Jurnal Rekayasa Elektrika Vol. 18, No. 1, Maret 2022, hal. 35-42
ISSN. 1412-4785; e-ISSN. 2252-620X, Terakreditasi RISTEKDIKTI No. 30/E/KPT/2018
DOI: 10.17529/jre.v18i1.23255

Breast Cancer Detection in Mammography Image using Convolutional Neural Network

Farell Fahrozi, Sugondo Hadiyoso, and Yuli Sun Hariyani
School of Applied Science, Telkom University, Bandung, Indonesia
Jl. Telekomunikasi No. 1, Terusan Buahbatu - Bojongsoang, Sukapura, Bandung 40257
e-mail: sugondo@telkomuniversity.ac.id

***Abstract*—Breast cancer is one non-contagious disease that tends to increase every year. This disease occurs almost entirely in women but can also occur in men. One way to detect this disease is by observing mammography images. However, mammography images often tend to be blurry with low quality so that it is possible to detect them incorrectly. Therefore, in this study, automatic classification of breast cancer on mammographic images was carried out using the Convolutional Neural Network (CNN). This proposed system uses the VGG16 architecture with a transfer learning system. The proposed method is then optimized using Adam optimizers and RMSprop optimizers. System testing results for normal, benign, and malignant classifications obtained an accuracy value of 85% - 94%, with the highest accuracy achieved using Adam's optimizers. With this proposed system, it is hoped that it can help in the clinical diagnosis of breast cancer.**

Keywords: ***breast cancer, mammography, CNN, classification***

***Abstrak*—Kanker payudara merupakan salah satu penyakit tidak menular yang cenderung meningkat setiap tahunnya. Penyakit ini terjadi hampir seluruhnya pada wanita tetapi juga dapat terjadi pada pria. Salah satu cara untuk mendeteksi penyakit ini adalah dengan mengamati gambar mamografi. Namun, gambar mamografi seringkali cenderung buram dengan kualitas rendah sehingga memungkinkan untuk mendeteksinya secara tidak benar. Oleh karena itu, pada penelitian ini dilakukan klasifikasi otomatis kanker payudara pada citra mamografi menggunakan *Convolutional Neural Network* (CNN). Sistem yang diusulkan ini menggunakan arsitektur VGG16 dengan sistem transfer learning. Metode yang diusulkan kemudian dioptimalkan menggunakan pengoptimal Adam dan pengoptimal RMSprop. Hasil pengujian sistem untuk klasifikasi normal, jinak, dan ganas diperoleh nilai akurasi sebesar 85% - 94%, dengan akurasi tertinggi dicapai dengan menggunakan *Adam's optimizer*. Dengan sistem yang diusulkan ini, diharapkan dapat membantu dalam diagnosis klinis kanker payudara.**

Kata kunci: ***kanker payudara, mamografi, CNN, klasifikasi***

I. Introduction

Breast cancer is a type of malignant tumor that develops in breast cells. Cancer cells will multiply faster than normal cells and accumulate, which then form lumps. Based on data presented by The Global Cancer Observatory (GCO), in 2020, there is a breast cancer rate of 44 per 100,000 population with an average death rate of 15.3 per 100,000 population [1]. This is a severe threat to the world's population. There is no sure way to prevent breast cancer. However, early detection is one effective way to get the proper treatment [2].

Mammography imaging is one of the gold standards in breast cancer detection [3]. However, high thoroughness is needed to observe tissue structure changes in medical images that may have low contrast. This really requires a high-cost time, especially if the screening is done on many images. The image processing approach in decision support systems or diagnosis has recently attracted attention as a solution to overcome these problems.

One of the popular methods in the case of image classification is the Neural Network which was later developed into the Convolutional Neural Network (CNN). Research related to neural network-based image classification is reported in the study [4]. In this study, a neural network approach was used to determine the number of calories based on food images. These studies generate an accuracy of 66% to 98%. Another research in image classification using CNN was also reported by Peryanto, et al [5]. Research by Suta, et al, also prove that the convolutional neural network method can also be used to detect a brain tumor on MRI images [6]. Amaliah designed a system to detect the location of a tumor or breast cancer with an accuracy of 88% [7].

In addition, several researchers have also tried to apply CNN to the field of medical image analysis [8-10] to help radiologists make correct judgments. To verify the effectiveness of the CNN model in clinical diagnosis.

Based on previous studies, it is showing that CNN has the potential for early detection of breast cancer based on mammographic images. CNN can divide the input matrix into tiny parts so that the detection of the resulting image

Received 01 November 2021; Revised 14 December 2021; Accepted 11 February 2022

is entirely accurate and detailed [11].

