The utilization of artificial intelligence applications to improve breast cancer detection and prognosis

Walaa M. Alsharif, MSc, PhD.

ABSTRACT

يواجه تصوير الثدي بعض التحديات مع الزيادة الحالية في طلبات التصوير والاورام التي يمكن أن تفوت خلال برامج فحص الثدي. يجري البحث عن إيجاد حلول للتخفيف من هذه التحديات من خلال التقدم الحديث في مجال تطبيقات الذكاء الاصطناعي لتعزيز كفاءة سير العمل ونتائج الرعاية الصحية للمرضى. بعض الدراسات طرحت تصاميم لبعض تطبيقات الذكاء الاصطناعي في مجال التصوير الطبي لاستخدامها في تحليل صور اشعة الثدي في مختلف تقنيات التصوير الطبي مثل التصوير الاشعاعي للثدي، التصوير بالموجات الصوتية والتصوير بالرنين المغناطيسي للكشف عن أورام الثدي، تصنيفها، تقييمها وتحديد مدى الخطر. هذه الورقة سوف تقدم شرح بخصوص نظام الكشف التقليدي بمساعدة الحاسوب، التطبيقات القائمة على الذكاء الاصطناعي في التصوير الطبي للثدي. بالإضافة لذلك، سوف تتم مناقشة التحديات والقيود المحتملة في إمكانية استخدام هذه التطبيقات.

Breast imaging faces challenges with the current increase in medical imaging requests and lesions that breast screening programs can miss. Solutions to improve these challenges are being sought with the recent advancement and adoption of artificial intelligent (AI)-based applications to enhance workflow efficiency as well as patient-healthcare outcomes. rtificial intelligent tools have been proposed and used to analyze different modes of breast imaging, in most of the published studies, mainly for the detection and classification of breast lesions, breast lesion segmentation, breast density evaluation, and breast cancer risk assessment. This article reviews the background of the Conventional Computer-aided Detection system and AI, AI-based applications in breast medical imaging for the identification, segmentation, and categorization of lesions, breast density and cancer risk evaluation. In addition, the challenges, and limitations of AI-based applications in breast imaging are also discussed.

Keywords: artificial intelligence, breast cancer, breast imaging

Saudi Med J 2023; Vol. 44 (2): 119-127
doi: 10.15537/smj.2023.44.2.20220611

From the Diagnostic Radiology Technology Department, College of Applied Medical Sciences, Taibah University, Al Madinah Al Munawwarah; and from the Society of Artificial Intelligence in Healthcare, Riyadh, Kingdom of Saudi Arabia.

Address correspondence and reprint request to: Dr. Walaa M. Alsharif, Diagnostic Radiology Technology Department, College of Applied Medical Sciences, Taibah University, Al Madinah Al Munawwarah, Kingdom of Saudi Arabia. E-mail: Wsheref@taibahu.edu.sa ORCID ID: https://orcid.org/0000-0001-7607-3255

Breast cancer is the second biggest cause of death for women, and early screening and identification can increase treatment choices and reduce mortality rates.[1-3] Consequently, several countries have applied breast screening programs for all women between the ages of 40 and 50. Over 42 million exams are undertaken across the world.[4-6] It was reported that breast screening programs can miss between 15% and 35% of cancers, either due to error or because the cancer was not detectable or perceptible to radiologists at the time of scanning.[7] Also, radiologists may face challenges in breast image interpretation, as breast images are subjected to several limitations. First, breast density can affect image sensitivity, which may result in breast cancer not being detected during the screening.[8,9] Second, false-positive findings can be generated, which leads to patient discontent and increased cost and workload. These false-positive findings can lead to unnecessary follow-ups and invasive diagnostic practices such as biopsies.[10,11]

Computer-aided detection (CAD) has been designed and used in breast imaging to aid radiologists and automate the early discovery and diagnosis of breast lesions.[12] Studies show that a solo interpretation plus CAD can be used as an substitute to double reading.[13,14] However, several studies highlight the low specificity of the CAD systems and no improvement in cost-effectiveness.[15,16] Increasing call-back rates/false-positive recalls were found when CAD had been used in breast screening.[17] This inaccuracy and unpredictability of the

CAD system has raised some doubts about whether artificial intelligence (AI) applications and recent advances in deep learning (DL) can help radiologists improve their performance in detecting breast cancer.[18] Machine learning (ML) and DL are both subsets of AI. The ML model requires more human interference and structured data to achieve results. It should be provided with data/inputs, image features (scuh as shape, width, and edge), applying a classification algorithm, and then the model will predict outputs. However, the DL model permits direct features extraction from the original unstructured data/inputs (such as image, text). Therefore, the DL model outweighs the ML model in the case of large datasets as no manual feature extraction is needed.[19] A DL-CAD tool has been recently developed and used for breast lesion detection and characterization (benign or malignant).[20-25] The literature refers to the potential role of DL-CAD regarding enhancing diagnostic accuracy and specificity in breast screening. Despite the AI strengths reported in the literature, several drawbacks of the existing AI applications have also been highlighted (Table 1).[26] This review aims to assess the available literature to evaluate the readiness of the existing AI-based applications for breast screening and to advise the research directions in this field.

Conventional Computer-Aided Detection system (CAD) and AI-based applications. The role of computers in assisting radiologists in their clinical practice is not new. In 1998, CAD systems utilizing traditional ML were developed and used as a second opinion to analyze patients' images in mammography and improve radiologists' performance.[27,28] The effectiveness of this instrument in breast cancer diagnosis has long been controversial. Literature shows that despite the positive impact of CAD systems on breast cancer screening, decreases in specificity and increases in recall rates are also noted. Several studies failed to ascertain the meaningful value of the CAD system in clinical practice. The relationship between CAD system usage, image interpretation accuracy, and recall rates has been documented in the literature.[15] Lehman et al[17] claim that there is no improvement in detection rate and/or prognostic characterization of breast cancer with the CAD system.

