Artificial Intelligence in Applied Sciences
Cutting-Edge Scientific Solutions

e-ISSN: 3108-4060
RESEARCH ARTICLE
Vol.1 / **No.**1 / 2025 / **pp.**13-19
https://doi.org/10.69882/adba.ai.2025073

Deep Learning for Automated Breast Cancer Detection in Ultrasound: A Comparative Study of Four CNN Architectures

Yigitcan Cakmak [*,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 is one of the most common malignancies among women globally, and it constitutes a significant public health problem in terms of morbidity and mortality. Since early-stage diagnosis significantly increases treatment success and survival rates, effective screening and diagnostic methods are of great importance. Various imaging modalities, such as mammography, ultrasonography (US), and magnetic resonance imaging, play a critical role in the detection of breast cancer. Ultrasound, in particular, is a valuable imaging method due to its non-ionizing nature, its accessibility, and its role as a complementary tool in dense breast tissue. In recent years, deep learning (DL) algorithms, particularly Convolutional Neural Networks (CNNs), have exhibited promising results in medical image analysis, especially in cancer detection. The aim of this research is to investigate and compare the four most common CNN architectures, ResNet50, DenseNet169, InceptionV3 and InceptionV4, for breast ultrasound images to classify breast cancer automatically. We have utilized publicly available breast ultrasound image datasets for the models and reported results in metrics of accuracy, precision, sensitivity, and F1-score. The InceptionV3 architecture had the best performance across the models examined with metrics of accuracy: 96.67%, precision: 96.55%, sensitivity: 96.38%, and F1-score: 96.41%. It was also noticed that the DenseNet169 model performed similarly to the InceptionV3 model but had substantially fewer parameters. The results of this study suggest that the InceptionV3 DL architecture may have significant potential for accuracy in the classification of cancer from breast ultrasound images and can contribute to the development of computer aided diagnosis systems for the early detection of breast cancer.

KEYWORDS
Breast cancer
Deep learning
Breast ultrasound
Image classification
Computer-aided diagnosis (CAD)

INTRODUCTION

Breast cancer is one of the leading cancers that affect women's health around the world and is the abnormal and unregulated growth of mammary epithelial cells (Kim *et al.* 2025; Xiong *et al.* 2025). The origin of breast cancer is a multifactorial process mediated by genetic susceptibility, hormones, lifestyle, and environmental factors (Obeagu and Obeagu 2024). Given that the prognosis for treatment response and survival rate improve drastically if the cancer is found at an early stage, better screening and diagnostic strategies are crucial. Therefore, it is important to study medical imaging techniques that are the least invasive way to characterize abnormal changes in the breast (Kiani *et al.* 2025; Alshawwa *et al.* 2024; Begum *et al.* 2024). Breast cancer screening programs have relied on mammography as the definitive tool of choice (Katsika *et al.* 2024; Trentham-Dietz *et al.* 2024). However, mammography may lack diagnostic sensitivity particularly with women with dense breast tissue and in women who are younger. This raises the need for complementary or alternative forms of imaging (Abeelh and AbuAbeileh 2024).

Ultrasonography is a valuable component in evaluating breast lesions, primarily due to its non-ionizing nature, availability, affordability, and real-time images (Iacob *et al.* 2024). Specifically, it has advantages for the evaluation of breast lesions, especially in determining whether suspicious findings on mammography are cystic or solid masses, and facilitates biopsy procedures. For women with dense breast parenchyma, ultrasonography is essentially an adjunct that improves mammographic diagnostic performance and provides clarity in graphically characterizing lesions (Gordon *et al.* 2025). However, ultrasonography has disadvantages, including operator-dependency that introduces inter-observer variability in the detection and interpretation of lesions. Additionally, ultrasonography is limited in its ability to detect microcalcifications. Research has been conducted to evaluate new ways to provide more objective and standardized analysis of ultrasound images (Vogel-Minea *et al.* 2025; Rana *et al.* 2024).

