Optimizing hyperparameters in multiview convolutional neural network for improved breast cancer detection in mammograms

Sisilia Anggraini, Tri Arief Sardjono, Nada Fitrieyatul Hikmah
Department of Biomedical Engineering, Faculty of Intelligent Electrical and Informatics Technology, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

Article Info	ABSTRACT
	High accuracy in breast cancer classification contributes to the effectiveness of early breast cancer detection. This study aimed to improve the multiview convolutional neural network (MVCNN) performance for classifying breast cancer based on the combined mediolateral (MLO) and craniocaudal (CC) views. The main contribution of this study is the development of a system, consisting of an effective image pre-processing method to create datasets using background removal techniques, and image enhancement. Also, a simplicity of preprocessing stage in the classifier machine, which does not require a feature extraction process. Furthermore, the performance of the classifier was improved by combining preprocessing dataset techniques and evaluating the best hyperparameter in MVCNN architecture. The digital dataset for screening mammography (DDSM) dataset was used for evaluation in this study. The best result from this proposed method achieved accuracy, precision, sensitivity, and specificity of 98.63%, 97.29%, 100%, and 97.29%. The evaluation results demonstrated the capability to improve classification performance. The method proposed in this work can be applied to the detection of breast cancer. *This is an open access article under the [CC BY-SA](#) license.*

Corresponding Author:

Nada Fitrieyatul Hikmah
Department of Biomedical Engineering, Faculty of Intelligent Electrical and Informatics Technology
Institut Teknologi Sepuluh Nopember
Keputih, Sukolilo, Surabaya, 60111, Indonesia
Email: nadafh@bme.its.ac.id

1. INTRODUCTION

Breast cancer has emerged as one of the most pressing concerns in the medical field, characterized by high incidence and mortality rates. According to data from the World Health Organization (WHO) in 2020, 7.8 million women were diagnosed with breast cancer, and the same year witnessed 684,996 deaths attributed to breast cancer [1], [2]. Recently, research focusing on breast cancer classification based on mammography images has been widely developed to create an optimal detection system for breast cancer. Anatomically, the breast comprises three main types of tissue: glandular tissue, fat, and connective tissue. When abnormal cells grow uncontrollably within the breast, this poses a significant threat, especially when these abnormal cells spread to glandular tissue [3]. Abnormal cells growing within the breast can be detected using breast imaging methods such as mammography [4], positron emission tomography (PET) [5], magnetic resonance imaging (MRI) [6], computed tomography (CT) [7], and ultrasonography [8]. Mammography is considered one of the most effective methods for detecting breast cancer tissue due to its higher sensitivity in identifying small changes or abnormalities in breast tissue compared to various other imaging methods [9]. However, unfortunately, breast cancer lumps often appear on mammography images with low contrast and appear blurry. Although breast cancer can be identified through radiologist interpretation of mammography

images, there is potential for errors due to human visual limitations and varying levels of subjectivity [10], [11]. Therefore, it is essential to develop an intelligent system capable of quickly and accurately detecting and diagnosing abnormalities [11]-[13].

According to Sickles *et al*. [14], radiology experts evaluate mammography results through five stages: identifying indications, analyzing breast tissue categories, evaluating significant findings, comparing with previous studies, and ultimately assigning a final assessment based on the breast imaging-reporting and data system (BI-RADS) category. During the mammography procedure, two different angles of the patient's breast are captured to identify potential abnormalities. The craniocaudal (CC) projection captures an image from top to bottom of the breast, displaying the medial and lateral outer aspects of the breast. On the other hand, the mediolateral (MLO) projection is used to visualize the entire breast at an angle of approximately 40 to 60 degrees from the side, showing lymph nodes and pectoral muscles [3]. Although breast cancer can be diagnosed by radiology experts, there is a possibility of inaccuracies in diagnosis due to human visual limitations and subjectivity. Hence, automatic breast cancer detection with the aid of computers is necessary to assist doctors and radiologists in the diagnostic process. Automatic analysis and diagnosis of breast cancer in mammography using computer-aided diagnosis (CAD) [15] not only reduces dependence on medical knowledge and doctor's experience but also provides objective and accurate suggestions to doctors. With the rapid advancement of machine learning techniques, several machine learning algorithms have been applied to breast cancer diagnosis in mammography to enhance breast cancer diagnostic performance [16], [17].

Previous studies have utilized deep learning to classify mammography images using the CC view. In this analysis, the CC view is used because it offers the visualization of most breast tissue without pectoral muscle representations in mammograms that can interfere with detection outcomes. The results of this research show that the proposed model can achieve a high classification performance, with an accuracy of 92.84% for distinguishing between cancer and normal cases in mammograms from the CC view [18]. In contrast, the study by Firdi *et al*. [19] addressed issues in the MLO view by employing pectoral muscle removal techniques. The results showed that this method improved breast cancer detection accuracy in mammography images from the MLO view. However, both of these studies focused on a single view and did not address the simultaneous use of both mammography views. Akilan [20] found specific differences, indicating that images taken from different views of the same object can provide complementary information. In this context, mammography is a modality that captures images from two views: CC and MLO. By harnessing both views in the deep learning process, a richer set of visual features can be obtained.

