Artificial Hype

ARTIFICIAL HYPE, L. DONOSO-BACH
AI IS THE NEW REALITY: THE 4TH HEALTHCARE REVOLUTION IN MEDICINE, I. A. WEISSMAN
ARTIFICIAL INTELLIGENCE IN HEALTHCARE: WHAT IS VERSUS WHAT WILL BE, R. PEARL
WILL AI LEAD TO JOB CUTS OR WILL THE TECH IMPROVE WORKING LIVES? L. FLORIDI ET AL.
AI OPPORTUNITIES FOR HEALTHCARE MUST NOT BE WASTED, L. FLORIDI
AI-BASED PREDICTION IN CLINICAL SETTINGS: CAN WE TRUST IT? W. LEODOLTER
ARTIFICIAL INTELLIGENCE: A NEXT WAY FORWARD FOR HEALTHCARE, R. CORBRIDGE
AI IN MEDICAL IMAGING MAY MAKE THE BIGGEST IMPACT IN HEALTHCARE, M. COLANGELO & D. KAMINSKIY
FUTURE OF ULTRASOUND: WHERE ARE WE GOING? P. SIDHU
THE AI-POWERED RADIOLOGIST, M. J. D. CANDAMIO
CLINICAL ULTRASOUND IN THE AGE OF ARTIFICIAL INTELLIGENCE, D. A. LICHTENSTEIN
AI APPLICATIONS IN BREAST IMAGING, J. TEUWEN ET AL.

NYC GO RED FOR WOMEN
MOVEMENT: STEM GOES RED, C. BEECHER

TRUST-ABUNDANT TEAM
PRINCIPLES, E. E. SULLIVAN

UNDERSTANDING BAD
COMMUNICATION, M. KEEN

VALUE-BASED HEALTHCARE AND
THE DOCTOR-PATIENT
RELATIONSHIP, M. GAFANOVICH

NURSING ON THE MOVE:
CROSS BORDER HIRING,
I. MEYENBURG-ALTWARG

HOW FOLLOWING STEPS FOR
QUALITY IMPACT HEALTHCARE
CONSUMERISM, P. FAGBENRO

THE “ONE STICK STANDARD” FOR
VASCULAR ACCESS, N. NIKHANI

CRITICAL ANALYSIS OF MRI-BASED
CLASSIFICATION SYSTEMS FOR
SPORT MUSCLE INJURIES,
J. ISERN-KEBSCHULL ET AL.

REVITALISING THE AGED HEART
THROUGH SPERMINDE-RICH DIET,
M. ABDELLATIF & S. SEDEJ

THE ROLE OF SOCIAL MEDIA IN
CARDIOLOGY, V. VASSILIOU

ACHIEVING ZERO AVOIDABLE PATIENT
DEATHS BY 2020, K. MCQUEEN

INNOVATION IN PAEDIATRIC
REHABILITATION, M. BEERI & E. BE’ERI
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 approximately 50 years. Recent 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.

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 areas 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.

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; Rodriguez-Ruiz et al. 2019).
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...
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 properly, 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 economically 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 variance 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. Implementing 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 heterogeneous 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**


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