Based on this review, this study designed a system architecture that can be used in the early detection of breast cancer based on mammographic images. The results of this study are expected to get the best model or architecture from CNN to detect breast cancer on mammographic images with high accuracy. In this study, the classification of breast cancer on mammographic images is carried out on three classes, including normal, benign, and malignant breast cancer. This system is expected to be able to assist in the early detection and analysis of the severity of breast cancer.

The structure of this paper is arranged as follows; section 2 contains explanations of materials and methods that were used in the findings in this research. Section 3 had the test results of the proposed system and included them with the discussion. Meanwhile, section 4 describes the conclusions and implications of this study.

II. Material and Methods

A. Breast Cancer Dataset

The mammography images used in this study were taken from an open dataset collected by Al-Dhabyani W et al. [12]. This mammographic image was taken of the breasts of women aged 25-75 years using ultrasound imaging in 2018. Mammography images were recorded from 600 female patients. The number of mammography images is 780, with a size of 500*500 pixels in PNG format. The ground truth image is also available in this database. The images consist of three classes, including normal with 133 images, benign with 437 images, and malignant with 210 images. Figure 1 is an example of mammography images including normal, benign cancer, and malignant cancer.

B. Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) was first introduced by Yann Lecun in 1988. CNN is one of the methods that started Deep Learning. In the 1950s Hubel and Wisel conducted experiments on one area of the cat's brain, the visual cortex. Hubel and Wisel found 2 types of visual cortex, namely simple cells and complex cells. Based on these observations, in the 1980s Kunihiko Fukushima designed Neocognitron which is a Hierarchical Multilayered Neural Network model. This model has been used in several cases such as the classification of handwritten characters (Handwritten Character Recognition). This model is the inspiration for the Convolutional Neural Network. In the case of image classification, CNN receives input or input images then processed and classified into certain categories (eg planes, ships, birds, cats, cows). CNN also has architectural variations such as ResNet [13], Inception [14], Xception [15], MobileNet [16], EfficientNet [17], NASNet [18], etc. The main components contained in CNN include input layer, convolutional layer, activation layer, pooling layer, flatten, fully connected layer, softmax as shown in Fig. 2.

C. Transfer Learning

Transfer learning is a technique by utilizing a model that has been trained on a dataset. The function that is often used for transfer learning is ImageNet. ImageNet is a visual database designed for visual object recognition purposes.

D. VGG16

VGG16 is a CNN architecture proposed by the Visual Geometry Group (VGG), University of Oxford. This model has also been tested with 14 million images in 1000 classes achieve an accuracy of 92.7% with a top 5 rating on ImageNet [20]. The fully connected layer in this model uses the ReLU activation function, except for the output layer using the softmax activation function to estimate the probability of each class/label. The softmax function can be seen in Equation 1.

$$Soft\max(z_1) = \frac{\exp(z_1)}{\sum \exp(z_1)}. \quad (1)$$

E. Flowchart of System Design

Figure 3 shows the workflow of the proposed system followed by a detailed explanation of each stage.

- Input

In this stage, the mammography dataset is uploaded to Google Drive and connected to Google Colab.

- Preprocessing

In this step, the image is resized to 150 x 150.

- Dataset Sharing

The next stage is dividing the dataset into two groups, namely training data and test data with the train test split parameters. In the simulation, 80% training data and others as test data.

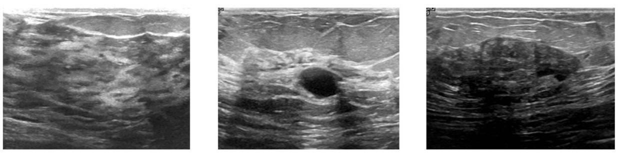

Fig. 1. Mammography images (normal, benign and malignant)

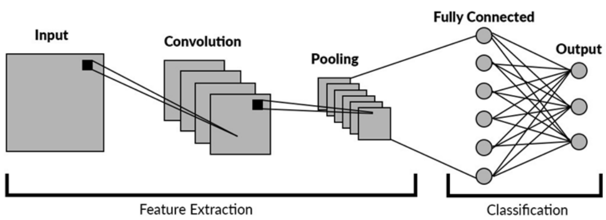

Fig. 2. Basic CNN modeling

- Pre-Trained Model

The model used in this study is based on the VGG16 model as presented in Fig. 4. The modified parameter is the last network layer, from the initial 1000 classes to 3 classes representing normal, benign and malignant.

- Training

The next stage is to train the model, before doing the training, some parameters must be used to compile the model that has been made. The parameters used are Adam and RMSprop optimizer, learning rate 0.01, and using loss categorical cross-entropy. Next, learning the model using 20 epochs with a batch size of 32.