Artificial intelligence (AI), and deep learning (DL) algorithms in particular, have generated paradigm shifting advances in medical image analysis in recent years (Pacal *et al.* 2025; Pacal and Attallah 2025a). DL architectures, such as convolutional neural networks (CNNs), have shown significant potential in many med-

Manuscript received: 29 May 2025,
Revised: 15 June 2025,
Accepted: 19 June 2025.

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

ical specialties, including radiology and pathology, due to their ability to automatically learn complex patterns and hierarchical features from massive image datasets (Ozdemir *et al.* 2025; Lubbad *et al.* 2024b; Pacal 2025). With respect to breast cancer, DL models have demonstrated excellent performance to achieve high levels of accuracy in the detection, classification, and segmentation of suspicious lesions from mammograms, ultrasound images, and magnetic resonance scans (İnce *et al.* 2025; Lubbad *et al.* 2024a). In this paper, we plan to train several CNN algorithms (e.g. ResNet50, DenseNet169, InceptionV3, InceptionV4) with breast ultrasound image dataset publicly available online and then analyze and compare the results to determine their potential to assist clinicians in breast cancer diagnosis (Pacal and Attallah 2025b; Cakmak *et al.* 2024; Bayram *et al.* 2025). The hope is that by performing a DL based analysis of ultrasound images we can assist with the early diagnosis process and also improve diagnostic accuracy (Cakmak and Pacal 2025; Zeynalov *et al.* 2025; Kurtulus *et al.* 2024).

The field of medicine is undergoing a transformative evolution through the integration of AI, particularly its sub-disciplines of DL and machine learning (ML) (Obuchowicz *et al.* 2024; Koçak *et al.* 2025). These technologies offer revolutionary advancements across a broad spectrum, ranging from the early diagnosis of diseases to the development of personalized treatment protocols, from drug discovery to the analysis of complex biological data (Li *et al.* 2024; Islam *et al.* 2024). Medical imaging, in particular, holds immense potential due to the capacity of AI algorithms to detect subtle patterns and anomalies imperceptible to the human eye and to rapidly process and interpret large volumes of data (Chambi *et al.* 2025; Meng *et al.* 2024). Disciplines such as radiology, pathology, and oncology are rapidly adopting these innovations with the promise of enhancing diagnostic accuracy, optimizing workflows, and ultimately improving patient outcomes. In the management of prevalent and serious health issues like breast cancer, the combination of AI with accessible imaging modalities such as ultrasound is opening promising avenues for early-stage detection and effective treatment strategies (Rajkumar *et al.* 2024).

In studies on breast cancer classification and segmentation, various AI approaches have gained prominence. Abhisheka et al., highlighting the importance of breast cancer in the healthcare sector, noted the insufficiency of traditional ML or DL models alone and, accordingly, proposed the Hybrid Breast Cancer Prediction System (HBCPS) model. This system combines deep CNN features (obtained via ResNet50) with handcrafted features (Histogram of Oriented Gradients - HOG) and uses a Support Vector Machine (SVM) for classification. The system also incorporates a Block-Matching and 3D filtering (BM3D) filter to reduce noise in Breast Ultrasound (BUS) images, achieving satisfactory results on the BUSI dataset, such as 89.02% accuracy and an AUC of 0.8717 (Abhisheka *et al.* 2025). Similarly, Latha *et al.* (2024) combined a scalable CNN architecture, EfficientNet-B7, with advanced data augmentation techniques to address low accuracy in minority classes, particularly malignant tumors. They also integrated eXplainable AI (XAI) techniques like Grad-CAM to enhance the interpretability of the model's predictions. With this approach, they achieved a high classification accuracy of 99.14%, significantly outperforming existing CNN-based approaches. These studies underscore the potential of both hybrid modeling and the integration of advanced CNN architectures with XAI techniques in breast cancer classification.