> **Disclosure.** Author has no conflict of interests, and the work was not supported or funded by any drug company.

The conventional CAD system has been established to aid radiologists, not to be used as a primary screening tool. It is designed and trained to detect specific features that radiologists look for, such as masses or classifications. This conventional system depends on manual features extraction by expert.[29] Although the conventional CAD system can achieve a high sensitivity compared to radiologists, high false-positive rates might increase.[30] Therefore, radiologists should screen and read the images as carefully as they would without CAD, and then use CAD as 'spell checker' following their own interpretation.[28] This tool has evolved over time from traditional approaches to modern DL methods. Deep learning-based CAD can interpret vast amounts of data, learn features from images, learn from mistakes, and improve performance over time.[31] This makes DL technology more robust to adapt different circumstances related to the type of scanners and patient population (when the training data are available).[28] Currently, conventional/artificial neural networks (CNNs/ANNs) are most commonly used in DL for pattern recognition tasks in images.

The human brain is made up of millions of neurons, so CNNs are really just a composition of perceptions, connected in different ways and operating on different activation functions. Conventionals basically process information in a similar way to the human brain: they have self-learning capabilities, and will learn from examples (such as inputs/images) and experience (such as training) without the need for manually designed features.[32] They are composed of fully connected layer, which creates the features map and passes it next to a pooling layer, where down-sampling takes place, then the outcome is passed to a fully connected layer to be assorted; and finally the output layer, which creates the analysis results of the data (Figure 1).[33] The addition of layers is determined by the application and the problem that needs to be solved. The more layers there are, the more feature extraction and data abstraction can be achieved.[34] The potential role of deep CNNs (DCNN) brings strong guarantees in improving the accuracy and promoting CAD as a clinical support system for medical imaging.[28]

AI-based applications in breast cancer screening. The DCNNs can be designed and applied to detect lesions (such as changes in size, shape, or texture), classifications (such as benign/malignant), and segmentation of organs and tumors.[18,35] This section will discuss the current advancement of DL for breast screening and cancer detection in different imaging modalities.

AI-based application and mammography. Mammography is a non-invasive method that is

often used to detect breast cancer.[36] It can successfully identify non-palpable masses and classify between benign lesions and malignant tumors.[37] The CAD system was developed to facilitate decision-making and to lessen the demand for numerous readers. Using the CAD system, on the other hand, has been linked to higher recall rates in breast cancer screening.[14,16] Notable advancements have been made in recent years on AI-based applications to analyze mammographic images to detect breast masses and calcifications.[38-40] A higher performance was shown by radiologists in cancer detection as measured by the area under the receiver operating characteristic curve (AUC) when using AI in reduced reading times (reduced by approximately 4.5%).[39] A comparable result was stated by Pacile et al.[40] This might imply that AI will take over the more routine and tedious cases, so that radiologists have sufficient time to focus on complex cases.[39] A dramatic

Table 1 - Strengths and drawbacks of artificial intelligence (AI)-based applications in radiology.

| Strengths	Drawbacks
Enhance analysis: Automated pathology screening, detection, and characterization.	Time and cost consuming in training and testing AI model.
Accurate classification: Categorize image based on abnormality (benign and malignant).	Ethical and legal issues: a. Ethics of data (How should we use, label and protect data?) b. Ethics of algorithm and trained model (How does the AI model make decisions? How can we diminish the risk of patient harm from privacy breaches? Who is responsible for mistakes resulting from the use of the AI model?) c. Ethics of practice (monitor and verify AI-driven autonomy)
Extract additional required information from previous detected pathology.	Biased predicted outcomes due to incomplete and/or unrepresentative data.
Offer a second opinion which increases confidence of the diagnosis.	Lack of interpretability can lead to a lack of trust and acceptance of AI models by physicians and patients.
Minimize interindividual variability, bias and time.	*Lack* of strong evidence and regulations to support the *use of an AI model.*
	Lack of standardized benchmarks which make it difficult to validate the performance of an AI model.
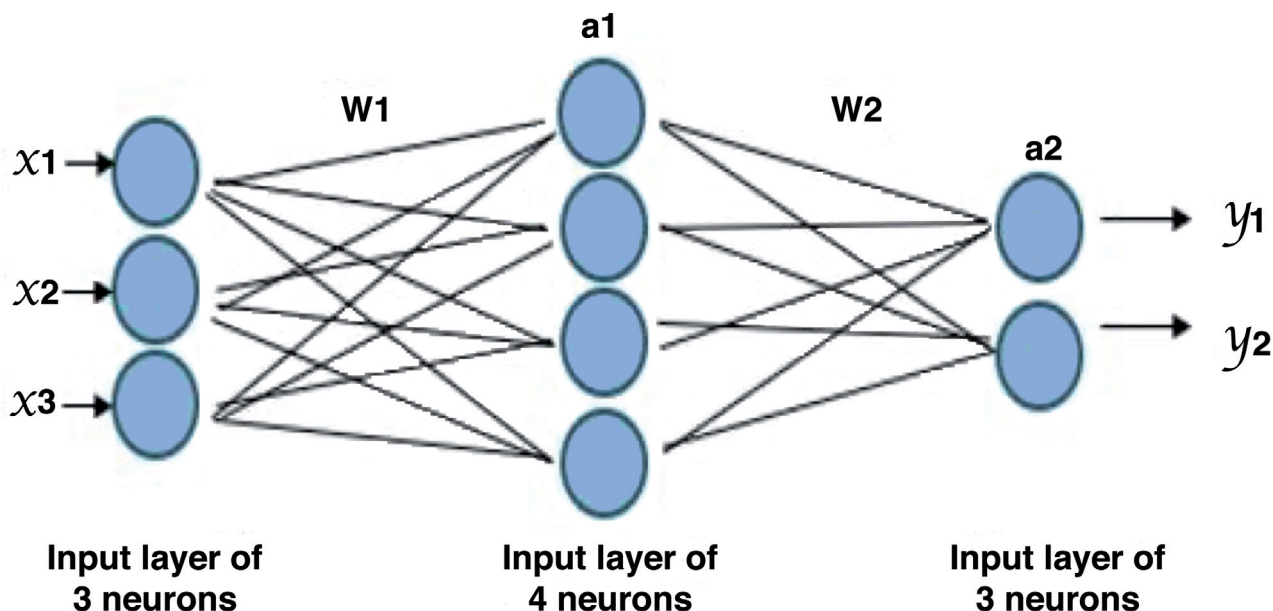
Figure 1 - Conventional/Artificial neural networks (CNN/ANN) composition

