Artificial intelligence as another set of eyes in breast cancer diagnosis

Syed M. Anwar, Ulas Bagci

Center for Research in Computer Vision (CRCV), University of Central Florida (UCF), Orlando, FL, USA

Correspondence to: Ulas Bagci. Center for Research In Computer Vision, University of Central Florida (UCF), Orlando, FL, USA. Email: ulasbagci@gmail.com.

Comment on: Rodríguez-Ruiz A, Krupinski E, Mordang JJ, *et al.* Detection of Breast Cancer with Mammography: Effect of an Artificial Intelligence Support System. Radiology 2019;290:305-14.

Received: 18 April 2019; Accepted: 29 April 2019; published: 08 May 2019.

doi: 10.21037/jmai.2019.04.03

View this article at: http://dx.doi.org/10.21037/jmai.2019.04.03

Breast cancer is the most common cancer in women worldwide and the second most common cancer overall, hence it is a significant public health concern. According to Global Health Estimates (1), over half a million women died in 2011 due to breast cancer. In 2018, there were over 2 million new breast cancer cases (2). In the United States, the average risk of a woman developing breast cancer sometime in her life is about 12%. One effective way to reduce breast-cancer related deaths is to use radiology imaging (particularly mammography) as a screening strategy (3). However, breast cancer screening suffers from low specificity, requiring an image-guided biopsy to reach a definite diagnosis, and often it imposes very inefficient workload due to double reading of examinations or missing tumor cases along with misinterpretations (3). Historically, some of these problems have been partially addressed by developing computer-aided diagnosis (CAD) tools using machine learning algorithms. Yet, the actual benefit of using CADs during the breast cancer screening procedure is still unclear.

More recently, scientific advances in artificial intelligence (AI) have revolutionized almost all fields spanning from computer vision to finance, robotics, other computing industries, and medicine. Major developments of the AI have been largely distilled from the specific computer algorithms, called "deep learning" (4), a rapidly growing subfield of machine learning. Since famous *AlexNet* paper in 2012 (4), more advanced neural network algorithms have been developed to solve more complex problems that have never been achieved before, thanks to: (I) updates on hardware technology for processing and storing large data sets and billions of parameters, and (II) the availability of more advanced computer algorithms (5).

In biomedical and clinical landscape, past few years have started to witness the explosion of deep learning based methods. For instance, deep learning has dramatically improved the state-of-the-art in many biomedical applications like segmentation of brain and lung tumors (6,7), automatic disease detection and diagnosis from X-ray, computed tomography (CT), and magnetic resonance imaging (MRI) (8), aiding pathologist in analyzing large-scale pathology images (9), and several other tasks at the human-expert level (10,11). Similar to CAD systems, the common goal in AI-powered imaging-based diagnosis tasks is to provide greater precision in diagnostic tasks (12).

In parallel to the tremendous successes of deep learning in the biomedical imaging in general, we started to see dramatic changes in radiology screening paradigms for breast cancer detection too. One of such advances, published in *Radiology Journal*, was developed by Rodriguez-Ruiz *et al.* (3) at Radbound Medical Center (Netherland), in collaboration with Emory University, Dutch Expert Centre for Screening, Lynn Women's Health and Wellness Institute (Boca Raton), ScreenPoint Medical BV, and Referenzzentrum Mammographie Munich. Authors compared radiologists' performances for reading mammographic examinations unaided versus aided (supported by an AI system) and revealed that radiologists improved their cancer detection performance at mammography when using an AI system. Results also indicated that this benefit was obtained without requiring additional reading time. More details: screening digital mammographic examinations from 240 women were included in the study (100 normal, 40 leading to false positive recalls, and 100 showing cancers). Scans were examined by 14 radiologists (mammography quality standards act-qualified) once with and once without AI support. Authors reported that sensitivity of the examinations was increased with AI support (86% *vs.* 83%) while specificity was not shown a significant change compared to unaided screening.

In a complementary study of the same authors (13), the performance of a stand-alone AI system was compared with that of radiologists in breast cancer screening for digital mammography. For this purpose, multi-center data were used. Unlike many other early-stage tools that have a limited implementation in the real-world clinical setting, the data from seven countries was curated by 101 radiologists. This broad experimental setting included a total of 2,652 exams and the stand-alone AI system was statistically similar to that of radiologists' interpretations. In quantitative evaluations, the area under the curve (AUC) for AI system (0.840) was found to be better than the average of radiologists (0.814). The sensitivity and specificity of the system was also found to be better than majority of the radiologists, but always worse than the best radiologist, which is not surprising. These results indicate that AI tools can be used in a much more broader settings that have never been used before in breast cancer diagnosis routine, but for this to be a regular clinical practice, there is still an expectation that a lot more experimentation should be done in both retrospective and prospective settings for independent validations. There is also a strong need for a benchmark dataset where the performance of these methods can be evaluated in detail and perhaps with a lesser bias in evaluations when comparing different AI solutions.

