Computers and Electronics in Medicine
Cutting-Edge Scientific Solutions

e-ISSN: 3023-8609
RESEARCH ARTICLE
Vol.3 / No.1 / 2026 / pp.94-98
https://doi.org/10.69882/adba.cem.20260111

Classification of Breast Cancer with Breast X-Ray Images via Convolution Neural Networks, Vision Transformers and AlexNet

Bora Basaran [*,1] **and Ali Burak Oncul** [α,2]

[*]Applied Data Science, TED University, 06420, Ankara, Türkiye, [α]Department of Computer Engineering, Faculty of Engineering and Architecture, Kastamonu University, 37150, Kastamonu, Türkiye.

ABSTRACT Breast cancer is one of the most dangerous types of cancer, affecting many people long-term and leading to death. For this reason, it has become a frequently studied and emphasized topic in the medical field. Furthermore, significant advances in computer science have attracted the attention of the medical world, and computer science has also been incorporated into this challenging disease process to address it. The rapid development of artificial intelligence in recent years has led to a rapid increase in research on breast cancer. Numerous AI-based models have been developed to prevent human errors and to assist and support researchers in decision-making. In this study, one of these models was developed, and three different deep learning (DL) models were proposed to classify breast cancer as breast cancer-negative and breast cancer-positive. The study was adapted for computer vision (CV) using a Kaggle dataset called Breast Cancer, consisting of 3,383 breast tumor mammography images; the labels are 0 and 1, respectively, and the image dimensions are 640 x 640 pixels. In this study, three models were trained to classify breast cancer images: a Convolutional Neural Network (CNN), VisionTransformers (ViT), and AlexNet, trained with 45, 75, and 50 epochs, respectively. HuggingFace Space was used with these three models to classify breast cancer. The HuggingFace web application provided breast cancer classification based on the three models. Performance metrics, accuracy, loss, and execution time outperformed the CNN model, achieving a more optimal execution time (807.82 seconds), accuracy (0.9544), and loss (0.1078). The model has achieved significant success in breast cancer, and with further refinements, it is anticipated that the model will be suitable for use as a decision support system.

KEYWORDS
Cancer
Computer vision processing
Convolutional neural network
Deep learning
Vision transformers

INTRODUCTION

Detection and Classification of breast cancer plays a vital role in monitoring the health-condition of cancer-patients in early stages of cancer. To monitor disease and flow, health-professionals keep track of the breast cancer patients' mammography x-ray images. The study has been tailored to lead health-professionals with a state-of-art solution to abstain from the wrong treatments and classifying breast cancer from breast images in early stage via classification of breast x-ray images via DL and CV. Their well-known models' and their implementations that are CNN, ViT and AlexNet have been investigated in this study. The breast X-ray image dataset have been utilized. Dataset comprises 3,383 mammogram images specifically highlighting breast tumors, organized within a structured folder format. Originally exported from Roboflow, a platform dedicated to computer vision projects it serves as a valuable resource for developing and evaluating deep learning models for breast tumor detection in mammographic images.

Manuscript received: 7 October 2025,
Revised: 15 December 2025,
Accepted: 17 December 2025.

[1]bora.basaran@tedu.edu.tr (**Corresponding author**)
[2]boncul@kastamonu.edu.tr.

Recent studies have demonstrated the effectiveness of deep learning and machine learning techniques in breast cancer classification, particularly in analyzing histopathological and radiological images. Convolutional neural networks (CNNs) have been widely applied to pathology and histology datasets, achieving high accuracy in both binary and multi-class classification tasks (Liu *et al.* 2022; Golatkar *et al.* 2018; Abunasser *et al.* 2023; Nguyen *et al.* 2019). These models are capable of automatically learning hierarchical features critical for distinguishing between malignant and benign cases. Hybrid approaches that integrate handcrafted features with dense neural layers have also shown promising results, improving model robustness and interpretability (Joseph *et al.* 2022). In addition, several studies have utilized patch-based strategies to capture localized tissue structures, enhancing classification accuracy through multi-size and discriminative input representations (Li *et al.* 2019). Transfer learning techniques have been employed to leverage pre-trained models and improve generalization in settings with limited labeled data (Saber *et al.* 2021).

