Early Breast Cancer Detection Using Deep Learning Classification

Atef Ben Miled
ATSSEE Laboratory, Science Faculty of Tunis, University of Tunis Manar, Tunis, Tunisia
atssse.fst2016@gmail.com (corresponding author)

Adnen Cherif
ATSSEE Laboratory, Science Faculty of Tunis, University of Tunis Manar, Tunis, Tunisia
adnane.cherif@fst.utm.tn

Received: 19 March 2025 | Revised: 10 April 2025 | Accepted: 15 April 2025

 DOI: https://doi.org/10.48084/etasr.11030

ABSTRACT

Breast cancer is one of the most prevalent and lethal forms of cancer among women. Early detection and correct analysis play crucial roles in improving patient results and increasing survival rates. However, conventional strategies for the screening and analysis of most breast cancers, including mammography, ultrasound, and biopsy, may be limited by their accuracy and specificity, mainly due to the neglect of fine-lesion cases. This paper describes a system for the early detection of breast cancer, based on a Deep Learning (DL) strategy to enhance the accuracy of breast cancer detection. The proposed system uses a Convolutional Neural Network (CNN) trained on the DDSM database of mammogram images to process and classify suspicious lesions with high precision. The DL model is optimized using advanced techniques, transfer learning, data augmentation, and the ResNet50 model to improve its performance and generalization skills and capabilities. The implementation results demonstrated significant precision (98%), especially in the detection of fine lesions and suspicious microcalcifications.

Keywords-computer-aided diagnosis; cancer detection; mammography images; deep learning; CNN classifier

I. INTRODUCTION

Breast cancer is one of the most common diseases that affect women, particularly those older than 45 years [1]. In 2020, nearly 2.3 million new cases of breast cancer were diagnosed, making it the most common cancer among girls and women [2]. However, despite its high incidence, scientific and medical advances have highlighted the importance of early detection to increase survival rates and reduce medical and social costs [3]. Breast cancer is characterized by the presence of a malignant tumor due to the extraordinary and anarchic multiplication of cells in tissue [4]. Several forms of breast cancer are described by the American College of Radiology (ACR) as abnormalities in mammographic images, including mass lesions, architectural distortions, and microcalcifications [5-6]. However, although mammography images are evaluated and interpreted by radiologists, previous studies have confirmed that between 10% and 30% of breast cancer cases are undiagnosed or undetected at the first level [7]. This justifies the need to use smart diagnostic guidelines and Artificial Intelligence (AI) tools to improve clinical results and detect the disease early.

Early work on Computer-Aided Diagnosis (CAD) systems for breast cancer detection dates back to the 1990s. Traditional CAD systems rely on handcrafted features and machine learning strategies, such as Support Vector Machines (SVMs) and Random Forest (RF), to analyze mammogram images and detect suspicious lesions [8-9]. However, these systems are limited by their dependence on manually engineered features and their inability to capture complex data. Later, in [10], a segmentation approach was presented to address the inherent ambiguities in the sensor data. This technique included spatial context information for pixel types, mirroring an essential element of human reasoning. With the current developments in Deep Learning (DL), CNNs have revolutionized the field of scientific image analysis, such as breast cancer detection [11]. DL models, such as CNNs, can process hierarchical representations in medical images, eliminating the need for manual feature engineering. One of the first advanced studies on this discipline was carried out in [12], where a CNN-based approach was presented to classify breast lesions on mammograms. This approach achieved an Area Under the Receiver Operating Characteristic Curve (AUC-ROC) of 0.87, outperforming conventional CAD methods. In [13], a transfer learning approach was presented, using pre-trained CNNs that achieved an AUC of 0.90 on a large-scale mammogram dataset. As models advance, multimodal strategies are explored, including multiple imaging modalities such as mammograms, ultrasound, and Magnetic Resonance Imaging (MRI), to improve the accuracy of breast cancer diagnosis. In [14], a multimodal DL framework used mixed mammogram and ultrasound statistics, achieving an AUC of 0.95.