improvement in the detection of breast cancer was noted among radiologists with less experience and trainees when they used the AI.[39,41] A reasonable explanation for this finding may have to do with a greater acceptance of AI-based applications by younger radiologists. In addition to this, experienced radiologists may ignore the flags produced by the AI more often than those with less experience.[41] Dembrower et al[42] studied the potential change in cancer detection and found that AI application can detect subtle tumors that were not identified previously without AI. Another study has recently been published in which AI was shown to be superior to radiologists in detecting cancer in cases of fatty and dense breast tissue.[43] Dahlblom et al[44] carried out a study to compare the performance of an AI tool with digital breast tomosynthesis (DBT) in the breast cancer detection. They demonstrated that 44% of the breast cancer detected only using DBT by radiologists were spotted with digital mammography by the AI tool.[44] Although almost half of the cases were identified with digital mammography by AI, further investigation with a larger sample may be warranted.

Researchers refer to the potential advantage of AI tools in finding tumors with marginal indications or false-negative interval cancers on earlier screening exams that have not yet been studied.[45-47] They discovered that AI tools could help experts in identifying up to 19.3% of the interval cancers with minimal signs of malignancy.[45] Numerous research compared AI performance to radiologist accuracy in clinical settings. Schaffter et al[48] stated that AI-based application combined with a radiologist showed higher accuracy in breast image screening in comparison with a single radiologist interpretation alone. This is in line with Salim et al,[49] who used three commercial AI systems and indicated that one AI application had higher sensitivity (81.9%) and the other two lesser sensitivity (67%, 67.4%) compared to the radiologist (77.4%).[49] Likewise McKinney et al's[49] study found that the AI tool beat the radiologist performance in sensitivity (56% versus [vs] 48%) and specificity (84% vs 81%).[46] This was in line with Hmida et al[50] and Sapate.[51] Even though the aforementioned studies highlight the potential benefit of AI in improving breast cancer screening and image interpretation, it is also important to note that mammographic images were examined in laboratory settings with a limited number of readers, which makes it difficult to generalize the findings to the clinical practice.

AI-based application and ultrasound (US). Ultrasound frequently serves as an additional scanning modality to mammography in breast screening programs. It is a common simple imaging technique involved in evaluation of palpable breast abnormalities and characterized breast masses.[52] Ultrasound scanning of the breast has several benefits related to other imaging modalities, including lower cost, lack of ionizing radiation, and the capability to assess images in real time.[53] However, radiologists face challenges in interpreting breast US images due to blurry borders, inherent low contrast, and high levels of shadowing.[54,55] Several researchers referred to intra-reader variability in interpreting breast US images, and increased false-positive findings.[56,57]

Recent advances in DL have accelerated the development of AI-based applications for the automated identification, segmentation, extraction, and classification of breast cancer from US images.[58-60] A study conducted by Shen et al[61] to identify malignant lesions in breast US images found that AI tools reach a higher area under the receiver operating characteristic curve (AUROC) and area under the precision-recall curve (AUPRC) than expert radiologists. In the same study, radiologists managed to reduce false positive rates by 37.3% and the number of biopsies requested by 27.8%, while maintaining the level of sensitivity.[61] Several studies proved that DL won in terms of classification and recognition of breast cancer based on US images.[62-64] A study carried out by Fujioka et al[59] highlighted the fact that DL with CNNs showed equal or higher AUC than radiologists in distinguishing benign from malignant breast lesions on US images.[65] Furthermore, it was found that DL systems showed high performance in breast lesion calcifications with an accuracy of 93.4%, a sensitivity of 88.6%, a specificity of 97.1%, and an area under the AUC of 0.947.[62] Similar results were reported by Becker et al[62] and Han et al,[64] in which AI systems achieved higher performance in differentiating breast lesions in less time with accuracy similar to that of radiologists. Another AI application was assessed by Mango et al,[66] who highlighted an improvement in the accuracy of US breast lesion assessment when combining radiologists' evaluation with the AI. There was a significant drop in the inter- and intra-observer variability.