The idea is supported by a study in [13], which revealed that classifying mammography images using both CC and MLO views simultaneously can achieve an accuracy of 82.02%. This study employed a multiview approach that combined the CC and MLO views using a multiview convolutional neural network (MVCNN) in conjunction with a multi-dilated convolutional neural network (MDCNN). Similarly, Gu *et al*. [21] found that developing an auto-diagnosis model using multi-view mammography and transfer learning techniques can improve breast tumor diagnosis accuracy. This model achieved the best results in classifying tumors with an area under the curve AUC of 0.97 and 0.98 using two different datasets. However, in both of these studies, image preprocessing stages to improve characteristics, contrast, and noise in mammography images were not yet integrated. Beeravolu *et al*. [12] found specific differences, indicating that the use of preprocessing stages on all images before feeding them into machine learning is an effective way to enhance accuracy and reduce computational time during training, validation, and testing. These findings align with the conclusions of previous research by Tavakoli *et al*. [22], emphasizing the importance of image preprocessing to obtain accurate training data. Their study achieved 94.68% accuracy and a 95% AUC by eliminating unwanted areas, such as artifacts and noise in mammography images, followed by contrast enhancement. Similarly, Hikmah *et al*. [23] in their research, developed a preprocessing framework for breast cancer detection using multi-view mammography images. The study has room for improvement in the detection method used, where the image preprocessing framework was not yet incorporated into machine learning processes such as convolutional neural networks (CNNs).

Although a previous study has successfully classified two mammography images from both CC and MLO views using machine learning algorithms, there is still room for a further comprehensive investigation to implement the use of an image preprocessing framework to improve characteristics, contrast, and noise in mammogram images before entering the machine learning algorithms. Furthermore, machine learning performance is highly dependent on various hyperparameters such as architecture, the number of filters, kernel size, and stride in convolution and pooling layers. Therefore, this study aims to classify mammography images by combining the use of an image preprocessing framework and enhancing classification performance by finding the best hyperparameters. Its specific aims include; i) evaluate the results of image preprocessing to create an effective dataset; ii) evaluate the impact of the number of filters, kernel size, and stride in convolution layers on accuracy; iii) evaluate the influence of kernel size and stride in pooling layers on accuracy; and iv) compare accuracy between the proposed method and the MVCNN

machine learning method without using image preprocessing stages. The contributions of the study include; i) the development of a machine learning architecture with the most effective hyperparameters based on a modified MVCNN for direct pattern recognition using preprocessed mammography images as input to the classifier and ii) the simplification of the preprocessing stage in the machine classifier, which does not require a feature extraction process.

This paper is structured as follows: section 2 presents the materials and the proposed method in comparison to other standard approaches. Section 3 describes the results of this study. Section 4 presents the discussion, while conclusions are presented in section 5.

2. METHOD

The overall system design can be seen in Figure 1. This system comprises an input dataset, followed by a preprocessing stage to remove artifact labels and enhance image quality. Subsequently, the processed images will be used as input for machine learning classification. Finally, the MVCNN is employed to classify breast cancer.
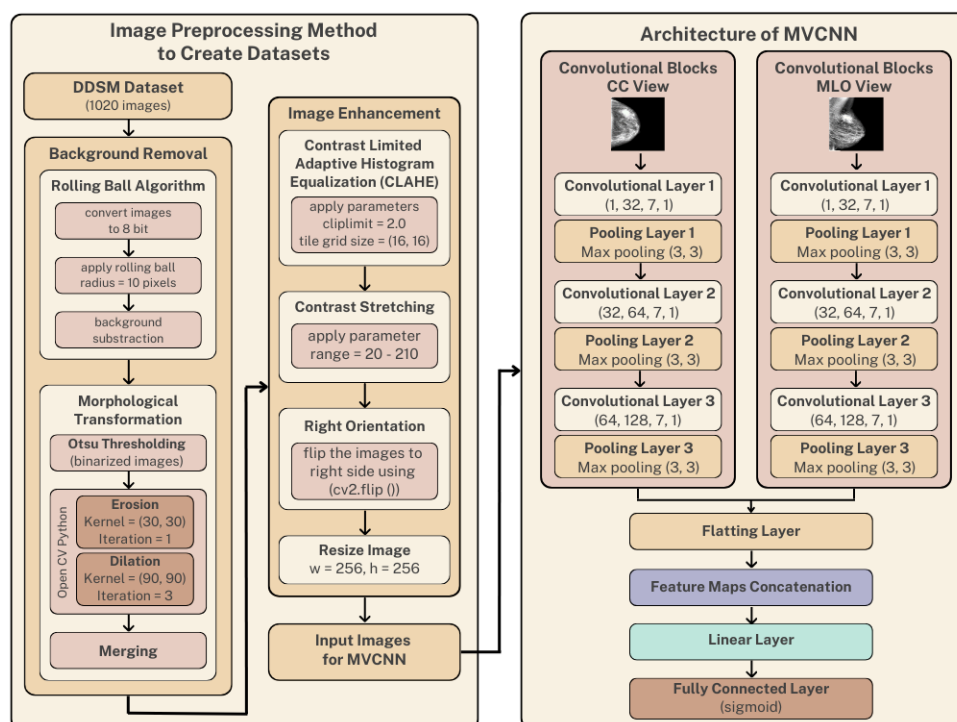

Figure 1. Diagram of system

2.1. Dataset

The research leverages a rich source of data: the digital dataset for screening mammography (DDSM) [24], a secondary open-source resource. This dataset encompasses 255 subjects, each contributing multiple mammogram images for a total of 1020. Remarkably, the dataset maintains a balanced distribution between healthy and cancerous cases, with 510 images classified as "normal" and 510 classified as "cancer." This balanced representation of both classes ensures robust training and evaluation of the cancer detection model.