With the increasing adoption of DL methods in clinical applications, their interpretability has become a critical issue. Researchers have explored strategies that include saliency maps, attention mechanisms, and Concept Activation Vectors (CAVs) to offer insights into the feature selection procedure of DL models for breast cancer diagnosis [15]. To address the challenge of limited labeled data in the medical domain, multimodal techniques have been used, in which supervised models are first tuned on mammograms or other medical image records [16]. Additionally, statistical augmentation strategies, including random cropping, flipping, and rotation, are employed to increase the diversity of training data and improve model performance [17]. Recent studies have explored the use of superior DL architectures, including attention-based models [18], Generative Adversarial Networks (GANs) [19], and federated learning approaches for collaborative model training [20]. Furthermore, researchers are investigating the integration of DL models with clinical decision support systems and the development of end-to-end breast cancer diagnosis pipelines [21].

II. MATERIALS AND METHODS

A. *Computer Aided Diagnosis (CAD) Methods*

There are two essential categories of image diagnostic support systems. The first is based on medical observation, where the radiologist intervenes in the interpretation of the results and the decision. The second type combines image processing with AI [22]. The proposed non-invasive strategy uses real-time images and AI to detect and classify images according to a pre-established database. Such systems aim to provide radiologists with confidence in their decisions regarding the characterization of a detected lesion as malignant or benign. Figure 1 shows the automatic detection of suspicious areas and the localization of potentially identified lesions. After the first step of image acquisition and editing, a preprocessing stage is used for filtering and end cropping, followed by segmentation based on the detection of suspect lesion regions. Feature extraction is used to feed the CNN classifier, which is very important because it is responsible for the final decision. The evaluation stage uses classification metrics to assess the detection results and determine the reliability of the system.

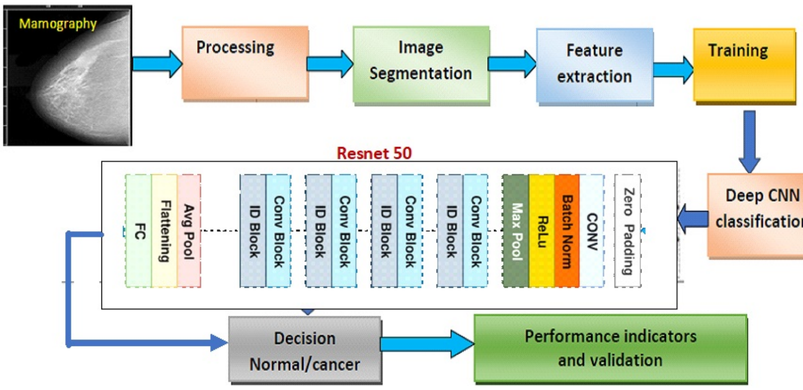

Fig. 1. CAD algorithm.

B. *Dataset*

This study used a large, multi-institutional dataset comprising mammograms, images, and clinical data from breast cancer screening and diagnostic centers. The dataset involves data collected from the Digital Database for Screening Mammography (DDSM), which is a collaborative effort involving the Massachusetts General Hospital, the University of South Florida, and Sandia National Laboratories [23]. This database consists of mammography images of the right and left breasts taken in two planes, the MLO (medial-lateral oblique) and CC (caudal) planes. The DDSM consists of 2500 cases of mammographic images. The acquisition was performed with an oblique position relative to the breast for MLO and a horizontal position for CC. Clinical information about patients, such as acquisition date, age, and breast tissue density, was also collected. The database contains approximately 2500 cases of healthy and pathological patients. Each case includes two images of each breast, as well as certain patient (age at the time of the study, ACR breast density index, subtlety of anomalies index, description of anomalies by ACR keywords) and image (scanner, spatial resolution) information.

