a tumor, and the process of spreading to other parts of the body is called metastasis (Fitri et al., 2024). Cancer can occur in various parts of the body and can have different types, depending on the type of cells involved and the location where the cancer develops (Rahayuwati et al., 2020). One common form of cancer is breast cancer (Hidayat et al., 2024).

Cancer is a form of cancer that arises from the proliferation of breast cells and is commonly found in women (Hu et al., 2021), although it can also occur in men (Momenimovahed & Salehiniya, 2019). In Indonesia, breast cancer ranks second in the number of deaths, accounting for approximately 20,052 lives or about 1.41% (Fauzi et al., 2020). The cancer is caused by the uncontrolled growth of malignant cells within the breast tissue (Bustamam et al., 2019), which can then spread to other parts of the body (Harafani, 2020). The number of fatal cases is expected to continue increasing over time if appropriate measures are not implemented (Singh, 2019). In the context faced by medical professionals, the increase in fatality rates can be attributed to the complexity of analyzing heterogeneous and variable medical data, thus limiting the ability of medical professionals to make accurate classifications. Therefore, a more progressive approach to handling is needed. By integrating data with AI (Artificial Intelligence) technology, it can enhance the accuracy of precise classification in diagnosing individual patients. In recent years, several studies have been conducted exploring deep learning methods that are

*Bayu Angga Wijaya



This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

considered effective in classifying breast cancer. Deep learning is an AI (Artificial Intelligence) technique that replicates the principle of the mind's activity in extracting significant patterns from vast datasets (Yudistira, 2021).

In the conducted research (Natakusumah & Ernastuti, 2022) an accuracy of 53-54% was achieved. In the previous research, researchers utilized a dataset of 400 breast images classified into 4 classes, labeled by pathology experts, and experimented with CNN14, CNN42, and CNN84. However, the accuracy was suboptimal in the previous study. Therefore, in the upcoming research, the method of transfer learning will be applied to the Convolutional Neural Network for breast cancer classification. Convolutional Neural Network (CNN) is a neural network architecture specifically designed for image processing tasks and pattern recognition in spatial data. With convolutional and pooling layers, CNNs effectively extract hierarchical features from image data, starting from local patterns to their ability to understand hierarchical features at various levels of abstraction (Iswantoro & Handayani UN, 2022). In image processing, transfer learning becomes an essential element, and combining CNN with transfer learning allows the model to leverage extensive knowledge about image features, overcome limitations of training data, and achieve good performance, even with relatively small datasets.

From the context provided, it can be concluded that CNN (Convolutional Neural Network) has demonstrated good performance. Therefore, the research aims to conduct different experiments by applying transfer learning to the CNN algorithm. The hope is that by combining data and implementing transfer learning on CNN, higher accuracy can be achieved compared to previous studies. This is expected to assist medical professionals in classifying breast cancer more accurately, provide efficiency in the diagnosis process, and ultimately enable faster and more effective treatment for patients.

LITERATURE REVIEW

The following are several literature reviews of research that have been found by the researchers. The research conducted by (Putra et al., 2022) regarding the classification of lung cancer patients using the ANN method. The research results show that ANN achieved a training accuracy of 92.79% with a precision of 86.98%, as well as a test accuracy of 95.12% with a precision of 90.23%. The research conducted by (Kholil et al., 2022) on the classification of poultry diseases using feces images with CNN showed high prediction rates for several poultry diseases. The study results indicated the following prediction accuracies using CNN with feces images: 95.40% for coccidiosis, 94% for healthy condition, 90.21% for Newcastle disease, and 96.50% for pullorum disease. The research conducted by (Roslidar et al., 2022) developed the BreaCNet model for breast thermal image classification using CNN to distinguish between normal and abnormal images. Testing showed that the application was able to classify images in less than 1 second with an accuracy of 85%. The study conducted by (Astuti et al., 2021) researched the enhancement of breast cancer detection by combining Naïve Bayes and Forward Selection, resulting in an increased accuracy rate of up to 96.49%. The research conducted by (Aji mahesa, 2022) investigated the analysis of breast cancer histology image classification using Ensemble CNN, demonstrating improved performance compared to individual models. With the ensemble approach, the balanced accuracy reached 86%, and the F1-score was 0.8682 for the VGG16 + MobileNet combination. Research by (Hartono et al., 2023) compared classification algorithms for breast cancer prediction using Ensemble Learning Bagging with data from SEER NCI on Kaggle.com, consisting of 16 attributes and 4025 records. The comparison results revealed that Random Forest was the best-performing algorithm in terms of accuracy. Ensemble Learning Bagging techniques improved algorithm performance, with an average increase of 1%. Incorporating additional algorithmic models such as AdaBoost, Logistic Regression, XGBClassifier, LGBMClassifier, and ExtraTreeClassifier, along with Ensemble Learning Heterogenous techniques, alongside Machine Learning development processes using the Streamlit framework, can enhance breast cancer prediction.

