ADBA Computational Systems and Artificial Intelligence
Cutting-Edge Scientific Solutions

e-ISSN: 3108-4079
RESEARCH ARTICLE
Vol.1 / No.1 / 2025 / pp.26-31
https://doi.org/10.69882/adba.csai.2025075

# Artificial Intelligence in Mammography: A Study of Diagnostic Accuracy and Efficiency

**Luaay Alswilem** [*,1] **and Nurettin Pacal** [α,2]
[*]Department of Computer Engineering, Faculty of Engineering, Igdir University, 76000, Igdir, Türkiye, [α]Department of Biology, Faculty of Arts and Sciences, Igdir University, 76000, Igdir, Türkiye.

**ABSTRACT** Breast cancer continues to be a considerable global health problem, highlighting the need for early and accurate diagnosis to improve patient outcomes. Although mammography is widely considered the gold standard for screening, its interpretation is not straightforward and varies among readers. Our study aimed to compare the performance and computational efficiency of three leading Convolutional Neural Network (CNN) architectures for classifying breast cancer automatically from mammogram images. We used a publicly available dataset consisting of 3,383 mammogram images, which were labeled as either Benign or Malignant, and we trained and evaluated three models: EfficientNetB7, EfficientNetv2-Small, and RexNet-200. We found the RexNet-200 architecture had the best performance across the performance metrics we measured, achieving the best accuracy (76.47%), precision (75.18%), and F1-score (77.44%). Even though EfficientNetB7 had a slightly better recall than the RexNet-200 model; the RexNet-200 model showed a more compelling accuracy-board balance in diagnosis. Furthermore, RexNet-200 had the best performance and lowest computational cost with a very low parameters count (13.81M) and lowest GFLOPS (3.0529) of the three models. Our study demonstrated that RexNet-200 had the best prospects for achieving the ideal balance of high diagnostic accuracy and economical use of resources. Therefore, RexNet-200 is a very promising candidate for incorporation into clinical decision support systems designed to assist radiologists in the early detection of breast cancer.

**KEYWORDS**
Breast cancer
Mammography
Deep learning
Computational efficiency
RexNet-200

## INTRODUCTION

Cancer is ranked as one of the most complex and lethal diseases in modern medicine, and is primarily characterized by the dysregulated proliferation of cells into abnormal growths called tumors. When healthy tissue is formed by regulated cell growth, division, and death to maintain homeostasis, this balance is disrupted by genetic mutations in cancer, continuously stimulating cells to form tumors (Kurtulus *et al.* 2024a; Bayram *et al.* 2025). Tumors harm local tissue and can spread into distant organs through the bloodstream or lymphatic system (referred to as metastasis). Such progression can compromise the function of vital organs and significantly increase mortality risk. Cancer is one of the worldwide leading causes of death, yielding millions of lost lives as a result. Appropriately cancer is the focus of research, with the understanding of the biology underpinning cancer and the promotion of early and accurate diagnosis to improve survival and efficacy of treatment (Lubbad *et al.* 2024a).

Breast cancer is the most common type of cancer, especially among women. It represents an immense proportion of new cases and deaths annually. Like many cancers, the appearance of breast cancer involves multiple motivations, including genetic, hormonal, individual lifestyle behaviors, and environmental exposures. While breast cancer is concerning, it is very treatable when detected early. An early diagnosis means that a tumor or tumors can be identified before it has a chance to invade contiguous tissue, thus requiring a much less aggressive and invasive treatment option. Additionally, early diagnosis has remarkably high five-year survival rates, greater than 90% in many cases (İnce *et al.* 2025; Lubbad *et al.* 2024b). Thus, public health campaigns and interventions directed toward greater awareness and regular screening have been effective in reducing breast cancer morbidity and mortality globally.