- Test

After the model has been trained, it will produce a new model. The model is then tested with test data.

F. Performance Evaluation of Proposed System

Evaluation is done by generating a confusion matrix. The matrix used to evaluate the model represents precision, recall, f1-score, and accuracy. The parameter used in the evaluation is the confusion matrix. To understand the matrix used, we will first define True Positive (TP), False Positive (FP), False Negative (FN) and True Negative (TN) as in the confusion matrix Table 1. TP is defined as the actual normal value with predictions as normal and TN were defined as true values of cancer with predictions of being cancer. While FN is the opposite of TP, namely the actual value is normal with a prediction as cancer and FP is the opposite of TN, which is the actual value of cancer with a normal prediction.

Accuracy is the ratio of true predictions to the overall data. Precision is True Positive (TP) with the amount of data that is predicted to be positive. Recall is a comparison between True Positive (TP) with the number of data that are actually positive. F1-score is the average of recall and precision. Equations 2, 3, 4 and 5 are the formulas for precision, recall, f1-score and accuracy.

$$\text{Precision}=\frac{TP}{TP+FP}, \tag{2}$$

$$\text{Recall}=\frac{TP}{TP+FN}, \tag{3}$$

$$\text{F1-score}=\frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}, \tag{4}$$

$$\text{Accuracy}=\frac{TP+TN}{TP+TN+FPFN}. \tag{5}$$

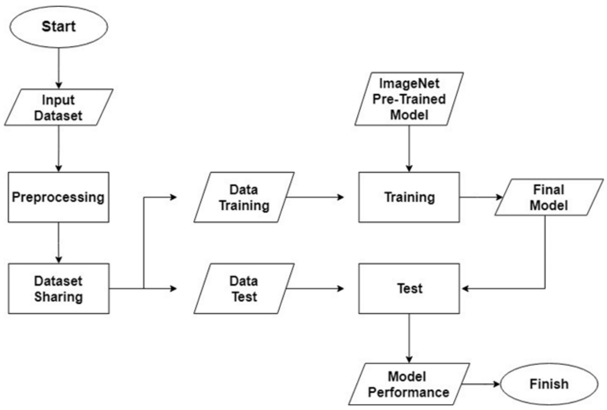

Fig. 3. Flowchart of the developed system

For the multi-class confusion matrix as shown in Table 2, the explanation is the same as the ordinary confusion matrix, which differs only in the number of classes. C-1 was defined as normal data with a normal prediction, C-5 was defined as benign data with a benign prediction, C-9 was defined as malignant data with a

Table 1. Confusion Matrix Multi-ClasConfusion Matrixg

		Prediction value		
		+ (Normal)	- (Cancer)	Malignant
True value	+ (Normal)	TP	FP	C-3
	- (Cancer)	FN	TN	C-6

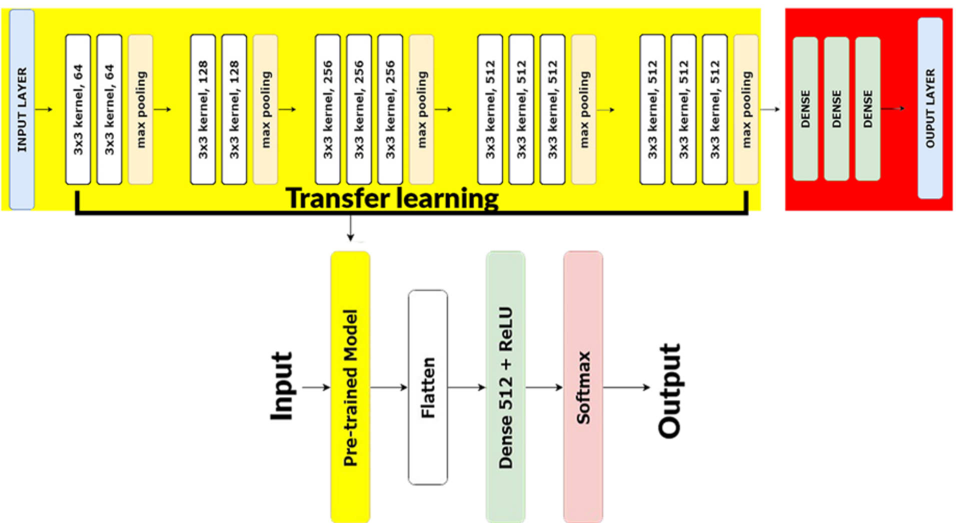

Fig. 4. VGG16 architecture on the proposed system

malignant prediction. C-2 and C-3 were defined as normal data with a prediction of benign and malignant. C-4 and C-6 were defined as benign data with a prediction of normal and malignant. C-7 and C-8 were defined as malignant data with normal and benign predictions.

