PRAGUE UNIVERSITY OF ECONOMICS AND BUSINESS

Acta Informatica Pragensia
2024, Volume 13, Issue 2, pp. 308–326
Special Issue: Future Trends of Machine Intelligence in Science and Industry
https://doi.org/10.18267/j.aip.244

Review Open Access

Advancements in Breast Cancer Diagnosis: A Comprehensive Review of Mammography Datasets, Preprocessing and Classification Techniques

Hama Soltani [1], Issam Bendib [1], Mohamed Yassine Haouam [1], Mohamed Amroune [1,2]

[1] Laboratory of Mathematics, Informatics and Systems, Larbi Tebessi University – Tebessa, Algeria
[2] University of Science and Technology Houari Boumediene, Algeria

Corresponding author: Hama Soltani (hama.soltani@univ-tebessa.dz)

Abstract

Breast cancer, a pervasive global health concern, necessitates early detection for an improved prognosis. Mammography, a pivotal screening tool, faces challenges in interpretation, motivating the integration of advanced computational models. This paper offers a comprehensive examination of breast cancer classification through mammography, focusing on machine learning (ML) and deep learning (DL) approaches. The discussion encompasses widely used mammography datasets, preprocessing techniques, data augmentation and diverse classification algorithms. Noteworthy datasets include LAMIS-DMDB, EMBED and INbreast. Preprocessing involves denoising and contrast enhancement, employing techniques such as Wiener filtering and histogram equalization. Data augmentation, a critical factor in handling small datasets, is explored using basic and advanced techniques, including generative adversarial networks. ML algorithms analyse entire mammograms, while DL techniques, notably convolutional neural networks, focus on localized regions of interest. Despite promising strides, challenges persist in obtaining high-quality datasets and ensuring model interpretability, as well as the strong similarities between cancer and non-cancer regions and irrelevant feature extraction. The paper concludes by outlining potential research directions to further transform breast cancer prognosis and treatment.

Keywords

Breast cancer; Lesion classification; Mammography datasets; Mammography classification techniques.

Citation: Soltani, H., Bendib, I., Haouam, M. Y., & Amroune, M. (2024). Advancements in Breast Cancer Diagnosis: A Comprehensive Review of Mammography Datasets, Preprocessing and Classification Techniques. *Acta Informatica Pragensia*, 13(2), 308–326. https://doi.org/10.18267/j.aip.244

Special Issue Editors: Hakim Bendjenna, Larbi Tebessi University – Tebessa, Algeria
Lawrence Chung, University of Texas at Dallas, USA
Abdallah Meraoumia, Larbi Tebessi University – Tebessa, Algeria

Academic Editor: Zdenek Smutny, Prague University of Economics and Business, Czech Republic

1 Introduction

Breast cancer remains one of the most common and serious diseases worldwide, affecting a significant proportion of the female population. According to the World Health Organization (WHO), it is estimated that over 2.3 million women were diagnosed with breast cancer and there were 685,000 deaths globally in 2020 (Arnold et al., 2022). Due to factors such as population growth and ageing, it is anticipated that by the year 2040, the incidence of breast cancer will rise to more than 3 million new cases annually, with the disease expected to cause over 1 million deaths each year (Arnold et al., 2022). Early detection and accurate diagnosis significantly improve prognosis and survival rates, emphasizing the importance of effective screening tools.

Breast cancer originates from abnormal cell proliferation, leading to lesions such as asymmetries, architectural distortion, microcalcifications and masses. Benign abnormalities, usually harmless, are monitored due to potential risks, as they might spread or apply pressure to vital structures. Malignant abnormalities, linked to breast cancer, are confirmed by biopsy when detected in mammograms. Understanding the severity of suspicious abnormalities is crucial for accurate diagnosis and intervention in breast cancer research (Loizidou et al., 2023).

Mammography is a well-established imaging technique for breast cancer screening and has been instrumental in the early detection of breast cancer (Lehman et al., 2017). It uses low-energy X-rays to visualize the internal structure of the breasts and can reveal abnormalities, including tumours, before they become palpable or symptomatic (Tabár et al., 2011). Despite its efficacy, mammographic interpretation is a challenging task due to the subtle and varying appearances of breast lesions, making the role of advanced computational models vital (Giger et al., 2013). The mammographic process involves capturing two distinct images for each breast: cranio-caudal (CC) and medio-lateral oblique (MLO) projections. In CC, the image is taken from above, while MLO captures a side view at an angle, revealing the pectoral muscle. The ratio of non-dense to dense tissue is quantifiable, enabling the assignment of breast density levels. These levels—almost entirely fatty, scattered fibroglandular density, heterogeneously dense and extremely dense—determine the difficulty of image assessment. Higher density poses challenges due to the visual similarity between dense normal and abnormal tissues (Loizidou et al., 2023).

While mammography has been a cornerstone in early breast cancer detection, its reliance on visual interpretation poses inherent challenges, including the subjective nature of assessments and the potential for human error in distinguishing between subtle abnormalities, a factor that underscores the need for advanced computational models in improving diagnostic accuracy.

Machine learning (ML) and particularly deep learning (DL) approaches have made significant contributions to breast cancer diagnosis using mammography, demonstrating promising performance in classifying cancerous and non-cancerous lesions (Burt et al., 2018; Shen et al., 2019). In the realm of ML, various algorithms such as support vector machines (SVM), decision trees, random forests, *k*-nearest neighbours (KNN) have been extensively studied and applied to enhance the accuracy of breast cancer classification of mammography images. SVM, for instance, has been effectively employed in differentiating between benign and malignant breast abnormalities (Bektas et al., 2018; Vidić et al., 2018). DL, a subset of ML characterized by its use of deep neural networks, has had a significant impact in the medical imaging field. DL models, particularly convolutional neural networks (CNN), have shown excellent performance in image-based tasks. CNN have the ability to automatically and adaptively learn spatial hierarchies of features, which has proven to be highly effective for mammogram analysis. The automatic feature learning ability of CNN eliminates the need for manual feature extraction, which is typically required in traditional ML models (Antropova et al., 2017). These models, trained on amounts of mammographic data, can make predictions and decisions based on patterns learned from the data. They reduce the possibility of human error often associated with the visual interpretation of mammographic images.

By improving the precision of lesion classification, these models enhance the overall diagnostic accuracy, facilitating early and effective intervention for patients (Rodríguez-Ruiz et al., 2019; Soltani et al., 2021). However, while the strides in AI-based breast cancer diagnosis are promising, challenges remain. For instance, these models often require extensive, high-quality, annotated datasets for training, which can be resource-intensive to gather. Another critical concern is the interpretability and transparency of these models. Understanding the “black-box” decision-making process of DL models remains a challenge that raises issues of trust and reliability (Holzinger et al., 2019). Despite these challenges, the potential of ML and DL in breast cancer diagnosis is immense, and investing in research and development in this area is very promising.

The paper is organized as follows: Section 2 outlines the methodology employed in this review, detailing the criteria and processes for selecting and evaluating the included studies. Section 3 focuses on mammography datasets, discussing registered knowledge, relevant datasets in the literature and the criteria for dataset selection. Section 4 describes the preprocessing methods and data augmentation approaches used to enhance dataset quality and variability. Section 5 presents the classification techniques utilized in the reviewed studies, examining their effectiveness and applications. Section 6 highlights the current challenges in the field and suggests potential directions for future research aimed at improving breast cancer prognosis and treatment. Finally, Section 7 concludes the paper with a summary of findings and implications for future research and clinical practice.

2 Methodology

Mammography image classification methodologies have been applied to mammogram datasets to enhance their efficacy. This study aims to consolidate the literature on mammogram datasets, preprocessing techniques and classification algorithms utilized in medical imaging applications, particularly in the domain of deep learning-based breast cancer diagnosis. Specifically, our investigation extensively examines mammogram classification, emphasizing two primary categorizations. The first category involves the classification of entire mammogram images based on malignancy, benignity, BI-RADS assessment or breast density evaluation. The second category focuses on the automatic segmentation and classification of specific regions of interest (ROIs) within the mammogram. These ROIs are designated as malignant or benign, or classified according to the type of detected abnormality, such as masses or calcifications.