An AI system equipped in US machines has also been recently proposed, which can offer immediate decision of benignity or malignancy in the static US images after marking the region of interest (ROI).[18] Kim et al[24] evaluated the diagnostic functioning of an AI system equipped in a US machine to discern between benign and malignant breast lesions. Accuracy was significantly higher, and the AUC was 0.72 compared to the radiologists. A year later in 2018, a similar performance

was indicated by Di Segni et al,[67] who indicted a higher sensitivity of >90% and specificity of 70.8% in the assessment of focal breast lesions. -Although some of the above AI-based applications are approved in some countries, until now there are yet no guides to endorse the AI-based applications alongside US in daily clinical practice.[68]

AI-based application and magnetic resonance imaging (MRI). Breast MRI shows high sensitivity in breast cancer detection.[68] Dynamic contrast material enhanced (DCE) MRI of the breast screening showed the highest sensitivity compared to the other imaging modalities (such as mammogram and US).[69] However, false-negative results can be obtained with a reported sensitivity of 90.9% for invasive cancer and 73% for ductal carcinoma.[70] Several studies reviewed breast MRI examinations retrospectively and found that some breast lesions had been either missed, misdiagnosed, or mismanaged.[71,72] This may be attributed to the various factors such as dense breast tissue, visual examination pattern, incorrect assessment, image quality, distraction, fatigue, workload, poor enhancement, or misdiagnose enhancement.[73-75] Automated detection of breast carcinomas in MRI images via AI-based applications (such as CNN) has been indicated for systematic diagnostic interpretation and identification of tumors on images stored on an archiving system.[76,77]

Literature reported the utilization of DL for MRI breast screening[75,78-81] and most popular role of the proposed AI-based applications have included tasks, such as detection, segmentation, and classification of lesions in MRI images.[75,78,82-84] A decent performance was reported by Truhn et al,[85] who used CNN to categorize segmented lesions as benign or malignant, reached AUC of 0.88, which was superior to radiomics analysis (AUC=0.81); however, this was inferior to the breast radiologist's interpretation (AUC = 0.98).[85] Similarly reported by Herent et al,[80] who used the DL model to identify and categorize breast lesions as benign or malignant, achieving an AUC of 0.816. Ayatollahi et al[86] promoted a DL model for the detection of breast lesions in ultrafast DCE-MRI sequences. This proposed model achieved a high detection rate of 0.90 (0.876-0.934), sensitivity of 0.95 (0.934–0.980), and a detection rate of benign lesions of 0.81 (0.751-0.871).

An improvement in radiologists' clinical performance with the assistance of an AI tool was also reported by Adachi et al.[81] A study carried out by Eskreis-Winkler et al[87] showed improvement in accuracy (92.8%), sensitivity (89.5%), and specificity (94.3%) of detection of breast cancer with the assistance of an AI tool. They also found that using DL tools could lead to a reduction

in reading time (3 and 45 seconds per case).[87] This was consistent with Jiang et al,[88] who evaluated the clinical performance of radiologists in detecting breast cancer at DCE MRI, and found that AUC showed an improvement in accuracy from 0.71 to 0.76 when AI was used.[69] This was reinforced recently by Wu et al,[89] who found that the proposed CNN model based on DCE MRI achieved diagnostic accuracy of 87.7%, precision of 91.2%, sensitivity of 86.1%, and AUC of 91.2%. In addition, it was highlighted that DL can be a favorable tool to increase the proficiency and accessibility of breast MRI.[88] Jing et al[88] found that the proposed AI model achieved an AUC of 0.81 with a 15.7% reduction in workload and a 16.6% reduction in scanning time.

In addition, AI and radiomics approach have gained popularity in medical imaging to facilitate disease diagnosis (such as breast lesions).[89,90] In an MRI-based radiomics and AI study, entropy of breast lesions found to be a worth parameters to differentiate between malignant and benign breast lesions.[90,91] Fusco et al[92] reported consistent findings. Another study carried out by Crivelli et al[93] revealed that radiomics values were lower than those suspected by expert breast radiologists. Thus, it is essential to notice that the promising evidence of radiomics still requires further evaluation and many issues need to be solved prior to it being ready to be implemented in clinical practice.

The above studies show that AI-based applications are a promising tool for breast image screening and cancer detection. However, several challenges and limitations need to be considered. Picture Archiving and Communication Systems (PACS) and the Digital Imaging and Communications in Medicine (DICOM) have ensured that datasets required to train and test AI-based applications are prepared for easy access and recovery. Yet, organizing the datasets (labeling, annotation, segmentation) represents a major issue in developing AI-based applications as trained professionals (clinical scientists and informaticians) are required and this is a time-consuming and high-cost process.[94-96] Comparing the clinical value of different AI-based applications is challenging due to the variation in the datasets, the approach of testing, and validation of the tool performance.[39,97,98] This is combined with the lack of high-quality, categorized, labeled datasets, which are representative as well as including a good distribution of abnormalities, demographics, and breast density.[18] Inappropriate datasets and poor image quality may limit the conspicuity of the breast lesion's characters or offer inadequate inputs for the AI system.[99,100]

Although the reviewed studies in this article reported encouraging findings of the proposed AI model's accuracy, most of these studies were retrospective, based on relatively small datasets, came from single institutions, and identified several methodological restrictions that negatively impact on the applicability and strength of the AI-based applications in the breast screening setting. Using small datasets from the same source may raise concerns regarding the depth, quality, and representativeness of the images that are used to teach the AI-based applications and increase the possibility of bias and overfitting. In addition, datasets enriched with malignant lesions or suspicious abnormalities were also used to train the AI tool. This approach may assist the feasibility of designing and developing an AI tool; conversely, the applicability and accuracy of the proposed AI tool performance will still be unclear or overestimated as the datasets may not reflect the real-world screening of detectable breast cancer. Several factors may affect the performance of the AI-based applications, in particular patient populations such as heterogeneity of the breast cancer risk factors, and imaging characters of the populations. Therefore, a larger validation dataset from diverse screening environments and populations is required to ensure that AI tools are ready for use in real clinical practice.