Other notable contributions in the literature have focused on improving segmentation accuracy and computational efficiency. Umer *et al.* (2024) proposed a U-shaped autoencoder-based CNN model featuring a multi-attention mechanism and a triple decoder, focusing on capturing multi-scale spatial features and highlighting the tumor region, particularly in BC segmentation from U/S images. Their proposed model achieved Dice scores of 90.45% and 89.13% on the UDIAT and BUSI datasets, respectively. On the other hand, Cai *et al.* (2024), as a solution to the challenges of high computational complexity and large model parameters in existing segmentation methods, developed the SC-Unext model. This model, based on the Unext network and inspired by the mechanisms of cellular apoptosis and division, not only improved segmentation performance but also reduced model parameters and computational resource consumption, achieving a 75.29% Dice score and 97.09% accuracy on the BUSI dataset; it also demonstrated fast inference times on CPUs. These studies demonstrate the importance of developing not only complex architectures but also efficient and lightweight models, especially for segmentation tasks.

Finally, the comparison of next-generation architectures and the development of holistic systems for clinical application also hold a significant place on the research agenda. Cai *et al.* (2024) compared Mamba-based models (VMamba and Vim) with traditional CNNs and Vision Transformers (ViTs), demonstrating that some Mamba-based architectures offer statistically significant performance improvements, particularly due to their ability to capture long-range dependencies in limited data. For instance, on dataset B, Mamba-based models were reported to provide an improvement of 1.98% in mean AUC and 5.0% in mean Accuracy. Nasiri-Sarvi *et al.* (2024) adopted an approach aimed at presenting the radiologist with both the tumor mask and its classification. They examined different DL models and identified the best-performing one, which achieved over 90% accuracy, 92% precision, 90% sensitivity, and a 90% F1-score on the BUSI dataset. This study emphasizes that DL architectures are effective in the classification and segmentation of ultrasound breast images and could be used in clinical trials in the near future. Such comparative studies and proposals for integrated systems further solidify the role of AI in breast cancer diagnosis and pave the way for its clinical adaptation (Gagliardi *et al.* 2024).

MATERIALS AND METHODS

Dataset

In this research, a publicly available dataset, the "Breast Ultrasound Images Dataset", was used to classify and analyze breast ultrasound images. This dataset was made available through the Kaggle platform by Sabah Saraki (Kaggle 2025), and contains ultrasound images which demonstrate different appearances of breast cancer. The dataset contains samples of ultrasound images grouped into three main classes based on pathologically confirmed diagnoses of benign tumors, malignant tumors, and normal breast tissue images. This variety gives a solid ground for evaluating the capability of the DL models to distinguish other tissue structures and lesion types.

In order to ensure a standardized and reproducible model development and evaluation process, the dataset, comprising a total of 780 samples (437 benign, 210 malignant, and 133 normal), was meticulously partitioned into training, validation, and testing subsets. This partitioning allocated 70% of the data (545 samples) to the training set, 15% (115 samples) to the validation set, and the remaining 15% (120 samples) to the test set. These proportions were selected to ensure the model is trained on sufficient data, while simultaneously allowing for a reliable assessment of its generalization capability and mitigating the risk of overfitting. Furthermore, a stratified sampling approach was employed to ensure that the class distribution within each subset precisely mirrors that of the original dataset, a crucial step to prevent the model from developing a bias towards any particular class. Consequently, the training set was composed of 305 benign, 147 malignant, and 93 normal samples; the validation set contained 65 benign, 31 malignant, and 19 normal samples; and the test set consisted of 67 benign, 32 malignant, and 21 normal samples. This dataset partitioning is also illustrated in Figure 1.

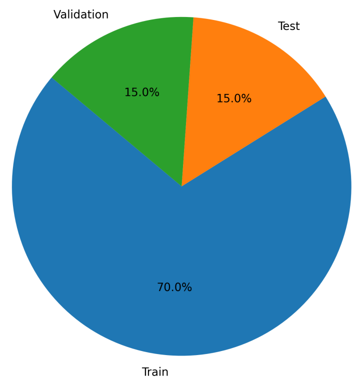

Figure 1 Distribution of the Breast Ultrasound Images Dataset into Training, Validation, and Test Sets (70%-15%-15%)

To better visualize the structure of the dataset and the types of images it contains, representative ultrasound images from each class (benign, malignant, and normal) are presented in Figure 2. As can be seen in Figure 2, benign lesions generally present with regular borders and a homogeneous internal echo, whereas malignant lesions may exhibit more irregular margins, spiculated extensions, and heterogeneous internal structures. Normal breast tissue images, in turn, show typical fibroglandular and adipose tissue patterns. These examples reflect not only the visual differences between the classes but also the inherent challenges of ultrasound imaging, such as speckle noise and low contrast. These visual representations help in understanding the fundamental morphological features that our models must learn and differentiate, and they provide an insight into the diversity of the dataset.