While there are many potential diagnostic modalities, mammography has remained a universal standard diagnostic tool for early breast cancer. Mammography involves low-dose X-ray imaging for the breast, creating radiographic images with sufficient detail to detect small masses that are not palpable, architectural distortion, and microcalcifications, which when clustered are precursory indicators of early malignancy (Ozdemir *et al.* 2025). However, mammogram image interpretation is complex and subjective, not to mention dependent on the detail and experience of the radiologist in addition to contingent considerations such as the clinician's workload, burnout, and nuanced participant features on an image. Misinterpretation may lead to important missed diagnoses and unbelievbly, false positives.These errors can lead to delayed treatment, unnecessary biopsies, and increased anxiety for patients

**Manuscript received:** 10 June 2025,
**Revised:** 13 July 2025,
**Accepted:** 17 July 2025.

[1]luaayalswilem3@gmail.com (**Corresponding author**)
[2]nurettin.pacal@igdir.edu.tr

(Pacal and Attallah 2025; Pacal *et al.* 2025).

In response to these challenges, there is an increase in the demand for objective, rapid and dependable decision-support systems for mammogram evaluation. In recent years, artificial intelligence (AI) has advanced rapidly with a strong focus on deep learning resulting in improved decision-support systems for medical image evaluation and also obtained excellent results. Compared to traditional methods of supervised learning, reliance on specifically trained artificial intelligence using convolutional neural networks (CNN) obtained a new standard for extracting multi-level visual features for evaluation and in some cases has outperformed human experts analyzing, interpreting and accurately correlating mammograms to radiological images (Cakmak *et al.* 2024; Zeynalov *et al.* 2025; Pacal 2025). CNN models can learn the subtle differences between normal tissue patterns and benign or malignant lesions to improve improve the classification of mammogram images. In this study, three of the current top-of-the-line AI deep learning architectures EfficientNetB7, EfficientNet-V2-Small, and RexNet-200 were used to classify breast mammograms into benign (0) or malignant (1) classifications. The challenge of the study was to evaluate and compare the performance and computational complexity of all three models to determine the best AI-advantages solution to support all radiologists and improve diagnostic accuracy (Cakmak and Pacal 2025; Kurtulus *et al.* 2024b).

Artificial Intelligence (AI), particularly its branches of deep learning (DL) and hybrid modeling, continues to redefine the landscape of medical image analysis, offering new capabilities in precision, scalability, and interpretability. These advancements have had a significant impact on breast cancer diagnostics, enabling not only early detection but also improved classification and segmentation accuracy. A recent hybrid framework combining preprocessing methods such as CLAHE, Gaussian blur, and sharpening with ensemble models including YOLOv5 and MedSAM has demonstrated outstanding results, achieving an accuracy of 99.7% and a processing time of only 0.75 seconds per mammographic image. Another novel architecture, HybMNet, integrates a self-supervised Swin Transformer with a CNN via a fusion mechanism, delivering AUC scores ranging from 0.864 to 0.889 on benchmark datasets like CMMD and INbreast (Chen and Martel 2025). EfficientNet-B7 has additionally been utilized with explainability techniques such as Grad-CAM for ultrasound breast cancer classification, achieving 99.14% accuracy while enhancing the interpretability of model decisions (Latha *et al.* 2024).

Similarly, combining Attention U-Net with EfficientNet-B7 for thermal imaging enabled more effective segmentation and classification of suspicious breast regions, demonstrating strong clinical potential (Mridha *et al.* 2023). Regarding segmentation, a deep learning ensemble model based on CNN and a pruned Extreme Learning Machine (HCPELM) showed improved recognition performance on the MIAS dataset, reaching 86% accuracy (Sureshkumar *et al.* 2024). Generative adversarial networks (GANs) have also emerged in this domain; one study proposed a conditional self-attention GAN (ExCSA-GAN) designed for mammographic image analysis, with a focus on increasing both performance and model transparency (Sreekala and Sahoo 2025). Furthermore, the importance of explainable AI (XAI) in medical diagnostics has been emphasized. One comprehensive review evaluated the use of XAI specifically in mammography, proposing tailored criteria for assessing model transparency and reliability in clinical settings (Ansari *et al.* 2025). In essence, the literature highlights a growing trend toward hybrid AI systems, self-supervised learning, and interpretable models. These directions reflect the need for accurate, efficient, and trustworthy tools in breast cancer diagnosis.