In conclusion, this review of the available research on AI-based applications and breast cancer screening provides insight into the value of AI tools combined with diverse imaging modalities in breast lesion detection and diagnosis. The shift from the conventional CAD system to the advanced AI tools such as DL-CAD has the potential to reduce false-positive findings, increase diagnostic accuracy, improve radiologist performance, and assist with decision-making. However, the current evidence regarding the use of AI-based applications in the detection of the breast cancer is not yet fully optimized due to a lack of standardized methodology and prospective studies, the possibility of bias, and a lack of depth and quality. Future randomized controlled trials and cohort studies in large-scale sample with high-quality evidence are required to consider the future use of AI-based applications in breast cancer screening.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018; 68: 394-424.

2. Lehman CD, Arao RF, Sprague BL, Lee JM, Buist DS, Kerlikowske K, et al. National performance benchmarks for modern screening digital mammography: update from the Breast Cancer Surveillance Consortium. *Radiology* 2017; 283: 49.

3. Black E, Richmond R. Improving early detection of breast cancer in sub-Saharan Africa: why mammography may not be the way forward. *Global Health* 2019; 15: 1-11.

4. Lee CH, Dershaw DD, Kopans D, Evans P, Monsees B, Monticciolo D, et al. Breast cancer screening with imaging: recommendations from the Society of Breast Imaging and the ACR on the use of mammography, breast MRI, breast ultrasound, and other technologies for the detection of clinically occult breast cancer. *J Am Coll Radiol* 2010; 7: 18-27.

5. Oeffinger KC, Fontham ET, Etzioni R, Herzig A, Michaelson JS, Shih Y-CT, et al. Breast cancer screening for women at average risk: 2015 guideline update from the American Cancer Society. *Jama* 2015; 314: 1599-1614.

6. Siu AL, Force UPST. Screening for breast cancer: US Preventive Services Task Force recommendation statement. *Ann Intern Med* 2016; 164: 279-296.

7. Houssami N, Hunter K. The epidemiology, radiology and biological characteristics of interval breast cancers in population mammography screening. *NPJ Breast Cancer* 2017; 3: 1-13.

8. Vourtsis A, Berg WA. Breast density implications and supplemental screening. *Eur Radiol* 2019; 29: 1762-1777.

9. Nelson HD, O'meara ES, Kerlikowske K, Balch S, Miglioretti D. Factors associated with rates of false-positive and false-negative results from digital mammography screening: an analysis of registry data. *Ann Intern Med* 2016; 164: 226-235.

10. Tosteson AN, Fryback DG, Hammond CS, Hanna LG, Grove MR, Brown M, et al. Consequences of false-positive screening mammograms. *JAMA internal medicine* 2014; 174: 954-961.

11. Alcusky M, Philpotts L, Bonafede M, Clarke J, Skoufalos A. The patient burden of screening mammography recall. *J Womens Health (Larchmt)* 2014; 23: S11-S19.

12. Jalalian A, Mashohor S, Mahmud R, Karasfi B, Saripan MIB, Ramli ARB. Foundation and methodologies in computer-aided diagnosis systems for breast cancer detection. *EXCLI J* 2017; 16: 113.

13. Bargalló X, Santamaría G, Del Amo M, Arguis P, Ríos J, Grau J, et al. Single reading with computer-aided detection performed by selected radiologists in a breast cancer screening program. *Eur J Radiol* 2014; 83: 2019-2023.

14. Henriksen EL, Carlsen JF, Vejborg IM, Nielsen MB, Lauridsen CA. The efficacy of using computer-aided detection (CAD) for detection of breast cancer in mammography screening: a systematic review. *Acta Radiologica* 2019; 60: 13-18.

15. Masud R, Al-Rei M, Lokker C. Computer-aided detection for breast cancer screening in clinical settings: scoping review. *JMIR Med Inform* 2019; 7: e12660.

16. Katzen J, Dodelzon K. A review of computer aided detection in mammography. *Clinical imaging* 2018; 52: 305-309.

17. Lehman CD, Wellman RD, Buist DS, Kerlikowske K, Tosteson AN, Miglioretti DL, et al. Diagnostic accuracy of digital screening mammography with and without computer-aided detection. *JAMA Intern Med* 2015; 175: 1828-1837.

18. Le E, Wang Y, Huang Y, Hickman S, Gilbert F. Artificial intelligence in breast imaging. *Clin Radiol* 2019; 74: 357-366.

19. Liu H, Lang B. Machine learning and deep learning methods for intrusion detection systems: A survey. *Appl Sci* 2019; 9: 4396.

20. Choi J-H, Kang BJ, Baek JE, Lee HS, Kim SH. Application of computer-aided diagnosis in breast ultrasound interpretation: improvements in diagnostic performance according to reader experience. *Ultrasonography* 2018; 37: 217-225.

21. Lee J, Kim S, Kang BJ, Kim SH, Park GE. Evaluation of the effect of computer aided diagnosis system on breast ultrasound for inexperienced radiologists in describing and determining breast lesions. *Med Ultrason* 2019; 21: 239-245.

22. Di Segni M, de Soccio V, Cantisani V, Bonito G, Rubini A, Di Segni G, et al. Automated classification of focal breast lesions according to S-detect: validation and role as a clinical and teaching tool. *J Ultrasound* 2018; 21: 105-118.

23. Bartolotta TV, Orlando A, Cantisani V, Matranga D, Ienzi R, Cirino A, et al. Focal breast lesion characterization according to the BI-RADS US lexicon: role of a computer-aided decision-making support. *Radiol Med* 2018; 123: 498-506.

24. Kim K, Song MK, Kim E-K, Yoon JH. Clinical application of S-Detect to breast masses on ultrasonography: a study evaluating the diagnostic performance and agreement with a dedicated breast radiologist. *Ultrasonography* 2017; 36: 3.