2.2. Image preprocessing method

Effective cancer detection relies on a well-prepared dataset, and this study employs a dedicated image preprocessing stage for optimization. The critical stage involves two primary steps: i) background Removal, systematically eliminating unwanted elements like labels or text annotations to prevent potential model misguidance and ii) image enhancement, applying techniques such as contrast adjustments, noise reduction, and sharpening for improved image quality. These enhancements enhance clarity and feature visibility, facilitating precise feature extraction and leading to more reliable cancer classification.

2.2.1. Background removal

The first stage is background removal, which eliminates label artifacts in mammography images. This is done by determining the foreground and background of the image. To extract objects from the background, thresholding is applied to the binary image using Otsu's thresholding method, as adapted from Firdi *et al.* [19]. Next, morphological transformations, are performed using erosion and dilation as transformation operations. The original image is then combined with the morphologically transformed image using a bitwise logic gate operator.

2.2.2. Image enhancement

The second stage is image enhancement using the contrast limited adaptive histogram equalization (CLAHE) method. A clip limit of 2.0 and a tile size grid of 16×16 is used. Subsequently, normalization is applied using contrast stretching to enhance the visibility of abnormalities in the mammography image. The upper and lower limits of the range are set to 20 and 210, respectively. After normalization, the image is processed to determine its orientation, as adapted from Firdi *et al.* [19]. The orientation of the image is determined by counting the number of black pixels in half of the width of the mammography image while flipping the image if it has more black pixels than half of the width. This ensures that both mammography images face the same direction (right). Finally, the image size is resized to 256×256 pixels.

2.3. Proposed MVCNN model

The proposed architecture in this research is an adaptation of a previous study [13] with modifications made by the author. The architecture incorporates modifications in the number of convolution layers and pooling layers used. It also utilizes different variations of filters, kernel sizes, and strides. Figure 1 illustrates the architecture employed in this research. The architecture takes two mammography images as input, the CC and MLO views. These two images are then processed through two convolution blocks, where each block represents a convolution block for the CC and MLO images, respectively. Each convolution block consists of three convolutional layers followed by a pooling layer. The convolutional layers employ different variations of filters with distinct kernel sizes and strides. The max pooling method is employed in this research, utilizing specific kernel sizes and strides. The ReLU activation function is applied to the output of each convolution operation in the convolutional layers. Subsequently, the outputs from both convolution blocks are fed into a flattening layer to transform the pooled results into a vector representation. These representations are then concatenated in a feature maps concatenation process. The concatenated feature maps are then passed through a linear layer to convert the combined features into detection outputs with dimensions corresponding to the number of classes to be detected. The output of the linear layer is further activated using the sigmoid activation function to generate a final output ranging between 0 and 1.

2.4. Experimental setting

The utilized dataset has undergone image preprocessing stages, resulting in a collection of images sized at 256×256 pixels. The settings for training, validation, and testing are configured as follows: the learning rate is set to 10^{-4}, the batch size is set to 8, the employed epochs are 150, and optimization employs the Adam algorithm. The proposed method is developed and tested using the Python programming language, utilizing Google Collaboratory as the editing and compiling platform, with an A100GPU hardware accelerator. The PyTorch framework is employed as the backend to implement the Python-based MVCNN architecture. This study conducts a comparison with other machine learning techniques, as previously undertaken in the research, such as in the case of MVMDCNN [13].

2.5. The performance: preprocessing images and MVCNN

To assess the performance of both the image preprocessing stage and the multiview detection system, a performance evaluation needs to be conducted for each stage. The performance evaluation for the image preprocessing stage can be done by measuring the image quality through a comparison between the processed images and the original images. This can be achieved by using parameters such as mean squared error (MSE) and peak signal noise ratio (PSNR).

$$MSE \ = \ \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}(G(i,j) - P(i,j))^2 \tag{1}$$

$$PNSR \ = \ 10 \, log_{10}\left(\frac{(MAX)}{\sqrt{MSE}}\right) \tag{2}$$

To evaluate the performance of the proposed model, a confusion matrix. The confusion matrix provides indicators such as accuracy, precision, sensitivity, and specificity, which are calculated using (3)-(6), respectively.

$$accuracy = \frac{TP+TN}{TP+TN+FN+FP} \tag{3}$$

$$precision = \frac{TP}{TP+FP} \tag{4}$$

$$sensitifity = \frac{TP}{TP+FN} \tag{5}$$

$$specificity = \frac{TN}{TN+FP} \tag{6}$$

The performance evaluation is crucial for assessing the performance of the trained detection model using the three different types of datasets.

3. RESULTS AND DISCUSSION

In this study, each image will follow a general system design. The process involves removing artifact labels, enhancing image quality, and ultimately classifying breast cancer in the processed images. The following section will discuss the results and delve into each stage.