## MATERIALS AND METHODS

### Dataset

This study was based on the publicly accessible "Breast Cancer Detection" dataset from the Kaggle platform, which was used to classify mammographic images into benign and malignant categories. The dataset features a total of 3383 pathologically verified mammograms, consisting of 2225 benign cases (class 0) and 1158 malignant cases (class 1), providing a comprehensive foundation for developing and validating deep learning models. To ensure the standardization of experiments and prevent overfitting, the dataset was meticulously divided using stratified sampling into three distinct subsets. The majority of the data, 70% (2367 images), was allocated for training the model. The remaining data was split equally, with 15% (506 images) used for validation and 15% (510 images) reserved for final testing. This stratified approach successfully preserved the class balance within each partition: the training set contains 1557 benign and 810 malignant samples, the validation set includes 333 benign and 173 malignant samples, and the test set comprises 335 benign and 175 malignant samples. A complete statistical breakdown of this data distribution is illustrated in Figure 1.

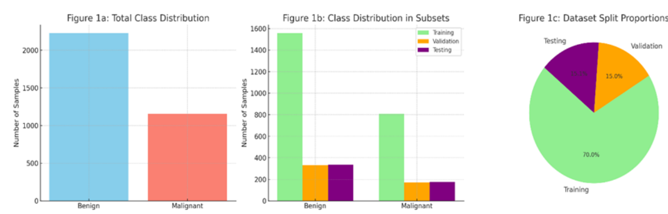


**Figure 1** Statistical Breakdown of the Dataset by Class, Subset, and Proportional Distribution

In addition to numerical analysis, visual examination of the mammographic data provides insight into the types of features present in each class. Representative images of benign and malignant mammograms are presented below. As shown in the examples, benign lesions typically appear as well-defined, rounded masses with smooth contours. In contrast, malignant tumors are more likely to demonstrate irregular shapes, spiculated (star-like) margins, and increased tissue density, often accompanied by microcalcification clusters, which are early indicators of malignancy. These visual distinctions, though significant, may be subtle and easily obscured due to challenges such as low image contrast, dense breast tissue, or overlapping anatomical structures. Such complexities highlight the importance of leveraging deep learning models capable of extracting discriminative visual patterns from complex mammographic data. Figure 2. Sample mammography images demonstrating benign (left) and malignant (right) characteristics.

### Data Augmentation

In order to enhance the generalization performance of the deep learning models and reduce the risk of overfitting - a common challenge faced in medical imaging tasks because of limited data. Our work utilized a comprehensive family of online data augmentation strategies during model training. The augmentations were applied dynamically, and randomly to each image in real-time during training, which maximized the exposure of the model to new visual stimulus and enhanced its robustness. The augmentation procedure included a number of spatial and color-based

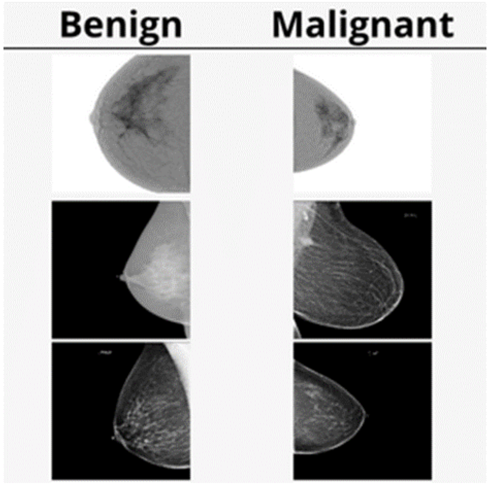


**Figure 2** Sample mammography images demonstrating benign (left) and malignant (right) characteristics

transformations. For all input images, the first transformation was a Random Resized Crop. The method randomly selected a location of the original image between a scale of 8% and 100% of the original image size, with an aspect ratio between 0.75-1.33. Next, the randomly selected crop was resized to an input resolution of 224×224 pixels while applying a random interpolation method to emphasize more variability in pixel representation. This cropping method has been shown to be an effective regularization method in the convolutional neural networks used in image classification, even when using small or imbalanced datasets (Faryna *et al.* 2024). Following the cropping, a Random Horizontal Flip was used with a probability of 0.5. The purpose of the random horizontal flip was to enable the model to learn invariance to left-right orientation, a transformation that is commonly used and clinically acceptable when a mammogram is analyzed (Islam *et al.* 2024).