Classical machine learning algorithms, such as Bayes classifiers and support vector machines, continue to be explored for their simplicity and effectiveness in structured data contexts (Bazila and Thirumalaikolundusubramanian 2018; Chen *et al.* 2023; Wu and Hicks 2021; Ara *et al.* 2021). Furthermore, recent work has proposed fully automated deep learning pipelines for breast cancer detection (Ghrabat *et al.* 2022), and efforts to detect metastatic cancer using large-scale deep models have demonstrated the clinical relevance of AI-based solutions (Wang *et al.* 2016). These collective advancements provide a solid foundation for developing and evaluating diverse architectures such as CNNs, Vision Transformers (ViTs), and adapted AlexNet models within the scope of breast cancer classification, as investigated in this study.

The study is divided into four parts. In Section II, the proposed method and three deep learning (DL) models used to effectively address breast cancer classification through three different model architectures are discussed. In Section III, the results and discussion are presented based on the performance of the three models, including outcomes obtained via Hugging Face Spaces. Finally, Section IV concludes the study, demonstrating the effectiveness of the proposed approach and providing suggestions for future work.

PROPOSED METHOD

It is known that medical professionals have investigate breast x-ray images to effectively keep track of the health-monitoring of patients that have signs of breast cancers. The study has been tailored via provide comparative analysis to the medical professionals in breast cancer treatment domain with the classification models that they have been carried out. The comparative study has been divided into three models that are CNN, ViT and AlexNet to provide the most successful model to abstain from the faulty treatments to the health-professionals. In this study three DL models have been trained to obtain the most successful model on classification of breast cancer-domain.

Three DL models CNN, ViT and AlexNet have been trained in this manner with different hyperparameters. The images that exist on the dataset are images with the 640x640 pixel size that are each Breast X-ray images which those datasets are generally known as mammography data. Since 224x224 pixel size images have been used on classification in health-care domain; the images have been resized to the intended pixel size. Three models CNN, ViT and AlexNet will be investigated in a comparative manner, and they will provide the most optimal model to successfully classification of breast cancer patients via negative and positive that the classes of dataset have been spitted into two parts that are 0 and 1 respectively. To comparison of performances on each model, metrics that are accuracy, loss and time comparison has been investigated and based on those metrics; models' have been deployed on HuggingFace space via option of selection of model. Via HuggingFace space, health professionals will be uploading their breast x-ray images and they will be monitoring the health-condition of their patients effectively and the results they will obtain will be successfully classified images which will be breast cancer negative and breast cancer positive respectively.

The first architecture used in this study is a lightweight Custom Convolutional Neural Network (CNN), built from scratch for binary classification of breast X-ray images. It processes 224×224×3 input images through three convolutional blocks, each comprising a Conv2D layer with ReLU activation and a max-pooling layer. The blocks use 32, 64, and 32 filters respectively, all with 3×3 kernels. After the final block, the feature maps are flattened into an 86528-dimensional vector, followed by a dense layer with 128 ReLU units. A final sigmoid-activated neuron outputs the binary classification, cancerous or non-cancerous. This model offers an efficient and interpretable design suitable for real-time or resource-constrained clinical applications.

The second model employs a Vision Transformer (ViT), introducing attention-based mechanisms for classifying 224×224×3 breast X-ray images. The image is divided into 196 non-overlapping 16×16 patches, each flattened and embedded into a 64-dimensional space with positional encoding. These embeddings pass through two Transformer Encoder blocks containing LayerNorm, Multi-Head Self-Attention, and Feed Forward Networks with residual connections. A global average pooling layer aggregates the information, followed by a dense layer (128 units, ReLU) and a final sigmoid-activated neuron for binary cancer prediction. ViT enables the model to capture global context and subtle spatial patterns beyond the reach of traditional CNNs.

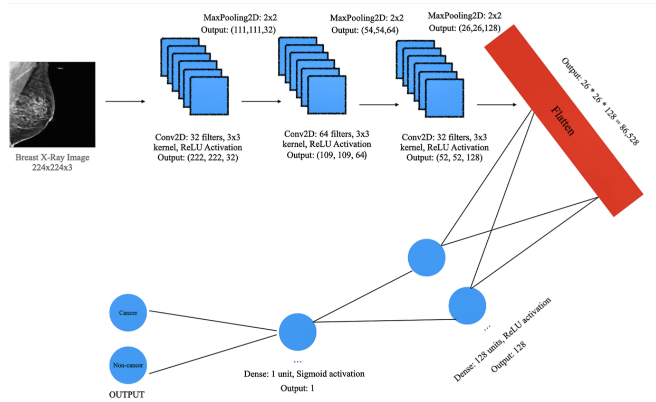
The third model adapts AlexNet, a classical deep CNN, for breast cancer classification. It processes 224×224×3 images through five convolutional layers with filters ranging from 96 to 384, using 11×11, 5×5, and 3×3 kernels, followed by max-pooling. The extracted features are flattened into a 9216-dimensional vector and passed through two fully connected layers with 4096 ReLU units and dropout (0.5). A final sigmoid-activated layer provides binary output. This adaptation highlights the effectiveness of traditional deep networks in medical imaging tasks when domain-optimized (Figure 1).