METHOD

The study employs Convolutional Neural Network (CNN) algorithms to classify breast cancer, utilizing deep learning methods for image data processing. Researchers focus on categorizing images into malignant, benign, or normal categories using CNN, emphasizing experimentation within the deep learning framework.

Work Procedure

To achieve the evaluation results, the following procedures will be implemented by applying CNN with Transfer Learning to perform breast cancer classification. The steps of the working procedure for the conducted research are as follows.

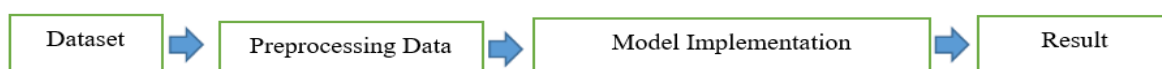


Fig. 1 Work Procedure

In Fig 1, the diagram illustrates the workflow procedure conducted in the research. The following is an explanation of the procedure.

*Bayu Angga Wijaya



This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Dataset

A dataset is a collection of data consisting of various pieces of information, observations, or facts that are organized and structured. In this study, the required dataset involves breast cancer data in the form of images, which will be the focus of the research. The image-based dataset used originates from <https://www.kaggle.com/datasets/anissaapril/kanker-payudara/>, which serves as the subject of the study.

Preprocessing Data

In this stage, the image data is converted into CSV format. Then, the data is cleaned to remove any empty data. After that, data normalization and data splitting will be performed. The split data will be trained by the model.

Model Implementation

In this step, a model is constructed to be used on preprocessed data. Here, the architecture of the model will be defined. Once the model architecture is created, the model will be trained with data.

Result

This is the final stage of all stages. In the result stage, all data have been trained and tested. The data will be evaluated with the transfer learning model from Convolutional Neural Network (CNN), which will yield accuracy results. After obtaining the accuracy, the Convolutional Neural Network model will classify the image data. Then, the image data will be categorized into a confusion matrix.

Tools and Materials

In the conducted research, the tool used for development and experimentation based on Python is Google Colab. Meanwhile, the materials used are image data obtained from an open-source platform, Kaggle, which includes labels for malignant, benign, and normal, as shown in Figure 1.

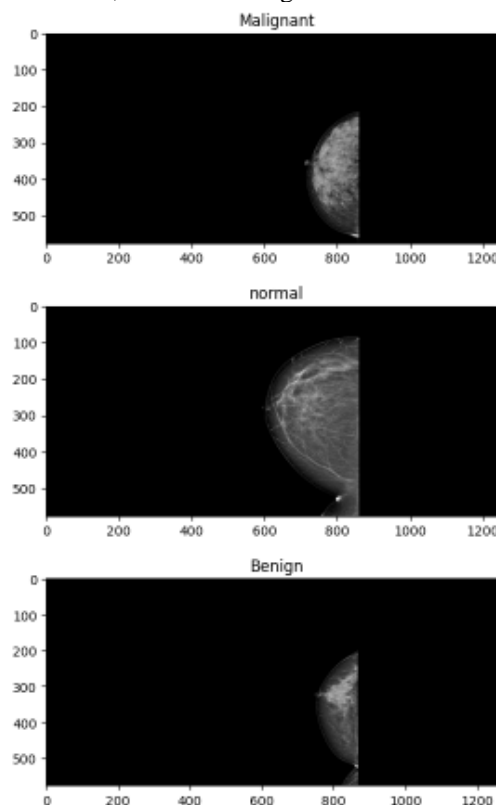


Fig. 2 The image data consists of breast cancer images

The proposed model

The research utilized transfer learning approach on AlexNet and ResNet50 network models to classify inputted breast cancer data