Vertical Flip was consiously excluded, since these transformations may produce anatomically impossible variations in medical images, and potentially create ambiguity in the learning (Shorten and Khoshgoftaar 2019). As for color based augmentation, a Color Jitter operation was also added, which increased the brightness, contrast, saturation, and hue of the image in a random way by a factor of 0.4 for each operation. These variations are meant to measure differences in imaging conditions and equipment based settings to encourage the model to learn features that are structural and pathological and not lighting differences (Zhang *et al.* 2024). All of the techniques used for data augmentation in general help to expose the model to a wider distribution of data, essentially augmenting the training set a more opportunistic approach without a need for collection of additional data. This is particularly important in medical imaging, which is generally cost prohibitive and time-consuming to obtain large, balanced, and annotated datasets. In addition, the random augmentation techniques help to provide increased robustness and generalizability of the model when applied to unseen clinical situations and conditions (Tajbakhsh *et al.* 2020).

### Model Architectures

The study addressed the task of classifying breast cancer from mammograms in which high-performing deep CNN models were applied with measurable complexity. Due to the difficulty obtaining labeled medical data, the transfer learning method was used to exploit pre-trained models from large datasets, e.g., ImageNet, to accelerate training, improve generalization, and reduce the chances of overfitting (Tajbakhsh *et al.* 2016). Three state-of-the-art CNN models were chosen based on performance in image classification: EfficientNetB7, EfficientNet-V2-Small, and RexNet-200. EfficientNet series is characterized by the systematic optimization of networks in depth, width, and spatial resolution in a way that maintains accuracy with respect lesser in model parameters.

The EfficientNetB7 model, as the largest of its size, usually performs well. EfficientNet-V2-Small, on the other hand, offers a lighter, faster model, despite requiring fewer parameters at lower cost without sacrificing comparable accuracy. It is ideal for scenarios where a combination of speed and performance is desired. The RexNet-200 model represents a more recent and deeper architecture with the advantages of deep networks and the benefit of gradient flow strategies that offer up more stable training and may yield, accuracy, precision, and recall. It usually requires less parameters at lower computational cost, than other larger models. The study critically employed three models, in order to compare their performance, cost in relation to their performance, to discover a good model for mammogram classification, enabling the potential for developing intelligent diagnostic tools for early breast cancer detection. The comparative study of these three architectures provided value in understanding how their different architecture impacts biological image representations and computer vision tasks, and contributed to the main objective of developing trustworthy, AI-assisted diagnostic tools for the early detection of breast cancer.

### Performance Metrics

The evaluation phase is an essential part of any classification project based on deep learning, as it provides objective standards of performance for quantifying the model's performance and assessing the success of the proposed methodology. Solid evaluation metrics help to discover dataset imbalances, biases, and modification opportunities for the models to potentially alleviate overfitting and underfitting. In this evaluation study, specifically on the classification of breast cancer from mammographic images, we utilized some standard metrics that have broad acceptance and application: Accuracy, Precision, Recall, and the F1-score. Each of those metrics provides unique viewpoints on model performance that are equally important as part of the understanding of the effectiveness of the classification. Accuracy represents the correctness of the model by the percentage of correct predictions relative to the number of samples evaluated.