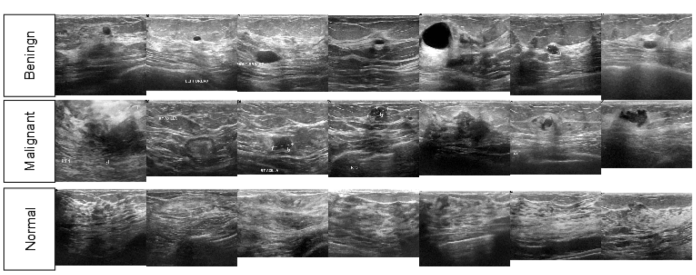

Figure 2 Sample Ultrasound Images Illustrating the Three Classes in the Breast Cancer Dataset: Benign, Malignant, and Normal.

Data Augmentation

To enhance the generalization capability of the DL models and to mitigate the problem of overfitting, a frequent challenge in limited datasets such as medical images, various on-the-fly data augmentation techniques were integrated into the training process in this study. Initially, as the focus was on the classification task, the mask.png files, which were included in the original dataset for segmentation purposes, were excluded from the analysis. During the training phase, the primary augmentation methods randomly applied to each image were as follows: images were first subjected to a "Random Resized Crop," where they were cropped to a random size with a scale ranging from 8% to 100% of the original area (scale: [0.08, 1.0]) and an aspect ratio between 0.75 and 1.33 (ratio: [0.75, 1.3333333333333333]), and subsequently resized to 224x224 pixels (img-size: 224) using a random interpolation method (train interpolation: random). Additionally, random horizontal flipping was applied to each image with a 50% probability (hflip: 0.5). For color-based augmentations, random alterations were made to the color properties of the images, including brightness, contrast, saturation, and hue, with a factor of 0.4 (color-jitter: 0.4). Vertical flipping was not utilized in this study (vflip: 0.0). These on-the-fly augmentation strategies were intended to ensure that the model encounters differentiated data samples during each training epoch, thereby preventing it from becoming overly dependent on the specific features of the training data and aiming for a more robust and reliable performance on unseen data (Wang *et al.* 2024; Mumuni *et al.* 2024).

Model Architectures

In this study, for the automatic classification of breast cancer from breast ultrasound images, well-established and widely recognized deep CNN architectures from the field of computer vision were utilized. In domains such as medical imaging, where the amount of labeled data is often limited, adopting a transfer learning approach rather than training a model from scratch presents significant advantages. Transfer learning enables the transfer of the rich feature extraction capabilities of models pre-trained on large-scale, general-purpose datasets (e.g., ImageNet) to a more specific and smaller target dataset. This approach aims to achieve faster model convergence, improved generalization performance, and a reduced risk of overfitting, particularly when working with limited data. Within the scope of this study, all selected CNN architectures were initialized with weights pre-trained on the ImageNet dataset and were subsequently subjected to a fine-tuning process on our target dataset comprising breast ultrasound images.

First, the ResNet50 architecture, based on the principle of residual learning, was employed. Developed by He et al., ResNet architectures addressed the vanishing gradient problem encountered in the training of very deep networks through the use of "residual blocks" containing "shortcut connections," which allow the input to be passed directly to subsequent layers. ResNet50, a 50-layer deep implementation of this structure, is frequently preferred as a robust baseline model for image classification tasks (He *et al.* 2016). Another architecture of choice was DenseNet169. Proposed by Huang et al., Densely Connected Networks (DenseNets) introduce a "dense connectivity" structure where each layer receives the feature maps from all preceding layers as input and passes on its own feature maps to all subsequent layers. This architecture strengthens feature propagation, encourages feature reuse, reduces the number of parameters, and improves gradient flow, making it particularly prominent for its parameter efficiency; DenseNet169 is a 169-layer version of this architecture (Huang *et al.* 2017).