RESULTS AND DISCUSSION

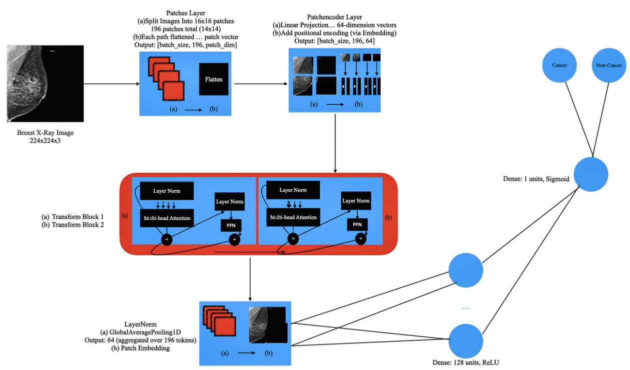
Carrying out three (DL) models provide health-professionals to choose to compare the results with three models' outcomes and their sight. Furtherly study has been conducted with the results of breast cancer negative and positive respectively on Hugging-Face space on a use of medical professionals. Both breast cancer negative and positive images have been investigated, and their potential results' have been analyzed in a comparative manner to provide valuable insight into medical and computer science integrated domain which solely focusses on cancer detection. In Study results' that have been observed that some results' have been classified as faulty even though high accuracy metrics' that have been employed. To abstain from the wrong classification, the corresponding output images as breast cancer negative and positive will be carried out with even the sight of the medical professionals for the most correct outcome that will be derived from the study which will be in use of medical professionals that the breast x-ray images that they have in their hand and patients'.

In this study for provide an example two breast x-ray images from the dataset have been selected to compare the performances of classification that one of the images are selected from breast cancer negative and that the other image has been selected from breast cancer positive and their classification performance on Hugging-Face space have been investigated. According to the HuggingFace results that they have been obtained in positive breast-cancer results based on breast x-ray images there are no misclassification on models CNN, ViT and AlexNet respectively. However, there is a minor fault classification in negative breast-cancer result in the AlexNet model; even though high classification accuracy that has been obtained in Training set. In that circumstance the sight of medical professional in chest major is a must that they are provided in (Figure 2).

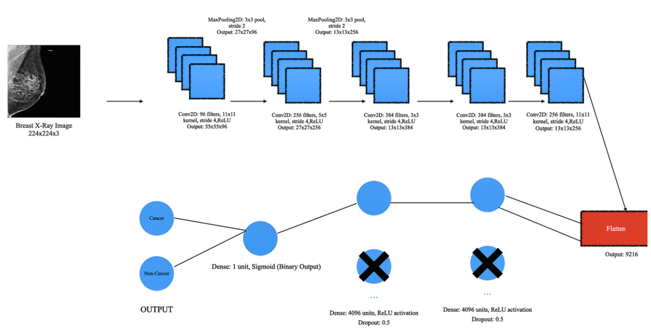
Due to the misclassification on breast-cancer-negative image on the AlexNet model must keep trackt of the models' performance with several metrics. Several metrics in this domain are based on classification metrics due to the classification study has been tailored that the metrics that are in use will be accuracy and loss respectively. Accuracy and losses of the system will keep track of during each of epochs during execution of training models re-



(a) CNN Architecture



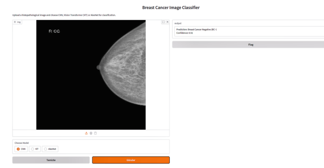
(b) ViT Architecture



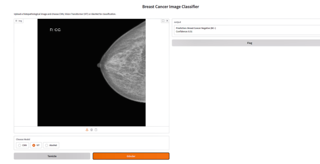
(c) AlexNet Architecture

Figure 1 Model Architectures of Models that they have trained during the Breast Cancer Classification study.