*Bayu Angga Wijaya



This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

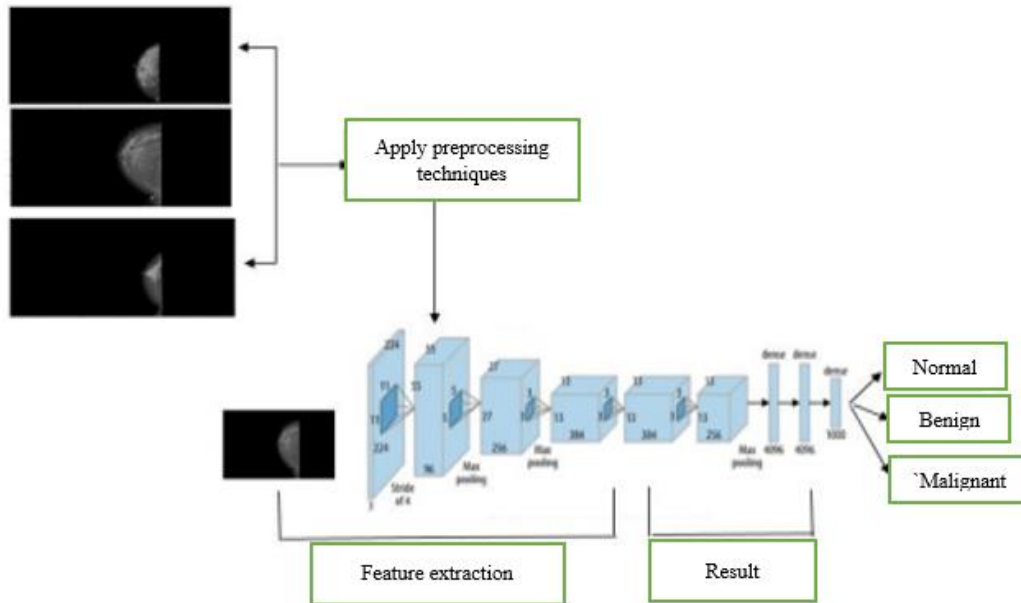


Fig. 3 AlexNet Model Framework

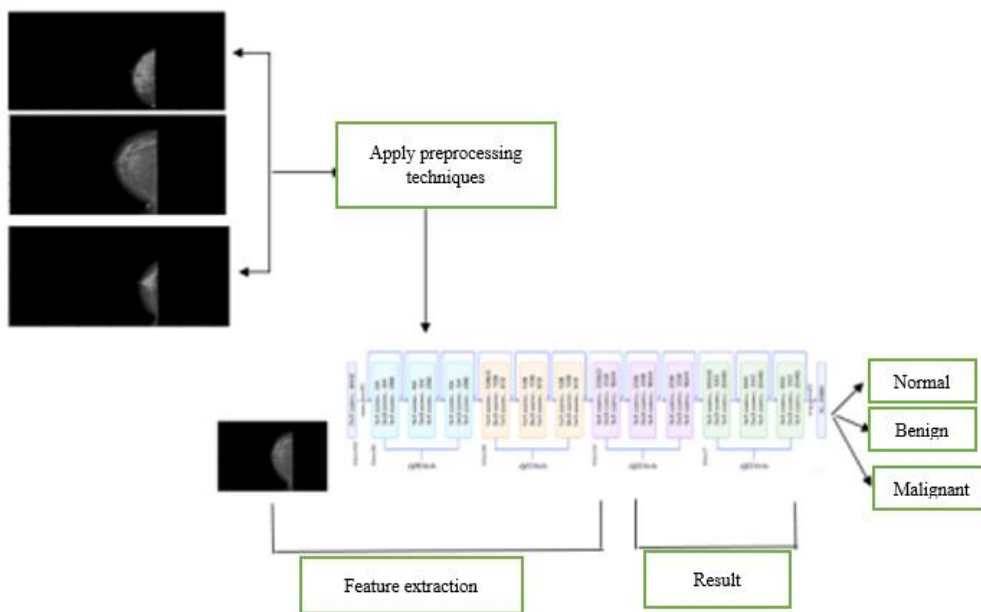


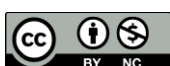
Fig. 4 ResNet50 Model Framework

Figures 2 and 3 depict the framework of the proposed transfer learning model. Transfer learning testing is an approach where a dataset is used as a starting point to address similar data problems. This method involves adjusting and updating parameters from the initial point to match the characteristics of the new dataset; the main goal is to leverage previously gained knowledge in learning and avoid starting the process from scratch. The research utilized transfer learning networks from Convolutional Neural Networks (CNN), namely the AlexNet and ResNet50 models.

RESULT

The purpose of this phase is to provide a visual understanding of the image dataset used in the research. One example image from each class (label) will be selected to represent the different categories within the dataset. The researchers downloaded the dataset from Kaggle, which consists of 531 images categorized into three labels: benign, malignant, and normal. These images will be used for training and validating a Convolutional Neural Network (CNN) model.

*Bayu Angga Wijaya