TABLE I. DDSM DATABASE

AI operations	Normal cases	Cancer cases
Training	800	800
Testing	150	150
Validation	50	50
Total	1000	1000

In the database, 400 cases, 300 for training and 100 for testing, represent three types of breast cancer: mass lesion, microcalcification, and architectural distortions. Since this study examined only one type of cancer, the number of cases was reduced.

C. *Image Segmentation*

Iterative segmentation methods are based on the use of Regions Of Interest (ROI), chosen either manually or automatically, as illustrated in Figure 2. These ROIs designate the regions within the image to be segmented. Based on similarity measures, each region is compared to its immediate spatial neighborhood [24]. Considering the regions obtained, the process is iterated until the regions likely to be merged are exhausted [25].

D. Feature Extraction

After segmentation, relevant features, such as texture, shape, and density, are extracted from the identified regions. These parameters, combined with the preprocessed images, form a structured dataset [26]. This set of parameters serves as the input layer parameters for the neural network. An appropriate CNN classification model is selected for training on this set. Training includes model validation and tuning, including hyperparameter optimization, to improve performance. Finally, the model is integrated into the clinical workflow, allowing radiologists to effectively use the diagnostic support system for a faster and more accurate evaluation of mammograms. A continuous feedback process, with the collection of feedback from healthcare professionals, contributes to the continuous improvement of the model by integrating new data and adjusting the algorithms if necessary. The use of DL algorithms, particularly CNNs, is effective for mammography image classification due to its ability to learn complex patterns from the data.

E. Proposed DL Architecture

The proposed CAD tool employs a CNN to analyze multimodal data. The CNN architecture, shown in Figure 3, consists of a feature extraction backbone and several task-specific heads for mammogram analysis and clinical data processing. Feature extraction is performed using a pre-trained ResNet-50 model that was fine-tuned on mammograms. The task-specific heads were designed to capture modality-specific patterns and consisted of several convolution and Fully Connected (FC) layers. The backpropagation algorithm was used to train the neural network.

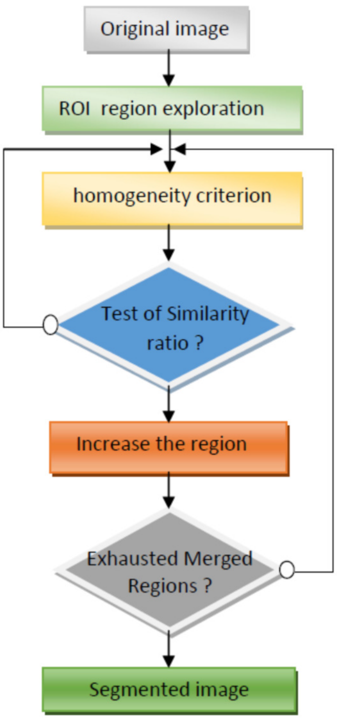

Fig. 2. Segmentation principle.

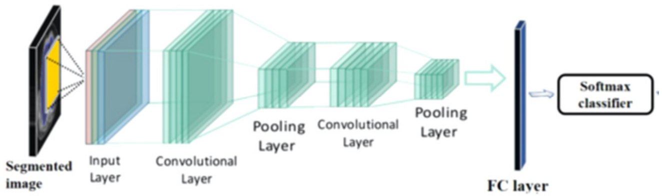

Fig. 3. Structure of a CNN classifier.