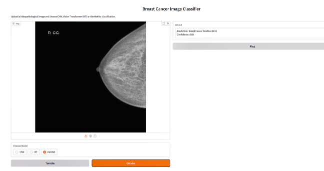
spectively on training and validation sets to observe how model behaves with the inputs that the inputs are breast x-ray images. Loss function of binary cross-entropy will be observed due to it is a binary classification study that the sigmoid activation function activates the neurons at the last layer of each model. Each models' accuracy and loss metrics have been played a vital role in breast-cancer classification study with the accuracies that are over 0,9 in each of the models' separately. Loss functions have been diminished based on epochs in each model which is a shown that the models behave well under several conditions in both validation and training splits. Additionally, validation performances of models have been trained; strong even though plot of them seems not



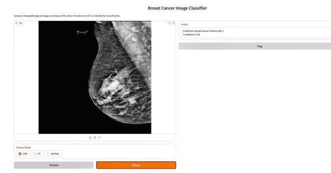
(a) CNN – Breast Cancer Negative



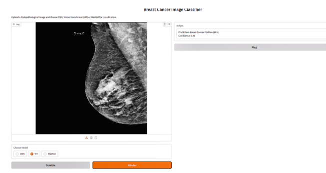
(b) ViT – Breast Cancer Negative



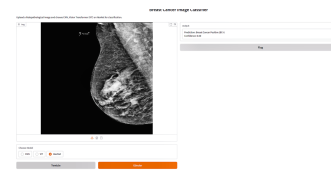
(c) AlexNet – Breast Cancer Negative



(d) CNN – Breast Cancer Positive



(e) ViT – Breast Cancer Positive



(f) AlexNet – Breast Cancer Positive

Figure 2 HuggingFace Breast Cancer Classification Space which one Breast Cancer Negative and one Breast Cancer Positive Images have been investigated.

that promising. Even though breast-cancer classification has been trained and the outcomes have been provided correct classification which will also lead health-professionals on the domain of cancer classification which will save many human-beings' life in early diagnosis of breast cancer. Accuracy and Loss plots of each model on training and validation sets have been provided in (Figure 3).

Table 1 presents a comprehensive comparison of the three deep learning models CNN, Vision Transformer (ViT), and AlexNet, in terms of classification accuracy, loss, training epochs, parameter counts, and execution time during breast X-ray image classification. Among the three models, the Custom CNN achieved the highest accuracy of 95.44% with a low loss of 0.1078 in just 45 epochs. This performance demonstrates its strong ability to extract relevant features from medical images using a lightweight architecture. Additionally, the CNN model required approximately 807.82 seconds of execution time and had around 11.1 million trainable parameters, striking a good balance between performance and computational efficiency.

The ViT model, on the other hand, while introducing advanced attention mechanisms, achieved a lower accuracy of 91.56% with a higher loss of 0.1987 despite training for 75 epochs. Notably, the ViT had the fewest trainable parameters (236,737) and the lowest model complexity, yet it required 1365.79 seconds to execute nearly double the time of the CNN. This indicates that although ViT excels at capturing long-range dependencies, it may not be as well-suited for small-scale medical datasets or classification tasks where local features dominate. AlexNet, the most complex model in terms of architecture and parameters (46.75 million trainable parameters), achieved an accuracy of 95.20%, which is comparable to CNN, with a slightly better loss of 0.1072.

However, this performance came at a significant computational cost, it recorded the highest execution time of 2597.42 seconds,

Table 1 Models and Their Performances have been analyzed in a comparative manner in Breast Cancer Classification study

Metrics / Model	CNN	ViT	AlexNet
Accuracy	0.9544	0.9156	0.9520
Epochs	45	75	50
Loss	0.1078	0.1987	0.1072
Trainable Parameters	11.169.089	236.737	46.751.105
Total Parameters	11.169.089	236.737	46.751.105
Execution Time	807.82	1365.79	2597.42