Our study centres on two main aspects: (1) published mammography datasets and (2) classification techniques. We conducted a systematic search using combinations of keywords such as "mammography dataset", "breast cancer classification", "deep learning", "mammogram preprocessing" and "data augmentation in mammogram" to identify relevant literature. The search was performed across multiple databases, including PubMed, IEEE Xplore, Google Scholar and Scopus. The search period was from November 2023 to January 2024. Additionally, we utilized recent tools such as Elicit and Scispace to make use of their advanced capabilities. Our initial search yielded a total of 220 articles and datasets. Articles that did not address datasets, mammogram image classification or preprocessing were excluded from our analysis. Regarding dataset selection, we exclusively focused on datasets directly sourced from medical imagery centres, disregarding those generated from other datasets or subjected to additional image processing techniques. This filtering process resulted in the exclusion of 155 articles and datasets, leaving us with 65 relevant articles and datasets for our final review. Our research primarily targets datasets and classification algorithms specific to mammogram images, thereby excluding articles pertaining to other imaging modalities such as CT scans, breast MRI, breast ultrasounds or histopathology. Additionally, we ignored datasets generated from other datasets to maintain the authenticity of our findings.

Comprehensive surveys on image classification methods, particularly in the context of medical images, exist in the scientific literature. For instance, Loizidou et al. (2023) proposed various mammogram image techniques to enhance tumour detection and classification accuracy. Another study focused on diagnosis methods in medical image analysis, particularly in breast cancer using histopathology images and deep learning techniques (Kalavathi & Swamy Das, 2023). Our paper aims to provide a comprehensive overview of mammography image analysis, with a specific focus on datasets and classification algorithms. We aim to elucidate both traditional and recent preprocessing and classification approaches to provide readers with valuable insights into this evolving field. Through a thorough examination of articles from diverse sources, including conferences, books and indexed journals, we aim to establish a solid foundation for further research in this domain.

3 Mammography Datasets

The advancement of breast cancer research and diagnosis heavily relies on the availability and quality of mammographic datasets. These datasets, comprising a collection of mammographic images alongside pertinent clinical information, serve as the backbone for developing and validating ML and DL algorithms aimed at detecting and classifying breast abnormalities. In this section, we embark on a comprehensive exploration of various notable mammogram datasets, elucidating their origins, characteristics and accessibility. By delving into the intricacies of each dataset, we aim to provide a nuanced understanding of the diverse resources available to researchers in the field of breast cancer detection and classification.

3.1 Registered knowledge in mammography datasets

Several datasets were identified and are outlined in Table 1. The selection process considered their relevance and the richness of mammographic imaging data provided, offering valuable resources for analysis. These datasets vary in key aspects, providing diverse characteristics. These differences can be classified into the following areas:

- **Year:** The year of creation or publication of the dataset, indicating its temporal context and relevance.
- **Origin:** The country or institution from which the dataset originates, providing insights into its source and potential biases.
- **Number of patients:** The total number of patients included in the dataset, reflecting the population diversity and sample size.
- **Number of samples:** The total count of mammographic images contained within the dataset, delineating the scale and scope of the resource.
- **Type of image:** The format or type of mammographic image utilized in the dataset, such as screen-film mammography (SFM) or full-field digital mammography (FFDM).
- **Image format:** The digital format in which the mammographic images are stored, facilitating interoperability and compatibility with analytical tools.
- **Resolution:** The spatial resolution of the mammographic images, denoting the level of detail captured and retained in the dataset.
- **Annotation/metadata:** The presence or absence of annotations, including ground truth annotations for abnormality localization or metadata annotations for clinical information.
- **Type of abnormality:** The specific types of breast abnormalities annotated or labelled within the dataset, providing insights into the pathological conditions addressed.
- **BI-RADS classification:** The breast imaging reporting and data system (BI-RADS) classification assigned to mammographic findings, aiding in standardized reporting and interpretation.

- **Clinical data availability:** The availability of accompanying clinical data, such as patient demographics, histopathological results and follow-up information, enhances the clinical relevance and utility of the dataset.

3.2 Relevant datasets in the literature

Through an in-depth analysis of the criteria of previous section, we aim to unravel the intricacies of each mammogram dataset, empowering researchers with the knowledge to make informed decisions regarding dataset selection and utilization in their breast cancer detection endeavours. Each dataset offers unique attributes and challenges for further examination, as shown below.

MIAS (Mammographic Image Analysis Society) database: The MIAS database is a smaller public dataset that includes 322 mammogram images from 161 patients. Each image in the MIAS database is also annotated with information about the presence or absence of abnormalities (Maitra et al., 2012).

Digital Database for Screening Mammography (**DDSM**): DDSM is one of the largest public-domain databases. It contains approximately 2620 studies, each of which includes two images of each breast, along with associated patient information and radiologist's annotations (Heath et al., 1998).

CBIS-DDSM (curated breast imaging subset of DDSM): An updated and curated version of the original DDSM, CBIS-DDSM includes a set of images and annotations more suitable for ML techniques. The dataset includes pathology-proven cancers and benign cases (Lee et al., 2017).

Breast Cancer Digital Repository (**BCDR**): This public dataset consists of a variety of images, including mammograms, ultrasound and MRI scans. Each case has associated patient information and clinical history (Ramos-Pollán et al., 2012).

INbreast: The INbreast dataset is a relatively new and robust mammogram database that includes 410 digitized film-screen mammograms from 115 patients. Images are provided with extensive annotations, such as masses, calcifications and architectural distortions (Moreira et al., 2012).

OPTIMAM mammography database (**OMI-DB**) is a comprehensive mammography image repository featuring over 145,000 cases, translating to over 2.4 million images. These images, primarily sourced from the UK's National Health Service Breast Screening Programme, comprise both unprocessed and processed full field digital mammograms (FFDM). The database includes images from various manufacturers, notably Hologic Inc. and General Electric (GE) Medical Systems, denoted as OMI-H and OMI-G, respectively. The dataset categorization utilizes BI-RADS ratings and mass conspicuity measures. Each case provides images from two perspectives of each breast: medio-lateral oblique (MLO) and cranio-caudal (CC), with additional views for cases showcasing suspected abnormalities (Halling-Brown et al., 2021). Agarwal et al. (2020) utilized only the processed FFDM from the database, which were annotated by experts to provide ground-truth data, yielding a total of 2145 cancer cases.

DSTSM-DB (Digital Subtraction of Temporally Sequential Mammograms): This dataset comprises 100 pairs of full-field digital mammograms sourced from various local hospitals from 2012 to 2020. These mammograms, acquired from women aged 38 to 83 years, represent two sequential screening rounds, yielding a total of 400 images. Careful curation was undertaken to ensure a diverse representation, encompassing varying characteristics such as the presence or absence of microcalcifications and different BI-RADS classifications. This deliberate diversity enables the development and assessment of algorithms designed to accurately detect and classify microcalcifications in accordance with the BI-RADS categorization system (Loizidou et al., 2020). In essence, the dataset serves as a valuable resource for researchers and practitioners in the realm of breast imaging. Its utility extends to the facilitation of algorithm development, validation and comparison, with a specific focus on enhancing the detection and classification of microcalcifications. By doing so, the DSTSM-DB dataset plays a pivotal role in advancing breast cancer diagnosis.

VinDr-Mammo is a unique dataset originating from Vietnam that enriches the diversity of publicly accessible mammography data with its extensive lesion-level annotations and breast-level assessments. The dataset comprises 5000 mammography examinations, each incorporating the four standard views and subject to double reading, with any disagreements settled through arbitration. Its primary objective is to evaluate the Breast Imaging Reporting and Data System (BI-RADS) and breast density at the level of the individual breast. Additionally, the dataset offers detailed information, including the classification, location and BI-RADS assessment of findings that are not benign (Nguyen et al., 2023).

