Artificial intelligence in breast ultrasonography

ULTRA SONO GRAPHY

Jaeil Kim[1]*, Hye Jung Kim[2]*, Chanho Kim[1], Won Hwa Kim[2]

[1]School of Computer Science and Engineering, Kyungpook National University, Daegu; [2]Department of Radiology, School of Medicine, Kyungpook National University, Kyungpook National University Chilgok Hospital, Daegu, Korea

REVIEW ARTICLE

https://doi.org/10.14366/usg.20117
pISSN: 2288-5919 • eISSN: 2288-5943
Ultrasonography 2021;40:183-190

Although breast ultrasonography is the mainstay modality for differentiating between benign and malignant breast masses, it has intrinsic problems with false positives and substantial interobserver variability. Artificial intelligence (AI), particularly with deep learning models, is expected to improve workflow efficiency and serve as a second opinion. AI is highly useful for performing three main clinical tasks in breast ultrasonography: detection (localization/segmentation), differential diagnosis (classification), and prognostication (prediction). This article provides a current overview of AI applications in breast ultrasonography, with a discussion of methodological considerations in the development of AI models and an up-to-date literature review of potential clinical applications.

Keywords: Artificial intelligence; Breast neoplasm; Ultrasonography; Convolutional neural network; Breast diseases

Received: July 30, 2020
Revised: November 10, 2020
Accepted: November 12, 2020

Correspondence to:
Won Hwa Kim, MD, PhD, Department of Radiology, School of Medicine, Kyungpook National University, Kyungpook National University Chilgok Hospital, 807 Hoguk-ro, Buk-gu, Daegu 41404, Korea

Tel. +82-53-200-2527
Fax. +82-53-200-3349
E-mail: greenoaktree9@gmail.com

*These authors contributed equally to this work.

How to cite this article:
Kim J, Kim HJ, Kim C, Kim WH. Artificial intelligence in breast ultrasonography. Ultrasonography. 2021 Apr;40(2):183-190.

Introduction

Breast cancer is the most common type of cancer in women in Korea according to data from the Korea National Cancer Incidence Database [1]. Its incidence rate has been increasing, with an annual percentage rise of 6%, and is expected to further increase in the next 10 years. The rise in cases may be related to reproductive/lifestyle factors and an aging society [2]. These epidemiological findings underscore the importance of effective and accurate diagnoses of breast cancer using mammography and ultrasonography, which lead to an increased workload for radiologists. Although mammography is known to reduce breast cancer mortality, mammography is limited as a diagnostic modality because of its wide variability in interpretation and diagnostic performance among radiologists. A benchmark study showed that considerable rates of radiologists in the Breast Cancer Surveillance Consortium had suboptimal performance measures in terms of the abnormal interpretation rate and specificity [3]. Moreover, mammography has an intrinsic problem of imperfect sensitivity due to obscured cancers, especially in women with dense breasts. As breast density notification legislation in the United States has been widely implemented, radiologists are now required to inform women with dense breasts that additional supplemental screening modalities (e.g., ultrasonography) may be necessary [4,5].

Ultrasonography is the mainstay modality for differentiating between benign and malignant breast masses and has been traditionally used in the diagnostic setting. Due to growing evidence

that ultrasonography can detect mammographically occult cancers, interest in the use of ultrasonography for screening has increased [6,7]. Moreover, ultrasonography has several advantages in comparison with other modalities (e.g., mammography, digital breast tomosynthesis, and magnetic resonance imaging). It is generally safer (non-ionizing), more economical, easy to use, and allows real-time guidance and monitoring. However, interobserver variability is high in the acquisition and interpretation of ultrasound images, even among experts [8], which contributes to a high rate of false positives, leading to unnecessary biopsies and surgical procedures.

Deep learning (DL), as a subset of artificial intelligence (AI), has made great strides toward the automated detection and classification of medical images. For mammography, as a modality with commercially available DL-based decision support systems, recent validation studies have shown that several AI systems can perform at the level of radiologists. These studies suggest that AI has the potential to democratize expertise in settings with a lack of experienced radiologists. In addition, the radiologists' workload is reduced by improving workflow efficiency, and AI systems prevent overlooked findings or interpretation errors by giving a second opinion. These advantages may translate to ultrasonography. In recent years, DL algorithms have been increasingly applied to breast ultrasonography, mostly in feasibility studies for automated detection, differential diagnosis, and segmentation [9–15]. The development of DL-based AI systems for ultrasonography is in its early stages relative to mammography, and ultrasonography has unique characteristics in terms of the development process. This article provides a current overview of AI applications in breast ultrasonography, along with a discussion of methodological considerations in the development of these applications and an up-to-date literature review of potential clinical applications.