In the case of the confusion matrix which has many classes, the equations used to find the values of precision,

Normal

$$TP = C\text{-}1 \qquad FN = C\text{-}2 + C\text{-}3$$
$$FP = C\text{-}4 + C\text{-}7 \qquad TN = C\text{-}5 + C\text{-}6 + C\text{-}8 + C\text{-}9 \qquad (6)$$

Benign

$$TP = C\text{-}5 \qquad FN = C\text{-}4 + C\text{-}6$$
$$FP = C\text{-}2 + C\text{-}8 \qquad TN = C\text{-}1 + C\text{-}3 + C\text{-}7 + C\text{-}9 \qquad (7)$$

Malignant

$$TP = C\text{-}9 \qquad FN = C\text{-}7 + C\text{-}8$$
$$FP = C\text{-}3 + C\text{-}6 \qquad TN = C\text{-}1 + C\text{-}2 + C\text{-}4 + C\text{-}5 \qquad (8)$$

Table 2. Confusion Matrix

		Prediction value		
		Normal	Benign	Malignant
True value	Normal	C-1	C-2	C-3
	Benign	C-4	C-5	C-6
	Malignant	C- 7	C-8	C-9

recall, f1-score and accuracy are using Equations 2, 3, 4 and 5. For the multi-class confusion matrix the values of TP, FN, FP, and TN are not yet known, so we must use the formula as in Equations 6, 7 and 8.

G. Testing Scenario

In the performance test of the proposed system, there are two test scenarios as described in the following sub-section. The goal is to test the robustness of the proposed system.

1. First Scenario

In the first scenario, the training was carried out with the same number of mammographic image datasets for each class. In this scenario, there are two class models, namely Model A with three test classes, namely the normal class with 133 data, benign with 133 data and malignant with 133 data. Furthermore, Model B has two test classes, namely normal class with 133 data and cancer with 133 data. the number of data is 133, cancer data is taken from benign class data 67 images and malignant 66 images. In this scenario, Adam system optimizers and RMSprop optimizers are used to compare the performance of each architecture.

2. Second Scenario

In the second scenario, the training is carried out with the number of mammographic image datasets in accordance with the original datasets taken from the source. This scenario is the same as the first scenario, which has two models. Model A with three test classes, namely normal class with 133 data, benign with 437 data and malignant with 210 data. In Model B, testing is carried out with two classes, namely normal class with 133 data and cancer with 647 data. cancer is taken from the data class benign and malignant. In this scenario, Adam system optimizers and RMSprop optimizers are used to compare the performance of each architecture.

III. Result and Discussion

This section discusses the performance results of the proposed system. This CNN algorithm simulation with VGG16 architecture uses google colab based on python3. The graphics processing core used is the NVIDIA Tesla P100-PCIE GPU with 16 GB of RAM. The dataset is divided into two models, namely Model A for the three test classes in normal class, benign class and malignant class and Model B for two testing classes in normal class and cancer class. The libraries used are Keras and TensorFlow, and there are additional libraries from “sklearn”.

For testing the second scenario, two models were made similar to the first scenario, only that the difference from the scenarios was the number of datasets being tested. Of the two scenarios, 20% of the total image data is used for test data and the remaining 80% is for training data. This test is carried out with two different optimizers, namely Adam optimizers and RMSprop optimizers to produce a difference from each model and scenario tested in order to get the best architecture.

To avoid overfitting, several parameters must be added, such as a learning rate of 0.01, with 755 decay steps, 0.9 decay rate. Furthermore, for the training process using batch size 32 with 20 epochs. The evaluation results from each test are calculated using Equations 2, 3, 4 and 5.

A. Effect of Optimizers on Accuracy

This section describes the differences between the results of the Adam optimizer and the RMSprop optimizer which were tested using the VGG16 architecture with a transfer learning system. There are two models used, namely Model A which is a testing model of three classes, namely normal, benign and malignant classes. Model B which is a test model of two classes, namely normal and cancer classes. The test results are presented in Table 3, Table 4, Table 5, Table 6, Table 7, Table 8, Table 9, Table 10.

Table 3 describes the value of precision, recall, and f1-score from Model A in the first scenario. Model A has a precision value of 0.88, which means that 88% of correct normal predictions of all correct normal predictions from each class, has a recall value of 0.92 which means that only 92% of normal correct predictions of all normal data, then has an f1-score of 0.90 which means 90% the results of the average values of recall and precision, these meanings apply to each model in each scenario.

