Artificial Hype

AI applications in breast imaging

Review of some state-of-the-art applications of artificial intelligence on mammography and MRI.

Computer aided imaging is not novel, having been around for 50 years. Developments have boosted the accuracy of computer-based analysis and breast imaging is at the forefront, as large databases are available, and radiologists tasks on images are relatively error prone.

Jonas Teuwen
Assistant Professor
Radboud University
Medical Centre
Nijmegen, the Netherlands

Netherlands Cancer Institute /
Antoni van Leeuwenhoek hospital
Amsterdam, the Netherlands

🐦 @jonasteuwen

Jonas.Teuwen@radboudumc.nl

Artificial intelligence is the buzzword of today for radiology. Still computer aid for imaging is not novel, but has been around for approximately 50 years. Recent developments have boosted the accuracy of computer-based analysis in many fields. Among these, breast imaging applications are at the forefront, both because they are very commonly used and therefore large databases are available, and the tasks of radiologists on these images are relatively error prone. In this article we highlight some of the applications of artificial intelligence on mammography and MRI.

Mammography

Reading mammograms in a screening setting is one of the most difficult tasks in radiology. Even in a double reader setting where two radiologists rate the same exam, breast cancer is missed relatively often. To improve upon this, computer-aided detection (CAD) systems were developed. It was assumed that if a CAD system displays suspicious areas, radiologists would not miss them. However, in practice the use of CAD marks to highlight suspicious lesions was far from perfect. The large amount of false positive findings marked by the CAD systems were considered to be a distraction and resulted in a perceived low reliability of the systems, and therefore limited use in clinical practice.

A further difficulty is that a study has shown that radiologists not so much miss suspicious areas, but that a correct classification of observed potential abnormalities is the actual problem. Consequently, a system that supports radiologists with the decision to refer a woman for further examination appears to be more effective than a classical CAD-system, which intends to reduce detection errors (Hupse et al. 2013a). The detection system used in this study was built before deep learning techniques were introduced into medical imaging but already came quantitatively close to the performance of radiologists (Hupse et al. 2013b). In contrast to these classical systems, which use carefully hand-crafted features designed to capture certain characteristics of lesions such as spiculation (Karssmeijer and Te Blake 1996), deep learning-based systems learn these features from the annotated data allowing them to surpass the diagnostic accuracy of the classical systems and achieve performances previously assumed to be only within the human realm.

Nikita Moriakov
Diagnostic Image Analysis Group
Radboud University
Medical Centre
Nijmegen, the Netherlands

Nikita.Moriakov@radboudumc.nl

> ❝READING MAMMOGRAMS IN A SCREENING SETTING IS ONE OF THE MOST DIFFICULT TASKS IN RADIOLOGY ❞

Current deep learning systems allow determining the probability of a suspicious region to be a carcinoma, whether it is a soft-tissue lesion or calcifications with high accuracy. Several research and commercially available AI-based systems are now available for mammography analysis. These systems have an accuracy that is on par with that of average, but dedicated, breast radiologists on heterogeneous datasets of mammograms (Ribli et al. 2018; Rodríguez-Ruiz et al. 2019).

Ritse Mann
Breast and Interventional
Radiologist
Head of Clinical Breast Research
Radiology Department

Radboud University
Medical Centre
Nijmegen, the Netherlands

Netherlands Cancer Institute /
Antoni van Leeuwenhoek Hospital
Amsterdam, the Netherlands