Methodological Considerations

Datasets

AI is a data-driven technology, and its performance is highly dependent on the quantity and quality of training data. To develop a robust AI model for ultrasonography, a multi-institutional large-scale dataset is required with a wide spectrum of diseases and non-disease entities, as well as images obtained from ultrasound devices from multiple vendors. Depending on the working conditions, the same lesion can be captured and interpreted differently, because more than 10 companies produce ultrasound equipment with various transducers and technical settings. In addition, ultrasound technology has evolved over the decades. Older ultrasonographic images are normally of lower resolution and have a higher noise level, while newer images are of higher resolution and have lower noise levels. Thus, AI algorithms trained with older images may not be externally valid for newer images.

The number of images used per patient in AI development has not been specified or standardized. Although most studies have used more than one image per patient for training and validation, specific details on the number of patients and images are needed based on the recently proposed Checklist for Artificial Intelligence in Medical Imaging [16]. Further studies or guidelines may be necessary to specify the structure and details of the dataset according to the type of dataset (training/validation/test) for a generalizable AI system, by minimizing selection and spectrum bias [17,18].

Image Preprocessing and Data Augmentation

An image annotation process involving manual delineation of the region of interest (ROI) of the lesion is usually required to train AI models in a supervised manner. The ROI can be automatically detected using various computer-aided segmentation techniques. However, human verification of the ROI is still required to guarantee the quality of training data. The need for a massive number (usually more than thousands) of annotated images is a barrier in the development of well-performing and robust AI systems, because the image annotation process is both time- and labor-intensive. In addition, manual annotation can be biased due to subjective prejudgment of the lesion character. To relax the requirements of manual ROI delineation in training data, weakly-supervised or semi-supervised methods are now emerging, in which unannotated images with only image-level labels (i.e., benignity and malignancy) are used for image classification and localization [19–21]. After image annotation, the images are usually cropped with a fixed margin around the ROI and resized. The margin is defined as the distance between the lesion boundary and the boundary of the cropped image itself. In a previous study, a 180-pixel margin showed the best performance [15]; however, cropping with variable margins (0–300 pixels) has been used in AI studies of breast ultrasonography. To input cropped images into AI models, it is necessary to resize cropped images to a fixed size, which also varies from study to study.

Data augmentation is commonly used to avoid overfitting and to increase the volume of training data [22,23]. Data augmentation is a process of creating new data (images) by manipulating the original data using a variety of augmentation strategies, including flipping, rotation, translation, and noise injection (Fig. 1). Even though resizing and data augmentation are essential steps for AI model training, these processes also carry the risk of reducing classification performance by altering some breast lesion attributes in ultrasonographic images. Byra et al. [13] suggested that images should not be rotated so that the longitudinal direction is shifted.

