International Journal Of Health, Engineering And Technology (IJHET) *E-ISSN 2829 - 8683*
Volume 5, Number 3, September 2026, *Page. 3323 - 3332*
Email : editorijhess@gmail.com

Implementation Of Convolutional Neural Network(CNN) Resnet-50 Architecture For Mammography Image Classification In Early Breast Cancer Detection

Akbar Muhammad Dzidan[1)], Bagus Satrio Waluyo Poetro[2)], Moh. Taufik Febriansah[3)]
[1,2,3)] Sultan Agung Islamic University

*Corresponding Author
E-mail: akbardzidan01@gmail.com

Abstract

Breast cancer is one of the leading causes of death among women. Early detection through mammography screening is crucial to reduce mortality rates. However, manual image interpretation by radiologists is susceptible to subjectivity and visual fatigue, which can lead to misdiagnoses (False Positives and False Negatives). This study aims to design and implement a deep learning model using a Convolutional Neural Network (CNN) with the ResNet-50 architecture to classify mammography images into three tissue conditions: Benign, Malignant, and Normal. The model was developed using a Transfer Learning approach based on ImageNet weights, with specific modifications made to the classifier head layer. The secondary dataset was sourced from Kaggle, resulting in a final total of 23,939 images after undergoing preprocessing stages that included background removal (Auto-Crop), dimensional resizing (224x224 pixels), normalization, and geometric data augmentation. Model training was optimized using a Two-Stage Training strategy consisting of Feature Extraction and Fine-Tuning—to overcome potential performance degradation caused by an imbalanced data distribution (class imbalance) and to maximize the recognition of pathological features. Testing results conducted on 4,542 images in the testing data demonstrated that the proposed model achieved an Overall Accuracy of 98%. The evaluation of the F1-Score metric recorded a value of 0.98 for the Benign and Malignant classes, and a perfect 1.00 for the Normal class. Furthermore, the model exhibited excellent sensitivity (Recall) with zero False Negative predictions for malignant cases detected as Normal. Overall, this modified ResNet-50 architecture is proven to be reliable, precise, and holds significant potential to be implemented as a Computer-Aided Diagnosis (CAD) system to support the efficiency of medical professionals' clinical decisions.

Keywords*: *Breast Cancer, Convolutional Neural Network, Mammography, ResNet-50

INTRODUCTION

CancerDadBreast cancer is one of the most common types of cancer found in women. According to data from the World Health Organization (WHO), breast cancer cases reach around 2 million each year and contribute to 684,996 deaths. By the end of 2020, approximately 7.8 million women had been diagnosed with breast cancer in the past five years. Early detection efforts are a crucial step in reducing the death rate from this disease. Over the past three decades, breast cancer mortality rates have shown a downward trend, in line with the implementation of early detection screening, increased understanding of the disease, and the development of increasingly effective therapeutic strategies.(Suparna & Sari, 2022)

Early breast cancer detection has the potential to significantly reduce the risk of death. Approximately 43% of cancer deaths can be prevented through regular early screening and efforts to avoid various triggers. Failure to treat breast cancer in its early stages is often fatal. Delayed detection causes the spread of cancer cells (metastasis) to other organs, making medical therapy ineffective, drastically increasing mortality rates, and multiplying the cost of treatment compared to managing cases in the early stages. In 2019–2020, cancer treatment funding through the Social Security Agency (BPJS) reached approximately IDR 7.6 trillion. Indonesia's National Breast Cancer Management Strategy establishes three main pillars: health promotion, early detection, and case management. These three pillars are aimed at achieving early detection in 80% of women aged 30–50, encouraging 40% of cases to be identified in stages 1 and 2, and ensuring patients receive treatment within a maximum of 90 days.(Rizky et al., 2022).

Early detection of breast cancer can be achieved through breast screening, a test that identifies changes or abnormal cell proliferation in breast tissue before clinical manifestations appear. Several methods can be used, including mammography, breast ultrasound, breast MRI, breast self-examination

(SADARI), and clinical breast examination (SADANIS). To date, mammography is the screening method with the strongest evidence of effectiveness..