The success of deep learning as a tool to build AI systems is pushing the performance even closer to humans in computer aided diagnosis and screening in radiology rooms. The subjectivity in terms of the underlying data (type of lesions, racial and age differences, device manufacturers) when training the deep learning models remains a challenge that need to be carefully addressed. A stand-alone AI system can supplement expert radiologist as a second reader which can translate in a reduction in reading time. However, for this to be adapted as a clinical practice, the system performance has to be further improved in terms of sensitivity and specificity for more diverse data set. This is necessary to persuade expert radiologists to adopt this technology as currently the performance of the system is found to be lower. For radiologists with lesser experience, the success of these systems is evident. Training of radiologists for working with these AI tools will also be a concern in clinical practice and most likely the screening time can see a further reduction, once radiologists are properly trained for using the system. The radiologists involved in this study were already aware that the data was enriched with malignant scans and to validate the effectiveness of the AI tool, more real world studies must be performed where such information is not already known.

In conclusion, AI tools have started to find a wider success particularly in computer aided diagnostics systems and generally in the field of medicine, predominantly due to the recent success of deep learning based methods. The gains in quantitative parameters seen by training these models on selected collection of data are still to be translated to clinical practice. For AI, to be used up to its full potential in clinical practice, more studies should be performed in real world settings. The increasing number of scans for diagnosing breast cancer generates tremendous workload for radiologists. For efficient screening and precise diagnosis, AI can play its role as shown in recent studies on breast cancer screening. Factors such as requirements from regulators, training of clinical personnel, and training bias in terms of pre- existing knowledge and subject profiles should be critically addressed. There is still space for better quantitative performance with novel and interpretable deep learning models. The future is bright for AI to act as another set of eyes for breast cancer screening and diagnosis in real world scenarios.

Acknowledgments

Funding: None.

Footnote

Provenance and Peer Review: This article was commissioned and reviewed by the Editorial Board Member Xiao Li (Department of Urology, Jiangsu Cancer Hospital & Jiangsu Institute of Cancer Research & Affiliated Cancer Hospital of Nanjing Medical University, Nanjing, China).

Conflicts of Interest: Both authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/jmai.2019.04.03). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. CA Cancer J Clin 2016;66:7-30.
2. How Common Is Breast Cancer? Available online: https://www.cancer.org/cancer/breast-cancer/about/how-common-is-breast-cancer.html
3. Rodríguez-Ruiz A, Krupinski E, Mordang JJ, et al. Detection of Breast Cancer with Mammography: Effect of an Artificial Intelligence Support System. Radiology 2019;290:305-14.
4. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems. 2012:1097-105.
5. Burt JR, Torosdagli N, Khosravan N, et al. Deep learning beyond cats and dogs: recent advances in diagnosing breast cancer with deep neural networks. Br J Radiol 2018;91:20170545.
6. Khosravan N, Bagci U. S4ND: Single-Shot Single-Scale Lung Nodule Detection. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham., 2018:794-802.
7. Hussain S, Anwar SM, Majid M. Segmentation of glioma tumors in brain using deep convolutional neural network. Neurocomputing 2018;282:248-61.
8. Mortazi A, Karim R, Rhode K, CardiacNET: segmentation of left atrium and proximal pulmonary veins from MRI using multi-view CNN. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham., 2007:377-85.
9. Janowczyk A, Madabhushi A. Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. J Pathol Inform 2016;7:29.
10. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature 2017;542:115-8.
11. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med 2019;25:44-56.
12. Anwar SM, Majid M, Qayyum A, et al. Medical Image Analysis using Convolutional Neural Networks: A Review. J Med Syst 2018;42:226.
13. Rodriguez-Ruiz A, Lång K, Gubern-Merida A, et al. Stand-Alone Artificial Intelligence for Breast Cancer Detection in Mammography: Comparison With 101 Radiologists. J Natl Cancer Inst 2019. [Epub ahead of print].

doi: 10.21037/jmai.2019.04.03
Cite this article as: Anwar SM, Bagci U. Artificial intelligence as another set of eyes in breast cancer diagnosis. J Med Artif Intell 2019;2:10.