EMBED (Emory breast imaging dataset) comprises 3.4 million mammographic images for both screening and diagnostic purposes. This extensive dataset includes two-dimensional and digital breast tomosynthesis mammograms from 116,000 women, ensuring a balanced representation of African American and white patients. It also features 40,000 annotated lesions, each linked to structured imaging descriptors, along with 56 ground-truth pathologic outcomes categorized into seven severity classes. Notably, interpreting radiologists annotated the regions of interest (ROIs), directly linking them to a single finding in approximately 80% of cases. The remaining 20% involved cases with multiple findings, necessitating manual linkage. The primary objective behind creating this dataset is to facilitate the development and validation of deep learning models for breast cancer screening, ensure equitable performance across various patient demographics and contribute to the reduction of healthcare disparities (Jeong et al., 2023).

CMMD (the Chinese Mammography Database)**:** A repository of mammography data, presenting researchers with invaluable insights into breast pathology. Comprising two distinct datasets, CMMD1 and CMMD2, this resource offers a nuanced exploration of breast pathologies and molecular subtypes. CMMD1, encompassing 1026 meticulously curated cases corresponding to 2214 mammographies, delineates biopsy confirmed instances of benign and malignant tumours. Conversely, CMMD2, comprising 1498 mammographies from 749 patients, delves into cases characterized by known molecular subtypes, enriching the dataset with comprehensive molecular profiling. In total, CMMD presents 3712 mammography images sourced from 1775 patients, facilitating a multifaceted examination of breast cancer etiology and classification. Acquired using state-of-the-art imaging systems—including the GE Senographe DS and Siemens Mammomat Inspiration—the CMMD dataset originates from esteemed medical institutions, including the Sun Yat-sen University Cancer Centre in Guangzhou and the Nanhai Affiliated Hospital of Southern Medical University in Fushan, China (H. Cai et al., 2023).

LAMIS-DMDB stands as a structured repository containing 2216 FFDM mammograms, offering an ample dataset for the advancement of computer-aided detection (CAD) systems in mammography. Comprising diverse findings, including 914 normal mammograms, 802 with masses or nodules, 196 with calcifications, 4 displaying architectural distortion and 300 images with multiple findings, this database provides comprehensive information on both benign and malignant cases. Collected over five years from two Algerian sources, its uniqueness lies in its exclusive focus on patients of Algerian origin residing in Algeria, making it the sole dataset of its kind on the African continent. This distinctiveness not only introduces balance and diversity to the available data but also paves the way for exploration into various research domains. In essence, the LAMIS-DMDB emerges as a valuable resource for the development and evaluation of AI algorithms in breast cancer detection and diagnosis, addressing the critical need for standardized public images in medical imaging research (Imane et al., 2024).

Table 1. *Available mammogram datasets and their characteristics.*

Dataset	Year	Origin	Number of patients	Number of images	Type of image	Image format	Resolution	Annotation / metadata	Type of abnormality	BI-RADS classification	Clinical Data	Availability
MIAS	1994	UK	161	322	SFM	PGM	1024*1024	Text file	Various	/	/	Public
DDSM	1996	USA	2620	Over 10,000	SFM	LJPEG	4084*3328	Ground truth (PNG) CSV file (metadata)	Mass, calcifications	/	✓	Public
BCDR	2008	Portugal	1010	3703	SFM FFDM	DICOM	1167*720	XML files	Various	✓	✓	Public
INBREAST	2012	Portugal	115	410	FFDM	DICOM	4084*3328 3328*2560	XML files, Ground truth (PNG) CSV file (metadata)	Various	✓	✓	Public
CBIS-DDSM	2016	USA	2620	Over 10,000	SFM	DICOM	4084*3328	Ground truth (PNG) CSV file (metadata)	Mass, calcifications	/	✓	Public
OMI-DB	2021	UK	172,282	3,072,878	FFDM	DICOM	Not specified	Private database	Various	Available only for processed set	✓	Private
DSTSM-DB	2021	Cyprus	100	400	FFDM	DICOM	4084*3328	Ground truth (PNG) CSV file (metadata)	Micro-calcifications	✓	✓	Public
VinDr-Mammo	2022	Vietnam	5000	20,000	FFDM	DICOM	3518*2800	CSV files (metadata, BBox)	Various	✓	✓	Upon request
EMBED	2023	USA	115,910	Total: 3.4M images (+58% mammograms)	FFDM	DICOM; PNG	Not specified	Ground truth & metadata	Various	✓	✓	Private (only 20% upon request)
CMMD	2023	China	1775	3712	FFDM	DICOM	2294*1914	CSV files (metadata)	Mass, calcifications	/	✓	Public
LAMIS-DMDB	2024	Algeria	588	2216	FFDM	DICOM	4084*3328	CSV files (metadata)	Various	✓	✓	Upon request

3.3 Dataset selection criteria and discussion

Several significant differences emerge among the available datasets, encompassing factors such as data volume (number of patients and images), image quality (resolution), annotation, imaging type and availability conditions. These variations contribute to a landscape of poor interoperability, hindering the development of clinically validated and generalizable tools for mammographic imaging. A key distinction among the datasets lies in their clinical use cases, often aligning with specific research objectives. For example, while the MIAS dataset may be suited for background tissue assessment, it may not meet the requirements of researchers focusing on lesion detection. Conversely, the INbreast dataset could be invaluable for research teams developing comprehensive mammography triage systems targeting binary outcomes such as malignancy detection. Furthermore, differentiation between film and digital scans plays a crucial role in dataset selection. Despite the predominance of digital mammography in developed regions, the inclusion of film mammograms remains crucial, especially for commercial enterprises seeking to expand mammographic screening to developing economies, where film scanners are still prevalent. Interoperability and standardization pose additional challenges, with datasets often adopting diverse labelling schemes and image formats. Standardization approaches, such as the use of a single lossless image format such as DICOM, could enhance interoperability among datasets and facilitate data sharing and analysis. Data volume also warrants consideration, as larger datasets such as OMI-DB offer greater utility for machine learning research, enabling the development of more robust and accurate models. However, smaller datasets such as INbreast may still hold value, especially when augmented with synthetic data generation techniques. It is also noteworthy that there is a lack of datasets of African origin (for women living in Africa), with the exception of the newly introduced LAMIS-DMDB dataset. This absence raises questions about the potential impact of tools developed using these datasets on women in Africa, where breast cancer incidence rates are significant. The introduction of the LAMIS-DMDB dataset from Africa not only addresses this gap but also holds promise in augmenting the effectiveness of mammographic imaging tools by providing data representative of the unique demographic and environmental factors prevalent in the region. In conclusion, addressing these disparities and utilizing the strengths of diversified datasets are crucial steps to advance the field of mammographic imaging. By fostering collaboration, standardization and data augmentation strategies, researchers can overcome existing barriers and accelerate the development of innovative solutions for breast cancer detection and diagnosis.

4 Data Preparation

The effectiveness of breast cancer diagnosis and classification relies heavily on the quality and diversity of the datasets used to train computational models. In this section, we delve into the intricate process of dataset preparation, where preprocessing techniques and data augmentation emerge as pivotal phases in refining and enriching the dataset landscape.

4.1 Preprocessing techniques

Image preprocessing is a critical step in breast cancer classification using mammograms. It helps improve the quality of the images and remove any noise or artifacts that could interfere with the classification process. Several preprocessing techniques frequently employed for breast cancer classification encompass: image denoising and contrast enhancement.

Image denoising refers to the process of removing noise from images. In the context of mammograms, noise can interfere with the ability to clearly see and interpret the images, which can potentially lead to difficulties in detecting abnormalities such as tumours. Therefore, denoising is an important step in the processing of mammogram images. It is typically used in the case of scanned mammograms. There are several techniques used for denoising mammogram images, including Wiener filtering, Gaussian filtering, median filtering, adaptive median filtering and hybrid median filtering (Hamed et al., 2018; Joseph et al., 2017).

The other crucial technique is contrast enhancement, which aims to improve the visibility of details in an image by increasing the difference in intensity between lighter and darker areas. This can help in the detection of abnormalities such as tumours or microcalcifications, which are often indicative of breast cancer. One of the techniques commonly used for contrast enhancement in mammograms is histogram equalization (HE). The main idea of HE is to expand the most frequent intensity values in an image to heighten the overall contrast. Adaptive HE (AHE), an enhancement over the HE, enhances contrast and edges in each image region by calculating multiple histograms. Contrast limited AHE (CLAHE), an improvement over AHE, limits noise amplification by clipping the histogram at a designated value (Tripathy & Swarnkar, 2020).