Mammography is an imaging method that utilizes low-energy X-rays to evaluate the structure of breast tissue. The examination is performed using standardized projections to identify abnormalities, particularly masses and calcifications. The sensitivity of mammography ranges from 63–95%, with sensitivity exceeding 95% for palpable lesions, around 50% for non-palpable lesions, 83–92% in women over 50, and can decrease to 35% in high-density breasts. Meanwhile, specificity ranges from 14–90%, reaching 90% for palpable lesions. Mammography remains the gold standard in breast cancer screening because it can detect malignancies at an early stage, including through the identification of microcalcifications before clinical manifestations appear.(Soekersi et al., 2022)However, manual interpretation of mammograms by radiologists has many weaknesses and limitations. In women with dense breast tissue, mammography is often less effective because white cancers are difficult to distinguish from dense, white glandular tissue. Radiologists may miss tumors due to their small size, difficult location, or concealment (false-negative results). Or, conversely, radiologists may interpret normal tissue as suspicious, causing patient anxiety and unnecessary additional procedures (false-positive results). Radiologist fatigue, impaired concentration, and variability in skill level can also contribute to inconsistent interpretation of results.(Nelda.A et al., 2024).

PeThe development of Artificial Intelligence (AI) is increasingly showing significant promise in the healthcare sector, particularly through the application of computer vision. In the process of interpreting medical images, AI technology can identify and extract image characteristics automatically and consistently, so that analysis results are not easily influenced by fatigue or human subjectivity. One deep learning method widely used in this field is the Convolutional Neural Network (CNN). CNN architecture has proven highly reliable in analyzing visual data due to its ability to capture spatial patterns, from edges and textures to complex structures in images. In mammography images, CNNs can be trained to distinguish between normal tissue, benign lesions, and malignant lesions with accuracy comparable to radiological assessment (Sanusi et al., 2023)..

Although CNNs have demonstrated promising performance, there are challenges in building deeper networks. Conceptually, increasing the number of layers is expected to expand the model's capacity to extract increasingly complex patterns and features. However, excessive network depth can trigger vanishing gradients, a condition where the gradient values responsible for weight updates shrink significantly, approaching zero during the backpropagation process. This condition makes it difficult for the initial layers to learn optimally, so that increasing network depth no longer results in performance improvements and can even degrade it. This problem was then addressed through the development of the Residual Network (ResNet) architecture by implementing skip or shortcut connections. This approach allows information to pass directly through one or more layers, thus maintaining gradient flow and allowing the network to be built up to 50 layers or more without experiencing performance degradation. In this study, the ResNet-50 architecture was used because it offers a good compromise between network depth and computational requirements. Furthermore, its ability to extract detailed image characteristics, including microcalcifications and lesion contours in mammography images, makes it relevant for the development of an automated breast cancer detection system.(Pramesti, 2023).

Based on the description above, a computer-aided diagnosis support system is needed that can assist radiologists in detecting breast cancer early from mammography images more accurately and efficiently. The application of deep learning with the ResNet-50 architecture is expected to be able to classify mammography images into normal and abnormal (cancer) categories by utilizing the ability of skip connections to capture complex features without experiencing performance degradation. Therefore, this study will build an automatic breast cancer detection model in mammography images using the CNN algorithm with the Residual Network (ResNet-50) architecture.

RESEARCH METHODS

This study uses a quantitative approach with experimental methods and applied research design to implement and test the performance of the ResNet-50 Convolutional Neural Network (CNN) architecture in classifying mammography images into benign, malignant, and normal classes. The research data is an open mammography image dataset obtained from Kaggle, then goes through the stages of data collection, pre-processing, augmentation, and division into training, validation, and testing. The pre-processing stage includes auto-crop to remove irrelevant backgrounds, image resizing to standardize image size, and pixel value normalization. Next, augmentation is carried out in the form of rotation, shift, shear, zoom, horizontal flip, and brightness adjustment to increase data diversity and reduce the risk of overfitting. The dataset division is done proportionally while maintaining the characteristics of the class distribution to test the model's generalization ability.