The research utilizes Google Colab as the platform for code editing and conducting machine learning and data science experiments. The dataset is directly downloaded from Kaggle via the Kaggle API integrated with Google Colab. After downloading, the dataset is used for testing and training a Convolutional Neural Network (CNN) model. The researchers employ transfer learning models such as AlexNet and ResNet50. In developing the deep learning model, data is crucial for training the model. Therefore, the data is split into two groups: training data for model training and validation data for model testing, with a ratio of 70% training and 30% validation. Pre-processing steps, including resizing images to 227 x 227 pixels, rotating images, and removing noise to sharpen the focus on test image data, are performed. All models in the research are trained using the Adam optimizer. The results of the overall model evaluation using transfer learning are shown in the figure below. Each epoch represents one iteration through the entire training dataset. In each epoch, the neural network is trained using the training data and then tested using the validation data to evaluate its performance.

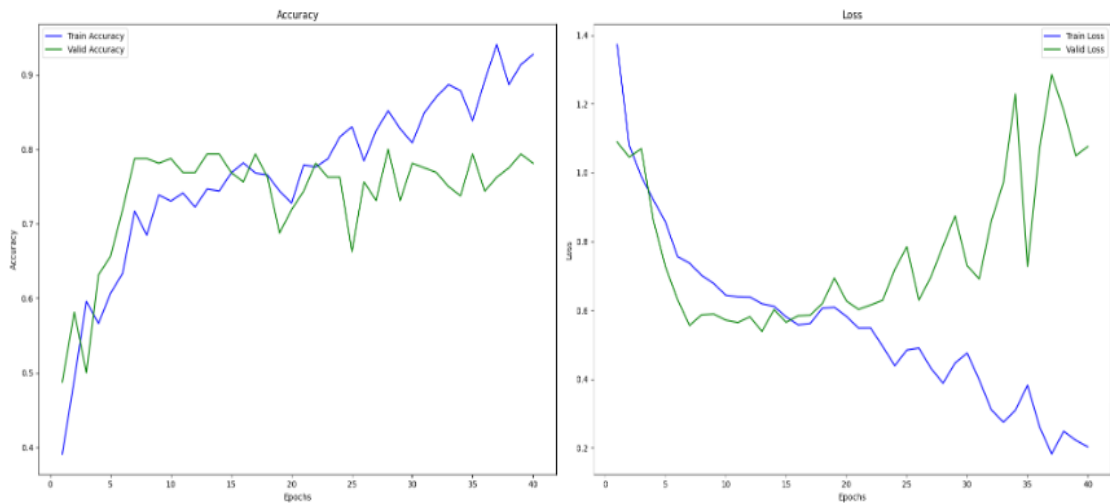


Fig. 1 Accuracy and Loss Graph of AlexNet

based on Figure 5 in the model training with 40 epochs, it is observed that the training accuracy reaches 0.9801 with a loss value of 0.2034. However, the validation accuracy is only 0.7812 with a validation loss of 0.5812.