4.2 Data augmentation

DL approaches in the medical imaging field have become increasingly important, but their success is heavily reliant on the size of the annotated training set, which is a time-consuming task for expert radiologists. Accessible biomedical image datasets are often small and obtaining large datasets is challenging due to privacy and legal concerns. Consequently, many supervised DL models are at risk of overfitting and struggle to produce generalizable results. To tackle this issue, data augmentation has gained popularity. It involves applying various transformations to augment the training set and improve model performance on new data. Various data augmentation techniques are applied to mammogram images to enhance the performance and robustness of DL models (Oza et al., 2022). These augmentation techniques can be broadly classified into two main groups: basic and advanced techniques (Oza et al., 2022). The basic techniques involve simple and commonly used transformations to create variations in the dataset such as geometric techniques (flipping, rotation, translate, etc...) and pixel-level techniques. On the other hand, advanced augmentation techniques are more complex and often involve more sophisticated transformations to introduce further diversity into the dataset. The dominant techniques used in this area are generative adversarial networks (GAN) (Jeong et al., 2022) and neural style transfer (NST) (Wang et al., 2020). Recently, GAN have demonstrated their efficacy in addressing the limitation of labelled data, known for their efficiency with fewer training data. GAN are applied in breast density estimation, high-resolution mammogram synthesis, effective tumour segmentation, shape analysis, feature extraction and image augmentation for mammogram classification (Gopal et al., 2020; Oyelade et al., 2022). Figure 1 represents the augmentation techniques used in the domain of breast cancer classification and diagnosis.

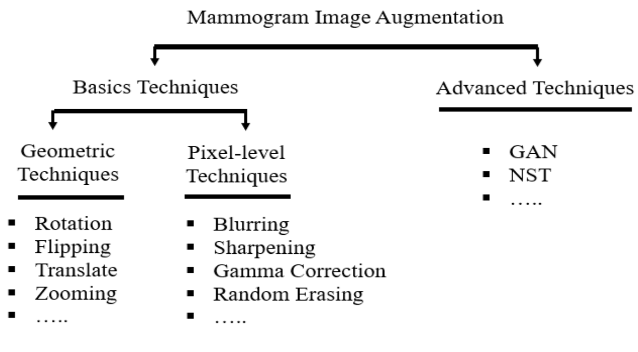

Figure 1. *Augmentation techniques in breast cancer classification and diagnosis.*

5 Classification Algorithms

The classification of mammograms for breast cancer diagnosis is a pivotal focus within the realm of medical imaging research, where various groups of researchers have adopted distinct strategies (Loizidou et al., 2023). One group pursues a comprehensive approach involving the classification of the entire mammogram image. This method focuses on various aspects, including the classification of abnormalities (normal/ abnormal, mass/ calcification), density degree and BI-RADS assessment, offering a holistic evaluation of breast tissue to identify potential abnormalities. In contrast, another group concentrates on the classification of specific regions of interest (ROIs) within the mammogram post-detection and segmentation. This targeted approach aims to differentiate between benign and malignant regions, focusing on localized areas identified as potential indicators of breast cancer. These divergent strategies, embraced by different groups of researchers, contribute to the ongoing evolution of mammogram classification, each providing unique insights and opportunities to refine breast cancer diagnostic capabilities. These distinct strategies adopted by different groups of researchers often involve the utilization of various ML and DL methods (Hassan et al., 2022; Nasser & Yusof, 2023).

5.1 ML-based methods

The application of machine learning (ML) in mammogram classification represents a pivotal stride in enhancing breast cancer diagnostics over the previous decades. ML algorithms, often employed by researchers adopting a comprehensive approach, play a crucial role in analysing entire mammogram images. These algorithms excel at recognizing intricate patterns and variations in mammogram image abnormalities. Through feature extraction and pattern recognition, ML models can classify mammograms based on characteristics such as shape, margin and intensity (Bektas et al., 2018; Ramos-Pollán et al., 2012).

Various studies have employed different machine learning techniques such as SVM classifier, ANN, KNN, fuzzy C-means and CNN to analyse mammography images for classification of breast cancer. The accuracy rates ranged from 86.96% to 99.67% in identifying benign and malignant tumours (Jalloul et al., 2023).

For instance, Vijayarajeswari et al. (2019) applied the Hough transform for mammogram classification, specifically focusing on microcalcifications and masses. The method employed both the Hough transform and SVM for feature extraction and classification, resulting in heightened accuracy. The effectiveness of the approach was confirmed through testing on a test dataset comprising 95 mammograms.

Ragab et al. (2021) introduced a computer-aided diagnosis (CAD) system for classifying breast cancer lesions in mammograms. The system conducted four different experiments to determine the optimal approach. The first experiment utilized end-to-end pre-trained fine-tuned deep CNN (DCNN) networks. The second experiment extracted deep features from DCNN, feeding them into a SVM classifier. The third experiment explored the fusion of deep features, demonstrating enhanced accuracy in SVM classifiers. Finally, the fourth experiment incorporated principal component analysis (PCA) to reduce the feature vector size and computational costs. Evaluations on the CBIS-DDSM and MIAS datasets showcased the system versatility, providing valuable insights into effective feature extraction and classification methods, ultimately aiding radiologists in breast cancer lesion diagnosis (Ragab et al., 2021). Table 2 presents a summary of ML algorithms and methods specifically designed for the classification of calcifications and masses in mammograms.

5.2 DL-based methods

Also, DL methods have emerged as a cutting-edge approach in mammogram classification, particularly in scenarios where researchers focus on classifying ROIs. DL techniques, such as CNN, demonstrate exceptional capabilities in extracting hierarchical features from complex mammogram images. The deep architecture of these models allows nuanced differentiation between benign and malignant regions within localized areas (Nasser & Yusof, 2023). The utilization of DL methodologies thus presents a promising avenue for refining the precision of breast cancer diagnosis, making use of the ability of neural networks to discern subtle details and patterns in mammographic data (Debelee et al., 2020; Zhu et al., 2023).

Table 2. *Summary of ML algorithms and methods designed for classification of calcifications and masses in mammograms.*

Paper	Dataset	Image type	No. of samples	Feature extraction	Classifier	Classes	Result (%)
Alam et al. (2019)	DDSM; OPTIMAM; MIAS	Scanned; FFDM	590	Based on shape, intensity, texture	Stack generalization	Malign / Benign	Accuracy (97.00)
George et al. (2019)	DDSM; OPTIMAM; MIAS	Scanned; FFDM	595	Multiscale connected chain	KNN	Malign / Benign	Accuracy (86.00)
Fanizzi et al. (2020)	BCDR	FFDM	260	Based on texture	Random forest	Malign / Benign	Accuracy (92.00)
Chen et al. (2020)	Private Data	FFDM	4160	CNN	Cascade adaboost	Malign / Benign	Accuracy (83.40)
H. Cai et al. (2019)	Private Data	FFDM	990	Handcrafted/ CNN	SVM	Malign / Benign	Accuracy (90. 20)
Danala et al. (2018)	Private data	Scanned	222	Based on shape and density	MLP	Malign / Benign	Accuracy (78.40)
Rahmani Seryasat & Haddadnia (2018)	DDSM	Scanned	295	Based on shape, border, tissue, fractal dimension	Ensembled	Malign / Benign	Accuracy (78.40)
Kermouni Serradj et al. (2022)	INbreast	FFDM	308	GLCM	SVM	Normal / Abnormal	Precision (98.20)
Loizidou et al. (2020)	Private Data	FFDM	320	Based on intensity and shape; GLCM	SVM EDT	Benign / Suspicious; Normal/Abnormal	Accuracy (99.55) Accuracy (91.02)
de Brito Silva et al. (2020)	DDSM	Scanned	794	Mass geometry, topology and shape	SVM	Malign / Benign	Accuracy (90.18)

Jabeen et al. (2023) introduced an automated framework for breast cancer classification. They enhanced image contrast through a novel haze-reduced local-global technique and employed dataset augmentation to amplify diversity, thereby enhancing the training capacity of the selected deep learning model, EfficientNet-b0. Deep features underwent extraction and fusion using a serial-based approach, further optimized by the equilibrium-Jaya controlled Regula Falsi feature selection algorithm. Notably, experimental results on the CBIS-DDSM and INbreast datasets showcased impressive accuracies of 95.4% and 99.7%.