3.1. Image preprocessing

The process of transforming images into an effective dataset through image processing involves two main phases: background removal and image enhancement. Background removal utilizes the rolling ball algorithm and morphological transformation involving processes such as otsu thresholding, erosion, dilation, and subsequent merging to convert the original images into binary images. The rolling ball algorithm is employed to smooth the image surface and mitigate sharp intensity disparities within the image, using a spherical element. The radius of this sphere influences the conversion of the original image into a binary representation. This study explores the optimal radius value, achieved by comparing image preprocessing outcomes using MSE and PSNR metrics, as shown in Table 1. In this study, it was found that the choice of the radius value in the rolling ball algorithm affects the conversion of images into binary images. A radius that is too small may result in binary images with significant noise or irrelevant small details being considered as objects. Conversely, a radius that is too large may cause binary images to lose important details or blur object boundaries. Therefore, testing is necessary to identify the optimal radius value, and the comprehensive findings of this radius optimization exploration are provided in Table 1, with the most effective value determined as 10. Upon achieving image surface homogenization, the binary transformation is accomplished through Otsu Thresholding. This method computes the threshold value for each image, with unique values designated for individual images. For instance, in Figure 2, the image labeled as subject A_1573_1 features a threshold value of 73. Subsequently, a pixel value adjustment is executed, with pixel values below 73 set to 0, while those exceeding 73 are set to 1. Following the binary image attainment, the subsequent phase involves the elimination of labels that manifest as artifacts within the image, facilitated by morphological transformations encompassing erosion and dilation operations. This research employs a kernel size of 30×30 for the erosion operation (with a single iteration) and a kernel size of 90×90 for the dilation operation (conducted over 3 iterations). These parameter choices contribute to an enhanced success rate in background removal. The erosion operation's purpose in this study is to separate adjacent objects, followed by the elimination of residual small objects stemming from the erosion procedure. Subsequently, a dilation operation enhances the object edges within the breast region. A final merging process amalgamates the original image with the binary image, yielding an artifact-free labeled image8. The outcomes of this second-stage image preprocessing are illustrated in Figure 2.

Subsequent to background removal, the focus shifts to image enhancement, achieved through CLAHE and contrast stretching techniques, aimed at augmenting image contrast. This enhancement process is followed by image orientation adjustment to the right and resizing. Mammography images often exhibit modest contrast, posing challenges in distinguishing between cancerous tissues and overlapping normal tissues. CLAHE is employed to ameliorate low-contrast images. The determination of optimal parameter values involves exploring clip limits and tile size grids through a comparison of PSNR and MSE metrics, as shown in Table 2. The experimentation entails testing CLAHE with clip limit parameters of 2.0, 6.0, and 12.0, along with tile size grid parameters of 8×8, 12×12, and 16×16. The outcomes of these experiments, leading to the identification of optimal tile size grid and clip limit values, are presented in Table 2. Within

this study, the most favorable results are achieved with a clip limit of 2.0 and a tile size grid of 16×16. Subsequently, the image, having undergone contrast enhancement via CLAHE, is normalized through contrast stretching to extend the intensity value range within the image. The testing of contrast stretching involves range parameters of 20-200, 20-208, and 20-210. The study determines the optimal range value within the 20-210 range, as shown in Table 3. After applying CLAHE and contrast stretching, image orientation is established by reflecting the image when black pixels exceed half the image width. The aligned image faces rightward. Finally, the image is resized to 256×256 pixels. The outcomes of this second-stage image preprocessing are illustrated in Figure 3.

Table 1. Experimental results to find the most optimal ball radius value

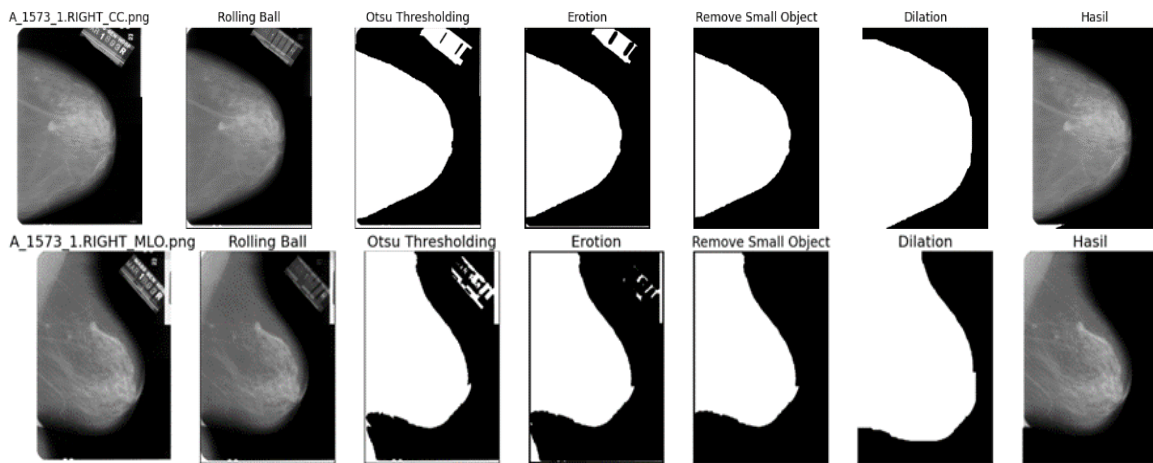| Ball radius	Mean MSE	Mean PSNR
10	34.7883	32.8589
15	42.7550	31.9527