Though conceptually simple and widely understood, just looking at accuracy may not be adequate in a setting with class imbalance. Precision measures the accuracy of our positive predictions by looking at the ratio of true positive predictions to total positive predictions. If our precision score is high, then we can be fairly confident that the positive predictions we are making are really positive predictions. High precision is extremely useful when we want to limit false positives (for example misdiagnosing a healthy patient). Recall (also known as sensitivity) looks at the model's ability to find all real positive cases (out of the total real positive cases), so it measures the percentage of true positive predictions out of all real positive predictions. High recall is paramount, especially in a medical setting, because missing real cases of disease can result in catastrophic consequences. The F1-score is a score that attempts to combine precision and recall into one measurement through the harmonic mean of precision and recall. It balances both false positives and false negatives into one single score. The advantage of the f1 score is that it gives a more comprehensive evaluative technique, particularly in cases of imbalanced data or when using either precision or recall is not sufficient. The definitions for all of these metrics are shown here:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{1}$$

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

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

$$F_1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

As a whole, these metrics create a complete assessment framework for evaluating the diagnostic capability of the developed deep learning models. Their use in conjunction provides an assessment of the metric that takes into account all aspects of the model, including its predictive accuracy and the potential harm caused by the model's errors - this is particularly critical in sensitive environments such as medical imaging and disease detection.

## RESULTS AND DISCUSSION

In this study, the performance and computational complexity of three different Convolutional Neural Network (CNN) architectures were comparatively evaluated for breast cancer classification from mammography images. The summarized results are presented in the corresponding performance table. Upon examining the evaluation metrics, it is evident that the RexNet-200 model outperformed the other architectures, achieving the highest accuracy of 76.47%. This model also delivered superior precision at 75.18%, a solid recall rate of 70.22%, and an F1-score of 77.44%. The detailed confusion matrix and training metrics for RexNet-200 are presented in Figure 3 and Figure 4, respectively. These figures illustrate the model's strong classification capabilities and stable training performance.

**Table 1** Performance and Complexity Comparison of CNN Models for Breast Mammogram Classification

| Model | Acc. (%) | Prec. (%) | Rec. (%) | F1 (%) | Params (M) | GFLOPs |
|---|---|---|---|---|---|---|
| EfficientNetB7 | 73.92 | 71.02 | 70.60 | 70.79 | 63.79 | 10.26 |
| EfficientNetV2-Small | 73.53 | 71.08 | 67.02 | 67.94 | 20.18 | 5.42 |
| RexNet-200 | 76.47 | 75.18 | 70.22 | 77.44 | 13.81 | 3.05 |

Following closely, EfficientNetB7 demonstrated a balanced performance with an accuracy of 73.92%, precision of 71.02%, recall of 70.60%, and an F1-score of 70.79%. The confusion matrix and training metrics for EfficientNetB7 can be found in Figure 5 and Figure 6, respectively, highlighting its effective discrimination between classes and consistent convergence during training.

EfficientNetv2-Small showed competitive but slightly lower results, with an accuracy of 73.53%, precision of 71.08%, and recall of 67.02%. Its confusion matrix and training metrics are displayed in Figure 7 and Figure 8, respectively. These demonstrate the model's reasonable classification performance and training stability, albeit with slightly reduced recall compared to the other models.

From a computational complexity perspective, RexNet-200 stands out as the most efficient model among the three, with only