Mahmood et al. (2022) introduced an approach based on CNN to significantly minimize human error in diagnosing malignant breast tissues. The approach involved preprocessing steps, synthetic data augmentation and transfer learning. Also, the validation included 322 raw mammogram images from MIAS and 580 from private datasets. The proposed approach achieved outstanding training accuracy (0.98), test accuracy (0.97), high sensitivity (0.99) and an AUC of 0.99 in classifying breast masses.

Soltani et al. (2023) introduced a novel approach by employing pre-trained CNN architectures, namely AlexNet, VGGNet, DenseNet and ResNet, to extract features from mammography patch images. This transfer learning strategy allows the model to utilize knowledge acquired during the solution of one problem for the classification of another. By utilizing these well-established architectures, the framework captured meaningful hierarchical features, which were then fed into a fully connected layer for the classification of malignant and benign cells.

Oza et al. (2023) introduced a pioneering deep-learning model for breast cancer classification. The model employed transfer learning with pretrained CNN architectures, including VGG-16, ResNet-50, Inception-V3 and Efficientnet-B7. The study demonstrated exceptional performance, particularly with a novel transfer-learning approach incorporating test time augmentation. Notably, the proposed method achieved a remarkable 99.97% accuracy on the MIAS dataset and 99.8% on CBIS-DDSM.

This innovative application of transfer learning contributes to the growing body of research aimed at improving the accuracy and efficiency of DL models in medical image analysis, particularly in the context of breast cancer diagnosis. The precise classification of mammograms using ML and DL models yields significant advantages, including the mitigation of annotation workload, enhanced utilization of contextual information, reduced call-back rates and diminished unnecessary tests, all while upholding a high level of sensitivity.

However, research into malignancy diagnosis encounters distinctive challenges. The scarcity of available breast images presents a notable obstacle, necessitating laborious and costly procedures for the acquisition of images from specific medical imaging centres to train and validate breast mass classification techniques. Moreover, imbalances within training datasets are prevalent, resulting in compromised model performance, particularly when dealing with smaller datasets. Additionally, Table 3 offers a comprehensive overview of DL algorithms and methods tailored for the classification of masses and calcifications in mammograms.

Table 3. *Summary of DL algorithms and methods designed for classification of calcifications and masses in mammograms.*

Paper	Dataset	Image Type	No. of samples	Feature Extraction	Classifier	Classes	Result (%)
Ribli et al. (2018)	Private data; INbreast; DDSM	FFDM; scanned	3222	/	Faster-RCNN	Malign / Benign/ Normal	AUC = 0.95
Rehman et al. (2021)	Private data; DDSM	FFDM; scanned	4145	/	FC-DSCNN	Malign / Benign	Accuracy (90.00)
Leong et al. (2022)	CIBS-DDSM	Scanned	1654	/	ResNet50	Malign / Benign	Accuracy (97.58)
Al-Masni et al. (2018)	DDSM	Scanned	600	/	FC-NN	Malign / Benign	Accuracy (97.00)
G. Cai et al. (2020)	Private data; INbreast	FFDM	722	/	CNN	Malign / Benign	Accuracy (93.70)
Gnanasekaran et al. (2020)	DDSM; MIAS	Scanned	1940	/	CNN	Malign /Benign /Normal	Accuracy (98.32)
Al-Antari et al. (2020)	DDSM; INbreast	Scanned; FFDM	1010	/	InceptionResNet-V2	Malign / Benign	Accuracy (97.50)
Zahoor et al. (2022)	MIAS; CBIS-DDSM; INbreast,	Scanned; FFDM	2014	NasNetMobile; MobileNetV2	MEWOA+CNN	Malign / Benign	Accuracy (99.70)
Baccouche et al. (2022)	CBIS-DDSM; private data; INbreast	Scanned; FFDM	1963	/	ResNet familly	Malign / Benign; BI-RADS Category; Shape Category	Accuracy (99.20); Accuracy (96.80); Accuracy (90.02)
Soltani et al. (2024)	INbreast	FFDM	108	/	Mask-RCNN+ DenseNet121	Malign / Benign/Normal	Accuracy (99.44)
Yu et al. (2024)	DDSM	Scanned	3500	VGG19-DF	DF-dRVFL	Malign / Benign	Accuracy (81.71)

6 Challenges and Potential Research Directions

6.1 Challenges

In the domain of breast cancer classification and diagnosis, current methodologies grapple with inherent complexities and challenges. Some examples follow.

- Breast cancer heterogeneity: Breast cancer is a heterogeneous disease at the molecular level, with diverse patterns of gene expression responsible for differences in tumour behaviour and prognosis (Zhang, 2023). This heterogeneity makes it difficult to develop a single diagnostic or treatment approach that works for all patients.
- Scarcity of available breast images: One of the challenges is the scarcity of data for training deep learning models, which can affect the performance and generalization of the models (Zakareya et al., 2023; Zhang, 2023).
- Imbalances within training datasets: Imbalances in the distribution of classes within training datasets can lead to biased models and affect the accuracy of classification (Zhang, 2023).
- Ambiguous imaging data: The meticulous classification of different breast abnormalities is challenging due to ambiguous imaging data and the indistinguishable appearance of some breast abnormalities from normal tissue (Tariq et al., 2021).

In summary, the elucidated challenges within breast cancer classification and diagnosis emphasize the critical necessity for innovative solutions. Addressing these obstacles is paramount for the progression of diagnostic precision.

6.2 Potential research directions

Recent studies have shown that deep learning-based breast cancer diagnosis has significant potential for future developments. Some potential future developments in deep learning-based breast cancer diagnosis are:

- Multi-modal deep learning: Multi-modal deep learning techniques that combine different types of medical images, such as mammography, ultrasound and magnetic resonance imaging, can improve the accuracy of breast cancer diagnosis (Gonzales Martinez & van Dongen, 2023; Sahu et al., 2023).
- Transfer learning: Transfer learning, which involves using pre-trained deep learning models for breast cancer diagnosis, can reduce the need for extensive manual annotation and feature extraction, leading to more efficient and accurate diagnosis (Nasser & Yusof, 2023).
- Explainable deep learning: Explainable deep learning methods can provide insights into the decision-making process of deep learning models, improving the interpretability and transparency of breast cancer diagnosis (Zakareya et al., 2023).
- Generative adversarial networks (GAN): GAN can be used to generate synthetic medical images, which can be used to augment the training datasets for deep learning models, addressing the issue of data scarcity (Nasser & Yusof, 2023).
- Integration of federate learning (FL): FL is a machine learning approach where an algorithm is trained across multiple decentralized edge devices or servers holding local data samples, without exchanging them. This approach can be particularly useful in healthcare, where privacy concerns and regulations often limit the ability to share patient data. FL allows data privacy, as raw patient data do not need to be shared between hospitals. Additionally, it enables improved diagnosis by learning from diverse datasets across different hospitals, potentially increasing predictive accuracy (Dhade & Shirke, 2024; Prayitno et al., 2021).

In conclusion, future developments in deep learning-based breast cancer diagnosis include multi-modal deep learning, transfer learning, GAN, FL and explainable deep learning. These developments have the potential to improve the accuracy, efficiency, security and interpretability of breast cancer diagnosis, ultimately contributing to better patient outcomes.

7 Conclusion

Our research provides a comprehensive exploration of mammography datasets, preprocessing and classification techniques crucial for breast cancer diagnosis. Recognizing breast cancer as a significant healthcare challenge, we emphasized the transformative role of ML and DL in diagnostics. We meticulously curated and analysed various mammogram datasets, including INbreast, LAMIS-DMDB, VinDR-Mammo and OMI-DB, offering insights for informed dataset selection. Our study also addressed data preparation, highlighting preprocessing methodologies and data augmentation techniques to enhance model robustness. The classification algorithm section compared diverse machine learning and deep learning methods, aiming to improve diagnostic accuracy and clinical efficacy. However, challenges such as dataset heterogeneity and privacy constraints present obstacles. We also discussed the complexities of federated learning and the need for innovations in communication protocols and data security.