F. CNN Classification

CNN classifiers have several structures depending on the multiple layers, such as VGG 16 and Resnet 50, composed of input, convolution, pooling, FC, and output layers, as illustrated in Figure 4. ResNet50 has 50 layers, which allows it to learn complex hierarchical features from input data. This depth enables it to capture patterns and nuances in mammography images, which are crucial for accurate classification tasks. Furthermore, ResNet-50 includes the concept of residual or skip connections, which mitigate the vanishing gradient problem commonly encountered in deep neural networks. These skip connections facilitate the flow of gradients during backpropagation, enabling more efficient training and convergence. Another key advantage of ResNet-50 is its ability to effectively combat overfitting, a common challenge in DL models. Skip connections in ResNet-50 promote feature reuse across different layers, effectively preventing the model from memorizing noise in the training data and promoting better generalization to unseen data. In general, CNNs have three types of layers: convolution, pooling, and FC layers. Figure 4 provides a detailed description of the Resnet50 architecture and operations. The convolution layer performs the calculation operations and is made up of the input data, a filter, and a feature map. The grouping or pooling layer is a subsampling layer that reduces the dimensionality and the number of input parameters. To do this, a filter on the entire input or an aggregation function with the values of the receptive field can be used, thus filling the output matrix. There are two main types of pooling: Max pooling and average pooling. Finally, in the FC layer, each node in the output layer connects directly to a node in the previous one. Therefore, this layer can perform classification based on the features extracted from the previous layers and their different filters. Convolution and pooling layers tend to use ReLU activation functions, and FC layers typically exploit a softmax activation function to classify inputs appropriately, resulting in a 0 to 1 probability.

In this work, breast cancer detection was implemented with a dataset of mammography images using Python, TensorFlow, and Keras. TensorFlow is an open-source machine learning framework developed by Google. This tool enables the training and deployment of CNNs for image classification tasks. Keras, a high-level neural network API, complements TensorFlow by providing a simplified interface for building and training DL models. Its intuitive design and abstraction layers make it easy to prototype and experiment with different network architectures, allowing rapid iteration and refinement of the detection system. The dataset contains many annotated mammography images, along with the corresponding ground truth labels that indicate the presence or absence of cancerous abnormalities.

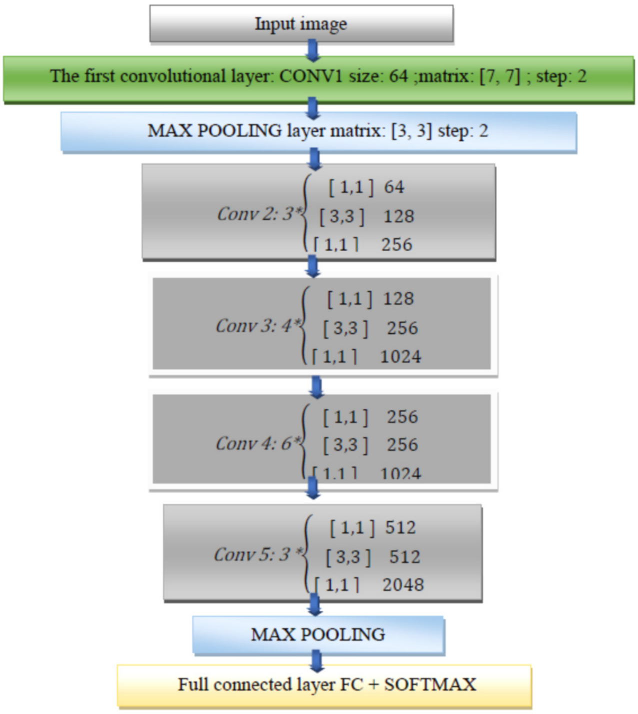

Fig. 4. CNN-ResNet 50 architecture.

G. . *Evaluation Metrics*

The performance of the proposed model was evaluated using precision, recall, sensitivity, specificity, error rate, accuracy, and F1-score.

$$\text{Precision} = \frac{t_p}{t_p+f_p} \tag{1}$$

$$\text{Recall} = \frac{t_p}{t_p+f_n} \tag{2}$$

$$\text{Sensitivity} = \frac{t_p}{t_p+t_n} \tag{3}$$

$$Specificity = \frac{t_n}{f_p+t_n} \tag{4}$$

$$\text{ErrorRate} = \frac{f_p+f_n}{total} \tag{5}$$

$$\text{Accuracy} = \frac{t_p+t_n}{t_p+t_n+f_p+f_n} \tag{6}$$

$$\text{F1}-\text{score} = 2\ \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \tag{7}$$

where t_p denotes the number of cases estimated as true positive, f_p denotes the number of cases estimated as false positives, f_n denotes the number of cases estimated as false negatives, t_n denotes the number of cases estimated as true negatives, and $total$ is the sum of all cases ($t_p + t_n + f_p + f_n$).