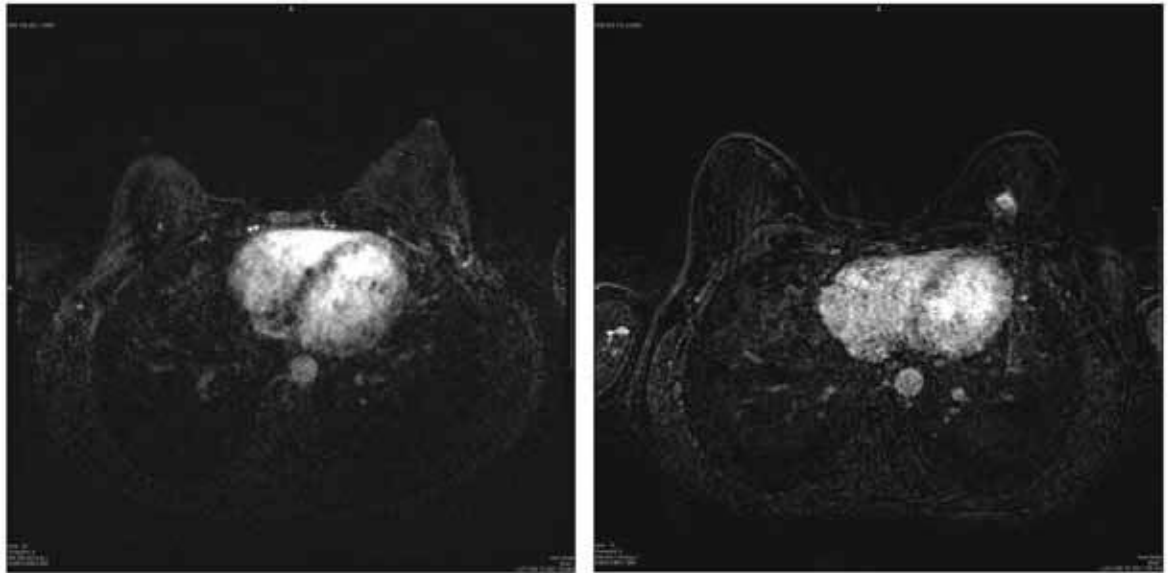🐦 @RitseMann

Ritse.Mann@radboudumc.nl

Figure 1. Two MRI acquisitions of suspicious areas which lead to a biopsy. Both images are of young women with a BRCA1 mutation. Left a TWIST acquisition with a BI-RADS 3 lesion centrodorsal in the right breast which was proven to be a small fibroadenoma. To the right a VIBE acquisition of a BI-RADS 4 lesion. The pathology was proven to be an inflamed cyst. Both lesions were declared not suspicious by the CAD system.

However, some radiologists still outperform even these AI-systems. This is likely due to the fact that not all available information is currently being used by these AI-systems. For instance, the temporal information provided by previous studies is not exploited with such systems. It is expected that the performance of AI-systems can be extended beyond the performance of an average breast radiologist by including these factors. In a newly funded project, systems will be designed which also take into account suspicious temporal changes.

It does not end there. In the coming years, and as is already happening in the United States, digital mammography will be replaced by digital breast tomosynthesis. While this system is more sensitive than mammography, this does not imply that this makes the task of recognition of suspicious areas easier. The larger amount of data and increased reading time further complicate this. Current research therefore also focuses on the applications of deep learning techniques to digital breast tomosynthesis. The main complication is that as the technique is rather new, no large screening datasets with proven malignancies and sufficient follow-up are available. This is certainly a problem for the deep learning algorithms as these derive the discriminative features from the (annotated) data itself, and therefore require large datasets to achieve a satisfactory performance. Because of this, researchers employ "transfer learning" techniques. In

this setting the system learns discriminative features on a different dataset such as mammography, and these features are subsequently transferred and fine-tuned for the tomosynthesis deep learning detection system on a tomosynthesis dataset. A recent study showed that an AI-based CAD system for DBT allows for faster reading without decreasing radiologists performance (Chae et al. 2018).

MRI

While mammography has shown to be a cost-effective method to reduce mortality of breast cancer over the past decades, it is known that in certain cases carcinomas tend to be less visible on mammograms. For instance, mammography is proven to be less sensitive for women with high mammography density (Wanders et al. 2017). This is not the case for breast MRI, where carcinomas can be detected with high sensitivity even for breasts with high density. In the DENSE trail (Emaus et al. 2015), which is to be presented at ECR 2019, women in the highest density category (ACR D) are invited for a complementary breast MRI. Next to studying the amount of screen detected carcinomas and the amount of false positives, the effect of breast MRI on the amount of interval cancers is also studied.

A study (Kuhl et al. 2017) has shown that with the addition of a breast MRI scan after a negative screening mammogram, an additional 15.5 carcinomas per 1000 can be detected. Unfortunately MRI
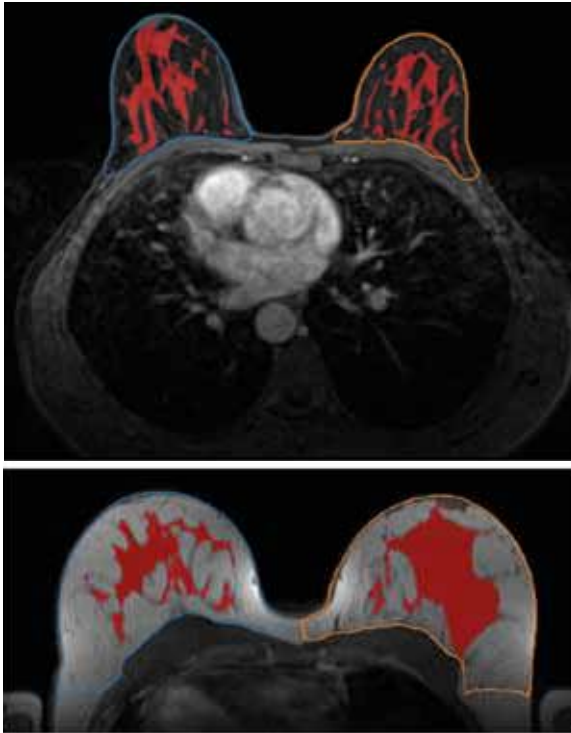
Figure 2. T1w acquisitions. Left is a fat saturated acquisition. The deep learning system is trained to segment the breast shape and the fibroglandular tissue. It is robust to changes in acquisition parameter settings and different sequences.

is not yet broadly applicable as a screening method for breast cancer due to the large associated costs, and for this reason it is only being used for women with an increased risk of breast cancer (Mann et al. 2008). The significantly increased reading time of a breast MRI exam compared to a mammogram adds to the limited applicability of breast MRI in a screening setting.