25. Choi JS, Han B-K, Ko ES, Bae JM, Ko EY, Song SH, et al. Effect of a deep learning framework-based computer-aided diagnosis system on the diagnostic performance of radiologists in differentiating between malignant and benign masses on breast ultrasonography. *Korean J Radiol* 2019; 20: 749-758.

26. Jiang L, Wu Z, Xu X, Zhan Y, Jin X, Wang L, et al. Opportunities and challenges of artificial intelligence in the medical field: Current application, emerging problems, and problem-solving strategies. *J Int Med Res* 2021; 49: 03000605211000157.

27. Ou WC, Polat D, Dogan BE. Deep learning in breast radiology: current progress and future directions. *Eur Radiol* 2021; 31: 4872-4885.

28. Chan HP, Hadjiiski LM, Samala RK. Computer-aided diagnosis in the era of deep learning. *Med Phys* 2020 ;47: e218-e27.

29. Kohli A, Jha S. Why CAD failed in mammography. *J Am Coll Radiol* 2018; 15: 535-537.

30. Le MT, Mothersill CE, Seymour CB, McNeill FE. Is the false-positive rate in mammography in North America too high? *Br J Radiol* 2016; 89: 20160045.

31. Chen X-W, Lin X. Big data deep learning: challenges and perspectives. *IEEE Access* 2014; 2: 514-525.

32. Agnes SA, Anitha J, Pandian S, Peter JD. Classification of mammogram images using multiscale all convolutional neural network (MA-CNN). *J Med Syst* 2020; 44: 1-9.

33. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015; 521: 436-444.

34. Rawat W, Wang Z. Deep convolutional neural networks for image classification: A comprehensive review. *Neural Comput* 2017; 29: 2352-449.

35. Balkenende L, Teuwen J, Mann RM, editors. Application of deep learning in breast cancer imaging. *Semin Nucl Med* 2022; 52: 584-596.

36. Welch HG, Prorok PC, O'Malley AJ, Kramer BS. Breast-cancer tumor size, overdiagnosis, and mammography screening effectiveness. New *N Engl J Med* 2016; 375: 1438-1447.

37. Alsheh Ali M, Eriksson M, Czene K, Hall P, Humphreys K. Detection of potential microcalcification clusters using multivendor for-presentation digital mammograms for short-term breast cancer risk estimation. *Med Phys* 2019; 46: 1938-1946.

38. Burt JR, Torosdagli N, Khosravan N, RaviPrakash H, Mortazi A, Tissavirasingham F, et al. Deep learning beyond cats and dogs: recent advances in diagnosing breast cancer with deep neural networks. *Br J Radiol* 2018; 91: 20170545.

39. Rodriguez Ruiz A, Krupinski E, Mordang J-J, Schilling K, Heywang-Kobrunner S, Sechopoulos I, et al. Detection of breast cancer with mammography: effect of an artificial intelligence support system 2019; 290(2):305-14

40. Pacilè S, Lopez J, Chone P, Bertinotti T, Grouin JM, Fillard P. Improving breast cancer detection accuracy of mammography with the concurrent use of an artificial intelligence tool. *Radiol Artif Intell* 2020; 2: e190208.

41. Watanabe AT, Lim V, Vu HX, Chim R, Weise E, Liu J, et al. Improved cancer detection using artificial intelligence: a retrospective evaluation of missed cancers on mammography. *J Digit Imaging* 2019; 32: 625-637.

42. Dembrower K, Wåhlin E, Liu Y, Salim M, Smith K, Lindholm P, et al. Effect of artificial intelligence-based triaging of breast cancer screening mammograms on cancer detection and radiologist workload: a retrospective simulation study. *Lancet Digital Health* 2020; 2: e468-e474.

43. Kim H-E, Kim HH, Han B-K, Kim KH, Han K, Nam H, et al. Changes in cancer detection and false-positive recall in mammography using artificial intelligence: a retrospective, multireader study. *Lancet Digital Health* 2020; 2: e138-e148.

44. Dahlblom V, Andersson I, Lång K, Tingberg A, Zackrisson S, Dustler M. Artificial intelligence detection of missed cancers at digital mammography that were detected at digital breast tomosynthesis. *Radiol Artif Intell* 2021; 3: e200299.

45. Lång K, Hofvind S, Rodríguez-Ruiz A, Andersson I. Can artificial intelligence reduce the interval cancer rate in mammography screening? *Eur Radiol* 2021; 31: 5940-5947.

46. McKinney SM, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, et al. International evaluation of an AI system for breast cancer screening. *Nature* 2020; 577: 89-94.

47. Mayo RC, Kent D, Sen LC, Kapoor M, Leung JW, Watanabe AT. Reduction of false-positive markings on mammograms: a retrospective comparison study using an artificial intelligence-based CAD. *Journal of digital imaging* 2019; 32: 618-624.

48. Schaffter T, Buist DS, Lee CI, Nikulin Y, Ribli D, Guan Y, et al. Evaluation combiend artificial intelligence and radiologist assessment to interpret screening mammograms. *JAMA Netw Open* 2020; 3: e20065-e.

49. Salim M, Wåhlin E, Dembrower K, Azavedo E, Foukakis T, Liu Y, et al. External evaluation of 3 commercial artificial intelligence algorithms for independent assessment of screening mammograms. *JAMA oncology* 2020; 6: 1581-1588.

50. Hmida M, Hamrouni K, Solaiman B, Boussetta S. Mammographic mass segmentation using fuzzy contours. *Comput Methods Programs Biomed* 2018; 164: 131-142.

