ARTIFICIAL INTELLIGENCE TO INTERPRET MAMMOGRAMS- ARE WE THERE YET?

Radiodiagnosis

Dr Dinesh Sethi Consultant, Dept of Radiodiagnosis, Amarleela Hospital, New Delhi.

Dr Namrita Sachdev* Associate Professor, Dept of Radiodiagnosis, ABVIMS & Dr RML Hospital, New Delhi.
*Corresponding Author

Dr Yashvant Singh Associate Professor, Dept of Radiodiagnosis, ABVIMS & Dr RML Hospital, New Delhi.

ABSTRACT

Breast cancer is one of the leading causes of cancer related mortality in women. Mammography is the most widely used imaging modality to detect breast cancer. Due to a large number of screening mammograms and a limited number of breast imaging radiologists available all over the world, the role of Artificial Intelligence in the form of Deep Learning algorithms is being explored to assist the radiologists in interpreting these mammograms.

KEYWORDS

Artificial intelligence, Machine learning, Deep learning, Mammography

INTRODUCTION

Breast cancer is one of the leading causes of cancer related mortality in women. Mammography is the most widely used imaging modality to detect breast cancer. It can be either screening or diagnostic.

Screening mammography helps in detection of early breast cancer in asymptomatic women, thereby reducing the mortality rate related to this cancer. The most common drawbacks of screening mammography are false positive and false negative cases. False positive cases are further subjected to non invasive imaging modalities like Ultrasound and MRI or invasive procedure like tissue biopsies. False negative mammograms show normal images although breast cancer is present. Dense breast is a common cause of false negative mammogram.

In the United States, most of the mammograms are reported by a single radiologist but most European countries and a few Asian countries follow the protocol of getting each mammogram reported by two radiologists in order to minimize the error. In the United States alone, the estimated number of mammograms done each year is about 40 million. Considering the staggering number of mammograms and the limited number of breast imaging radiologists available, it becomes important to have some sort of automated system to reduce the burden and increase the accuracy and efficiency.

Now with the emergence of Artificial Intelligence (AI) in the field of medical imaging, the radiologists have a strong ally to assist them for interpreting mammograms. Deep Learning, a subset of Artificial Intelligence, is a highly evolved system inspired by the neurons of human brain. Deep Learning system works in multiple layers. It uses data of images to create an algorithm that can eventually be used to analyze new images provided to it. In various studies throughout the world, AI algorithms have been used to interpret mammograms.

REVIEW OF DIFFERENT STUDIES

A retrospective study was performed by Alejandro Rodriguez-Ruiz et al¹ to detect Breast Cancer on Digital mammography. A total of 2652 exams were analyzed by Artificial Intelligence. The results were then compared with the interpretations by 101 radiologists. It was found that the area under the ROC curve was 0.840 for the AI system and 0.814 for the average of radiologists. Artificial Intelligence technique had an area under curve (AUC) higher than 61.4% of radiologists. This study concluded that the AI system showed an accuracy comparable to an average radiologist in interpreting the Digital Mammograms.

Another retrospective study was done by Hyo-Eun Kim et al². They developed and validated the algorithm with 170230 digital mammograms collected from five institutions in South Korea, the USA and the UK. AI algorithm was developed by using ResNet 34, a type of deep convolutional neural networks (CNN) architecture. The AU-ROC for the performance of radiologists without the assistance of AI was 0.810 and with the AI assistance it was shown to be 0.881. This suggested significant improvement in the performance of radiologists when assisted by AI.

McKinney et al³ used an AI system for a mammography based screening programme to predict the risk of developing breast cancer. A Deep Learning AI algorithm was used to assess the risk of breast cancer based on mammogram images. The results were compared with the interpretation of six radiologists. According to this study the AU-ROC for the AI based system was 0.871, whereas the AU-ROC for the radiologist based interpretation was 0.750. Thus this study showed an overall increase of 0.115 in the AU-ROC metrics with the use of AI, suggesting that AI system had a better predictive value compared to the human experts in this field.