The ResNet-50 model was designed using a Transfer Learning approach with ImageNet pre-trained weights and equipped with GlobalAveragePooling2D, Dense Layer, Dropout, and L2 regularization, while the output layer uses the Softmax activation function to produce multi-class classification. The training process uses the Adam optimizer and the Categorical Cross-Entropy function through a Two-Stage Training strategy, namely the Feature Extraction stage by freezing the base layer and the Fine-Tuning stage by opening some of the model weights to adjust the features to the characteristics of the mammography image. The training process was controlled using EarlyStopping and ReduceLROnPlateau. Model evaluation was carried out on test data using the Confusion Matrix and the Accuracy, Precision, Recall, and F1-Score metrics. This analysis was used to determine the model's ability to distinguish each class, especially in detecting malignant cases, as well as assess the model's robustness and generalization ability to data with class imbalance.

RESULTS AND DISCUSSION

Experimental Scenario and Preprocessing

Dataset Distribution and Class Imbalance Problem

This study used a dataset of mammography images categorized into three main classes,that isBenign, Malignant, and Normal. Based on the data exploration results, the class distribution in the dataset shows a significant class imbalance. In the Data Training phase, the Malignant class dominated with 10,427 samples, while the Normal class was the minority class with only 1,296 samples.

This imbalance is a reflection of the real challenges in medical datasets, where samples of certain pathological cases are sometimes more or less obtained thansampleNormal. Despite the imbalance, the split ratio is maintained proportionally in the Testing Data to test the model's robustness. The use of the F1-Score and Macro Average metrics in the subsequent evaluation stage is expected to provide objective and unbiased performance measurements even when the model is faced with an uneven class distribution.

Image Preprocessing Stages

In the preprocessing stage, the raw image is executed through three main transformations to make it ready for use by the model. The details of these three transformation stages are illustrated in the figure below.

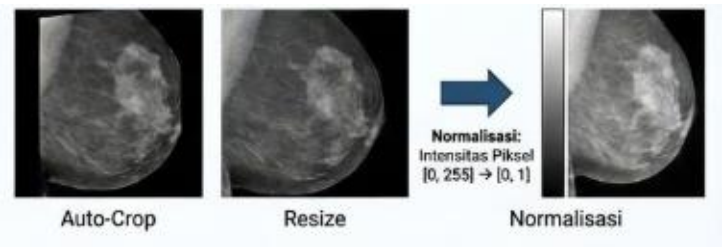

Figure 1. Details of Image Pre-Processing Transformation Stages

Based on the steps in Figure 1, there are three main processes. First, Auto-Crop automatically removes empty background areas; for example, the original image with 40% black area is reduced so that the matrix fully focuses on the Region of Interest (ROI) of breast anatomical tissue. Second, the cropped image is resized (Image Resizing) to a standard 224 x 224 pixel matrix, which is an absolute requirement for the input layer dimension in the ResNet-50 architecture. Third, pixel intensities are normalized (Data Normalization) into the range [0, 1] to standardize the input scale and accelerate the convergence of the optimization function without destroying the clinical details of the image.

Next, the image is passed through the built-in preprocessing function of the ResNet50 architecture which includes normalizing pixel values to a certain range and adjusting the dimensions of the image matrix. The Contrast Limited Adaptive Histogram Equalization (CLAHE) technique is intentionally skipped at this stage because the original dataset used is known to have undergone this contrast enhancement process during the data production stage, so additional contrast manipulation is no longer necessary.

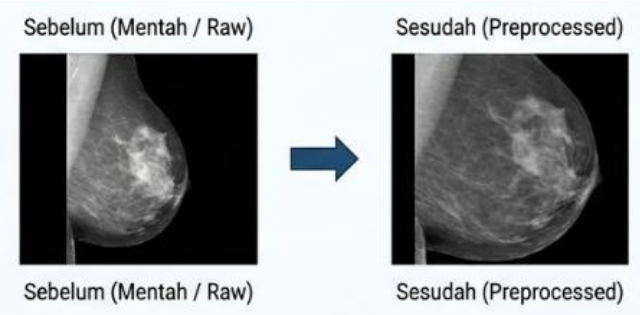

Figure 2. Comparison of images before (raw) and after pre-processing

Figure 2 shows the final result of how the raw image (left) is transformed into a processed image (right). Through cropping and adjustment, the image is now free of unnecessary background and more focused on the breast tissue area, ready for processing by the model architecture.