51. Sapate SG, Mahajan A, Talbar SN, Sable N, Desai S, Thakur M. Radiomics based detection and characterization of suspicious lesions on full field digital mammograms. *Comput Methods Programs Biomed* 2018; 163: 1-20.

52. Akkus Z, Cai J, Boonrod A, Zeinoddini A, Weston AD, Philbrick KA, et al. A survey of deep-learning applications in ultrasound: Artificial intelligence–powered ultrasound for improving clinical workflow. *J Am Coll Radiol* 2019; 16: 1318-1328.

53. Feig S. Cost-effectiveness of mammography, MRI, and ultrasonography for breast cancer screening. *Radiol Clin North Am* 2010; 48: 879-891.

54. Huang Q, Luo Y, Zhang Q. Breast ultrasound image segmentation: a survey. *Int J Comput Assist Radiol Surg* 2017; 12: 493-507.

55. Huang Q-H, Lee S-Y, Liu L-Z, Lu M-H, Jin L-W, Li A-H. A robust graph-based segmentation method for breast tumors in ultrasound images. *Ultrasonics* 2012; 52: 266-275.

56. Crystal P, Strano SD, Shcharynski S, Koretz MJ. Using sonography to screen women with mammographically dense breasts. *AJR Am J Roentgenol* 2003; 181: 177-182.

57. Lazarus E, Mainiero MB, Schepps B, Koelliker SL, Livingston LS. BI-RADS lexicon for US and mammography: interobserver variability and positive predictive value. *Radiology* 2006; 239: 385-391.

58. Gao Y, Geras KJ, Lewin AA, Moy L. New frontiers: an update on computer-aided diagnosis for breast imaging in the age of artificial intelligence. *AJR Am J Roentgenol* 2019; 212: 300.

59. Fujioka T, Mori M, Kubota K, Oyama J, Yamaga E, Yashima Y, et al. The utility of deep learning in breast ultrasonic imaging: a review. *Diagnostics (Basel)* 2020; 10: 1055.

60. Al-Dhabyani W, Gomaa M, Khaled H, Aly F. Deep learning approaches for data augmentation and classification of breast masses using ultrasound images. *Int J Adv Comput Sci Appl* 2019; 10: 1-11.

61. Shen Y, Shamout FE, Oliver JR, Witowski J, Kannan K, Park J, et al. Artificial intelligence system reduces false-positive findings in the interpretation of breast ultrasound exams. *Nature communications* 2021; 12: 1-13.

62. Becker AS, Mueller M, Stoffel E, Marcon M, Ghafoor S, Boss A. Classification of breast cancer in ultrasound imaging using a generic deep learning analysis software: a pilot study. *Br J Radiol* 2018; 91: 20170576.

63. Zhang Q, Xiao Y, Dai W, Suo J, Wang C, Shi J, et al. Deep learning based classification of breast tumors with shear-wave elastography. *Ultrasonics* 2016; 72: 150-157.

64. Han S, Kang H-K, Jeong J-Y, Park M-H, Kim W, Bang W-C, et al. A deep learning framework for supporting the classification of breast lesions in ultrasound images. *Phys Med Biol* 2017; 62: 7714.

65. Fujioka T, Katsuta L, Kubota K, Mori M, Kikuchi Y, Kato A, et al. Classification of breast masses on ultrasound shear wave elastography using convolutional neural networks. *Ultrason Imaging* 2020; 42: 213-220.

66. Mango VL, Sun M, Wynn RT, Ha R. Should we ignore, follow, or biopsy? Impact of artificial intelligence decision support on breast ultrasound lesion assessment. *AJR Am J Roentgenol* 2020; 214: 1445.

67. Di Segni M, de Soccio V, Cantisani V, Bonito G, Rubini A, Si Segni G, et al. Automated classification of focal breast lesions according to S-detect: validation and role as a clinical and teaching tool. *J Ultrasound* 2018; 21: 105-108.

68. Bitencourt A, Naranjo ID, Gullo RL, Saccarelli CR, Pinker K. AI-enhanced breast imaging: Where are we and where are we heading? *Eur J Radiol* 2021; 142: 109882.

69. Jiang Y, Edwards AV, Newstead GM. Artificial intelligence applied to breast MRI for improved diagnosis. *Radiology* 2021; 298: 38-46.

70. Bluemke DA, Gatsonis CA, Chen MH, DeAngelis GA, DeBruhl N, Harms S, et al. Magnetic resonance imaging of the breast prior to biopsy. *JAMA* 2004; 292: 2735-2742.

71. Vreemann S, Gubern-Merida A, Lardenoije S, Bult P, Karssemeijer N, Pinker K, et al. The frequency of missed breast cancers in women participating in a high-risk MRI screening program. *Breast Cancer Res Treat* 2018; 169: 323-331.

72. Yamaguchi K, Schacht D, Newstead GM, Bradbury AR, Verp MS, Olopade OI, et al. Breast cancer detected on an incident (second or subsequent) round of screening MRI: MRI features of false-negative cases. *AJR Am J Roentgenol* 2013; 201: 1155-1163.

73. Kuhl CK, Schrading S, Bieling HB, Wardelmann E, Leutner CC, Koenig R, et al. MRI for diagnosis of pure ductal carcinoma in situ: a prospective observational study. The Lancet 2007;370(9586):485-92.

74. Schnall MD, Blume J, Bluemke DA, DeAngelis GA, DeBruhl N, Harms S, et al. Diagnostic architectural and dynamic features at breast MR imaging: multicenter study. *Radiology* 2006; 238: 42-53.

75. Sheth D, Giger ML. Artificial intelligence in the interpretation of breast cancer on MRI. *J Magn Reson Imaging* 2020; 51: 1310-1324.