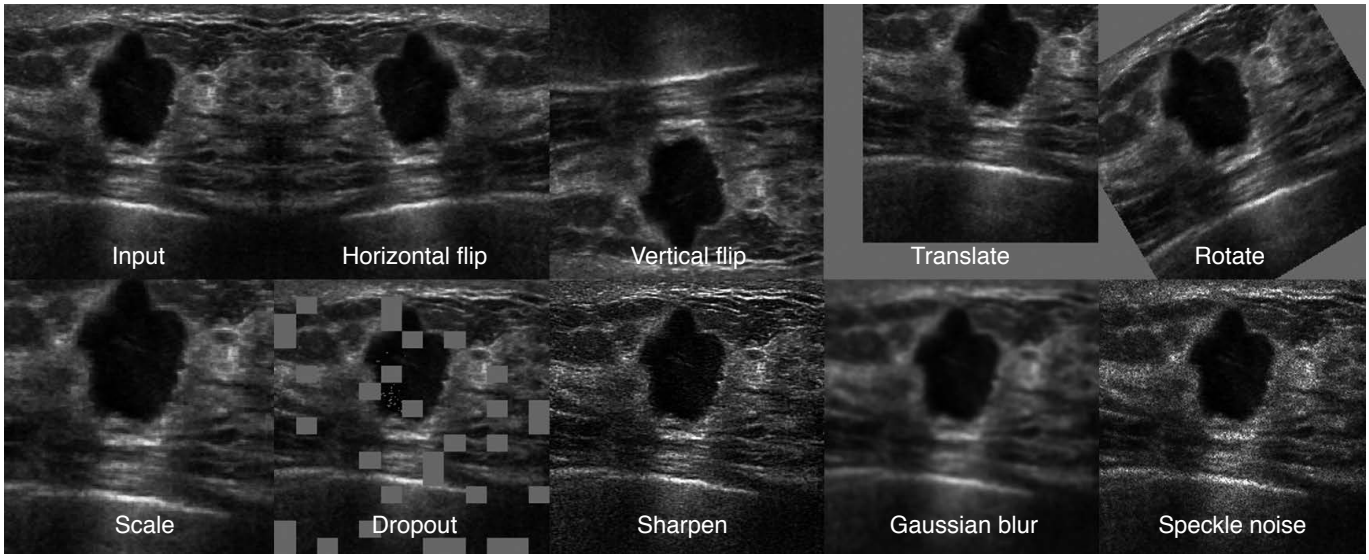

Fig. 1. **Examples of ultrasonographic image augmentation for convoluted neural network architectures.** Augmentation methods in the first row and scale in the second row show geometric transformations of ultrasound images. The rest of the figures present photometric methods to augment training datasets with random changes in image appearance.

For example, the posterior acoustic shadowing of a breast mass, which is one of the signs of malignancy, can be located anteriorly by longitudinal flipping [13].

Explainable AI

Since decisions derived from AI systems affect clinical decisions and/or outcomes, there is a need to understand the AI decision-making process. Considering the deep nature of current AI techniques with hundreds of layers and millions of parameters, this is a "black box" problem, in which AI output lacks explainability and justification. Thus, eXplainable AI (XAI) is now a crucial trend in the deployment of responsible AI models in medical imaging. One of the well-known approaches in XAI is the class activation mapping (CAM), which provides weighted feature maps of each class at the last convolution layers. CAM helps to understand the decision-making process by mapping the output back to the input image to see which parts of the input image were discriminative for the output [24–26]. With a CAM, breast lesions can be recognized (localized) in ultrasonographic images, and this localization is relevant for classification (Figs. 2, 3). Recent approaches and related issues have been reviewed [27–30]. However, much work remains to be done in the clinical interpretation of the explainable outputs provided by AI models.

Potential Clinical Applications

In breast ultrasonography, AI performs three main clinical tasks: detection (localization or segmentation), differential diagnosis (classification), and prognostication (prediction).

Detection (Localization or Segmentation)

As with other medical imaging modalities, DL-based lesion detection on breast ultrasonography is mostly performed by convolutional neural networks (CNNs). The CNN-based detection methods have shown superior accuracy in object detection compared with conventional computerized methods (i.e., radial gradient index filtering and multifractal filtering). For hand-held ultrasound (HHUS) images, Yap et al. [31] reported the performance of DL models with three different DL-based methods (a patch-based LeNet, a U-Net, and a transfer learning approach with FCN-AlexNet). They found that the transfer-learned FCN-AlexNet outperformed the other methods, with a true positive fraction (TPF) ranging from 0.92 to 0.98. Kumar et al. [32] proposed an ensemble model of multi U-Net models for automated segmentation of suspicious breast masses seen on ultrasonography, and reported that the model showed a TPF of 0.84.

DL research using HHUS has intrinsic limitations because still images are obtained from HHUS after a decision is made on whether and how to capture a certain portion of a lesion or an anatomical structure; thus, the process of image acquisition is highly dependent on the human imager. Hence, the clinical need for AI applications using HHUS necessarily extends to real-time detection. Zhang et al. [33] embedded a lightweight neural network, which was trained using a knowledge distillation technique to transfer knowledge