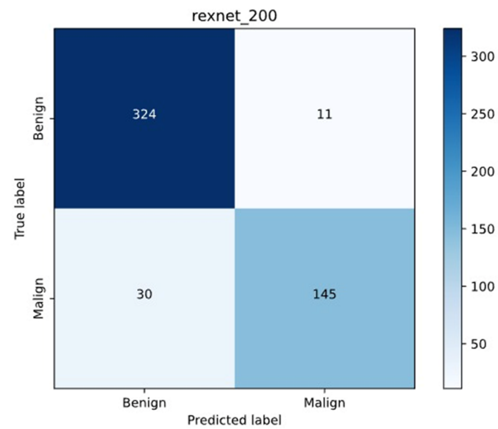


**Figure 3** Confusion Matrix RexNet-200

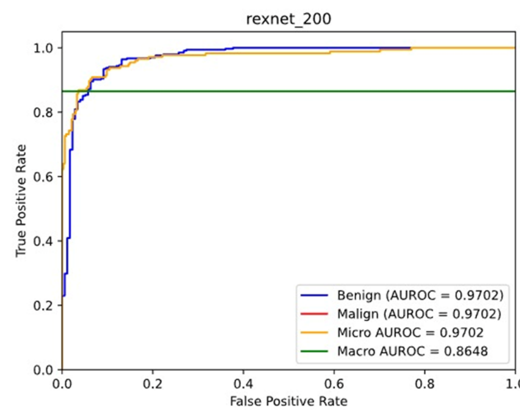


**Figure 4** ROC Curve for RexNet-200

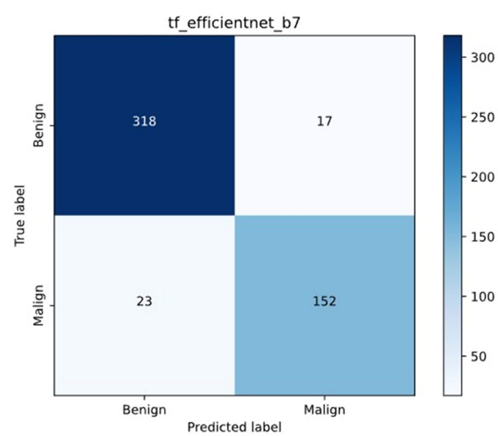


**Figure 5** Confusion Matrix for Efficientnet-B7

13.81 million parameters and 3.05 GFLOPs, indicating a favorable trade-off between performance and resource usage. Conversely, EfficientNetB7, while delivering good accuracy, is the most complex model with 63.79 million parameters and 10.26 GFLOPs. EfficientNetv2-Small falls in the moderate range regarding complexity and computational demand.Clinically, these findings have important implications. In breast cancer diagnosis, maintaining a balance between recall and precision is critical: high recall min-