76. Meyer-Bäse A, Morra L, Meyer-Bäse U, Pinker K. Current status and future perspectives of artificial intelligence in magnetic resonance breast imaging. *Contrast Media Mol Imaging* 2020; 2020.

77. Dalmış MU, Vreemann S, Kooi T, Mann RM, Karssemeijer N, Gubern-Mérida A. Fully automated detection of breast cancer in screening MRI using convolutional neural networks. *J Med Imaging (Bellingham)* 2018; 5: 014502.

78. Antropova NO, Abe H, Giger ML. Use of clinical MRI maximum intensity projections for improved breast lesion classification with deep convolutional neural networks. *Journal of Medical Imaging* 2018; 5: 014503.

79. Reig B, Heacock L, Geras KJ, Moy L. Machine learning in breast MRI. *J Magn Reson Imaging* 2020; 52: 998-1018.

80. Herent P, Schmauch B, Jehanno P, Dehaene O, Saillard C, Balleyguier C, et al. Detection and characterization of MRI breast lesions using deep learning. *Diagn Interv Imaging* 2019; 100: 219-225.

81. Adachi M, Fujioka T, Mori M, Kubota K, Kikuchi Y, Xiaotong W, et al. Detection and diagnosis of breast cancer using artificial intelligence based assessment of maximum intensity projection dynamic contrast-enhanced magnetic resonance images. *Diagnostics (Basel)* 2020; 10: 330.

82. Zhang Y, Chen J-H, Chang K-T, Park VY, Kim MJ, Chan S, et al. Automatic breast and fibroglandular tissue segmentation in breast MRI using deep learning by a fully-convolutional residual neural network U-net. *Acad Radiol* 2019; 26: 1526-1535.

83. Zhou J, Luo LY, Dou Q, Chen H, Chen C, Li GJ, et al. Weakly supervised 3D deep learning for breast cancer classification and localization of the lesions in MR images. *J Magn Reson Imaging* 2019; 50: 1144-1151.

84. Zhou J, Zhang Y, Chang KT, Lee KE, Wang O, Li J, et al. Diagnosis of benign and malignant breast lesions on DCE-MRI by using radiomics and deep learning with consideration of peritumor tissue. *J Magn Reson Imaging* 2020; 51: 798-809.

85. Truhn D, Schrading S, Haarburger C, Schneider H, Merhof D, Kuhl C. Radiomic versus convolutional neural networks analysis for classification of contrast-enhancing lesions at multiparametric breast MRI. *Radiology* 2019; 290: 290-297.

86. Ayatollahi F, Shokouhi SB, Mann RM, Teuwen J. Automatic breast lesion detection in ultrafast DCE-MRI using deep learning. *Med Phys* 2021; 48: 5897-5907.

87. Eskreis-Winkler S, Onishi N, Pinker K, Reiner JS, Kaplan J, Morris EA, et al. Using deep learning to improve nonsystematic viewing of breast cancer on MRI. *J Breast Imaging* 2021; 3: 201-207.

88. Jing X, Wielema M, Cornelissen LJ, van Gent M, Iwema WM, Zheng S, et al. Using deep learning to safely exclude lesions with only ultrafast breast MRI to shorten acquisition and reading time. *Eur Radiol* 2022: 1-10.

89. Wu Y, Wu J, Dou Y, Rubert N, Wang Y, Deng J. A deep learning fusion model with evidence-based confidence level analysis for differentiation of malignant and benign breast tumors using dynamic contrast enhanced MRI. *Biomed Signal Process Control* 2022; 72: 103319.

90. Vicini S, Bortolotto C, Rengo M, Ballerini D, Bellini D, Carbone I, et al. A narrative review on current imaging applications of artificial intelligence and radiomics in oncology: focus on the three most common cancers. *Radiol Med* 2022: 1-18: 819-836.

91. Parekh VS, Jacobs MA. Integrated radiomic framework for breast cancer and tumor biology using advanced machine learning and multiparametric MRI. *NPJ Breast Cancer* 2017; 3: 1-9.

92. Fusco R, Piccirillo A, Sansone M, Granata V, Runulotta MR, Patrosino T, et al. Radiomics and artificial intellidence analysis with textural metrics extracted by contrast-enhanced mammography in the breast lesions classification. *Diagnostics (Basel)* 2021; 11: 815.

93. Crivelli P, Ledda RE, Parascandolo N, Fara A, Soro D, Conti M. A new challenge for radiologists: radiomics in breast cancer. *Biomed Res Int* 2018 ;2018.

94. Bi WL, Hosny A, Schabath MB, Giger ML, Birkbak NJ, Mehrtash A, et al. Artificial intelligence in cancer imaging: clinical challenges and applications. *CA Cancer J Clin* 2019; 69: 127-57.

95. Gilbert F, Smye S, Schönlieb C-B. Artificial intelligence in clinical imaging: a health system approach. *Clin Radiol* 2020; 75: 3-6.

96. Tang A, Tam R, Cadrin-Chênevert A, Guest W, Chong J, Barfett J, et al. Canadian Association of Radiologists white paper on artificial intelligence in radiology. *Can Assoc Radiol J* 2018; 69: 120-135.

97. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine* 2019; 25: 44-56.

98. Mendelson EB. Artificial intelligence in breast imaging: potentials and limitations. *AJR Am J Roentgenol* 2019; 212: 293-299.

99. Reiner B. Contextualizing causation of uncertainty in medical reporting. *J Am Coll Radiol* 2018; 15: 325-327.

100. Sabottke CF, Spieler BM. The effect of image resolution on deep learning in radiography. *Radiol Artif Intell* 2020; 2: e190015.