Figure 2. Result of the background removal stage

Table 2. Optimization results of clip limit and tile size grid in CLAHE method for image B_3041_1.LEFT_MLO

Clip limit	Tile Size Grid	MSE	PSNR
2.0	8×8	59.1349	30.4124
	12×12	56.6601	30.5981
	16×16	**54.4589**	**30.7701**
6.0	8×8	71.4264	29.5922
	12×12	71.0187	29.6171
	16×16	70.7257	29.6350
12.0	8×8	97.5797	28.2372
	12×12	96.9621	28.2648
	16×16	97.2975	28.2498

Table 3. Optimization results of range in contrast stretching method for image B_3041_1.LEFT_MLO

Range	Mean MSE	Mean PSNR
20-200	64.8473	29.8799
20-208	64.8819	30.0096
20-210	**64.3624**	**30.0445**

3.2. Effect number of filters, kernel size, and stride in the convolutional layer on accuracy

In the convolutional layer, the number of filters, kernel size, and stride influence the accuracy of machine learning training. In this study, three variations of the number of filters are employed: [16, 32, 64], [30, 60, 120], and [32, 64, 128]. Each filter variation is explored concurrently with kernel size and stride value. The tested kernel size variations are 3×3, 5×5, and 7×7, all with a stride of 1. The experimental results are presented in Table 4. After evaluating the effects of the number of filters, kernel size, and stride on accuracy, the results demonstrate that the [32, 64, 128] filter variation exhibits higher accuracy compared to

the other variations. It is evident that as the kernel size increases, the obtained accuracy also increases, and the gap between the training and validation loss values becomes narrower. This indicates a close correspondence between these two losses, signifying that the model effectively generalizes the validation data. This experiment underscores that larger kernel sizes significantly impact the model's performance in generalizing validation data, leading to higher accuracy outcomes.
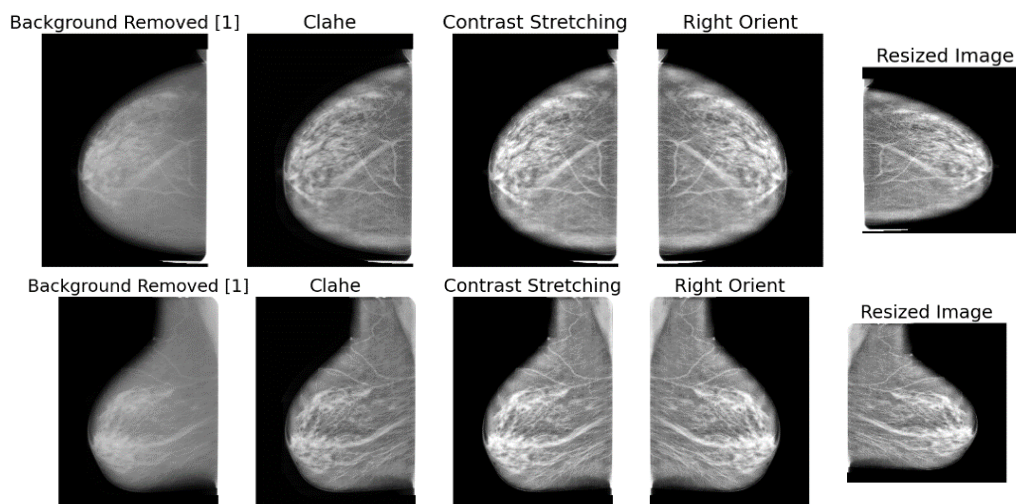

Figure 3. Result of the image enhancement stage

Table 4. Hyperparameter exploration results in the convolutional layer

Stride	Filters	Kernel size	Loss training	Loss validation	Loss testing	Accuracy testing (%)
1	[16, 32, 64]	3 × 3	0.3321	0.4408	0.3761	91.78
		5 × 5	0.3252	0.4427	0.3700	93.15
		7 × 7	0.3220	0.3866	0.3569	93.15
	[30, 60, 120]	3 × 3	0.3305	0.4672	0.3711	94.52
		5 × 5	0.3190	0.4289	0.3593	94.52
		7 × 7	0.3190	0.4164	0.3559	94.52
	[32, 64, 128]	3 × 3	0.3390	0.4264	0.3700	93.15
		5 × 5	0.3190	0.4421	0.3667	94.52
		7 × 7	**0.3162**	**0.4060**	**0.3569**	**95.89**

3.3. Effect number of filters, and kernel size in the pooling layer on accuracy

In this study, the search for the optimal kernel size and stride values in the pooling layer is conducted using the best parameters previously obtained from the exploration of variations in the number of filters, kernel size, and stride in the convolutional layer. The filter variation employed is [32, 64, 128] with a kernel size of 7×7 and stride of 1. Within the pooling layer, exploration is performed on several parameters, namely, kernel sizes of 2×2 and 3×3, along with stride parameters of 1, 2, and 3. The experimental results are presented in Table 5. It is evident that within the pooling layer, the kernel size shows no significant difference in accuracy between the use of 2×2 and 3×3 kernels. However, in terms of training, validation, and testing loss measurements, the employment of a 3×3 kernel yields the best outcomes. This is indicated by the alignment of both training and validation losses. As for the stride parameter in the pooling layer, larger stride values correspond to higher accuracy. From these results, it is inferred that a kernel size of 3×3 and a stride of 3 are the most optimal values within the pooling layer.