The study also evaluated two models from the Inception architecture family, developed by Google, which are capable of capturing features at multiple scales simultaneously. InceptionV3, through its "Inception modules," applies convolutional filters of different sizes (e.g., 1x1, 3x3, 5x5) and pooling operations in parallel at the same layer level and concatenates their outputs. This structure allows the model to analyze complex visual patterns at various scales, while performance is optimized through techniques such as factorizing larger convolutions into smaller ones and using auxiliary classifiers. InceptionV4, as an advancement over InceptionV3, aims to deliver improvements in both performance and computational efficiency by presenting the Inception modules in a more uniform and simplified structure. This model is characterized by deeper and more optimized Inception blocks (Szegedy *et al.* 2016).

These four distinct CNN architectures (ResNet50, DenseNet169, InceptionV3, and InceptionV4) were selected to compare their effectiveness in the task of differentiating between benign, malignant, and normal tissue classes in breast ultrasound images. The distinct architectural approaches and feature extraction strategies of each model are expected to approach this challenging medical image classification problem from different perspectives, thereby providing valuable insights into which architecture or architectural features are more suitable for this specific task. The performance of the models is carefully analyzed using various evaluation metrics, and the results contribute to the literature on the development of deep learning-based automated systems for the early diagnosis of breast cancer.

Evaluation Metrics

Assessing how well DL models work is a vital process to assess their usefulness, provide rationale for relevant decisions, and support data-driven decisions. Performance evaluation criteria can fulfill many important roles such as assessing the effectiveness of a classification models, helping them to be optimized, revealing errors or biases in the data, comparing models, and detecting overfitting. This paper focuses specifically on performance metrics for breast cancer classification, at the same time, we have decided to utilize standard evaluation criteria that are clearly entrenched in the academic literature.

The basic metrics that are used in this project (accuracy, precision, recall, and F1-score) are important in not only DL but other areas. Accuracy can be defined as the number of correctly classified instances over the total number of instances, giving insight into the performance as a whole. Precision (true positives / (true positives + false positives)) tells how reliable the model is in classifying positive instances; if the model has a high precision, it means there are few if any false positives. Recall tells us about the number of actual positives correctly identified the measure of completeness. The F1-score is defined as the harmonic mean of precision and recall, thus making it a single measure of performance that balances the trade-off between false positives and false negatives. While these definitions may seem complicated, they can also be defined mathematically:

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Number of total predictions}} \tag{1}$$

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \tag{2}$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \tag{3}$$

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

RESULTS AND DISCUSSION

In this section, we present and analyze the performance results of the different CNN architectures we evaluated for the purposes of classifying breast ultrasound images. We compared the performance of the ResNet50, DenseNet169, InceptionV3 and InceptionV4 models used in this work using some fundamental classification performance metrics: Accuracy, Precision, Recall, and F1 score - as well as Quantification metrics including the number of parameters (Params M) and GFLOPs (Giga Floating Point Operations per Second), which estimate the complexity of models and computational resources required for both model training and inference. We consider that exploring such metrics is vital to gaining insight into models' diagnostic performance and important use cases.

The results obtained are summarized in Table 1. Table 1 illustrates the performance metrics reached by each model on the test dataset, along with information on model complexity. These data show the strengths and weaknesses of the various architectures and demonstrate the trade-off between performance and computational expense.

Table 1 Comparative Performance and Complexity of CNN Models for Breast Ultrasound Image Classification

Model	Acc. (%)	Prec. (%)	Rec. (%)	F1 (%)	Params (M)	GFLOPs
Inception V3	96.67	96.55	96.38	96.41	21.79	5.67
DenseNet 169	94.17	92.71	95.43	93.91	12.49	6.72
Inception V4	94.17	92.01	95.97	93.60	41.15	12.25
ResNet 50	90.83	90.29	89.08	89.10	23.51	8.26

The data from Table 1 clearly illustrates that the InceptionV3 model showed the best performance. With an accuracy of 96.67%, precision of 96.55%, sensitivity of 96.38% and F1-score of 96.41%, InceptionV3 was the most capable of successfully classifying breast ultrasound images. It is expected that InceptionV3 performs so well because of the architecture's ability to capture features at different scales and learn complex patterns. Also worth noting, is that InceptionV3 (21.79 million parameters, 5.6719 GFLOPs) delivered the best results from a model complexity standpoint not because it was the most complex model. Having the lowest GFLOPs value means that it was performing at a high level while using a relatively low amount of computational cost.