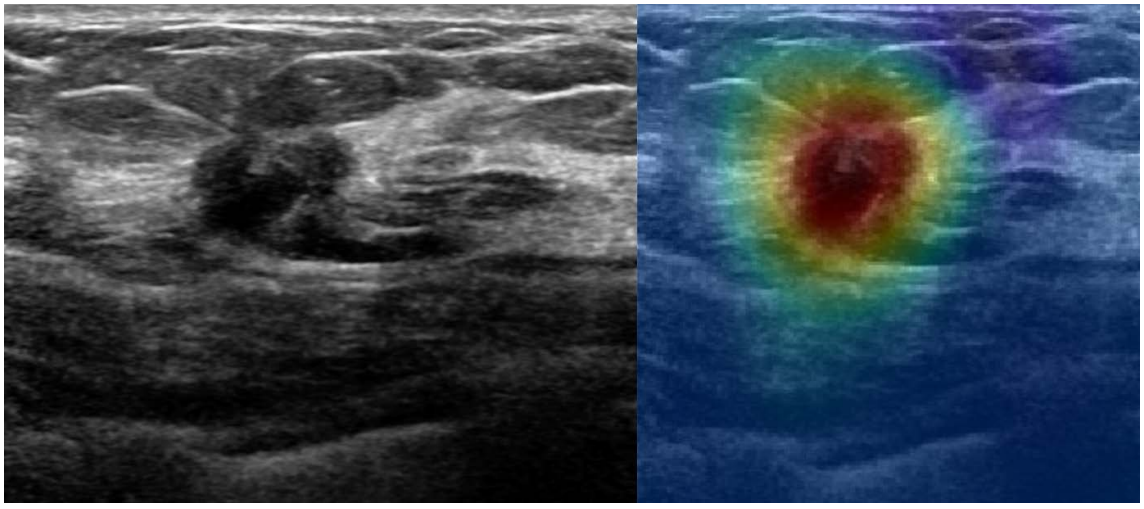

Fig. 2. **Class activation mapping (CAM) for a malignant breast mass on an ultrasound image.** The left is a pre-processing image and the right is an image of the overlapping CAM using a convolutional neural network (CNN). It can be seen that the CNN recognized the malignant mass well, and the probability of malignancy predicted by the CNN model was 99.25%.

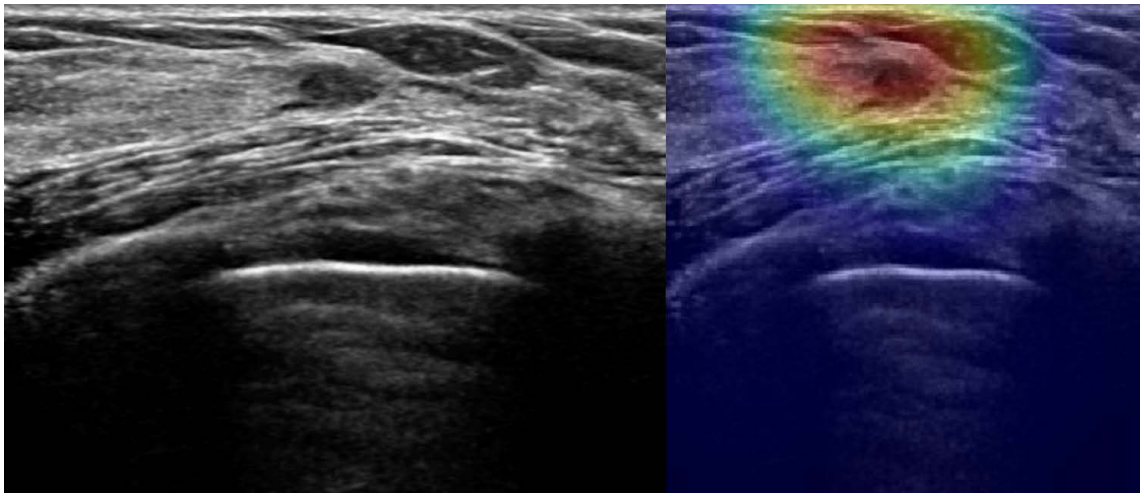

Fig. 3. **Class activation mapping (CAM) for a benign breast mass on an ultrasound image.** The left is a pre-processing image and the right is an image of the overlapping CAM using a convolutional neural network (CNN). It can be seen that the CNN recognized the benign mass well, and the probability of benignity predicted by the CNN model was 99.25%.

from deeper models to the shallow network, into ultrasonographic equipment. They reported successful test performance using real-time equipment at 24 frames per second [33]. This is the only study so far to implement real-time automated detection using single-source data, but it lacked performance metrics; thus, further studies are required to ascertain the potential clinical applications of AI to HHUS.

Automated breast ultrasonography (ABUS) is the modality of AI-powered lesion detection that is most strongly expected to assist radiologists in performing initial screenings and reducing the need for observational oversight because the thousands of images per patient generated by ABUS necessitate a prolonged interpretation time. Several computer-aided detection (CAD) algorithms, including the commercial software QV CAD (QView Medical, Los Altos, CA, USA), have been developed [34–37]. Studies have shown that the QV CAD system, with marks potentially malignant lesions, was helpful for radiologists (especially less-experienced radiologists) to improve cancer detection and reduce interpretation time [38]. Moon et al. [39] and Chiang et al. [40] proposed a 3-D CNN with a sliding window method and achieved high sensitivity (91%–100%) with a false-positive rate per case of 3.6%–21.6%.