III. RESULTS

A. *Graphical Interface*

A graphical interface was designed for processing and diagnosis. The initial step involves filling out a form with the patient's name or code, followed by her age. Subsequently, the user selects whether the mammography image was taken from the right or left side, along with its acquisition type. Finally, the density is selected to differentiate between milk-producing lobules or tumors. When the program is running, a window displays four images, as illustrated in Figure 5: the original image, the image identified representing the regions of interest by threshold segmentation, the segmented image indicating the location of the tumor, and the classified image showing the class it belongs to. Additionally, the program provides classification reports. This graphical interface serves as a comprehensive tool for early diagnosis, allowing healthcare professionals to input essential patient information and receive detailed diagnostic results. The form fields are strategically designed to capture key patient attributes, such as name/code and age, which are crucial for accurate diagnosis and tracking of patient history.

The consideration of breast density in the diagnostic process is vital, as it helps differentiate between normal breast tissue and potential tumors, contributing to more accurate and reliable diagnoses. The identified image, produced through threshold segmentation, highlights regions of interest within the mammography image, guiding healthcare professionals to areas that require closer inspection. The segmented image accurately delineates the location of any detected tumors within the breast, facilitating targeted treatment planning and intervention. Finally, the classification image shows the categorization of the detected abnormalities, providing information for treatment decision-making and prognosis assessment.

B. *Metrics and Validation*

In addition to visual outputs, the program offers quantitative metrics, including learning curves, classification reports, and accuracy scores, which provide assessments of the diagnostic process's performance. These metrics enable healthcare professionals to evaluate the efficacy of the CAD algorithm and make informed decisions regarding patient management. Figure 6 shows the evolution of the error rate across epochs. The training losses were around 5%, whereas the test error rate reached 14%.

C. *Classification Indicators and Confusion Matrix*

The CAD dashboard offers also a table of results containing all indicators and the confusion matrix, as illustrated below in Figure 7. The following confusion matrix is an example of a case study conducted on 144 cases but with only 114 cases retained after processing and feature extraction due to the quality of certain images or incompatible recording and coding formats. The values show that only 2 cases are false detected there are no false negative cases, which gives an accuracy of more than 98%.

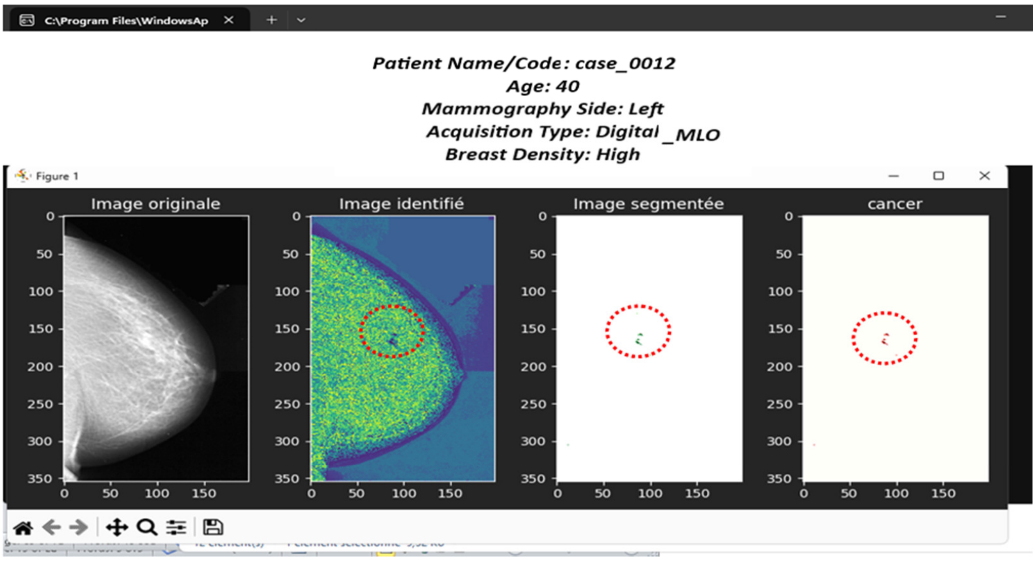

Fig. 5. Example of a patient with breast cancer with micro-calcification.