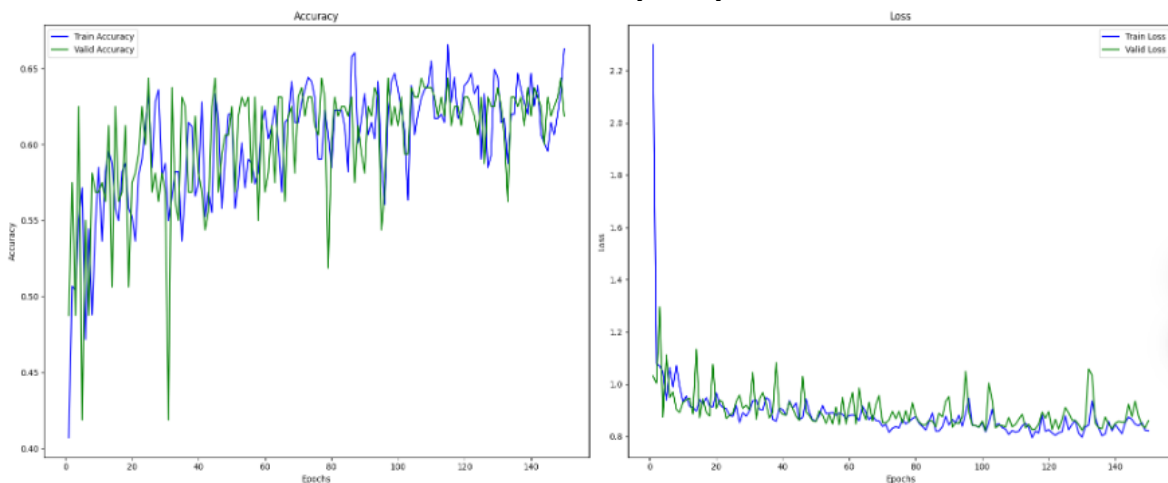


Fig. 2 Accuracy and Loss Graph of ResNet50

Based on Figure 6 in the model training with 150 epochs, the final training accuracy value is 0.6407 with a loss value of 0.8224. Meanwhile, for the validation value, the last accuracy is 0.6201 with a validation loss of 0.8603.

Based on the experiments conducted, the selection of the number of epochs in AlexNet and ResNet50 depends on the model's complexity, dataset size, and desired convergence rate. AlexNet, which is a relatively simpler model, achieves good performance with 40 epochs. With its simpler structure, AlexNet tends to converge faster compared to more complex models. Conversely, ResNet50, with a deeper and more complex structure, requires more time to converge. Therefore, ResNet50 uses 150 epochs to train its model. The use of different numbers of epochs is based on balancing the computational time required, the need for regularization, and performance goals

*Bayu Angga Wijaya



on the specific dataset. Thus, the selection of the number of epochs is an important aspect in developing an effective model.

After implementing the model and training it on the data, the next step is to examine the evaluation results of the model's performance with the visualization of the confusion matrix. Below are the confusion matrices for the 2 models.