Differential Diagnosis (Classification)

Differential diagnosis refers to the process of distinguishing a particular disease from others, and in the context of breast imaging, it usually refers to a distinction between benignity and malignancy. In clinical practice, the Breast Imaging Reporting and Data System (BI-RADS) developed by the American College of Radiology is used to standardize the reporting of breast ultrasound interpretation. Although the BI-RADS provides a systemic approach for lesion characterization and assessment, interobserver and intraobserver variability has been a subject of intense scrutiny, and AI solutions are expected to provide more reliable diagnoses [41,42].

Byra et al. [13] presented a CNN model with a transfer learning strategy using the VGGNet-19 which was pretrained on the ImageNet data set and fine-tuned on 882 breast ultrasound images with a matching layer to classify breast lesions as benign or malignant. The area under the receiver operating characteristic curve (AUC) of the better-performing CNN model was significantly greater than the highest AUC value for the radiologists (0.936 vs. 0.882). Other DL studies with CNN variants trained from scratch have demonstrated comparable or even higher diagnostic performance relative to radiologists [9,12,43]. Despite these promising results, further studies are warranted to prove the clinical utility of AI-powered classification systems. Most of the algorithms developed in up-to-date studies were trained through images obtained from a limited number of institutions and ultrasound systems; therefore, the developed algorithms do not necessarily perform well in different circumstances. Furthermore, ultrasound still images taken by an imager that capture a certain portion of the lesion, instead of viewing the whole lesion range, may contribute to under-representation or exaggeration of the ground-truth characteristics.

Han et al. [15] employed a transferred GoogLeNet model on 7,408 breast ultrasound images (4,254 benign and 3,154 malignant), which showed an AUC >0.9. This model is a component algorithm of S-Detect, which is a commercial CAD program embedded in ultrasound equipment that provides an automatic analysis based on BI-RADS descriptors. It has been implemented in the RS80A ultrasound machine (Samsung Medison Co. Ltd., Seoul, Korea). With its feature extraction technique and support vector machine classifier, it predicts the final assessment of breast masses in a dichotomized form (possibly benign or possibly malignant). Choi et al. [44] found that significant improvements in AUC were seen with CAD (0.823–0.839 vs. 0.623–0.759), especially for less-experienced radiologists. However, this program has limited applications with other vendors and requires user-defined lesion annotation, which is also a time-consuming process.

Differential diagnosis concurrent with detection or segmentation using DL models has been reported in various studies. Yap et al. [45] simultaneously performed both localization and classification using their model, and obtained sensitivity of 0.80–0.84 and 0.38–0.57 and dice scores of 0.72–0.76 and 0.33–0.76 for benign and malignant masses, respectively. Shin et al. [46] proposed a semi-supervised method, for which a small number of extensively annotated images and a larger number of image-level labeled images were used for model training, and they reported a 4.5 percentage point improvement in the correct localization measure, compared with the conventional fully-supervised methods with only the extensively annotated images, although quantitative metrics for classification were not presented. Kim et al. [47] introduced a weakly-supervised deep network with box convolution to detect suspicious regions of breast masses with various sizes and shapes in AI-based differential diagnosis. The box convolution method (accuracy, 89%) more accurately classified breast masses than conventional CNN models (86%–87%), by learning the clinically relevant features of the masses and their surrounding areas. Moreover, their proposed network provided more robust discrimination localization than conventional methods (AUC, 0.89 vs. 0.75–0.78). Although these efforts toward a framework for both localization and classification of breast lesions with less manual annotation are still on the level of feasibility studies, without clear performance metrics or validation through a multi-reader study design, they will provide the groundwork for AI application in ABUS or real-time ultrasonography.

Prognostication (Prediction)

Prognostication in breast cancer patients is usually conducted to predict the histopathological characteristics of the tumor before surgery or treatment, as well as treatment response and survival time. AI technologies involving prognostication have been rarely studied, because they are now beginning to be understood mostly at the level of detection and diagnosis. Several studies have attempted to predict axillary nodal status, which is of clinical significance because it guides treatment selection (i.e., the type of axillary surgery) [48]. Zhou et al. [49] found that the best-performing CNN model yielded a satisfactory prediction, with an AUC of 0.89–0.90, a sensitivity of 82%–85%, and a specificity of 72%–79% and this model outperformed three experienced radiologists in the receiver operating characteristic space. Another study reported that the DL radiomics of conventional ultrasonography and shear wave elastography, combined with clinical parameters, showed the best performance in predicting axillary lymph node metastasis, with an AUC of 0.902 [50]. In addition, this model could also discriminate between a low and heavy metastatic axillary nodal burden, with an AUC of 0.905.