The DenseNet169 model also achieved highly competitive results. With 94.17% accuracy, 92.71% precision, 95.43% sensitivity, and a 93.91% F1-score, it exhibited the second-best performance after InceptionV3. The most striking feature of DenseNet169 is its model complexity; with 12.49 million parameters, it has the lowest parameter count among the evaluated models, and with 6.7169 GFLOPs, it has the second-lowest GFLOPs value after InceptionV3 (there may be an error in the table, as the GFLOPs for InceptionV3 is lower). This indicates that, as a result of its dense connectivity structure that enhances feature propagation and increases parameter efficiency, DenseNet169 offers a favorable performance-to-efficiency balance. DenseNet169 could be an attractive alternative, especially for scenarios where computational resources are constrained.

The InceptionV4 model, despite having an accuracy rate of 94.17% similar to DenseNet169, along with 92.01% precision, 95.97% sensitivity, and a 93.60% F1-score, is the model with the highest complexity and computational cost among those evaluated, at 41.15 million parameters and 12.2450 GFLOPs. The fact that it did not surpass InceptionV3, despite having a deeper and more complex structure, suggests that for this specific task and dataset, increased complexity does not invariably translate to better performance. ResNet50, in contrast, exhibited a more modest performance compared to the other three models, with 90.83% accuracy and an 89.10% F1-score. Although it is a strong baseline model, it lagged behind the other more modern and complex architectures used in this study. It possesses a moderate level of complexity with 23.51 million parameters and 8.2634 GFLOPs.

The findings of this study indicate that the InceptionV3 architecture offers a compelling combination of high diagnostic accuracy and balanced computational efficiency. In contrast, DenseNet169 presents itself as a potent alternative for resource-constrained environments, owing to its lower parameter size and computational cost. These findings represent a major contribution to the choice of DL architectures in the context of developing automated solutions for the early diagnosis of breast cancer, and possibilities for future involvement of real-world clinical applications. In all cases, the choice of architecture must be assessed relative to the intended application's requirements (e.g., maximum accuracy versus fast inference time). To further assess the classification performance of the best performing InceptionV3, its confusion matrix is shown in Figure 3.

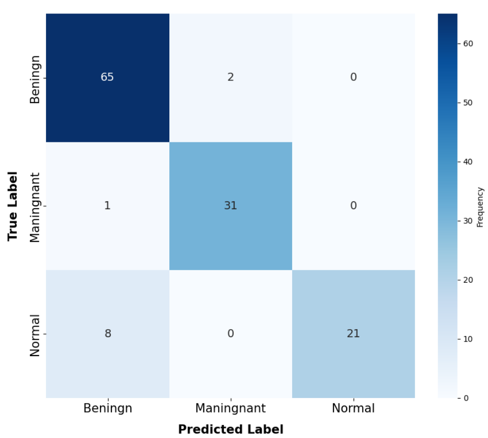

Figure 3 Confusion Matrix of the InceptionV3 Model for Breast Ultrasound Image Classification.

CONCLUSION

In this study, the performance of four different deep Convolutional Neural Network (CNN) architectures (ResNet50, DenseNet169, InceptionV3, and InceptionV4) was comprehensively compared and evaluated for the classification of breast cancer from breast ultrasound (US) images. The findings clearly demonstrated that the InceptionV3 model exhibited the highest classification performance compared to the other evaluated architectures, with superior metric values of 96.67% accuracy, 96.55% precision, 96.38% sensitivity, and a 96.41% F1-score. This high performance can be attributed to the Inception architecture's ability to effectively capture multi-scale features and learn complex visual patterns, while it is also noteworthy that the model offers a relatively efficient computational cost with 21.79 million parameters and 5.6719 GFLOPs.