Table 5. Hyperparameter exploration result in the pooling layer

Kernel size	Stride	Loss training	Loss validation	Loss testing	Accuracy testing (%)
2 × 2	1	0.7962	0.7924	0.8049	49.32
	2	0.3162	0.4060	0.3569	95.89
	3	0.3202	0.3833	0.3407	97.26
3 × 3	1	0.8133	0.7924	0.8049	49.32
	2	0.3190	0.3931	0.3585	93.15
	3	**0.3223**	**0.3625**	**0.3277**	**97.26**

3.4. Breast cancer classification

The detection system used in this study is the MVCNN. MVCNN serves as a suitable machine-learning approach for classification tasks. The evaluation of the detection system involves the assessment of hyperparameters within the convolutional and pooling layers. Based on the testing results of the detection system using the MVCNN architecture on the second dataset, the results shown in Table 6 demonstrate exceptional performance. The comparison graph of training and validation losses in Figure 4(a) indicates a rapid decline as epochs progress, with convergence approaching a stable point. The graph demonstrates a near-stabilized convergence, signifying that the model has reached an optimal state. Moreover, the proximity between training and validation losses suggests good generalization capability on validation data. Similarly, Figure 4(b) illustrates converging or near-stable accuracy improvement. Furthermore, the evaluation metrics outlined in Table 7 exhibit outstanding performance. The accuracy reaches 98.63%, reflecting a high success rate in classifying objects within the dataset. Precision attains 97.29%, indicating accurate positive prediction. Sensitivity reaches 100%, demonstrating the model's ability to correctly identify all true positive objects. Specificity attains 97.29%, highlighting the model's accurate classification of all true negative objects. As evident from Figure 4(c), out of the total 73 tested samples, 36 normal samples (true negative) and 36 abnormal samples (true positive) were correctly classified, with only 1 abnormal sample misclassified as normal (false negative). This showcases the model's high accuracy in object classification.

Table 6. Results using the best hyperparameter

Loss training	Loss validation	Loss testing	Acc (%)	Duration (s)
0.3163	0.3369	0.3245	98.63	45.29

Table 7. Evaluation metrics during training

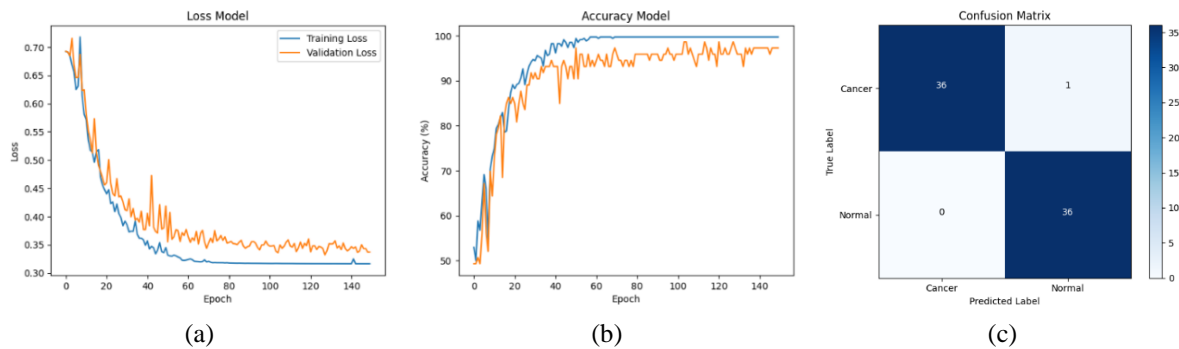| Accuracy	Precision	Sensitivity	Specificity
0.9863	0.9729	1.0	0.9729

(a)	(b)	(c)

Figure 4. Comparing simulation results of: (a) loss graph during training, (b) accuracy graph during training, and (c) confusion matrix results for testing

This study involves a comparison to assess the performance of the proposed method against the previous method that employs the same approach, namely a modified version of the MVMDCNN method [13], using the best-explored hyperparameters evaluated in this research. As shown in Table 8, it can be observed that the modified MVMDCNN method achieves an accuracy of 95.89% in classifying breast images into two categories: normal or cancerous. This method also exhibits a sensitivity of 97.29%, indicating its proficiency in accurately recognizing positive cancer cases among all actual positive cases. However, its specificity is 94.44%, suggesting slight difficulty in correctly identifying negative cases. On the other hand, the proposed method demonstrates superior overall performance. With an accuracy of 98.63%, it can classify these images with greater precision. Furthermore, its sensitivity is 100%, showcasing the method's ability to accurately identify all positive cases. Additionally, its specificity is also high at 97.29%, demonstrating the method's capability to accurately identify negative cases. Both the previous and proposed methods exhibit the same level of precision at 97.29%, highlighting their equivalent abilities in providing accurate positive predictions.