The authors report the average values of precision, recall, f1-score and accuracy in Table 11 and Table 12, the first scenario using Adam's optimizers in Model A has a precision of 85%, recall 85%, f1-score 85% and accuracy 85%. Model B has 94% precision, 91% recall, 92% f1-score and 93% accuracy. Furthermore, the first scenario using RMSprop optimizers on Model A has a precision of 82%, recall 81%, f1-score 80% and accuracy 80%. Model B has 75% precision, 70% recall, 66% f1-score and 67% accuracy.

For the second scenario, using Adam's optimizers in Model A, it has 82% precision, 81% recall, 81% f1-score and 83% accuracy. Model B has 96% precision, 82% recall, 87% f1-score and 94% accuracy. Then the second scenario using RMSprop optimizers on Model A has 78% precision, 80% recall, 77% f1-score and 79% accuracy. Model B has 95% precision, 75% recall, 81% f1-score and 91% accuracy.

From the two scenarios proposed with different optimizers, in terms of the average value of precision, recall, f1-score and accuracy, it can be concluded that the Adam optimizers have higher performance than the RMSprop optimizers.

B. Performance Test Based on Curve

Figure 5, Figure 6, Figure 7, and Figure 8 are learning curves using the first scenario and the second scenario with the Adam optimizer and RMSprop optimizer. The upper curve image uses Model B data which consists of 2 classes, namely normal and cancer. For the lower curve image using Model A data which consists of 3 classes, namely normal, benign and malignant classes.

From the learning process result curve above, it can be concluded that the Adam optimizer has a fast learning or training process and has an accurate validation accuracy value, the evidence can be seen from the green training accuracy curve and the red validation accuracy curve looks stable, there is no significant decrease to the resulting curve.

C. The Results of the two scenarios tested

From the two scenarios that have been tested, we can see the results of the heatmap confusion matrix that are issued from each scenario tested. Figure 9 and Fig. 10 are the heatmap confusion matrix as a result of the scenarios being tested.

Performance evaluation of the proposed model is also carried out by comparing it with several previous studies. Table 13 shows a summary of comparative studies. Comparative studies actually cannot be carried out directly because in general they use different mammographic

Table 3. Detail Precision, Recall, F1-score using Model A on first scenario (optimizers Adam)

Class	Prec.	Recall	F1-score
Normal	0.88	0.92	0.90
Benign	0.85	0.79	0.82
Malignant	0.81	0.85	0.83

Table 4. Detail Precision, Recall, F1-score using Model B, first scenario (optimizers Adam)

Class	Prec.	Recall	F1-score
Normal	0.89	1.00	0.94
Cancer	1.00	0.83	0.90

Table 5. Detail Precision, Recall, F1-score using Model A, first scenario (optimizers RMSprop)

Class	Prec.	Recall	F1-score
Normal	0.74	1.00	0.85
Benign	0.95	0.66	0.78
Malignant	0.77	0.77	0.77

Table 6. Detail Precision, Recall, F1-score using Model B, first scenario (optimizers RMSprop)

Class	Prec.	Recall	F1-score
Normal	0.93	0.45	0.61
Cancer	0.56	0.96	0.71

Table 7. Detail Precision, Recall, F1-score using Model A, second scenario (optimizers Adam)

Class	Prec.	Recall	F1-score
Normal	0.75	0.75	0.75
Benign	0.85	0.88	0.86
Malignant	0.86	0.79	0.82

Table 8. Detail Precision, Recall, F1-score using Model B, second scenario (optimizers Adam)

Class	Prec.	Recall	F1-score
Normal	1.00	0.64	0.78
Cancer	0.93	1.00	0.96

Table 9. Detail Precision, Recall, F1-score using Model A, second scenario (optimizers RMSprop)

Class	Prec.	Recall	F1-score
Normal	0.55	0.86	0.67
Benign	0.88	0.78	0.82
Malignant	0.91	0.76	0.83

Table 10. Detail Precision, Recall, F1-score using Model B, second scenario (optimizers RMSprop)

Class	Prec.	Recall	F1-score
Normal	1.00	0.50	0.67
Cancer	0.90	1.00	0.95

Table 11. Average Precision, Recall, F1-score and Accuracy in the first scenario

Class	Prec.	Recall	F1-score	Accuracy
		Optimizers Adam		
Model A	0.85	0.85	0.85	0.85
Model B	0.94	0.91	0.92	0.93
		Optimizers RMSprop		
Model A	0.82	0.81	0.80	0.80
Model B	0.75	0.70	0.66	0.67