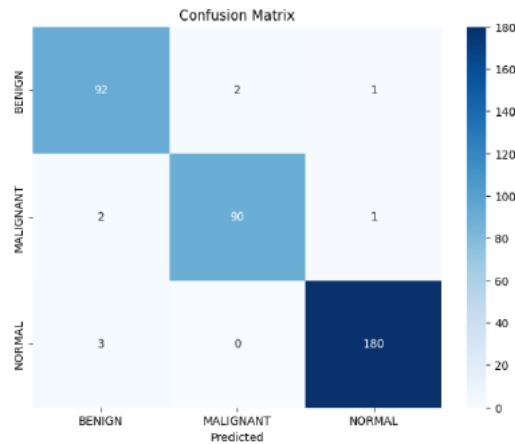


Fig. 3 Confusion Matrix of AlexNet

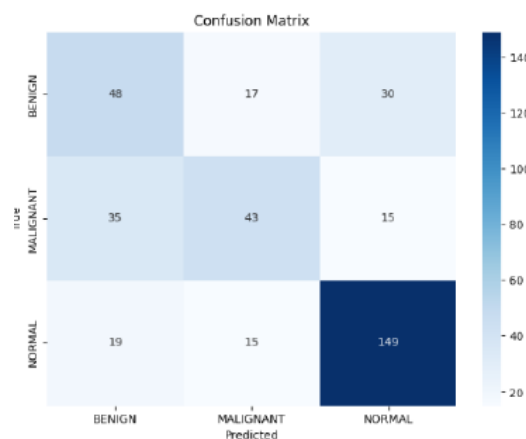


Fig. 4 Confusion Matrix of ResNet50

Based on Figure 7, it is apparent that the Confusion Matrix of the AlexNet Model shows quite promising results in classifying medical images. For the benign category, there are 92 images classified correctly with only 3 misclassifications. Meanwhile, for the malignant category, there are 90 images correctly classified with just 3 misclassifications as well. As for the normal category, the model successfully classified 180 images correctly, with only 3 misclassifications.

Moving on to Figure 8, it illustrates the Confusion Matrix of the ResNet50 model, where we can observe that 48 benign images are predicted correctly while 47 are misclassified. Similarly, for malignant images, 43 are correctly predicted while 50 are misclassified. Meanwhile, in the normal category, 149 images are predicted correctly while 24 are misclassified. These results indicate the performance of the ResNet50 model in classifying medical images, showing varying degrees of accuracy across different categories.

To calculate the accuracy from the Fig 7 and Fig 8 confusion matrix provided, we can use the following formula for accuracy:

$$Accuracy = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}}$$

Fig. 9 Formula Accuracy

*Bayu Angga Wijaya



Accuracy for AlexNet 98.1%. thus, the accuracy can be calculated as:

$$Accuracy = \frac{362}{369} = 0.981$$

Fig. 10 Formula Accuracy ResNet50

The explanation Fig 10 number of correct predictions = 92 (benign) + 90 (malignant) + 180 (normal) = 362 and total number of predictions = 92 + 2 + 1 + 3 + 90 + 1 + 0 + 0 + 180 = 369. Next, regarding the accuracy formula of ResNet50

$$Accuracy = \frac{240}{371} = 0.647$$

Fig. 11 Formula Accuracy ResNet50

The explanation Fig 11 number of correct predictions = 48 (benign) + 43 (malignant) + 149 (normal) = 240 and total number of predictions = 48 + 17 + 30 + 35 + 43 + 15 + 19 + 15 + 149 = 371

Table 1. Model Accuracy Comparison

Model	Training Data		Validation Data	
	Loss	Accuracy	Loss	Accuracy
AlexNet	0.2034	0.9801	0.5812	0.7812
ResNet50	0.8224	0.6407	0.8603	0.6201

The Table 1 contains the training results of two models, namely AlexNet and ResNet50, using training and validation data. The evaluated models are AlexNet and ResNet50. Training data shows the loss and accuracy values on the training data, while validation data displays the loss and accuracy values on the validation data. For AlexNet, the loss on the training data is 0.2034 with an accuracy of 0.9801, while on the validation data, the loss is 0.5812 with an accuracy of 0.7812. Meanwhile, for ResNet50, the loss on the training data is 0.8224 with an accuracy of 0.6407, and on the validation data, the loss is 0.8603 with an accuracy of 0.6201. This table provides an overview of the relative performance of both models on the training and validation data, where AlexNet demonstrates better performance in terms of accuracy and loss on both datasets.

DISCUSSIONS

In the exploration of medical image classification, selecting the right model plays a crucial role in obtaining accurate results. In the previous stage, the analysis of the Confusion Matrix and accuracy comparison tables for each model has been explained. Now, let's delve deeper into the performance comparison between two main models, namely AlexNet and ResNet50. Through the matrix report, we will discuss a comparison of the Confusion Matrix results from both models to observe how they perform in classifying medical images. Figures can be seen in Fig 9 and Fig 10.

```

Classification Report:
      precision    recall  f1-score   support

   BENIGN       0.95      0.97      0.96         95
  MALIGNANT     0.98      0.97      0.97         93
    NORMAL     0.99      0.98      0.99        183
    
```

Fig. 12 Classification Report Results for AlexNet Model

```

Classification Report:
      precision    recall  f1-score   support

   BENIGN       0.42      0.50      0.46         38
  MALIGNANT     0.55      0.27      0.36         44
    NORMAL     0.73      0.87      0.80         78
    
```