Data Augmentation

To minimize the risk of premature aging in the learning process (overfitting) and increase the model's robustness to variations in real-world medical images, a data augmentation strategy was applied to the training data. The applied geometric and photometric transformations included random rotation, width/height shift, shear, zoom, horizontal flip, and brightness adjustment. This augmentation forced the model to learn tumor morphological features that were invariant to the position, scale, and illumination of the mammography scanner.

Model Architecture and Training Scenarios

Base Model Specifications

This experiment implements the ResNet50 architecture as a base model. The model's parameter weights are initialized using pre-trained weights from the ImageNet dataset. Transfer learning aims to accelerate model convergence by leveraging the extraction capabilities of low-level features, such as edges, shapes, and textures, previously learned by large-scale models.

Training Stages (Transfer Learning)

The network weight optimization process is carried out through a two-stage training scheme:

Stage 1 (Feature Extraction)

All layers in the base ResNet50 model are frozen. Training focuses only on the newly added classifier head. The classifier head consists of a GlobalAveragePooling2D layer, followed by a Dense (256 neurons), Dropout, Dense (128 neurons), Dropout, and finally a Dense output layer (3 neurons with a Softmax activation function). This stage runs for 10 epochs with a learning rate of 1×10^{-4}. The main goal of this stage is to adapt the random weights in the new classifier to match the feature representation of ResNet50 without destroying the existing pre-trained features.

Stage 2 (Fine-Tuning):

After the classifier head reaches stability, the parameter locks on the last 50 layers of ResNet50 are unfrozen to be retrained along with the classifier. This fine-tuning phase runs for 10 epochs using

a much smaller learning rate of 1×10^{-5}. This low learning rate ensures that weight changes are gradual and specific to recognizing the microscopic characteristics of mammography images without triggering catastrophic forgetting.

Regularization and Callbacks Strategy

To maintain performance stability, L2 regularization with a penalty coefficient of 0.001 is inserted in each Dense layer, alongside a Dropout layer with a value of 0.3. In addition, two dynamic callbacks are integrated during the training process: EarlyStopping with a patience parameter of 10 to stop iterations if there is no decrease in validation loss, and ReduceLROnPlateau (multiplier of 0.5 and patience of 4) to automatically reduce the learning rate when the learning graph begins to plateau.

Training Results and Internal Validation

Analysis of Accuracy and Loss Function Graphs

The convergence evaluation process and model performance during the training period are monitored periodically through the accuracy curve (Training and Validation Accuracy) and the loss function curve (Training and Validation Loss). A comprehensive learning graph combining Phase 1 (Feature Extraction) and Phase 2 (Fine-Tuning) is presented in Figure 4.1 below:

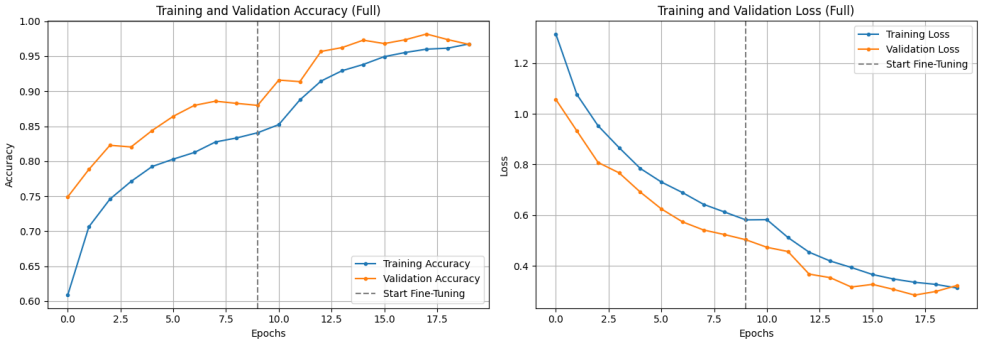

Figure 3. Graph*Accuracy and Loss Model*

Based on*Visualizing* the movement of the curve in Figure 3, there are several transitional learning dynamics that are crucial to analyze:

Phase 1 Analysis: Feature Extraction (Epoch 0 - 9):

In the first 10 epochs, where all layers of the ResNet50 base model are frozen, the graph shows a stable and progressive learning trend. The Training Accuracy value creeps up from an initial figure of around 0.61 to reach around 0.84 in the 9th epoch. This pattern is linear with the Training Loss curve which is successfully suppressed consistently from 1.30 to 0.58. In this phase, the Validation Accuracy value is consistently above the training accuracy. This healthy phenomenon indicates that the ImageNet pre-trained weights provide highly informative universal features, supported by the Dropout (0.3) regulation which works effectively to prevent early memory interference (overfitting).

Transition Analysis and Phase 2: Fine-Tuning (Epoch 10 - 19):

The vertical dotted line on the graph indicates the start of Phase 2 (Start Fine-Tuning), which is the unlocking of parameters in the last 50 layers of ResNet50 along with a decrease in the learning rate to 1×10^{-5}. Shortly after this transition occurs (between epochs 9 and 10), the learning curve responds with a significant spike in accuracy, where the validation accuracy jumps from 0.88 to exceed 0.91 in just one epoch. This spike proves that unlocking the top layers of the base model successfully gives the model the freedom to recalibrate its abstract weights to become much more specific in recognizing the pathological cellular morphology characteristics of mammography.

Final Conditions for Model Convergence:

Entering the final half of training (epochs 12 to 19), the two curves converged harmoniously. The final Training Accuracy and Validation Accuracy values converged at an optimal point in the range of 96% to 98%. A similar alignment was also shown in the loss function graph, where the Training Loss and Validation Loss reached a stable low point around 0.31. The absence of a wide gap between the training and validation curves at the end of the epoch confirmed that the model avoided overfitting or underfitting and possessed very robust internal generalization capabilities.

Performance on Internal Validation Data Through Confusion Matrix and Classification Report

To analyze the model's prediction accuracy in more detail and identify patterns of misclassification between classes, an evaluation was conducted using a confusion matrix and a classification report. The chaos matrix plots the amount of real data against the predicted results, while the classification report provides quantitative metrics such as precision, recall, and f1-score for each pathological class.

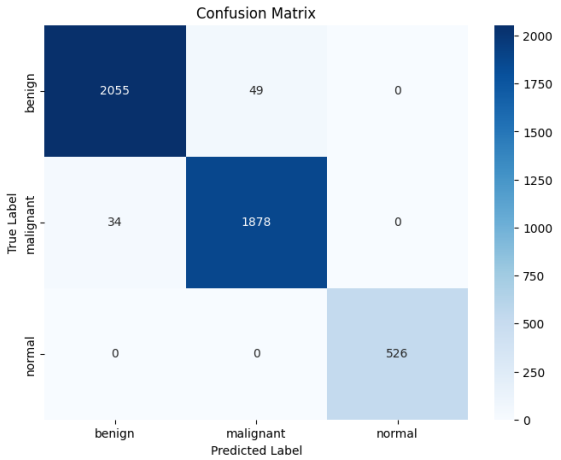

Figure 4.*Confusion Matrix*

The results of a comprehensive quantitative evaluation of the classification report on internal validation data are presented in Table 1 below:

Table 1. Classification Report

	precision	recall	f1-score	support
benign	0.98	0.98	0.98	2.104
malignant	0.97	0.98	0.98	1,912
normal	1.00	1.00	1.00	526
accuracy			0.98	4,542
macro avg	0.99	0.99	0.99	4,542
weighted average	0.98	0.98	0.98	4,542

Based on the quantitative data mapped in Table 1 and aligned with the Confusion Matrix data in Figure 4, the model classification performance for each class can be described as follows:

Benign Class:

The Benign class has the largest number of test samples (support), namely 2,104 images. The model recorded a very balanced and high precision, recall, and f1-score value at 0.98. Of the total samples, the model successfully predicted 2,055 samples correctly (True Positive). The classification error margin for this class is very low, where only 49 samples were incorrectly detected as Malignant, and no samples (0) were incorrectly classified into the Normal class.