While MRI has a high sensitivity for the detection of breast cancer, it also associated with a percentage wise similar increase in the number of false positive findings that further complicates the application of MRI in a screening setting. Next to this, several studies (Yamaguchi et al. 2013; Pages et al. 2012) have shown that between 47 and 58% of the earlier detected carcinomas were already visible in earlier screening rounds. One of our studies shows that in retrospect almost one third of all cancers was already visible and actionable on an earlier MRI (Vreemann et al. 2018a). However, one should balance this against the positive predictive value (PPV) ranging from 14 to 37% depending on the MRI-screening indication, being higher in patients at higher risk, and the programme based high overall sensitivity of screening with MRI (90%) (Vreemann et al. 2018b).

Our results show that a number of cancers that are missed by the radiologist can be detected by the CAD-system. In this study, we looked at all cancers that were classified as negative in a previous screening round (BI-RADS 1 or 2) but in retrospect were visible when these were detected upon follow up examination one year later. Such a system can therefore support the radiologist by denoting the suspicious areas after a negative classification. Our results show that for these cases 70% sensitivity can be reached with one false-positive finding per scan (Dalmiş et al. 2019; **Figure 1**).

The current state-of-the-art breast MRI protocol consists of multiple sequences and lasts about 15 minutes. To make MRI more available for a screening setting, the costs of the technique should be lowered and therefore a lot of research is going into abbreviated MRI-protocols. In such an abbreviated protocol the pre- and post-contrast T1 acquisitions are acquired in the earlier phases after the administration of the contrast agent but the later T1w, T2w and DWI acquisitions, which often occur in the complete protocol, are left out and the final decision is made on the basis of the available morphological information out of the first post-contrast subtraction.

To decrease the reading time, the Maximal Intensity Projection (MIP) of this volume is studied. In case there is a suspicious area, the complete volume is considered (Kuhl et al. 2014). While this significantly reduces the average reading time, studies have also shown that the use of the MIP can increase the number of reading and interpretation errors (Mango et al. 2015). A CAD deep learning system, which we have developed for this purpose, uses all images of this abbreviated protocol, and alerts the reader when potential findings that are not evident in the MIP images are present.

In a diagnostic setting we want to use a different CAD-system to support clinicians in deciding whether or not to acquire a biopsy. Just as with the previous system, this system can be used with an abbreviated protocol where the CAD-system assigns a malignancy score to a radiologist labelled region. As this system predicts the likelihood of malignant biopsy results, it has the potency to reduce the number of biopsies. Our results show that while maintaining high sensitivity it is possible to reduce at least 20% of all biopsies (Dalmiş et al. 2019).

To be able to use such a system in the clinic, we need to ensure that the model is robust to scanner variations which is quite pronounced for MRI scanners. We are developing methods that are robust against such variations. One typical example is to determine

the breast density on breast MRI. To do this prop-
erly, we need to have an accurate segmentation of
both the breast shape and the fibroglandular tissue.
Previous methods would build different models for
different sequences, making these models less usable
for epidemiological studies. In **Figure 2** we provide an
output of such a robust segmentation model.

Based on these positive results, and the success in
creating robust deep learning systems, we expect that
deep learning will contribute significantly to increase
the application areas of MRI and make it an economi-
cally viable breast cancer screening method. Not only
will this open more ways to detect breast cancer earlier
and reduce mortality, but will also decrease the vari-
ance in performance between radiologists and improve
the screening programme as a whole by supporting
less experienced radiologists with their decisions. Still
it should be noted that the applications of AI for breast
MRI have not yet left the research domain. Imple-
menting these in clinical practice and proving their
efficiency will be a major task for the future. Surely,
there are exciting times ahead for the applications of
AI in breast imaging. ∎

KEY POINTS

✓ Reading mammograms in a screening setting is one of the most difficult tasks in radiology.

✓ A system that supports radiologists with the decision to refer a woman for further examination appears to be more effective than a classical CAD-system.

✓ AI systems are now available for mammography analysis with accuracy that is on par with that of average, but dedicated, breast radiologists on hetero-geneous datasets of mammograms.