AI with DL is expected to further reveal information that human

experts cannot recognize and integrate imaging features and clinical variables. Further studies may provide insights into whether and how AI-aided predictions of clinical outcomes can be made with superior and reliable accuracy compared with human-crafted features. For now, very limited DL studies have been published on predictions of tumor response to neoadjuvant chemotherapy using magnetic resonance imaging [51,52].

Summary

AI has tremendous potential to contribute to workflow efficiency and the reduction of interobserver variability in breast ultrasonography. The studies reviewed in this article have reported potential clinical applications of AI for breast cancer detection, characterization, and prognostication using ultrasonography. However, some methodological considerations should be carefully considered for the development of more robust and responsible AI systems.

ORCID: Jaeil Kim: https://orcid.org/0000-0002-9799-1773; Hye Jung Kim: https://orcid.org/0000-0002-0263-0941; Chanho Kim: https://orcid.org/0000-0002-7743-040X; Won Hwa Kim: https://orcid.org/0000-0001-7137-9968

Author Contributions

Conceptualization: Kim J, Kim HJ, Kim WH. Data acquisition: Kim J, Kim HJ, Kim C, Kim WH. Data analysis or interpretation: Kim J, Kim HJ, Kim C, Kim WH. Drafting of the manuscript: Kim J, Kim HJ, Kim WH. Critical revision of the manuscript: Kim C, Kim WH. Approval of the final version of the manuscript: all authors.

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

Acknowledgments

This work was supported by a National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. 2020R1C1C1006453, No.2019R1G1A1098655) and Ministry of Education (No. 2020R1I1A3074639), and by the AI-based image analysis solution development program for diagnostic medical imaging devices (20011875, Development of AI based diagnostic technology for medical imaging devices) funded by the Ministry of Trade, Industry & Energy (MOTIE, Korea) and the Korean Society of Breast Imaging & Korean Society for Breast Screening (KSBI & KSFBS-2017-01).

References

1. Jung KW, Won YJ, Hong S, Kong HJ, Lee ES. Prediction of cancer incidence and mortality in Korea, 2020. Cancer Res Treat 2020;52:351-358.
2. Lee JE, Lee SA, Kim TH, Park S, Choy YS, Ju YJ, et al. Projection of breast cancer burden due to reproductive/lifestyle changes in Korean women (2013-2030) using an age-period-cohort model. Cancer Res Treat 2018;50:1388-1395.
3. 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;283:49-58.
4. Price ER, Hargreaves J, Lipson JA, Sickles EA, Brenner RJ, Lindfors KK, et al. The California breast density information group: a collaborative response to the issues of breast density, breast cancer risk, and breast density notification legislation. Radiology 2013;269:887-892.
5. Rhodes DJ, Radecki Breitkopf C, Ziegenfuss JY, Jenkins SM, Vachon CM. Awareness of breast density and its impact on breast cancer detection and risk. J Clin Oncol 2015;33:1143-1150.
6. Brem RF, Lenihan MJ, Lieberman J, Torrente J. Screening breast ultrasound: past, present, and future. AJR Am J Roentgenol 2015;204:234-240.
7. Vourtsis A, Berg WA. Breast density implications and supplemental screening. Eur Radiol 2019;29:1762-1777.
8. Berg WA, Blume JD, Cormack JB, Mendelson EB. Operator dependence of physician-performed whole-breast US: lesion detection and characterization. Radiology 2006;241:355-365.
9. Fujioka T, Kubota K, Mori M, Kikuchi Y, Katsuta L, Kasahara M, et al. Distinction between benign and malignant breast masses at breast ultrasound using deep learning method with convolutional neural network. Jpn J Radiol 2019;37:466-472.
10. Tanaka H, Chiu SW, Watanabe T, Kaoku S, Yamaguchi T. Computer-aided diagnosis system for breast ultrasound images using deep learning. Phys Med Biol 2019;64:235013.
11. Cao Z, Duan L, Yang G, Yue T, Chen Q. An experimental study on breast lesion detection and classification from ultrasound images using deep learning architectures. BMC Med Imaging 2019;19:51.
12. Ciritsis A, Rossi C, Eberhard M, Marcon M, Becker AS, Boss A. Automatic classification of ultrasound breast lesions using a deep convolutional neural network mimicking human decision-making. Eur Radiol 2019;29:5458-5468.
13. Byra M, Galperin M, Ojeda-Fournier H, Olson L, O'Boyle M, Comstock C, et al. Breast mass classification in sonography with transfer learning using a deep convolutional neural network and color conversion. Med Phys 2019;46:746-755.
14. Park HJ, Kim SM, La Yun B, Jang M, Kim B, Jang JY, et al. A computer-aided diagnosis system using artificial intelligence for