Fig. 13 Classification Report Results for ResNet50 Model

*Bayu Angga Wijaya



From the visualized figure 9 and 10, it is evident that AlexNet outperforms ResNet50 in classifying breast cancer. Here, based on the classification report results of the AlexNet model, it can be seen that the precision for benign is 95%, recall is 97%, f1-score is 96%, with a support of 95. For malignant, the precision is 98%, recall is 97%, f1-score is 97%, with a support of 183. Additionally, for normal, the precision is 99%, recall is 98%, f1-score is 99%, with a support of 183. Meanwhile, the classification report results for the ResNet50 model show that the precision for benign is 47%, recall is 51%, f1-score is 49%, with a support of 95. For malignant, the precision is 57%, recall is 46%, f1-score is 51%, with a support of 93. Additionally, for normal, the precision is 77%, recall is 81%, f1-score is 79%, with a support of 183. Up to this point, the research has not achieved the expected level of success in developing Convolutional Neural Network (CNN) models using transfer learning with AlexNet and ResNet50. After training stages with 371 training data, both models were tested with 160 validation data to evaluate their performance accuracy. However, the results indicate that AlexNet outperforms ResNet50.

CONCLUSION

The study applies Transfer Learning approach using Convolutional Neural Network (CNN) to classify breast images into categories of breast cancer or normal based on mammography images. The utilized Transfer Learning models are AlexNet and ResNet50. From the analysis results, it is evident that AlexNet outperforms with an accuracy rate of 97,57%, while ResNet50 achieves only 65% accuracy. In this case, the accuracy of the AlexNet model is already optimal, whereas ResNet50's accuracy is not yet optimal. For the ResNet50 model, improvements can be made by increasing the volume of data in the dataset and adding additional layers during model training.

REFERENCES

- Aji mahesa, G. (2022). Klasifikasi Citra Histologi Kanker Payudara Menggunakan Metode Ensemble CNN. *Jurnal Repositor*, 4(3), 373–384. <https://doi.org/10.22219/repositor.v4i3.1497>
- Astuti, L. W., Saluza, I., Faradilla, F., & Alie, M. F. (2021). Optimalisasi Klasifikasi Kanker Payudara Menggunakan Forward Selection pada Naive Bayes. *Jurnal Ilmiah Informatika Global*, 11(2). <https://doi.org/10.36982/jiig.v11i2.1235>
- Bustamam, A., Bachtiar, A., & Sarwinda, D. (2019). Selecting features subsets based on support vector machine-recursive features elimination and one dimensional-naïve bayes classifier using support vector machines for classification of prostate and breast cancer. *Procedia Computer Science*, 157, 450–458. <https://doi.org/10.1016/j.procs.2019.08.238>
- Fauzi, A., Supriyadi, R., & Maulidah, N. (2020). Deteksi Penyakit Kanker Payudara dengan Seleksi Fitur berbasis Principal Component Analysis dan Random Forest. *Jurnal Infortech*, 2(1), 96–101. <https://doi.org/10.31294/infortech.v2i1.8079>
- Fitri, R., Nurman, M., & Ningsih, N. F. (2024). *Terapi Guided Imagery Untuk Menurunkan Skala Nyeri Pada Klien Di Ruang Dahlia Rsud Arifin Achmad. 1.*
- Harafani, H.-. (2020). Forward Selection pada Support Vector Machine untuk Memprediksi Kanker Payudara. *Jurnal Infortech*, 1(2), 131–139. <https://doi.org/10.31294/infortech.v1i2.7398>
- Hartono, R., Sumaryana, Y., & Nurfaizi, A. (2023). Analisa Perbandingan Kinerja Algoritma Klasifikasi Untuk Prediksi Penyakit Kanker Payudara. *Jurnal Teknologi Informasi*, 7(1), 116–124. <http://jurnal.una.ac.id/index.php/jurti/article/view/3418>
- Hidayat, R., Khadafi, M., & Davi, M. (2024). *Klasifikasi Tumor Payudara Pada Citra Ultrasonografi Menggunakan Multi-fitur Tekstur dan Support Vector Machine. IX(2)*, 8805–8811.
- Hu, C., Hart, S. N., Gnanaolivu, R., Huang, H., Lee, K. Y., Na, J., Gao, C., Lilyquist, J., Yadav, S., Boddicker, N. J., Samara, R., Klebba, J., Ambrosone, C. B., Anton-Culver, H., Auer, P., Bandera, E. V., Bernstein, L., Bertrand, K. A., Burnside, E. S., ... Couch, F. J. (2021). A Population-Based Study of Genes Previously Implicated in Breast Cancer. *New England Journal of Medicine*, 384(5), 440–451. <https://doi.org/10.1056/nejmoa2005936>
- Iswantoro, D., & Handayani UN, D. (2022). Klasifikasi Penyakit Tanaman Jagung Menggunakan Metode Convolutional Neural Network (CNN). *Jurnal Ilmiah Universitas Batanghari Jambi*, 22(2), 900. <https://doi.org/10.33087/jiubj.v22i2.2065>
- Khandezamin, Z., Naderan, M., & Rashti, M. J. (2020). Detection and classification of breast cancer using logistic regression feature selection and GMDH classifier. *Journal of Biomedical Informatics*, 111(October),