Malignant Class:

In the Malignant class with a sample of 1,912 images, the model produced a precision value of 0.97 and a recall of 0.98, resulting in a robust f1-score of 0.98. This high recall value proves that the model successfully identified 1,878 samples accurately from the total number of existing malignancy cases. There was a minor margin of error of 34 samples that were incorrectly predicted as Benign. The most crucial clinical indicator is that there were no malignancy cases (0 samples) that were incorrectly detected as Normal, which means the risk of a missed diagnosis (False Negative) that endangers patient safety was successfully minimized by this artificial neural network.

Normal Class (Healthy):

The model demonstrated absolute specificity and sensitivity in the Normal class (526 samples), achieving a perfect precision, recall, and f1-score of 1.00 (100% accurate). This demonstrates that the model successfully predicted all Normal data samples without a single error. No normal breast images were misdiagnosed as benign or malignant, thus clinically minimizing the risk of patient anxiety due to false positive diagnoses.

Overall, the developed ResNet50 model achieved an overall accuracy of 0.98 from a total validation test dataset of 4,542 images. This achievement was reinforced by a macro average value of 0.99 and a weighted average of 0.98 across all evaluation metrics. The stable and even concentration of metric values above 0.97 empirically proves that the model does not experience bias or prediction inequities caused by oversampling in the training data, but has successfully generalized the unique characteristics of each mammography image class with precision in this internal testing domain.

Visual Analysis of Image Prediction Results (Inference Output)

To qualitatively verify the model's spatial interpretation capabilities, direct inference testing was conducted on several representative validation image samples from each class. The visual detection results, along with their true and predicted labels, are presented in Figure 5 below:

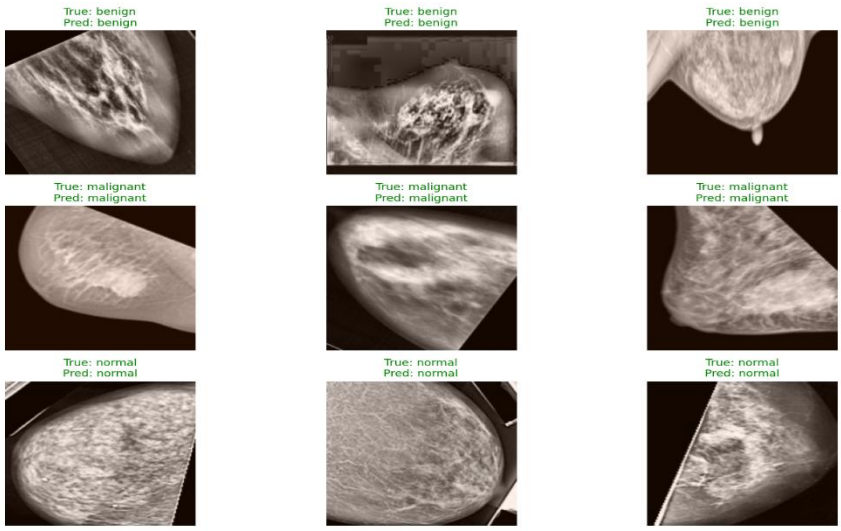

Figure 5. Model Prediction Results on Mammography Validation Images

Based on the visualization of the model output sample in Figure 4.3, the analysis of the detection characteristics can be described as follows:

Benign Class (First Row):

The model successfully identified breast tissue images with benign lesion characteristics accurately (True: benign, Pred: benign). It can be seen in the first and second image samples, despite the complex gland density patterns and varying grayscale gradations, the model is still able to localize the mass features with well-defined margins without being confused by the surrounding structural noise.

Malignant Class (Second Row):

In the sample image of the malignancy case (True: malignant, Pred: malignant), the model demonstrated very high sensitivity to the architecture of invasive tumor masses. The model successfully recognized tissue architecture distortions, irregular lesion edges (spiculates/ill-defined margins), and areas of high opacity, which are key clinical indicators of both early and advanced breast cancer.