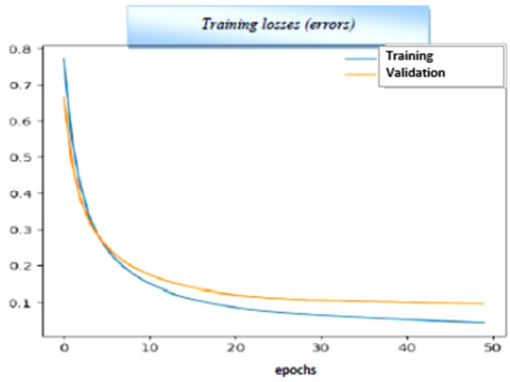

Fig. 6. Training and validation error rates.

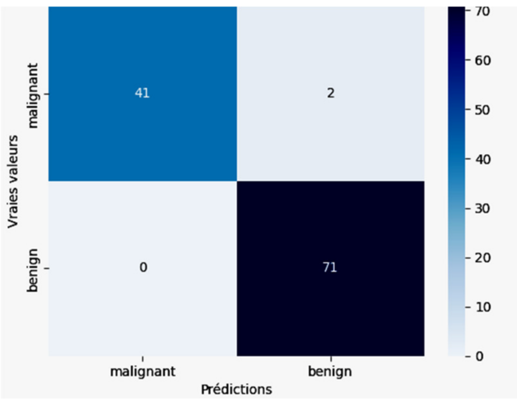

Fig. 7. Confusion matrix.

TABLE II. COMPARISON OF PERFORMANCE INDICATORS

Reference	Precision	Recall	Accuracy	F1-score	CNN
[27]	76%	69%	75%	75%	VGG16
[28]	98%	-	97%	94%	TTCNN
This study	98%	92%	96%	95%	Resnet50

IV. DISCUSSION

The proposed breast cancer diagnosis system achieved an overall accuracy of 96% in the detection of malignant lesions using DDSM data. The CNN-based classification system achieved a sensitivity of 93%, indicating a high rate of correct identification of cancer cases while maintaining an acceptable specificity of 90%. Compared to traditional CAD systems, the proposed DL approach significantly improved diagnostic performance. The model's segmentation mechanism was effective, allowing it to adaptively focus on relevant features such as microcalcifications and irregular mass shapes. The proposed DL system demonstrated the potential to significantly improve early breast cancer detection accuracy, leading to better patient outcomes and fewer unnecessary biopsies. Furthermore, integrating clinical risk factors into CNN input layers could increase disease diagnosis rates and precision. The results of the proposed model were compared with other research studies using the same DDSM database for the detection of breast cancer, as shown in Table II.

V. CONCLUSION

This study presented a CAD tool for early breast cancer detection and diagnosis using a CNN. The system uses a pre-trained CNN (ResNet50) applied to a large dataset of mammograms and clinical data from DDSM to identify visual characteristics associated with malignant and benign lesions. The proposed DL method showed good efficiency in detecting small lesions and calcified regions. However, it faces limitations in detecting lesions in dense breasts and its sensitivity to imaging artifacts. To overcome these issues, multimodal approaches can be adopted by combining mammographic data with ultrasound and MRI images. The main contribution of this approach is the potential to detect early breast cancer disease and microcalcifications, reducing the need for expensive, invasive, and painful chemical treatments.

REFERENCES

[1] J. Ferlay *et al.*, "Global Cancer Observatory: Cancer Today," *International Agency for Research on Cancer*. https://gco.iarc.who.int/today/.