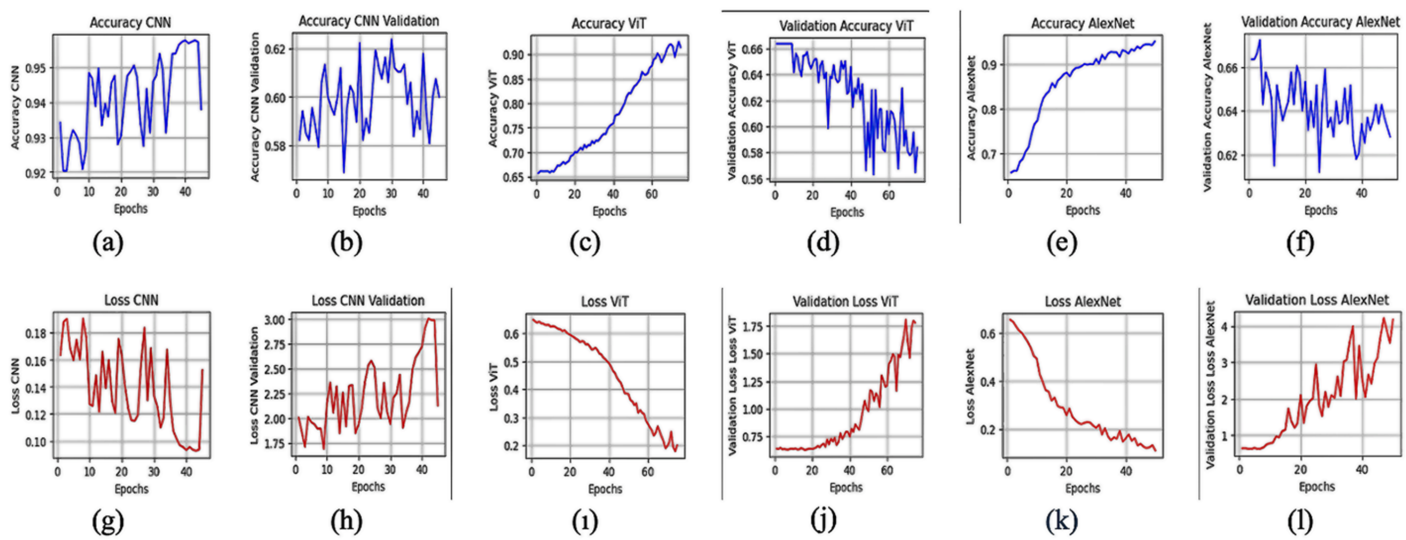


Figure 3 Accuracy and Loss Visualizations of models that they have been trained on Breast Cancer Classification study. (a) CNN – Accuracy (Train), (b) CNN – Accuracy (Validation), (c) ViT – Accuracy (Train), (d) ViT – Accuracy (Validation), (e) AlexNet – Accuracy (Train), (f) AlexNet – Accuracy (Validation), (g) CNN – Loss (Train), (h) CNN – Loss (Validation), (ı) ViT – Loss (Train), (j) ViT – Loss (Validation), (k) AlexNet – Loss (Train), (l) AlexNet – Loss (Validation).

more than three times that of CNN. This highlights the trade-off between deep architectural complexity and practical applicability in real-time or resource-limited clinical environments. In summary, the CNN model consistently outperformed both the Vision Transformer (ViT) and AlexNet architectures across multiple evaluation metrics, including accuracy, computational efficiency, and execution time, establishing it as the most optimal choice for the breast cancer classification task in this study.

The CNN's ability to effectively capture spatial hierarchies and local features contributed to its superior classification performance, making it well-suited for medical image analysis where precision is critical. While the ViT model demonstrated promise with its relatively minimal number of trainable parameters, its higher execution overhead and longer processing times present practical challenges for real-time or resource-constrained diagnostic environments. On the other hand, AlexNet, despite achieving comparable accuracy to CNN, proved to be computationally demanding and slower, limiting its feasibility for rapid clinical deployment. These findings highlight the trade-offs between model complexity, accuracy, and operational efficiency, emphasizing the CNN's balance as ideal for breast cancer detection workflows that require both reliability and speed. Future work could explore hybrid models or lightweight transformer variants to potentially combine the strengths of these architectures while mitigating their respective limitations.

CONCLUSION

In this study, three deep learning models trained from scratch that are respectively CNN, ViT and AlexNet have been analyzed in a comparative manner with the Breast X-ray images to classify the input images based on breast cancer negative and positive. The study has been conducted with a dataset of breast cancer x-ray images that are in use for computer vision and deep learning studies that the dataset consists of 3,383 breast x-ray images. Study has been tailored via using three deep learning models' and their respective implementations seperately and after training of models' HuggingFace space environment has been used to deployment of the study which the study has been classified several breast x-ray images based on breast cancer classification which classify images breast cancer negative and breast cancer positive respectively. To compare the performance of three proposed models' performance metrics of accuracy, loss and execution time have been analyzed in both training and validation sets respectively.