*Bayu Angga Wijaya



This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

103591. <https://doi.org/10.1016/j.jbi.2020.103591>

- Kholil, M., Waspada, H. P., & Akhsani, R. (2022). Klasifikasi Penyakit Infeksi Pada Ayam Berdasarkan Gambar Feses Menggunakan Convolutional Neural Network. *SINTECH (Science and Information Technology) Journal*, 5(2), 198–204. <https://doi.org/10.31598/sintechjournal.v5i2.1179>
- Momenimovahed, Z., & Salehiniya, H. (2019). Epidemiological characteristics of and risk factors for breast cancer in the world. *Breast Cancer: Targets and Therapy*, 11, 151–164. <https://doi.org/10.2147/BCTT.S176070>
- Natakusumah, G. P., & Ernastuti, E. (2022). Implementasi Metode CNN Multi-Scale Input dan Multi-Feature Network untuk Dugaan Kanker Payudara. *JOINTECS (Journal of Information Technology and Computer Science)*, 7(2), 43. <https://doi.org/10.31328/jointecs.v7i2.3637>
- Putra, T. D., Utami, E., & P.Kurniawan, M. (2022). Klasifikasi penderita kanker Paru Paru Menggunakan Algoritma Artificial Neural Network (ANN). *Explore*, 12(2), 13. <https://doi.org/10.35200/explore.v12i2.568>
- Rahayuwati, L., Rizal, I. A., Pahria, T., Lukman, M., & Juniarti, N. (2020). Pendidikan Kesehatan tentang Pencegahan Penyakit Kanker dan Menjaga Kualitas Kesehatan. *Media Karya Kesehatan*, 3(1), 59–69. <https://doi.org/10.24198/mkk.v3i1.26629>
- Roslidar, R., Syahputra, M. R., Muharrar, R., & Arnia, F. (2022). Adaptasi Model CNN Terlatih pada Aplikasi Bergerak untuk Klasifikasi Citra Termal Payudara. *Jurnal Rekayasa Elektrika*, 18(3), 185–192. <https://doi.org/10.17529/jre.v18i3.8754>
- Singh, B. K. (2019). Determining relevant biomarkers for prediction of breast cancer using anthropometric and clinical features: A comparative investigation in machine learning paradigm. *Biocybernetics and Biomedical Engineering*, 39(2), 393–409. <https://doi.org/10.1016/j.bbe.2019.03.001>
- Wijaya, B. A., Perisman Jaya Gea, Gea, A. D. ., Alvianus Sembiring, & Christian Mitro Septiano Hutagalung. (2023). Satellite Images Classification using MobileNet V-2 Algorithm. *Sinkron : Jurnal Dan Penelitian Teknik Informatika*, 7(4), 2316-2326. <https://doi.org/10.33395/sinkron.v8i4.12949>
- Yudistira, N. (2021). Peran Big Data dan Deep Learning untuk Menyelesaikan Permasalahan Secara Komprehensif. *EXPERT: Jurnal Manajemen Sistem Informasi Dan Teknologi*, 11(2), 78. <https://doi.org/10.36448/expert.v11i2.2063>

*Bayu Angga Wijaya



This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.