[2] M. Arnold *et al.*, "Current and future burden of breast cancer: Global statistics for 2020 and 2040," *The Breast*, vol. 66, pp. 15–23, Dec. 2022, https://doi.org/10.1016/j.breast.2022.08.010.

[3] F. Bray, A. Jemal, N. Grey, J. Ferlay, and D. Forman, "Global cancer transitions according to the Human Development Index (2008–2030): a population-based study," *The Lancet Oncology*, vol. 13, no. 8, pp. 790–801, Aug. 2012, https://doi.org/10.1016/S1470-2045(12)70211-5.

[4] A. Oliver *et al.*, "A review of automatic mass detection and segmentation in mammographic images," *Medical Image Analysis*, vol. 14, no. 2, pp. 87–110, Apr. 2010, https://doi.org/10.1016/j.media.2009.12.005.

[5] "ACR Breast Imaging Reporting & Data System (BI-RADS®)." https://www.acr.org/Clinical-Resources/Clinical-Tools-and-Reference/Reporting-and-Data-Systems/BI-RADS.

[6] M. P. Sampat, M. K. Markey, and A. C. Bovik, "Computer-Aided Detection and Diagnosis in Mammography," in *Handbook of Image and Video Processing*, Elsevier, 2005, pp. 1195–1217.

[7] A. M.. Khuzi, R. Besar, W. Wan Zaki, and N. Ahmad, "Identification of masses in digital mammogram using gray level co-occurrence matrices," *Biomedical Imaging and Intervention Journal*, vol. 5, no. 3, Jul. 2009, Art. no. e17, https://doi.org/10.2349/biij.5.3.e17.

[8] M. L. Giger, Z. Huo, M. A. Kupinski, and C. J. Vyborny, "Computer-Aided Diagnosis in Mammography,"F in *Handbook of Medical Imaging*, vol. 2, SPIE Press, 2009.

[9] W. Qian, L. P. Clarke, B. Zheng, M. Kallergi, and R. Clark, "Computer assisted diagnosis for digital mammography," *IEEE Engineering in Medicine and Biology Magazine*, vol. 14, no. 5, pp. 561–569, Sep. 1995, https://doi.org/10.1109/51.464772.

[10] W. Eziddin, "Segmentation itérative d'images par propagation de connaissances dans le domaine possibiliste : application à la détection de tumeurs en imagerie mammographique," Ph.D. dissertation, Télécom Bretagne, Université de Bretagne Occidentale, 2012.

[11] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, Dec. 2017, https://doi.org/10.1016/j.media.2017.07.005.

[12] D. Ribli, A. Horváth, Z. Unger, P. Pollner, and I. Csabai, "Detecting and classifying lesions in mammograms with Deep Learning," *Scientific Reports*, vol. 8, no. 1, Mar. 2018, Art. no. 4165, https://doi.org/10.1038/s41598-018-22437-z.

[13] T. Kyono, F. J. Gilbert, and M. van der Schaar, "MAMMO: A Deep Learning Solution for Facilitating Radiologist-Machine Collaboration in Breast Cancer Diagnosis." arXiv, Oct. 30, 2018, https://doi.org/10.48550/arXiv.1811.02661.

[14] L. Shen, W. Zhao, D. Capaldi, J. Pauly, and L. Xing, "A geometry-informed deep learning framework for ultra-sparse 3D tomographic image reconstruction," *Computers in Biology and Medicine*, vol. 148, Sep. 2022, Art. no. 105710, https://doi.org/10.1016/j.compbiomed.2022.105710.

[15] W. Changyan Bs *et al.*, "Deep Learning on Ultrasound Imaging for Breast Cancer Diagnosis and Treatment: Current Applications and Future Perspectives," *Adcanced Ultrasound in Diagnosis and Therapy,* 2, 2023, Art. no. 91, https://doi.org/10.37015/AUDT.2023.230012.