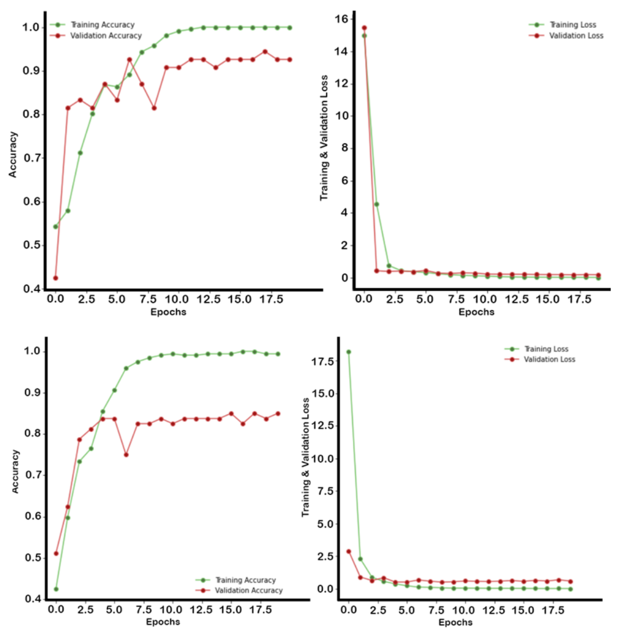

Fig. 5. The curve of the learning process using Adam's optimizers (first scenario)

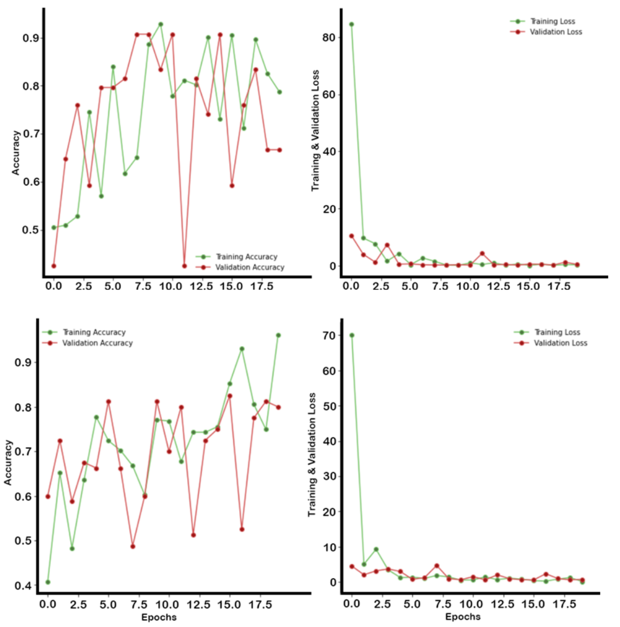

Fig. 6. The curve of the learning process using the RMSprop optimizers (first scenario)

Table 12. Average Precision, Recall, F1-score and Accuracy in the second scenario

Class	Prec.	Recall	F1-score	Accuracy
		Optimizers Adam		
Model A	0.82	0.81	0.81	0.83
Model B	0.96	0.82	0.87	0.94
		Optimizers RMSprop		
Model A	0.78	0.80	0.77	0.79
Model B	0.95	0.75	0.81	0.91

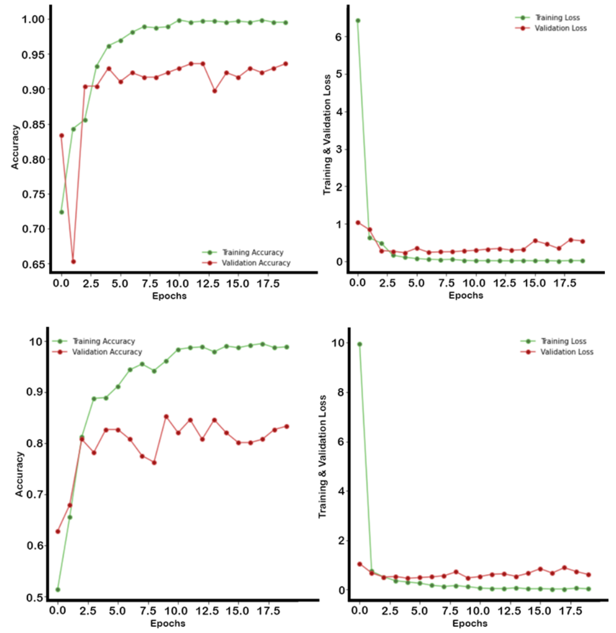

Fig. 7. The curve of the learning process using Adam's optimizers (second scenario)