Looking ahead, our study suggests future research directions, including harmonizing heterogeneous datasets to improve model generalizability and advancing secure multi-party computation to enhance federated learning. Ultimately, our study aims to drive advancements in breast cancer diagnosis and classification, using AI to improve healthcare delivery and patient outcomes globally.

Additional Information and Declarations

Acknowledgments: This work would not have been possible without financial support of Laboratory of Mathematics, Informatics and Systems, Larbi Tebessi University – Tebessa.

Conflict of Interests: The authors declare no conflict of interest.

Author Contributions: H.S.: Conceptualization, Methodology, Writing – Original draft. I.B.: Data curation, Writing – Review & editing. M.Y.H.: Data curation, Writing – Review & editing. M.A.: Supervision.

References

Agarwal, R., Díaz, O., Yap, M. H., Lladó, X., & Martí, R. (2020). Deep learning for mass detection in Full Field Digital Mammograms. *Computers in Biology and Medicine*, 121, Article 103774. https://doi.org/10.1016/j.compbiomed.2020.103774

Alam, N., Denton, E. R. E., & Zwiggelaar, R. (2019). Classification of Microcalcification Clusters in Digital Mammograms Using a Stack Generalization Based Classifier. *Journal of Imaging*, 5(9), 76. https://doi.org/10.3390/jimaging5090076

Al-antari, M. A., Han, S.-M., & Kim, T.-S. (2020). Evaluation of deep learning detection and classification towards computer-aided diagnosis of breast lesions in digital X-ray mammograms. *Computer Methods and Programs in Biomedicine*, 196, 105584. https://doi.org/10.1016/j.cmpb.2020.105584

Al-masni, M. A., Al-antari, M. A., Park, J.-M., Gi, G., Kim, T.-Y., Rivera, P., Valarezo, E., Choi, M.-T., Han, S.-M., & Kim, T.-S. (2018). Simultaneous detection and classification of breast masses in digital mammograms via a deep learning YOLO-based CAD system. *Computer Methods and Programs in Biomedicine*, 157, 85–94. https://doi.org/10.1016/j.cmpb.2018.01.017

Antropova, N., Huynh, B. Q., & Giger, M. L. (2017). A deep feature fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets. *Medical Physics*, 44(10), 5162–5171. https://doi.org/10.1002/mp.12453

Arnold, M., Morgan, E., Rumgay, H., Mafra, A., Singh, D., Laversanne, M., Vignat, J., Gralow, J. R., Cardoso, F., Siesling, S., & Soerjomataram, I. (2022). Current and future burden of breast cancer: Global statistics for 2020 and 2040. *Breast*, 66, 15–23. https://doi.org/10.1016/j.breast.2022.08.010

Baccouche, A., Garcia-Zapirain, B., & Elmaghraby, A. S. (2022). An integrated framework for breast mass classification and diagnosis using stacked ensemble of residual neural networks. *Scientific Reports*, 12(1), 12259. https://doi.org/10.1038/s41598-022-15632-6

Bektas, B., Emre, I. E., Kartal, E., & Gulsecen, S. (2018). Classification of Mammography Images by Machine Learning Techniques. In *2018 3rd International Conference on Computer Science and Engineering (UBMK),* (pp. 580–585). IEEE. https://doi.org/10.1109/UBMK.2018.8566380

Burt, J. R., Torosdagli, N., Khosravan, N., RaviPrakash, H., Mortazi, A., Tissavirasingham, F., Hussein, S., & Bagci, U. (2018). Deep learning beyond cats and dogs: recent advances in diagnosing breast cancer with deep neural networks. *British Journal of Radiology*, 91, 20170545. https://doi.org/10.1259/bjr.20170545

Cai, G., Guo, Y., Chen, W., Zeng, H., Zhou, Y., & Lu, Y. (2020). Computer-aided detection and diagnosis of microcalcification clusters on full field digital mammograms based on deep learning method using neutrosophic boosting. *Multimedia Tools and Applications*, 79(23–24), 17147–17167. https://doi.org/10.1007/s11042-019-7726-x

Cai, H., Huang, Q., Rong, W., Song, Y., Li, J., Wang, J., Chen, J., & Li, L. (2019). Breast Microcalcification Diagnosis Using Deep Convolutional Neural Network from Digital Mammograms. *Computational and Mathematical Methods in Medicine*, 2019, 1–10. https://doi.org/10.1155/2019/2717454

Cai, H., Wang, J., Dan, T., Li, J., Fan, Z., Yi, W., Cui, C., Jiang, X., & Li, L. (2023). An Online Mammography Database with Biopsy Confirmed Types. *Scientific Data*, 10(1), 123. https://doi.org/10.1038/s41597-023-02025-1

Chen, K.-C., Chin, C.-L., Chung, N.-C., & Hsu, C.-L. (2020). Combining Multi-classifier with CNN in Detection and Classification of Breast Calcification. In *Future Trends in Biomedical and Health Informatics and Cybersecurity in Medical Devices*, (pp. 304–311). https://doi.org/10.1007/978-3-030-30636-6_42

Danala, G., Aghaei, F., Heidari, M., Wu, T., Patel, B., & Zheng, B. (2018). Computer-aided classification of breast masses using contrast-enhanced digital mammograms. In K. Mori & N. Petrick (Eds.), *Medical Imaging 2018: Computer-Aided Diagnosis*, (p. 91). SPIE. https://doi.org/10.1117/12.2293136

de Brito Silva, T. F., de Paiva, A. C., Silva, A. C., Braz Júnior, G., & de Almeida, J. D. S. (2020). Classification of breast masses in mammograms using geometric and topological feature maps and shape distribution. *Research on Biomedical Engineering*, 36(3), 225–235. https://doi.org/10.1007/s42600-020-00063-x

Debelee, T. G., Schwenker, F., Ibenthal, A., & Yohannes, D. (2020). Survey of deep learning in breast cancer image analysis. *Evolving Systems*, 11(1), 143–163. https://doi.org/10.1007/s12530-019-09297-2

Dhade, P., & Shirke, P. (2024). Federated Learning for Healthcare: A Comprehensive Review. *Engineering Proceedings*, 59(1), 230. https://doi.org/10.3390/engproc2023059230

Fanizzi, A., Basile, T. M. A., Losurdo, L., Bellotti, R., Bottigli, U., Dentamaro, R., Didonna, V., Fausto, A., Massafra, R., Moschetta, M., Popescu, O., Tamborra, P., Tangaro, S., & La Forgia, D. (2020). A machine learning approach on multiscale texture analysis for breast microcalcification diagnosis. *BMC Bioinformatics*, 21(S2), 91. https://doi.org/10.1186/s12859-020-3358-4

George, M., Chen, Z., & Zwiggelaar, R. (2019). Multiscale connected chain topological modelling for microcalcification classification. *Computers in Biology and Medicine*, 114, 103422. https://doi.org/10.1016/j.compbiomed.2019.103422

Giger, M. L., Karssemeijer, N., & Schnabel, J. A. (2013). Breast image analysis for risk assessment, detection, diagnosis, and treatment of cancer. *Annual Review of Biomedical Engineering*, 15, 327–357. https://doi.org/10.1146/annurev-bioeng-071812-152416

Gnanasekaran, V. S., Joypaul, S., Meenakshi Sundaram, P., & Chairman, D. D. (2020). Deep learning algorithm for breast masses classification in mammograms. *IET Image Processing*, 14(12), 2860–2868. https://doi.org/10.1049/iet-ipr.2020.0070

Gonzales Martinez, R., & van Dongen, D.-M. (2023). Deep learning algorithms for the early detection of breast cancer: A comparative study with traditional machine learning. *Informatics in Medicine Unlocked*, 41, 101317. https://doi.org/10.1016/j.imu.2023.101317

Gopal, A., Gandhimaruthian, L., & Ali, J. (2020). Role of General Adversarial Networks in Mammogram Analysis: A Review. *Current Medical Imaging*, 16(7), 863–877. https://doi.org/10.2174/1573405614666191115102318