[16] T. L. Le, M. H. Bui, N. C. Nguyen, M. T. Ha, A. Nguyen, and H. P. Nguyen, "Transfer learning for deep neural networks-based classification of breast cancer X-ray images," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 12, no. 1, Dec. 2024, Art. no. 2275708, https://doi.org/10.1080/21681163.2023.2275708.

[17] Y. Mo *et al.*, "HoVer-Trans: Anatomy-Aware HoVer-Transformer for ROI-Free Breast Cancer Diagnosis in Ultrasound Images," *IEEE Transactions on Medical Imaging*, vol. 42, no. 6, pp. 1696–1706, Jun. 2023, https://doi.org/10.1109/TMI.2023.3236011.

[18] M. Bobowicz *et al.*, "Attention-Based Deep Learning System for Classification of Breast Lesions—Multimodal, Weakly Supervised Approach," *Cancers*, vol. 15, no. 10, May 2023, Art. no. 2704, https://doi.org/10.3390/cancers15102704.

[19] Y. Jiménez-Gaona, D. Carrión-Figueroa, V. Lakshminarayanan, and M. José Rodríguez-Álvarez, "Gan-based data augmentation to improve breast ultrasound and mammography mass classification," *Biomedical Signal Processing and Control*, vol. 94, Aug. 2024, Art. no. 106255, https://doi.org/10.1016/j.bspc.2024.106255.

[20] M. J. Sheller *et al.*, "Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data," *Scientific Reports*, vol. 10, no. 1, Jul. 2020, Art. no. 12598, https://doi.org/10.1038/s41598-020-69250-1.

[21] S. M. McKinney *et al.*, "International evaluation of an AI system for breast cancer screening," *Nature*, vol. 577, no. 7788, pp. 89–94, Jan. 2020, https://doi.org/10.1038/s41586-019-1799-6.

[22] N. Behar and M. Shrivastava, "A Novel Model for Breast Cancer Detection and Classification," *Engineering, Technology & Applied Science Research*, vol. 12, no. 6, pp. 9496–9502, Dec. 2022, https://doi.org/10.48084/etasr.5115.

[23] M. Heath, K. Bowyer, D. Kopans, R. Moore, and P. Kegelmeyer, "The digital database for screening mammography," in *Proceedings of the Fifth International Workshop on Digital Mammography*, 2001, pp. 212–218.

[24] S. M. Shaaban, M. Nawaz, Y. Said, and M. Barr, "An Efficient Breast Cancer Segmentation System based on Deep Learning Techniques," *Engineering, Technology & Applied Science Research*, vol. 13, no. 6, pp. 12415–12422, Dec. 2023, https://doi.org/10.48084/etasr.6518.

[25] M. Adel, M. Rasigni, S. Bourennane, and V. Juhan, "Statistical Segmentation of Regions of Interest on a Mammographic Image," *EURASIP Journal on Advances in Signal Processing*, vol. 2007, no. 1, Dec. 2007, Art. no. 049482, https://doi.org/10.1155/2007/49482.

[26] H. C. Burrell *et al.*, "Screening interval breast cancers: mammographic features and prognosis factors.," *Radiology*, vol. 199, no. 3, pp. 811–817, Jun. 1996, https://doi.org/10.1148/radiology.199.3.8638010.

[27] Y. Khourdifi, A. El Alami, M. Zaydi, Y. Maleh, and O. Er-Remyly, "Early Breast Cancer Detection Based on Deep Learning: An Ensemble Approach Applied to Mammograms," *BioMedInformatics*, vol. 4, no. 4, pp. 2338–2373, Dec. 2024, https://doi.org/10.3390/biomedinformatics4040127.

[28] S. Maqsood, R. Damaševičius, and R. Maskeliūnas, "TTCNN: A Breast Cancer Detection and Classification towards Computer-Aided Diagnosis Using Digital Mammography in Early Stages," *Applied Sciences*, vol. 12, no. 7, Mar. 2022, Art. no. 3273, https://doi.org/10.3390/app12073273.