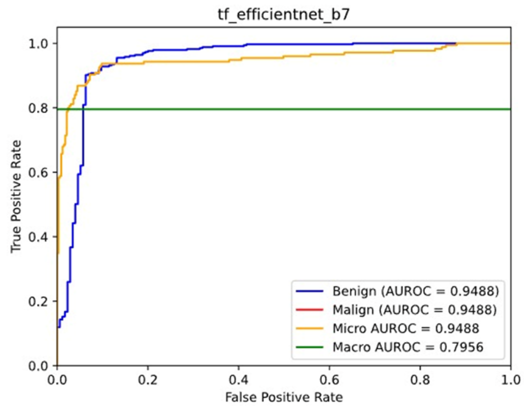


**Figure 6** ROC Curve for Efficientnet-B7

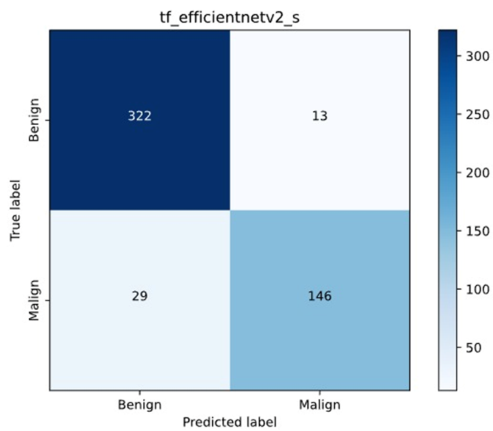


**Figure 7** Confusion Matrix for Efficientnetv2-s

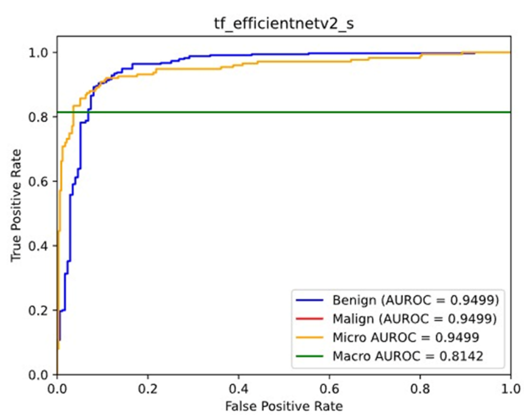


**Figure 8** ROC Curve for Efficientnetv2-s

imizes the risk of missing malignant cases, while high precision reduces false positives and avoids unnecessary anxiety and invasive procedures for patients. RexNet-200's ability to achieve the highest scores in both precision and recall metrics, combined with its lower computational cost, positions it as an ideal candidate for practical deployment in clinical decision support systems. Its efficiency also enables faster inference and easier integration into resource-constrained healthcare environments.Therefore, RexNet-200 emerges as the most promising architecture for developing an accurate, efficient, and clinically viable breast cancer diagnosis support tool.

## CONCLUSION

This study aimed to comparatively evaluate the performance and efficiency of three widely-used Convolutional Neural Network (CNN) architectures EfficientNetB7, EfficientNetv2-Small, and RexNet-200 for the classification of breast cancer from mammography images. The results clearly demonstrated that the RexNet-200 model outperformed the other architectures in terms of both diagnostic accuracy and computational efficiency. RexNet-200 achieved the highest accuracy of 76.47%, while also maintaining a relatively low model complexity with 13.81 million parameters. This outcome, particularly when compared to the more complex EfficientNetB7 model, which has 63.79 million parameters yet slightly lower accuracy, supports the idea that increased model complexity does not necessarily guarantee better performance for this task.

The success of RexNet-200 can be attributed to its efficient architecture that balances feature learning capacity with resource requirements, enabling more effective classification of breast cancer from mammography images. While these findings are encouraging, it is important to recognize the limitations of the dataset used. Future work should aim to validate these models on larger, more diverse clinical datasets, explore different preprocessing and data augmentation techniques, and incorporate Explainable Artificial Intelligence (XAI) methods to improve model transparency and trustworthiness.This research represents a significant step toward developing high-performance and computationally efficient AI models with strong potential for integration into clinical workflows as reliable decision support tools for radiologists.

### Ethical standard

The authors have no relevant financial or non-financial interests to disclose.

### Availability of data and material

The data that support the findings of this study are available from the corresponding author upon reasonable request.

### Conflicts of interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

## LITERATURE CITED

Ansari, Z. A., M. M. Tripathi, and R. Ahmed, 2025 The role of explainable ai in enhancing breast cancer diagnosis using machine learning and deep learning models. Discover Artificial Intelligence **5**: 75.

Bayram, B., I. Kunduracioglu, S. Ince, and I. Pacal, 2025 A systematic review of deep learning in mri-based cerebral vascular occlusion-based brain diseases. Neuroscience .

Cakmak, Y. and I. Pacal, 2025 Enhancing breast cancer diagnosis: A comparative evaluation of machine learning algorithms using the wisconsin dataset. Journal of Operations Intelligence **3**: 175–196.

Cakmak, Y., S. Safak, M. A. Bayram, and I. Pacal, 2024 Comprehensive evaluation of machine learning and ann models for breast cancer detection. Computer and Decision Making: An International Journal **1**: 84–102.

Chen, H. and A. L. Martel, 2025 Enhancing breast cancer detection on screening mammogram using self-supervised learning and a hybrid deep model of swin transformer and convolutional neural networks. Journal of Medical Imaging **12**: S22007–S22007.

Faryna, K., J. van der Laak, and G. Litjens, 2024 Automatic data augmentation to improve generalization of deep learning in h&e stained histopathology. Computers in Biology and Medicine **170**: 108018.

İnce, S., I. Kunduracioglu, B. Bayram, and I. Pacal, 2025 U-net-based models for precise brain stroke segmentation. Chaos Theory and Applications **7**: 50–60.

Islam, T., M. S. Hafiz, J. R. Jim, M. M. Kabir, and M. Mridha, 2024 A systematic review of deep learning data augmentation in medical imaging: Recent advances and future research directions. Healthcare Analytics **5**: 100340.

Kurtulus, I. L., M. Lubbad, O. M. D. Yilmaz, K. Kilic, D. Karaboga, *et al.*, 2024a A robust deep learning model for the classification of dental implant brands. Journal of Stomatology, Oral and Maxillofacial Surgery **125**: 101818.