the diagnosis and characterization of breast masses on ultrasound: added value for the inexperienced breast radiologist. Medicine (Baltimore) 2019;98:e14146.
15. Han S, Kang HK, Jeong JY, Park MH, Kim W, Bang WC, et al. A deep learning framework for supporting the classification of breast lesions in ultrasound images. Phys Med Biol 2017;62:7714-7728.
16. Norgeot B, Quer G, Beaulieu-Jones BK, Torkamani A, Dias R, Gianfrancesco M, et al. Minimum information about clinical artificial intelligence modeling: the MI-CLAIM checklist. Nat Med 2020;26:1320-1324.
17. Park SH, Han K. Methodologic guide for evaluating clinical performance and effect of artificial intelligence technology for medical diagnosis and prediction. Radiology 2018;286:800-809.
18. Yu AC, Eng J. One algorithm may not fit all: how selection bias affects machine learning performance. Radiographics 2020;40:1932-1937.
19. Jager M, Knoll C, Hamprecht FA. Weakly supervised learning of a classifier for unusual event detection. IEEE Trans Image Process 2008;17:1700-1708.
20. Oquab M, Bottou L, Laptev I, Sivic J. Is object localization for free? Weakly-supervised learning with convolutional neural networks. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition; 2015 Jun 7-12; Boston, MA, USA. Piscatawa, NJ: Institute of Electrical and Electronics Engineers, 2015;685-694.
21. Cheplygina V, de Bruijne M, Pluim JP. Not-so-supervised: a survey of semi-supervised, multi-instance, and transfer learning in medical image analysis. Med Image Anal 2019;54:280-296.
22. Gupta V, Demirer M, Bigelow M, Little KJ, Candemir S, Prevedello LM, et al. Performance of a deep neural network algorithm based on a small medical image dataset: incremental impact of 3D-to-2D reformation combined with novel data augmentation, photometric conversion, or transfer learning. J Digit Imaging 2020;33:431-438.
23. Hussain Z, Gimenez F, Yi D, Rubin D. Differential data augmentation techniques for medical imaging classification Tasks. AMIA Annu Symp Proc 2017;2017:979-984.
24. Lee H, Yune S, Mansouri M, Kim M, Tajmir SH, Guerrier CE, et al. An explainable deep-learning algorithm for the detection of acute intracranial haemorrhage from small datasets. Nat Biomed Eng 2019;3:173-182.
25. Zhou B, Khosla A, Lapedriza A, Oliva A, Torralba A. Learning deep features for discriminative localization. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition; 2016 Jun 27-30; Las Vegas, NV, USA. Piscatawa, NJ: Institute of Electrical and Electronics Engineers, 2016;2921-2929.
26. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: visual explanations from deep networks via gradient-based localization. In: 2017 IEEE International Conference on Computer Vision; 2017 Oct 22-29; Venice, Italy. Piscatawa, NJ: Institute of Electrical and Electronics Engineers, 2017;618-626.
27. Molnar C. A guide for making black box models explainable [Internet]. Morrisville, NC: Lulu, 2019 [cited 2019 Jul 30]. Available from: https://christophm.github.io/interpretable-ml-book.
28. Gilpin LH, Bau D, Yuan BZ, Bajwa A, Specter M, Kagal L. Explaining explanations: an overview of interpretability of machine learning. In: 2018 IEEE 5th International Conference on Data Science and Advanced Analytics; 2018 Oct 1-3; Turin, Italy. Piscatawa, NJ: Institute of Electrical and Electronics Engineers, 2019.
29. Ras G, van Gerven M, Haselager P. Explainable methods in deep learning: users, values, concerns and challenges. Preprint at https://arxiv.org/abs/1803.07517 (2018).
30. Hamon R, Junklewitz H, Sanchez I. Robustness and explainability of artificial intelligence. Luxembourg: Publications Office of the European Union, 2020.
31. Yap MH, Pons G, Marti J, Ganau S, Sentis M, Zwiggelaar R, et al. Automated breast ultrasound lesions detection using convolutional neural networks. IEEE J Biomed Health Inform 2018;22:1218-1226.
32. Kumar V, Webb JM, Gregory A, Denis M, Meixner DD, Bayat M, et al. Automated and real-time segmentation of suspicious breast masses using convolutional neural network. PLoS One 2018;13:e0195816.
33. Zhang X, Lin X, Zhang Z, Dong L, Sun X, Sun D, et al. Artificial intelligence medical ultrasound equipment: application of breast lesions detection. Ultrason Imaging 2020;42:191-202.
34. Ikedo Y, Fukuoka D, Hara T, Fujita H, Takada E, Endo T, et al. Development of a fully automatic scheme for detection of masses in whole breast ultrasound images. Med Phys 2007;34:4378-4388.
35. Chang RF, Chang-Chien KC, Takada E, Huang CS, Chou YH, Kuo CM, et al. Rapid image stitching and computer-aided detection for multipass automated breast ultrasound. Med Phys 2010;37:2063-2073.
36. Moon WK, Lo CM, Chen RT, Shen YW, Chang JM, Huang CS, et al. Tumor detection in automated breast ultrasound images using quantitative tissue clustering. Med Phys 2014;41:042901.
37. Moon WK, Shen YW, Bae MS, Huang CS, Chen JH, Chang RF. Computer-aided tumor detection based on multi-scale blob detection algorithm in automated breast ultrasound images. IEEE Trans Med Imaging 2013;32:1191-1200.
38. Yang S, Gao X, Liu L, Shu R, Yan J, Zhang G, et al. Performance and reading time of automated breast US with or without computer-aided detection. Radiology 2019;292:540-549.
39. Moon WK, Huang YS, Hsu CH, Chang Chien TY, Chang JM, Lee SH, et al. Computer-aided tumor detection in automated breast ultrasound using a 3-D convolutional neural network. Comput Methods Programs Biomed 2020;190:105360.
40. Chiang TC, Huang YS, Chen RT, Huang CS, Chang RF. Tumor detection in automated breast ultrasound using 3-D CNN and prioritized candidate aggregation. IEEE Trans Med Imaging 2019;38:240-249.
41. Abdullah N, Mesurolle B, El-Khoury M, Kao E. Breast imaging