Normal Class (Third Row):

The model correctly classifies healthy breast tissue without the presence of pathological masses (True: normal, Pred: normal). The images in this row display the intact pattern of normal fibrogranular

breast parenchyma. The success of this prediction demonstrates the model's robust specificity, allowing it to distinguish between dense breast tissue and a true tumor mass.

The green color in the labeling text for all image samples confirms the consistent prevalence of True Positives. These visual inference results align the qualitative evidence with the quantitative findings in the previous graphs and matrices, while also strengthening the conclusion that the developed ResNet50 model has a deep understanding of radiological anatomical features in this dataset domain.

Analysis of Clinical Implications of Misclassification

Although the proposed ResNet-50 model achieved an overall accuracy rate of 98%, a thorough analysis of the remaining error margin is crucial to evaluate the feasibility of implementing this system in a real-world clinical setting. Based on the Confusion Matrix in Figure 5, there are two types of cross-classification errors (misclassifications) occurring between the Benign and Malignant classes.

Cross False Positive Analysis (Predicted Malignant, Actual Benign)

There were 49 benign tissue image samples that were incorrectly predicted by the model as malignant. In medical terminology, this case is categorized as overdiagnosis. In visual computing, this error is generally triggered by the presence of complex cysts, coarsely calcified fibroadenomas, or overlapping dense breast tissue that morphologically form an opacity pattern or irregularly edged shadow that closely resembles the characteristics of malignant tumors. From a clinical perspective, this type of error (false positive) is considered a "tolerable" error rate. The main impact is temporary psychological anxiety in patients and the need for further diagnostic procedures (such as biopsy or MRI). Although cost-inefficient, this does not directly endanger the patient's life.

Cross False Negative Analysis (Predicted Benign, Actual Malignant)

Conversely, 34 malignant image samples failed to be identified and were predicted as benign. This error is the most dangerous type of underdiagnosis in the Computer-Aided Diagnosis (CAD) domain. In feature extraction, this failure can occur if the cancer cells are still at a very early stage (in situ) where the size of the microcalcifications is too small, or the tumor lesion has smooth edges (circumscribed margins) that are visually identical to benign tumors. The clinical implications of this cross-false negative are very fatal, because it can cause patients to miss the golden time to receive oncological treatment, which leads to the spread of cancer (metastasis) and an increased risk of mortality.

Despite 34 cases of inter-lesion underdiagnosis, the model's absolute reliability is demonstrated by its ability to reduce the False Negative rate to zero (0) when faced with completely normal breast tissue. This means that not a single cancer case was considered "Healthy/Normal" by the system. This fact confirms that the developed ResNet-50 model has an exceptionally robust first-line screening sensitivity in separating diseased breasts from completely healthy ones.

Comparative Analysis of Performance with Previous Research

To measure the success and reliability of the proposed architecture, a comparative analysis was conducted between the performance of the ResNet-50 model in this study and several previous studies focused on breast cancer classification using various image modalities and deep learning algorithms. A summary of the performance comparison is presented in Table 2.

Table 2. Comparison of Model Performance with Previous Research

Researcher / Year	Image Modality (Dataset)	Methods & Architecture	Accuracy
Lestandy, et al. (2022)	Histopathology (Kaggle)	Custom CNN	80.00%
Idawati, et al. (2024)	Ultrasonography / USG (BUSI)	CNN VGG-16	78.87%
Nelda, et al. (2024)	Mammography (INbreast)	CNN ResNet-50 (Modified)	97.00%

Purnama, et al. (2025)	Mammography (Buleleng Regional Hospital, etc.)	MV-CNNs (ConvNeXt)	50.00%
Arwoko & Ariyani (2026)	Ultrasonography / USG (BUSI)	CNN MobileNet	89.00%
This Research (2026)	Mammography (Kaggle - 23,939 images)	CNN ResNet-50 (Modified + Two-Stage Training)	98.00%

Based on Table 2, it can be seen that the implementation of the CNN ResNet-50 architecture withapproachThe modifications to the classifier head layer proposed in this study successfully recorded the highest performance.with an accuracy of 98.00% and an F1-Score of 98.00%.