Halling-Brown, M. D., Warren, L. M., Ward, D., Lewis, E., Mackenzie, A., Wallis, M. G., Wilkinson, L. S., Given-Wilson, R. M., McAvinchey, R., & Young, K. C. (2021). OPTIMAM mammography image database: A large-scale resource of mammography images and clinical data. *Radiology: Artificial Intelligence*, 3(1). https://doi.org/10.1148/ryai.2020200103

Hamed, G., Marey, M., Amin, S. E.-S., & Tolba, M. F. (2018). A Proposed Model for Denoising Breast Mammogram Images. In *2018 13th International Conference on Computer Engineering and Systems (ICCES),* (pp. 652–657). IEEE. https://doi.org/10.1109/ICCES.2018.8639307

Hassan, N. M., Hamad, S., & Mahar, K. (2022). Mammogram breast cancer CAD systems for mass detection and classification: a review. *Multimedia Tools and Applications*, 81(14), 20043–20075. https://doi.org/10.1007/s11042-022-12332-1

Heath, M., Bowyer, K., Kopans, D., Moore, R., Kegelmeyer, P., Moore, R., Chang, K. & Munishkumaran, S. (1998). Current Status of the Digital Database for Screening Mammography. *Digital Mammography,* (pp. 457–460). Springer. https://doi.org/10.1007/978-94-011-5318-8_75

Holzinger, A., Langs, G., Denk, H., Zatloukal, K., & Müller, H. (2019). Causability and explainability of artificial intelligence in medicine. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 9(4), e1312. https://doi.org/10.1002/widm.1312

Imane, O., Mohamed, A., Lazhar, R. F., Hama, S., Elhadj, B., & Conci, A. (2024). LAMIS-DMDB: A new full field digital mammography database for breast cancer AI-CAD researches. *Biomedical Signal Processing and Control*, 90, 105823. https://doi.org/10.1016/j.bspc.2023.105823

Jabeen, K., Khan, M. A., Balili, J., Alhaisoni, M., Almujally, N. A., Alrashidi, H., Tariq, U., & Cha, J. H. (2023). BC2NetRF: Breast Cancer Classification from Mammogram Images Using Enhanced Deep Learning Features and Equilibrium-Jaya Controlled Regula Falsi-Based Features Selection. *Diagnostics*, 13(7). https://doi.org/10.3390/diagnostics13071238

Jalloul, R., Chethan, H. K., & Alkhatib, R. (2023). A Review of Machine Learning Techniques for the Classification and Detection of Breast Cancer from Medical Images. *Diagnostics*, 13(14), 2460. https://doi.org/10.3390/diagnostics13142460

Jeong, J. J., Tariq, A., Adejumo, T., Trivedi, H., Gichoya, J. W., & Banerjee, I. (2022). Systematic Review of Generative Adversarial Networks (GANs) for Medical Image Classification and Segmentation. *Journal of Digital Imaging*, 35(2), 137–152. https://doi.org/10.1007/s10278-021-00556-w

Jeong, J. J., Vey, B. L., Bhimireddy, A., Kim, T., Santos, T., Correa, R., Dutt, R., Mosunjac, M., Oprea-Ilies, G., Smith, G., Woo, M., McAdams, C. R., Newell, M. S., Banerjee, I., Gichoya, J., & Trivedi, H. (2023). The EMory BrEast imaging Dataset (EMBED): A Racially Diverse, Granular Dataset of 3.4 Million Screening and Diagnostic Mammographic Images. *Radiology: Artificial Intelligence*, 5(1). https://doi.org/10.1148/ryai.220047

Joseph, A. M., John, M. G., & Dhas, A. S. (2017). Mammogram image denoising filters: A comparative study. In *2017 Conference on Emerging Devices and Smart Systems (ICEDSS),* (pp. 184–189). IEEE. https://doi.org/10.1109/ICEDSS.2017.8073679

Kalavathi, R., & Swamy Das, M. (2023). A Review on Deep Learning Approaches for Histopathology Breast Cancer Classification. In *Proceedings of Fourth International Conference on Computer and Communication Technologies,* (pp. 367–375). Springer. https://doi.org/10.1007/978-981-19-8563-8_35

Kermouni Serradj, N., Messadi, M., & Lazzouni, S. (2022). Classification of Mammographic ROI for Microcalcification Detection Using Multifractal Approach. *Journal of Digital Imaging*, 35(6), 1544–1559. https://doi.org/10.1007/s10278-022-00677-w

Lee, R. S., Gimenez, F., Hoogi, A., Miyake, K. K., Gorovoy, M., & Rubin, D. L. (2017). A curated mammography data set for use in computer-aided detection and diagnosis research. *Scientific Data*, 4(1), 170177. https://doi.org/10.1038/sdata.2017.177

Lehman, C. D., Arao, R. F., Sprague, B. L., Lee, J. M., Buist, D. S. M., Kerlikowske, K., Henderson, L. M., Onega, T., Tosteson, A. N. A., Rauscher, G. H., & Miglioretti, D. L. (2017). National performance benchmarks for modern screening digital mammography: Update from the Breast Cancer Surveillance Consortium. *Radiology*, 283(1), 49–58. https://doi.org/10.1148/radiol.2016161174

Leong, Y. S., Hasikin, K., Lai, K. W., Mohd Zain, N., & Azizan, M. M. (2022). Microcalcification Discrimination in Mammography Using Deep Convolutional Neural Network: Towards Rapid and Early Breast Cancer Diagnosis. *Frontiers in Public Health*, 10. https://doi.org/10.3389/fpubh.2022.875305

Loizidou, K., Elia, R., & Pitris, C. (2023). Computer-aided breast cancer detection and classification in mammography: A comprehensive review. *Computers in Biology and Medicine*, 153, 106554. https://doi.org/10.1016/j.compbiomed.2023.106554

Loizidou, K., Skouroumouni, G., Nikolaou, C., & Pitris, C. (2020). An Automated Breast Micro-Calcification Detection and Classification Technique Using Temporal Subtraction of Mammograms. *IEEE Access*, 8, 52785–52795. https://doi.org/10.1109/ACCESS.2020.2980616

Loizidou, K., Skouroumouni, G., Pitris, C., & Nikolaou, C. (2021). Digital subtraction of temporally sequential mammograms for improved detection and classification of microcalcifications. *European Radiology Experimental*, 5(1), 40. https://doi.org/10.1186/s41747-021-00238-w

Mahmood, T., Li, J., Pei, Y., Akhtar, F., Rehman, M. U., & Wasti, S. H. (2022). Breast lesions classifications of mammographic images using a deep convolutional neural network-based approach. *PLOS ONE*, 17(1), e0263126. https://doi.org/10.1371/journal.pone.0263126

Maitra, I. K., Nag, S., & Bandyopadhyay, S. K. (2012). Technique for preprocessing of digital mammogram. *Computer Methods and Programs in Biomedicine*, 107(2), 175–188. https://doi.org/10.1016/j.cmpb.2011.05.007

Moreira, I. C., Amaral, I., Domingues, I., Cardoso, A., Cardoso, M. J., & Cardoso, J. S. (2012). INbreast. *Academic Radiology*, 19(2), 236–248. https://doi.org/10.1016/j.acra.2011.09.014

Nasser, M., & Yusof, U. K. (2023). Deep Learning Based Methods for Breast Cancer Diagnosis: A Systematic Review and Future Direction. *Diagnostics*, 13(1), 161. https://doi.org/10.3390/diagnostics13010161

Nguyen, H. T., Nguyen, H. Q., Pham, H. H., Lam, K., Le, L. T., Dao, M., & Vu, V. (2023). VinDr-Mammo: A large-scale benchmark dataset for computer-aided diagnosis in full-field digital mammography. *Scientific Data*, 10(1), 277. https://doi.org/10.1038/s41597-023-02100-7

Oyelade, O. N., Ezugwu, A. E., Almutairi, M. S., Saha, A. K., Abualigah, L., & Chiroma, H. (2022). A generative adversarial network for synthetization of regions of interest based on digital mammograms. *Scientific Reports*, 12(1), 6166. https://doi.org/10.1038/s41598-022-09929-9

Oza, P., Sharma, P., & Patel, S. (2023). Breast lesion classification from mammograms using deep neural network and test-time augmentation. *Neural Computing and Applications,* (in press). https://doi.org/10.1007/s00521-023-09165-w