Table 8. Classification performance comparison with other studies

Method	Dataset	Number of classes	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)
TVNN [25]	DDSM	2	94.70	-	-	-
VGG16+MVML-GL [22]	DDSM	2	95	-	-	-
ResNet50+CvAM [20]	DDSM	2	86.2	-	-	-
MVMDCNN [13]	DDSM	2	82.2	-	-	-
MVMDCNN Modified [13]	DDSM	2	95.89	97.29	94.44	94.73
Purposed method	DDSM	2	98.63	100	97.29	97.29

4. CONCLUSION

This study aimed to improve performance using an image preprocessing framework, simplify the architecture, and determine the best hyperparameters based on the MVCNN algorithm in classifying breast cancer from CC and MLO views. The study establish that the accuracy of the proposed method outperformed other classifiers. In this work, preprocessing stages are applied to mammogram images, involving background removal and image contrast enhancement. The advantage of the proposed method is the simultaneous use of two mammography images from CC and MLO views, optimized through image preprocessing stages to create an effective input dataset, thus enhancing the system's ability to classify breast cancer with a higher level of accuracy and reducing the need for manual intervention. Findings indicate that filter variations [32, 64, 128] with a 7×7 kernel size and a stride of 1 are sufficient in the convolution layers. In the pooling layers, a 3×3 kernel size with a stride of 3 yields the best results, as indicated by the alignment of loss in both training and validation. Classification using MVCNN based on mammogram images and the selection of the best hyperparameter variations successfully achieved accuracy, sensitivity, specificity, and precision of 95.31%, 95.31%, 98.8%, and 95.55%, respectively. Future studies should focus on developing a more comprehensive classification approach by utilizing deep learning techniques within the MVCNN architecture to further enhance performance. Also, the proposed method could be implemented in computer-based detection systems for breast cancer diagnosis.

REFERENCES

[1] World Health Organization (WHO), "Fact sheets detail breast cancer," WHO.int. [Online]. Available: https://www-who-int.translate.goog/news-room/fact-sheets/detail/breast-cancer. (accessed: Feb, 7 2023).
[2] H. Sung *et al*., "Global cancer statistics 2020: globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA Cancer Journal for Clinicians*, vol. 71, no. 3, pp. 209–249, 2021, doi: 10.3322/caac.21660.
[3] W. A. Berg and J. W. T. Leung, "Diagnostic Imaging: Breast," in *Diagnostic Imaging Series*, 2019, Ed. 3, pp. 4-6.
[4] K. Kerlikowske *et al*., "Cumulative advanced breast cancer risk prediction model developed in a screening mammography population," *JNCI: Journal of the National Cancer Institute*, vol. 114, no. 5, pp. 676–685, 2022, doi: 10.1093/jnci/djac008.
[5] Y. Satoh *et al*., "Deep learning for image classification in dedicated breast positron emission tomography (dbpet)," *Annals of Nuclear Medicine*, vol. 36, no. 4, pp. 401-410, 2022, doi: 10.1007/s12149-022-01719-7.
[6] C. Bie *et al*., "Deep learning-based classification of preclinical breast cancer tumor models using chemical exchange saturation transfer magnetic resonance imaging," *NMR Biomed*, vol. 35, no. 2, p. e4626, 2022, doi: 10.1002/nbm.4626.
[7] J. Koh, Y. Yoon, S. Kim, K. Han, and E. K. Kim, "Deep learning for the detection of breast cancers on chest computed tomography," *Clin. Breast Cancer.*, vol. 22, no. 1, pp. 26-31, 2022, doi: 10.1016/j.clbc.2021.04.015.
[8] Z. Niu *et al*., "The value of contrast-enhanced ultrasound enhancement patterns for the diagnosis of sentinel lymph node status in breast cancer: systematic review and meta-analysis," *Quant. Image. Med. Surg.*, vol. 12, no. 2, pp. 936–948, 2022, doi: 10.21037/qims-21-416.
[9] N. M. Basheer and M. H. Mohammed, "Segmentation of breast masses in digital mammograms using adaptive median filtering and texture analysis," *International Journal of Recent Technology and Engineering*, vol. 2, no. 1, pp. 39-43, 2013, doi: 10.1155/2022/8576768.
[10] A. K. Singh and B. Gupta, "A novel approach for breast cancer detection and segmentation in a mammogram," *Procedia Computer Science*, vol. 54, pp. 676-682, 2015, doi: 10/1016/j.procs.2015.06.079.
[11] A. P. Singh and B. Singh, "Texture Features Extraction in Mammograms Using Non-Shannon Entropies, Machine Learning and Systems Engineering," *Lecture Notes in Electrical Engineering*, Dordrecht, Springer, pp. 341-351, 2010, doi: 10.1007/978-90-481-9419-3_26.
[12] A. R Beeravolu, S. Azam, M. Jonkman, B. Shanmugam, K. Kannorpatti, and A. Anwar, "Preprocessing of Breast Cancer Images

to Create Datasets for Deep-CNN," *IEEE Access*, vol. 9, pp. 33438-33463, 2021, doi: 10.1109/ACCESS.2021.3058773.

[13] L. Sun, J. Wang, Z. Hu, Y. Xu, and Z. Cui, "Multi-View Convolutional Neural Networks for Mammographic Image Classification," *IEEE Access*, vol. 7, pp. 126273-126282, 2019, doi: 10.1109/ACCESS.2019.2939167.