reporting and data system lexicon for US: interobserver agreement for assessment of breast masses. Radiology 2009;252:665-672.
42. 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.
43. 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.
44. Choi JS, Han BK, 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.
45. Yap MH, Goyal M, Osman FM, Martí R, Denton E, Juette A, et al. Breast ultrasound lesions recognition: end-to-end deep learning approaches. J Med Imaging (Bellingham) 2019;6:011007.
46. Shin SY, Lee S, Yun ID, Kim SM, Lee KM. Joint weakly and semi-supervised deep learning for localization and classification of masses in breast ultrasound images. IEEE Trans Med Imaging 2019;38:762-774.
47. Kim C, Kim WH, Kim HJ, Kim J. Weakly-supervised US breast tumor characterization and localization with a box convolution network. In: Proc SPIE 11314, Medical Imaging 2020: Computer-Aided Diagnosis; 2020 Feb 16-19; Houston, TX, USA. Bellingham, WA: SPIE, 2020;1131419.
48. Chang JM, Leung JW, Moy L, Ha SM, Moon WK. Axillary nodal evaluation in breast cancer: state of the art. Radiology 2020;295:500-515.
49. Zhou LQ, Wu XL, Huang SY, Wu GG, Ye HR, Wei Q, et al. Lymph node metastasis prediction from primary breast cancer US images using deep learning. Radiology 2020;294:19-28.
50. Zheng X, Yao Z, Huang Y, Yu Y, Wang Y, Liu Y, et al. Deep learning radiomics can predict axillary lymph node status in early-stage breast cancer. Nat Commun 2020;11:1236.
51. Ha R, Chin C, Karcich J, Liu MZ, Chang P, Mutasa S, et al. Prior to initiation of chemotherapy, can we predict breast tumor response? Deep learning convolutional neural networks approach using a breast MRI tumor dataset. J Digit Imaging 2019;32:693-701.
52. Ha R, Chang P, Karcich J, Mutasa S, Van Sant EP, Connolly E, et al. Predicting post neoadjuvant axillary response using a novel convolutional neural network algorithm. Ann Surg Oncol 2018;25:3037-3043.