The DenseNet169 architecture also stood out as a promising alternative for resource-constrained environments, drawing attention with its 94.17% accuracy rate and particularly its low parameter counts of 12.49 million. While InceptionV4 could not surpass InceptionV3 despite its high complexity, ResNet50 yielded more modest results. This study demonstrates that InceptionV3 is a strong candidate for the classification of breast US images in terms of both high diagnostic accuracy and acceptable computational efficiency. The obtained results offer valuable insights for the selection of appropriate DL architectures for the development of automated systems for the early diagnosis of breast cancer and underscore the potential for the integration of these technologies into future clinical applications. Validating these models on larger and more diverse datasets, investigating the impact of different data augmentation strategies and fine-tuning techniques, and integrating eXplainable AI (XAI) methods to enhance model interpretability represent critical next steps for advancing research in this field. Ultimately, such deep learning-based approaches have great potential to support the decision-making processes of radiologists, thereby improving the accuracy and efficiency of breast cancer diagnosis.

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

Abeelh, E. A. and Z. AbuAbeileh, 2024 Comparative effectiveness of mammography, ultrasound, and mri in the detection of breast carcinoma in dense breast tissue: a systematic review. Cureus **16**.

Abhisheka, B., S. K. Biswas, B. Purkayastha, and S. Das, 2025 Integrating deep and handcrafted features for enhanced decision-making assistance in breast cancer diagnosis on ultrasound images. Multimedia Tools and Applications pp. 1–23.

Alshawwa, I. A., H. Q. El-Mashharawi, F. M. Salman, M. N. A. Al-Qumboz, B. S. Abunasser, *et al.*, 2024 Advancements in early detection of breast cancer: Innovations and future directions .

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 .

Begum, M. M. M., R. Gupta, B. Sunny, and Z. L. Lutfor, 2024 Advancements in early detection and targeted therapies for breast cancer; a comprehensive analysis. Asia Pacific Journal of Cancer Research **1**: 4–13.

Cai, F., J. Wen, F. He, Y. Xia, W. Xu, *et al.*, 2024 Sc-unext: A lightweight image segmentation model with cellular mechanism for breast ultrasound tumor diagnosis. Journal of Imaging Informatics in Medicine **37**: 1505–1515.

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.

Chambi, E. A., D. G. Alzamora, and A. A. Salas, 2025 Ultrasonic image processing for the classification of benign and malignant breast tumors: Comparative study of convolutional neural network architectures. Engineering Proceedings **83**: 15.

Gagliardi, M., T. Ruga, E. Vocaturo, and E. Zumpano, 2024 Predictive analysis for early detection of breast cancer through artificial intelligence algorithms. In *International Conference on Innovations in Computational Intelligence and Computer Vision*, pp. 53–70, Springer.

Gordon, P. B., L. J. Warren, and J. M. Seely, 2025 Cancers detected on supplemental breast ultrasound in women with dense breasts: update from a canadian centre. Canadian Association of Radiologists Journal p. 08465371251318578.

He, K., X. Zhang, S. Ren, and J. Sun, 2016 Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778.

Huang, G., Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, 2017 Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4700–4708.

Iacob, R., E. R. Iacob, E. R. Stoicescu, D. M. Ghenciu, D. M. Cocolea, *et al.*, 2024 Evaluating the role of breast ultrasound in early detection of breast cancer in low-and middle-income countries: a comprehensive narrative review. Bioengineering **11**: 262.

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, M. R., M. M. Rahman, M. S. Ali, A. A. N. Nafi, M. S. Alam, *et al.*, 2024 Enhancing breast cancer segmentation and classification: An ensemble deep convolutional neural network and u-net approach on ultrasound images. Machine Learning with Applications **16**: 100555.

Kaggle, 2025 Breast ultrasound images dataset(busi).

Katsika, L., E. Boureka, I. Kalogiannidis, I. Tsakiridis, I. Tirodimos, *et al.*, 2024 Screening for breast cancer: a comparative review of guidelines. Life **14**: 777.

Kiani, P., H. Vatankhahan, A. Zare-Hoseinabadi, F. Ferdosi, S. Ehtiati, *et al.*, 2025 Electrochemical biosensors for early detection of breast cancer. Clinica Chimica Acta **564**: 119923.