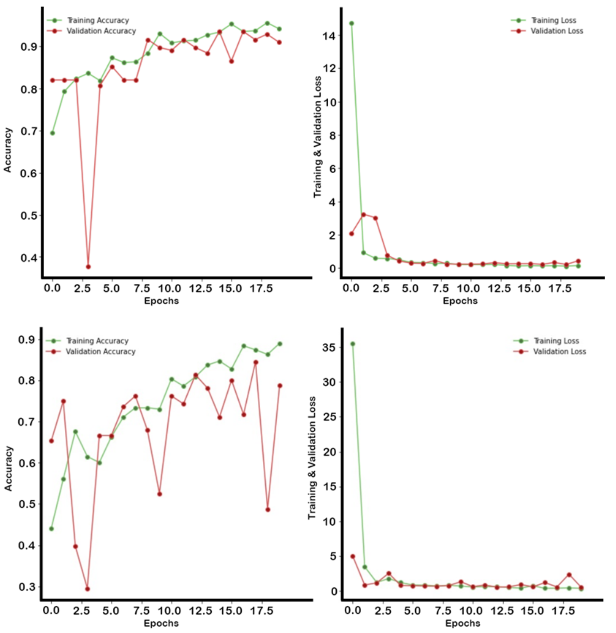

Fig. 8. The curve of the learning process using the RMSprop optimizers (second scenario)

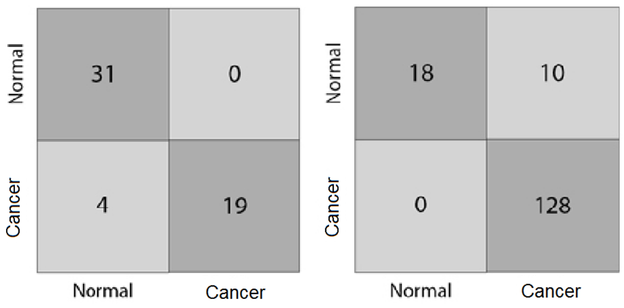

Fig. 9. Scenario 1 and Scenario 2 Heatmap Confusion Matrix 2 Class

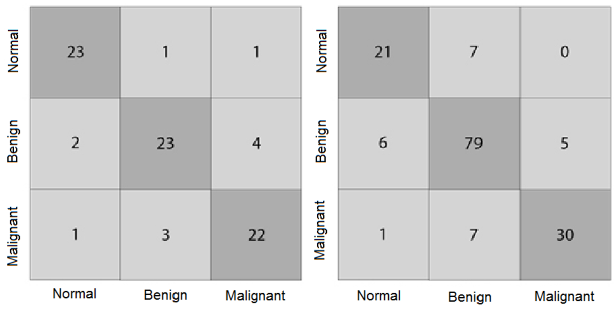

Fig. 10. Scenario 1 and Scenario 2 Heatmap Confusion Matrix 3 Class

datasets but at least their study objectives are the same in classification. In addition, some studies classify cases of two or three classes. From Table 13, it is known that the performance of the proposed model in this study outperforms previous studies, both in the case of two-class classification and three-class classification.

IV. Conclusion

In this study, a method for detecting breast cancer on mammographic images has been proposed. Simulations were carried out with two classes and three classes of mammography images. The proposed CNN network architecture is based on VGG16 with parameter modification. The test data used in each model is 20% of the total number of image data to be trained. The training process begins with a learning rate of 0.01, with 755 decay steps, 0.9 decay rate. The performance evaluation of the proposed method by calculating precision, recall, f1-score, and accuracy. Best performance is obtained by using Adam optimizers. The results of this study are quite good because the average value of precision, recall, f1-score, and accuracy achieved is up to 94% and can be considered for verification and validation by clinicians so that in the future it is expected to be applied to support the clinical diagnosis of breast cancer in larger population.

References

[1] The Global Cancer Observatory, "Cancer Incident in Indonesia," 2020.

[2] X. Chen et al., "Non-invasive early detection of cancer four years before conventional diagnosis using a blood test," Nat. Commun., vol. 11, no. 1, pp. 1–10, 2020.

[3] M. Funke, "Bildgebende diagnostik des dammakarzinoms: ein update," *Radiologe*, vol. 56, no. 10, pp. 921–938, 2016.

[4] P. Gede and S. Cipta, "Prediksi citra makanan menggunakan convolutional neural network untuk menentukan besaran kalori makanan," *J. Teknol. Inf. dan Komput*., vol. 6, no. 1, pp. 30–38, 2020.

[5] A. Peryanto, A. Yudhana, and R. Umar, "Klasifikasi citra menggunakan convolutional neural network dan K-fold cross validation," *J. Appl. Informatics Comput*., vol. 4, no. 1, pp. 45–51, 2020.

[6] I. B. L. M. Suta, R. S. Hartati, and Y. Divayana, "Diagnosa tumor otak berdasarkan citra MRI (Magnetic Resonance Imaging)," Maj. Ilm. Teknol. Elektro, vol. 18, no. 2, 2019.