Oza, P., Sharma, P., Patel, S., Adedoyin, F., & Bruno, A. (2022). Image Augmentation Techniques for Mammogram Analysis. *Journal of Imaging*, 8(5). https://doi.org/10.3390/jimaging8050141

Prayitno, Shyu, C.-R., Putra, K. T., Chen, H.-C., Tsai, Y.-Y., Hossain, K. S. M. T., Jiang, W., & Shae, Z.-Y. (2021). A Systematic Review of Federated Learning in the Healthcare Area: From the Perspective of Data Properties and Applications. *Applied Sciences*, 11(23), 11191. https://doi.org/10.3390/app112311191

Ragab, D. A., Attallah, O., Sharkas, M., Ren, J., & Marshall, S. (2021). A framework for breast cancer classification using Multi-DCNNs. *Computers in Biology and Medicine*, 131, 104245. https://doi.org/10.1016/j.compbiomed.2021.104245

Rahmani Seryasat, O., & Haddadnia, J. (2018). Evaluation of a New Ensemble Learning Framework for Mass Classification in Mammograms. *Clinical Breast Cancer*, 18(3), e407–e420. https://doi.org/10.1016/j.clbc.2017.05.009

Ramos-Pollán, R., Guevara-López, M. A., Suárez-Ortega, C., Díaz-Herrero, G., Franco-Valiente, J. M., Rubio-del-Solar, M., González-de-Posada, N., Vaz, M. A. P., Loureiro, J., & Ramos, I. (2012). Discovering Mammography-based Machine Learning Classifiers for Breast Cancer Diagnosis. *Journal of Medical Systems*, 36(4), 2259–2269. https://doi.org/10.1007/s10916-011-9693-2

Rehman, K. ur, Li, J., Pei, Y., Yasin, A., Ali, S., & Mahmood, T. (2021). Computer Vision-Based Microcalcification Detection in Digital Mammograms Using Fully Connected Depthwise Separable Convolutional Neural Network. *Sensors*, 21(14), 4854. https://doi.org/10.3390/s21144854

Ribli, D., Horváth, A., Unger, Z., Pollner, P., & Csabai, I. (2018). Detecting and classifying lesions in mammograms with Deep Learning. *Scientific Reports*, 8(1), 4165. https://doi.org/10.1038/s41598-018-22437-z

Rodríguez-Ruiz, A., Krupinski, E., Mordang, J. J., Schilling, K., Heywang-Köbrunner, S. H., Sechopoulos, I., & Mann, R. M. (2019). Detection of breast cancer with mammography: Effect of an artificial intelligence support system. *Radiology*, 290(3), 305–314. https://doi.org/10.1148/radiol.2018181371

Sahu, A., Das, P. K., & Meher, S. (2023). Recent advancements in machine learning and deep learning-based breast cancer detection using mammograms. *Physica Medica*, 114, 103138. https://doi.org/10.1016/j.ejmp.2023.103138

Shen, L., Margolies, L. R., Rothstein, J. H., Fluder, E., McBride, R., & Sieh, W. (2019). Deep Learning to Improve Breast Cancer Detection on Screening Mammography. *Scientific Reports*, 9(1). https://doi.org/10.1038/s41598-019-48995-4

Soltani, H., Amroune, M., Bendib, I., & Haouam, M. Y. (2021). Breast Cancer Lesion Detection and Segmentation Based On Mask R-CNN. In *2021 International Conference on Recent Advances in Mathematics and Informatics (ICRAMI),* (pp. 1–6). IEEE. https://doi.org/10.1109/ICRAMI52622.2021.9585913

Soltani, H., Amroune, M., Bendib, I., Haouam, M.-Y., Benkhelifa, E., & Fraz, M. M. (2023). Breast lesions segmentation and classification in a two-stage process based on Mask-RCNN and Transfer Learning. *Multimedia Tools and Applications*, 83, 35763–35780. https://doi.org/10.1007/s11042-023-16895-5

Tabár, L., Vitak, B., Chen, T. H. H., Yen, A. M. F., Cohen, A., Tot, T., Chiu, S. Y. H., Chen, S. L. S., Fann, J. C. Y., Rosell, J., Fohlin, H., Smith, R. A., & Duffy, S. W. (2011). Swedish two-county trial: Impact of mammographic screening on breast cancer mortality during 3 decades. *Radiology*, 260(3), 658–663. https://doi.org/10.1148/radiol.11110469

Tariq, M., Iqbal, S., Ayesha, H., Abbas, I., Ahmad, K. T., & Niazi, M. F. K. (2021). Medical image based breast cancer diagnosis: State of the art and future directions. *Expert Systems with Applications*, 167, 114095. https://doi.org/10.1016/j.eswa.2020.114095

Tripathy, S., & Swarnkar, T. (2020). Unified Preprocessing and Enhancement Technique for Mammogram Images. *Procedia Computer Science*, 167, 285–292. https://doi.org/https://doi.org/10.1016/j.procs.2020.03.223

Vidić, I., Egnell, L., Jerome, N. P., Teruel, J. R., Sjøbakk, T. E., Østlie, A., Fjøsne, H. E., Bathen, T. F., & Goa, P. E. (2018). Support vector machine for breast cancer classification using diffusion-weighted MRI histogram features: Preliminary study. *Journal of Magnetic Resonance Imaging*, 47(5), 1205–1216. https://doi.org/10.1002/jmri.25873

Vijayarajeswari, R., Parthasarathy, P., Vivekanandan, S., & Basha, A. A. (2019). Classification of mammogram for early detection of breast cancer using SVM classifier and Hough transform. *Measurement*, 146, 800–805. https://doi.org/10.1016/j.measurement.2019.05.083

Wang, S., Huo, J., Ouyang, X., Che, J., Xue, Z., Shen, D., Wang, Q., & Cheng, J.-Z. (2020). mr^{2}NST: Multi-resolution and Multi-reference Neural Style Transfer for Mammography. In *Predictive Intelligence in Medicine,* (pp. 169–177). Springer. https://doi.org/10.1007/978-3-030-59354-4_16

Yu, X., Ren, Z., Guttery, D. S., & Zhang, Y.-D. (2024). DF-dRVFL: A novel deep feature based classifier for breast mass classification. *Multimedia Tools and Applications,* 83, 14393–14422. https://doi.org/10.1007/s11042-023-15864-2

Zahoor, S., Shoaib, U., & Lali, I. U. (2022). Breast Cancer Mammograms Classification Using Deep Neural Network and Entropy-Controlled Whale Optimization Algorithm. *Diagnostics*, 12(2), 557. https://doi.org/10.3390/diagnostics12020557

Zakareya, S., Izadkhah, H., & Karimpour, J. (2023). A New Deep-Learning-Based Model for Breast Cancer Diagnosis from Medical Images. *Diagnostics*, 13(11), 1944. https://doi.org/10.3390/diagnostics13111944

Zhang, X. (2023). Molecular Classification of Breast Cancer: Relevance and Challenges. *Archives of Pathology & Laboratory Medicine*, 147(1), 46–51. https://doi.org/10.5858/arpa.2022-0070-RA

Zhu, Z., Wang, S.-H., & Zhang, Y.-D. (2023). A Survey of Convolutional Neural Network in Breast Cancer. *Computer Modeling in Engineering & Sciences*, 136(3), 2127–2172. https://doi.org/10.32604/cmes.2023.025484

Editorial record: The article has been peer-reviewed. First submission received on 13 February 2024. Revisions received on 14 May 2024, 20 May 2024 and 21 June 2024. Accepted for publication on 28 June 2024. The editors coordinating the peer-review of this manuscript were Hakim Bendjenna, Lawrence Chung, Abdallah Meraoumia, and Zdenek Smutny. The editor in charge of approving this manuscript for publication was Zdenek Smutny.

Special Issue: Future Trends of Machine Intelligence in Science and Industry. Selected papers from the National Conference on Artificial Intelligence: From Theory to Practice (NCAI'2023).

Acta Informatica Pragensia is published by Prague University of Economics and Business, Czech Republic.

ISSN: 1805-4951