[14] E. A. Sickles, C. J. D'Orsi, and L. W. Bassett, "ACR BI-RADS Atlas" in *ACR BI-RADS Atlas Breast Imaging Reporting and Data Systems*, Reston, American College Radiology, 2013, vol. 10, no.3, pp. 247-249.

[15] G. E. Park, B. J. Kang, S. H. Kim, and J. Lee, "Retrospective review of missed cancer detection and its mammography findings with artificial-intelligence-based, computer-aided diagnosis," *Diagnostics (Basel)*, vol. 12, no. 2, p. 387, 2022, doi: 10.3390/diagnostics12020387.

[16] G. Liang *et al*., "Joint 2D-3D Breast Cancer Classification," *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, San Diego, CA, USA, pp. 692-696, 2019, doi: 10.1109/BIBM47256.2019.8983048.

[17] E. Michael, M. He, L. Hong, and S. Qi, "An optimized framework for breast cancer classification using machine learning," *BioMed Res. Int.*, vol. 2022, 2022 doi: 10.1155/2022/8482022.

[18] G. Altan, "Deep Learning-based Mammogram Classification for Breast Cancer," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 8, no. 4, pp. 171–176, 2020, doi: 10.18201/ijisae.2020466308.

[19] N. P. Firdi, T. A. Sardjono, and N. F. Hikmah, "Using Pectoral Muscle Removers in Mammographic Image Process to Improve Accuracy in Breast Cancer," *Journal of Biomimetics, Biomaterials and Biomedical Engineering*, vol. 55, pp. 131-142, 2022, doi: 10.4028/p-35cy9o.

[20] T. Akilan, "A foreground inference network for video surveillance using multi-view receptive field," *arXiv preprint arXiv:1801.06593,* 2018. [Online]. Available: https://arxiv.org/abs/1801.06593.

[21] X. Gu, Z. Shi, and J. Ma, "Multi-view Learning for Mammogram Analysis: Auto-Diagnosis Models for Breast Cancer," *IEEE International Conference on Smart Internet of Things (SmartIoT)*, pp. 149-153, 2018, doi: 10.1109/SmartIoT.2018.00035.

[22] N. Tavakoli, M. Karimi, A. Norouzi, N. Karimi, S. Samavi, and S. M. R. Soroushmehr, "Detection of Abnormalities in Mammograms Using Deep Features," *J. Ambient Intell. Humanized Comput.*, pp. 1–13, 2019, doi: 10.1007/s12652-019-01639-x.

[23] N. F. Hikmah, T. A. Sardjono, W. D. Mertiana, N. P. Firdi, and D. Purwitasari, "An Image Processing Framework for Breast Cancer Detection Using Multi-View Mammographic Images," *EMITTER Int'l J. of Engin. Technol.*, vol. 10, no. 1, pp. 136-152, 2022, doi: 10.24003/emitter.v10i1.695.

[24] M. Heath, K. Bowyer, D. Kopans, R. Moore, and P. Kegelmeyer, "The digital database for screening mammography," in *Proc. 5th Int. Workshop Digit*. Mammography, 2000, pp. 212-218.

[25] H. Li, J. Niu, D. Li, and C. Zhang, "Classification of breast mass in two-view mammograms via deep learning," *IET Image Processing*, vol. 15, no. 2, pp. 454-67, 2021, doi: 10.1049/ipr2.12035.

BIOGRAPHIES OF AUTHORS

Sisilia Anggraini was born in Merauke on September 12, 2001. She attended SMA Negeri 1 Merauke for her senior high school education and earned her Bachelor's degree in Biomedical Engineering from the Sepuluh Nopember Institute of Technology (ITS), Indonesia, in 2023. During her studies, she specialized in medical imaging and image processing. She was an active member of the Catholic Student Family organization, serving as a business staff and holding the position of vice-chairperson at the KMK Entrepreneur Day event in 2021. Additionally, she was involved in the annual departmental event committee and held the role of secretary coordinator for BEACON 2021. She can be contacted at email: sisilia.19073@mhs.its.ac.id.

Tri Arief Sardjono earned his bachelor's degree in Electrical Engineering from the Sepuluh Nopember Institute of Technology (ITS), Indonesia, in 1994. He then pursued his Master's degree in the field of Informatics Engineering from the Bandung Institute of Technology (ITB) and received his master's degree in 1999. Lastly, he continued his education for a Ph.D. in Biomedical Engineering at the University of Groningen, Netherlands, and earned his doctoral degree in 2008. Currently, he works as a lecturer at the Department of Biomedical Engineering, Faculty of Electrical Technology and Intelligent Informatics, Sepuluh Nopember Institute of Technology (ITS), Indonesia. His research interests encompass medical imaging, ultrasound imaging, and analysis. He can be contacted at email: sardjono@bme.its.ac.id.

Nada Fitrieyatul Hikmah earned her Bachelor's degree in Biomedical Engineering from Airlangga University (UNAIR) in 2012. She then pursued her Master's education in Electrical Engineering-Electronics, specializing in Biomedical Engineering, at the Sepuluh Nopember Institute of Technology (ITS) and earned her master's degree in 2016. Currently, she works as a lecturer at the Department of Biomedical Engineering, Faculty of Electrical Technology and Intelligent Informatics, ITS, Indonesia. Her research interests include cardiac engineering, signal processing, and medical imaging. She can be contacted at email: nadafh@bme.its.ac.id.