Specifically, when compared with research that uses the same architecture and image modality, namely research(Nelda.A et al., 2024), the proposed model showsimprovementaccuracy of 1.00%. While the improvement may seem marginal in percentage terms, this achievement has much stronger technical significance.(Nelda.A et al., 2024)While the model in this study was trained and validated on a much larger and more varied dataset (INbreast), a total of 23,939 images after expansion, the model's ability to maintain 98% accuracy on the large-scale dataset demonstrates the robustness of the Two-Stage Training (Feature Extraction and Fine-Tuning) strategy.

Furthermore, when compared with other architectures such as the VGG-16 by(Idawati et al., 2024), which yields an accuracy of 78.87% or MobileNet by(Arwoko & Ariyani, 2026)With an accuracy of 89.00%, the superiority of ResNet-50 is very clear. The skip connection mechanism in ResNet allows the extraction of spatial texture features from complex medical images without experiencing the vanishing gradient problem, a problem that often hinders conventional networks. This is also in contrast to research(Purnama et al., 2025), which uses ConvNeXt but suffers from severe overfitting with an accuracy of only 50.00%.

In addition to being superior in accuracy, the application of the Transfer Learning strategy (utilizing ImageNet pre-trained weights) makes the ResNet-50 training time parameters much more efficient compared topracticemodel from scratch. In terms of inference time, ResNet-50 is slightly heavier than lightweight architectures like MobileNet, but its skip connection capability ensures its computations continue in milliseconds per image. This speed meets real-time operational standards for use as a second opinion in Computer-Aided Diagnosis (CAD) systems.

CONCLUSION

Based on the research results, the implementation and testing of the Convolutional Neural Network (CNN) model with the ResNet-50 architecture were successfully applied for mammography image classification. The use of Transfer Learning, classifier head modification, and step-by-step training resulted in effective and stable feature extraction. The preprocessing and augmentation stages also increased the model's resilience to data variation and imbalance. With the skip connection mechanism, ResNet-50 was able to overcome vanishing gradients and recognize complex pathological patterns, resulting in excellent classification performance in distinguishing benign, malignant, and normal conditions.

REFERENCES

Arwoko, H., & Ariyani, S. (2026). *Ultrasound Image Classification of Breast Cancer Using MobileNet*. https://doi.org/https://doi.org/10.32528/justindo.v11i1.4860

Idawati, Rini, D. P., Primanita, A., & Saputra, T. (2024). *Klasifikasi Kanker Payudara Menggunakan Metode Convolutional Neural Network (CNN) dengan Arsitektur VGG-16*. https://doi.org/https://doi.org/10.30865/json.v5i3.7553

Nelda.A, A. R., Setiawan, R., Hendrawan, M. A., & Nugroho, D. B. (2024). *Deteksi Kanker Payudara Hasil Citra Mammografi menggunakan Metode Convolutional Neural Network (CNN) Arsitektur ResNet-50*.

Pramesti, D. D. (2023). *Kasifikasi kanker payudara berdasarkan citra mammogram menggunakan metode convolutional neural network (cnn) model nasnet mobile*.

Purnama, G., Permana, A. A. J., & Kertiasih, N. K. (2025). Klasifikasi Citra Medis Mamografi Berbasis Multi-View Convolutional Neural Networks. *Jurnal Pendidikan Teknologi Dan Kejuruan*.

Rizky, A., Putri, W., & Yudhana, A. (2022). Klasifikasi Kanker Payudara Menggunakan Metode Digital Mammogram. *Jurnal Teknik Informatika Dan Sistem Informasi*, *9*(4).

Soekersi, H., Azhar, Y., & Akbari, K. S. (2022). Peran Mammografi Untuk Skrining Kanker Payudara: Sebuah Tinjauan Pustaka. *Journal Of The Indonesian Medical Association*.

Suparna, K., & Sari, L. M. K. K. (2022). *Kanker Payudara: Diagnostik, Faktor Risiko, Dan Stadium. Ganesha Medicine*. https://doi.org/https://doi.org/10.23887/gm.v2i1.47032