Kim, J., A. Harper, V. McCormack, H. Sung, N. Houssami, *et al.*, 2025 Global patterns and trends in breast cancer incidence and mortality across 185 countries. Nature Medicine pp. 1–9.

Koçak, B., A. Ponsiglione, A. Stanzione, C. Bluethgen, J. Santinha, *et al.*, 2025 Bias in artificial intelligence for medical imaging: fundamentals, detection, avoidance, mitigation, challenges, ethics, and prospects. Diagnostic and interventional radiology **31**: 75.

Kurtulus, I. L., M. Lubbad, O. M. D. Yilmaz, K. Kilic, D. Karaboga, *et al.*, 2024 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.

Li, X., L. Zhang, J. Yang, and F. Teng, 2024 Role of artificial intelligence in medical image analysis: A review of current trends and future directions. Journal of Medical and Biological Engineering **44**: 231–243.

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.

Meng, X., J. Ma, F. Liu, Z. Chen, and T. Zhang, 2024 An interpretable breast ultrasound image classification algorithm based on convolutional neural network and transformer. Mathematics **12**: 2354.

Mumuni, A., F. Mumuni, and N. K. Gerrar, 2024 A survey of synthetic data augmentation methods in machine vision. Machine Intelligence Research **21**: 831–869.

Nasiri-Sarvi, A., M. S. Hosseini, and H. Rivaz, 2024 Vision mamba for classification of breast ultrasound images. In *Deep Breast Workshop on AI and Imaging for Diagnostic and Treatment Challenges in Breast Care*, pp. 148–158, Springer.

Obeagu, E. I. and G. U. Obeagu, 2024 Breast cancer: A review of risk factors and diagnosis. Medicine **103**: e36905.

Obuchowicz, R., M. Strzelecki, and A. Piórkowski, 2024 Clinical applications of artificial intelligence in medical imaging and image processing—a review.

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

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

Rajkumar, R., S. Gopalakrishnan, K. Praveena, M. Venkatesan, K. Ramamoorthy, *et al.*, 2024 Darknet-53 convolutional neural network-based image processing for breast cancer detection. Mesopotamian Journal of Artificial Intelligence in Healthcare **2024**: 59–68.

Rana, A. S., J. Rafique, and H. Riffat, 2024 Advances in breast ultrasound imaging: Enhancing diagnostic precision and clinical utility. In *Latest Research on Breast Cancer-Molecular Insights, Diagnostic Advances and Therapeutic Innovations*, IntechOpen.

Szegedy, C., V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, 2016 Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2818–2826.

Trentham-Dietz, A., C. H. Chapman, J. Jayasekera, K. P. Lowry, B. M. Heckman-Stoddard, *et al.*, 2024 Collaborative modeling to compare different breast cancer screening strategies: a decision analysis for the us preventive services task force. JAMA **331**: 1947–1960.

Umer, M. J., M. Sharif, and M. Raza, 2024 A multi-attention triple decoder deep convolution network for breast cancer segmentation using ultrasound images. Cognitive Computation **16**: 581–594.

Vogel-Minea, C. M., W. Bader, J.-U. Blohmer, V. Duda, C. Eichler, *et al.*, 2025 Best practice guidelines–degum recommendations on breast ultrasound. Ultraschall in der Medizin-European Journal of Ultrasound **46**: 245–258.

Wang, Z., P. Wang, K. Liu, P. Wang, Y. Fu, *et al.*, 2024 A comprehensive survey on data augmentation. arXiv preprint arXiv:2405.09591 .

Xiong, X., L.-W. Zheng, Y. Ding, Y.-F. Chen, Y.-W. Cai, *et al.*, 2025 Breast cancer: pathogenesis and treatments. Signal transduction and targeted therapy **10**: 49.

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.

How to cite this article: Cakmak, Y., and Pacal, N. Deep Learning for Automated Breast Cancer Detection in Ultrasound: A Comparative Study of Four CNN Architectures. *Artificial Intelligence in Applied Sciences*, 1(1), 13-19, 2025.