✓ Deep learning will contribute significantly to increase the application areas of MRI and make it an economically viable breast cancer screening method.

✓ AI will open more ways to detect breast cancer earlier and reduce mortality; decrease the variance in performance between radiologists and improve the screening programme as a whole by supporting less experienced radiologists with their decisions.

REFERENCES

Chae EY, Kim HH, Jeong JW, Chae SH, Lee S, Choi YW (2018) Decrease in interpretation time for both novice and experienced readers using a concurrent computer-aided detection system for digital breast tomosynthesis. European Radiology. Available from: ncbi.nlm.nih.gov/m/pubmed/30547203/

Emaus MJ, Bakker MF, Peeters PH, et al. (2015) MR Imaging as an additional screening modality for the detection of breast cancer in women aged 50-75 years with extremely dense breasts: The DENSE trial study design. Radiology 277: 527-37

Dalmış MU, Vreemann S, Kooi T, Mann RM, Karssemeijer N, Gubern-Mérida A (2018) Fully auto-mated detection of breast cancer in screening MRI using convolutional neural networks. J Med Imaging (Bellingham) 5(1): 014502

Dalmış MU, Gubern-Mérida A, Vreemann S, Bult P, Karssemeijer N, Mann R, Teuwen J (2019) Artificial Intelligence-Based Classification of Breast Lesions Imaged With a Multiparametric Breast MRI Protocol With Ultrafast DCE-MRI, T2, and DWI. Investigative Radiology. Available from: ncbi.nlm.nih.gov/pubmed/30652985

Hupse R, Samulski M, Lobbes MB, et al. (2013) Computer aided detection of masses at mammog-raphy: interactive decision support versus prompts. Radiology 266: 123-9

Hupse R, Samulski M, Lobbes M, et al. (2013) Standalone computer-aided detection compared to radiologists' performance for the detection of mammographic masses. Eur Radiol 23: 93-100

Karssemeijer N, Te Brake GM (1996) Detection of stel-late distortions in mammograms. IEEE Trans Med Imaging 15(5): 611-9

Kim EK, Kim HE, Han K, Kang BJ, Sohn YM, Woo OH, Lee CW (2018) Applying Data-driven Imaging Biomarker in Mammography for Breast Cancer Screening: Preliminary Study. Sci Rep. 8(1): 2762

Kuhl CK, Schrading S, Strobel K, et al. (2014) Abbreviated breast magnetic resonance imaging (MRI): first post- contrast subtracted images and max-imum-intensity projection-a novel approach to breast cancer screening with MRI. J Clin Oncol 32: 2304-10

Kuhl CK, Strobel K, Bieling H, et al. (2017) Supplemental breast MR imaging screening of women with average risk of breast cancer. Radiology 283: 361-70

Mango VL, Morris EA, David Dershaw D, et al. (2015) Abbreviated protocol for breast MRI: are multiple sequences needed for cancer detection? Eur J Radiol 84: 65- 70

Mann RM, Kuhl CK, Kinkel K, Boetes C (2008) Breast MRI: guidelines from the European Society of Breast Imaging. Eur Radiol 18: 1307-18

Pages EB, Millet I, Hoa D, et al. (2012) Undiagnosed breast cancer at MR imaging: analysis of causes. Radiology 264: 40-50

Ribli D, Horváth A, Unger Z, Pollner P, Csabai I (2018) Detecting and classifying lesions in mammograms with Deep Learning. Sci Rep. 2018

Rodríguez-Ruiz A, Krupinski E, Mordang J-J, Schilling K, Heywang-Köbrunner SH, Sechopoulos I, Mann RM (2019) Detection of Breast Cancer with Mammography: Effect of an Artificial Intelligence Support System. Radiology (2): 305-314

The frequency of missed breast cancers in women participating in a high-risk MRI screening program. Breast Cancer Res Treat 169(2): 323-331

Vreemann S, Gubern-Merida A, Schlooz-Vries MS, et al. (2018) The influence of risk category and screen-ing round on the performance of an MRI and mam-mography screening program in BRCA mutation car-riers and other women at increased risk. Radiology 286(2): 443-451

Vreemann S, Gubern-Merida A, Lardenoije S, Bult P, Karssemeijer N, Pinker K, Mann RM (2018)

Vreemann S, Gubern-Merida A, Lardenoije S, Bult P, Karssemeijer N, Pinker K, Mann RM (2018)

Wanders JO, Holland K, Veldhuis WB, et al. (2017) Volumetric breast density affects performance of digital screening mammography. Breast Cancer Res Treat 162: 95-103

Yamaguchi K, Schacht D, Newstead GM, et al. (2013) Breast cancer detected on an incident (second or subsequent) round of screening MRI: MRI features of false-negative cases. AJR Am J Roentgenol 201: 1155-63