According to the outcomes of the results, medical professionals should keep track of the information of CNN preferable, apart from CNN; they may rely on the trust of the ViT model. However as discussed in the begining of the results section AlexNet model have proven some misclassifications even though its high accuracy and losses obtained during training. Accuracy and loss performances of the models' proven the suggestion that the medical Professionals should rely on the CNN models' trust in the classification of breast cancer x-ray images based on cancer. Outcomes have been proven that the 0,9544 accuracy and 0,1078 loss performance observed in CNN model which outperformed both ViT and AlexNet models respectively. In further studies It should be analyzed that the AlexNet model should be trained with higher number of epochs with more amount of execution time or in comparison, different models' apart from AlexNet should be investigated to performance comparison on breast cancer domain.

Acknowledgments

Preliminary findings of this study were presented as an oral abstract entitled "Evrişimsel Sinir Ağları, Görüntü Dönüştürücüler ve AlexNet ile Meme X-Ray Görüntüleriyle Meme Kanserinin Sınıflandırılması" at the 5th International Turkish World Engineering and Science Congress (Abstract No: 172092).

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

Abunasser, B. S., M. R. J. Al-Hiealy, I. S. Zaqout, and S. S. Abu-Naser, 2023 Convolution neural network for breast cancer detection and classification using deep learning. Asian Pacific journal of cancer prevention: APJCP **24**: 531.

Ara, S., A. Das, and A. Dey, 2021 Malignant and benign breast cancer classification using machine learning algorithms. In *2021 international conference on artificial intelligence (ICAI)*, volume 2021, pp. 97–101, IEEE Islamabad, Pakistan.

Bazila, B. B. A. and P. Thirumalaikolundusubramanian, 2018 Comparison of bayes classifiers for breast cancer classification. Asian Pacific Journal of Cancer Prevention **19**: 2917–2922.

Chen, H., N. Wang, X. Du, K. Mei, Y. Zhou, *et al.*, 2023 Classification prediction of breast cancer based on machine learning. Computational intelligence and neuroscience **2023**: 6530719.

Ghrabat, M. J. J., Z. A. Hussien, M. S. Khalefa, Z. A. Abduljabba, V. O. Nyangaresi, *et al.*, 2022 Fully automated model on breast cancer classification using deep learning classifiers. Indonesian Journal of Electrical Engineering and Computer Science **28**: 183–191.

Golatkar, A., D. Anand, and A. Sethi, 2018 Classification of breast cancer histology using deep learning. In *International conference image analysis and recognition*, pp. 837–844, Springer.

Joseph, A. A., M. Abdullahi, S. B. Junaidu, H. H. Ibrahim, and H. Chiroma, 2022 Improved multi-classification of breast cancer histopathological images using handcrafted features and deep neural network (dense layer). Intelligent Systems with Applications **14**: 200066.

Li, Y., J. Wu, and Q. Wu, 2019 Classification of breast cancer histology images using multi-size and discriminative patches based on deep learning. IEEE Access **7**: 21400–21408.

Liu, M., L. Hu, Y. Tang, C. Wang, Y. He, *et al.*, 2022 A deep learning method for breast cancer classification in the pathology images. IEEE Journal of Biomedical and Health Informatics **26**: 5025–5036.

Nguyen, C. P., A. H. Vo, and B. T. Nguyen, 2019 Breast cancer histology image classification using deep learning. In *2019 19th international symposium on communications and information technologies (ISCIT)*, pp. 366–370, IEEE.

Saber, A., M. Sakr, O. M. Abo-Seida, A. Keshk, and H. Chen, 2021 A novel deep-learning model for automatic detection and classification of breast cancer using the transfer-learning technique. IEEe Access **9**: 71194–71209.

Wang, D., A. Khosla, R. Gargeya, H. Irshad, and A. H. Beck, 2016 Deep learning for identifying metastatic breast cancer.

Wu, J. and C. Hicks, 2021 Breast cancer type classification using machine learning. Journal of personalized medicine **11**: 61.

How to cite this article: Basaran, B., and Oncul, A. B. Classification of Breast Cancer with Breast X-Ray Images via Convolution Neural Networks, Vision Transformers and AlexNet. *Computers and Electronics in Medicine*, 3(1), 94-98, 2026.