Table 13. Comparison with previous works

No	Study	Method	Class	Accuracy
1	Setiawan and Putra [21]	K-Means, GLCM, dan Support Vector Machine (SVM)	2	80%
2	Fuad and Setiawan [22]	Statistic and K-NN	3	66.7%
3	Wisudawati [23]	GLCM and SVM	2	83.59%
4	Our work	CNN	2 3	94% 85%

[7] A. Amaliah and I. Puspita, "Deteksi lokasi tumor payudara menggunakan algoritma morfologi dan multilevel threshold," *J. Keteknikan dan Sains,* vol. 1, no. 2, pp. 63–68, 2018.

[8] G. Litjens, T. Kooi, B. E. Bejnordi, A. Setio, F. Ciompi, M. Ghafoorian, "A survey on deep learning in medical image analysis," Medical Image Analysis, vol 42, pp. 60–88, 2017.

[9] B. Sahiner, A. Pezeshk, L. M. Hadjiiski, X. Wang, K. Drukker, K. H. Cha KH, "Deep learning in medical imaging and radiation therapy", *Medical Physics*, vol. 46, no 1, pp. e1–e36, 2019.

[10] M. Ahmad, M. J. Khan, A. Yousaf, S. Ghuffar, K. Khurshid, "Deep Learning: A breakthrough in medical imaging," *Current Medical Imaging Reviews*. Vol. 16, no. 8, pp. 946-956, 2020.

[11] L. Tobias, A. Ducournau, F. Rousseau, G. Mercier, and R. Fablet, "Convolutional Neural Networks for object recognition on mobile devices: A case study," in Proceedings - International Conference on Pattern Recognition, 2016, pp. 3530–3535.

[12] W. Al-Dhabyani, A. Fahmy, M. Gomaa, and H. Khaled, "Deep learning approaches for data augmentation and classification of breast masses using ultrasound images," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 5, pp. 618–627, 2019.

[13] M. Gao, J. Chen, H. Mu, and D. Qi, "A transfer residual neural network based on resnet-34 for detection of wood knot defects," *Forests*, vol. 12, no. 2, pp. 1–16, 2021.

[14] T. Pandit, A. Kapoor, R. Shah, and R. Bhuva, "Understanding inception network architecture for image classification," Experiments For Architectural Basis of Convolutional Neural Networks For Image Recognition, Tech. Report, February 2020.

[15] S. K. Venu, "An ensemble-based approach by fine-tuning the deep transfer learning models to classify pneumonia from chest X-ray images," in ICAART 2021 - Proceedings of the 13th International Conference on Agents and Artificial Intelligence, 2021, vol. 2, pp. 390–401.

[16] D. Sinha and M. El-Sharkawy, "Thin MobileNet: An Enhanced MobileNet Architecture," in 2019 IEEE 10th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference, UEMCON 2019, 2019, pp. 0280–0285.

[17] M. Tan and Q. V. Le, “EfficientNet: Rethinking model scaling for convolutional neural networks,” in 36th International Conference on Machine Learning, ICML 2019, 2019, vol. 2019-June, pp. 10691–10700.

[18] F. Yilmaz and A. Demir, “Cutting Effect on classification using nasnet architecture,” in TIPTEKNO 2020 - Tip Teknolojileri Kongresi - 2020 Medical Technologies Congress, TIPTEKNO 2020, 2020, pp. 1–3.

[19] V. H. Phung and E. J. Rhee, “A High-accuracy model average ensemble of convolutional neural networks for classification of cloud image patches on small datasets,” Appl. Sci., vol. 9, no. 21, pp. 1–16, 2019.

[20] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” in 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings, 2015, pp. 1–14.

[21] K. Novar Setiawan and I. M. Suwija Putra, “Klasifikasi citra mammogram menggunakan metode K-Means, GLCM, dan Support Vector Machine (SVM),” *J. Ilm. Merpati* (*Menara Penelit. Akad. Teknol. Informasi*), vol. 6, no. 1, pp. 13, 2018.

[22] M. Fuad and W. Setiawan, “Klasifikasi kanker payudara menggunakan ekstraksi ciri metode statistik,” *Multinetics*, vol. 2, no. 2, pp. 54–58, 2016.

[23] L. M. Wisudawati, “Klasifikasi tumor jinak dan tumor ganas pada citra mammogram menggunakan gray level co-cccurrence matrix (GLCM) dan Support Vector Machine (SVM),” *J. Ilm. Inform. Komput*., vol. 26, no. 2, pp. 176–186, 2021.