Kurtulus, I. L., M. Lubbad, O. M. D. Yilmaz, K. Kilic, D. Karaboga, *et al.*, 2024b A robust deep learning model for the classification of dental implant brands. Journal of Stomatology, Oral and Maxillofacial Surgery **125**: 101818.

Latha, M., P. S. Kumar, R. R. Chandrika, T. Mahesh, V. V. Kumar, *et al.*, 2024 Revolutionizing breast ultrasound diagnostics with efficientnet-b7 and explainable ai. BMC Medical Imaging **24**: 230.

Lubbad, M., D. Karaboga, A. Basturk, B. Akay, U. Nalbantoglu, *et al.*, 2024a Machine learning applications in detection and diagnosis of urology cancers: a systematic literature review. Neural Computing and Applications **36**: 6355–6379.

Lubbad, M. A., I. L. Kurtulus, D. Karaboga, K. Kilic, A. Basturk, *et al.*, 2024b A comparative analysis of deep learning-based approaches for classifying dental implants decision support system. Journal of Imaging Informatics in Medicine **37**: 2559–2580.

Mridha, K., T. Sarker, S. D. Bappon, and S. M. Sabuj, 2023 Attention u-net: A deep learning approach for breast cancer segmentation. In *2023 International Conference on Quantum Technologies, Communications, Computing, Hardware and Embedded Systems Security (iQ-CCHESS)*, pp. 1–6, IEEE.

Ozdemir, B., E. Aslan, and I. Pacal, 2025 Attention enhanced inceptionnext based hybrid deep learning model for lung cancer detection. IEEE Access .

Pacal, İ., 2025 Diagnostic analysis of various cancer types with artificial intelligence .

Pacal, I., O. Akhan, R. T. Deveci, and M. Deveci, 2025 Nextbrain: Combining local and global feature learning for brain tumor classification. Brain Research p. 149762.

Pacal, I. and O. Attallah, 2025 Inceptionnext-transformer: A novel multi-scale deep feature learning architecture for multimodal breast cancer diagnosis. Biomedical Signal Processing and Control **110**: 108116.

Shorten, C. and T. M. Khoshgoftaar, 2019 A survey on image data augmentation for deep learning. Journal of big data **6**: 1–48.

Sreekala, K. and J. Sahoo, 2025 Enhancing mammogram classification using explainable conditional self-attention generative adversarial network. Expert Systems with Applications p. 128640.

Sureshkumar, V., R. S. N. Prasad, S. Balasubramaniam, D. Jagannathan, J. Daniel, *et al.*, 2024 Breast cancer detection and analytics using hybrid cnn and extreme learning machine. Journal of Personalized Medicine **14**: 792.

Tajbakhsh, N., L. Jeyaseelan, Q. Li, J. N. Chiang, Z. Wu, *et al.*, 2020 Embracing imperfect datasets: A review of deep learning solutions for medical image segmentation. Medical image analysis **63**: 101693.

Tajbakhsh, N., J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, *et al.*, 2016 Convolutional neural networks for medical image analysis: Full training or fine tuning? IEEE transactions on medical imaging **35**: 1299–1312.

Zeynalov, J., Y. Çakmak, and İ. Paçal, 2025 Automated apple leaf disease classification using deep convolutional neural networks: A comparative study on the plant village dataset. Journal of Computer Science and Digital Technologies **1**: 5–17.

Zhang, H., Q. Chen, B. Xue, W. Banzhaf, and M. Zhang, 2024 P-mixup: Improving generalization performance of evolutionary feature construction with pessimistic vicinal risk minimization. In *International Conference on Parallel Problem Solving from Nature*, pp. 201–220, Springer.

***How to cite this article:*** Alswilem, L., and Pacal, N. Artificial Intelligence in Mammography: A Study of Diagnostic Accuracy and Efficiency. *Computational Systems and Artificial Intelligence*, 1(1), 26-31, 2025.