Watanabe et al⁴ used an AI algorithm called cmAssistTM to analyze the mammograms from 122 patients. This study was performed with the help of seven radiologists who interpreted the mammograms without and with the help of AI system. It was found that the use of AI system increased the cancer detection rate (CDR) of the radiologists. The percentage increase in CDR with the use of AI system was between 6% to 64 % (mean 27%). This indicated a significant improvement when AI system was used by the radiologists.

Screening Digital Mammogram images were analyzed by Rodriguez et al⁵ in a retrospective study. They used mammograms of 240 women (age range 39 years to 79 years). Out of these 100 had cancer, 100 were normal and 40 had false-recall. These images were interpreted by 14 radiologists expert in breast image reporting. They analyzed the images once on their own without any support and then once with the help of AI system. The study showed that the area under curve (AOC) was higher when the radiologists were assisted by AI system. The AOC was 0.87 without AI help and 0.89 with AI help. The radiologist assisted with AI also showed an increase in sensitivity (86%) as compared to interpretation without AI (83%). The P value was found to be 0.002. This study suggested that the cancer detection capability of the radiologists was improved with the help of AI system.

In a retrospective study, Adam Yala et al⁶ developed a Deep Learning model to triage mammograms as cancer-free. They used 223109 screening mammograms for this study. The Deep Learning model was evaluated on the test data. Radiologists, when working alone, achieved a sensitivity of 90.6% and specificity of 93.5%. When the radiologists were assisted by the Deep Learning model, the sensitivity was reported to be 90.1% and the specificity 94.2%. Thus it was shown that the DL simulated workflow increased the specificity, without a significant reduction in the sensitivity.

In one of the biggest study of its kind by Schaffter et al⁷, 144 231 screening mammograms from the United States were used to train and validate an AI algorithm. Another set of 166578 mammograms from Sweden was used for validation. 126 teams were formed from 44 countries comprising of 1100 radiologists. AI algorithms were refined and a final algorithm called challenge ensemble method (CEM) was created. Specificity of this CEM was compared with the specificity of radiologists at a given sensitivity. The dataset was divided into KPW dataset and KI dataset. Specificity was compared using a fixed radiologist sensitivity.

With a fixed KPW radiologist sensitivity of 85.9%, the specificity of CEM was 76.1% and the specificity of radiologist was 90.5%. With the KPW data, when the radiologist was assisted with the CEM (referred to as CEM+R) the specificity was found to be 92%.

With a fixed KI radiologist sensitivity of 77.1%, the specificity of CEM was 92.5% and the specificity of radiologist was 96.7%. With the KI data, the specificity for CEM+R was found to be 98.5%.

Thus this study concluded that the specificity of CEM (the most evolved model of AI in this case) was much lower than that of radiologist. But when a single radiologist is assisted with CEM (CEM+R), there is an improvement in specificity.

CONCLUSION

So, are we there yet? Most of the studies have revealed that when human experts are assisted by AI system, their performance is enhanced, translating into increased accuracy and reduced recall rate. At the moment, AI algorithm as a standalone method is not superior to the radiologist. But then, these Deep Learning systems are still evolving. The AI systems need a lot of data to produce a useful algorithm. Since a plethora of images are being generated for the purpose of diagnosis, these algorithms are bound to get better with each passing day. And in all likelihood, the most potent option would still be a combination of a AI and radiologist, i.e., technology and human expert.

REFERENCES

- 1 Alejandro Rodriguez-Ruiz, Kristina Lång, Albert Gubern-Merida, Mireille Broeders, Gisella Gennaro, Paola Clauser, Thomas H Helbich, Margarita Chevalier, Tao Tan, Thomas Mertelmeier, Matthew G Wallis, Ingvar Andersson, Sophia Zackrisson, Ritse M Mann, Ioannis Sechopoulos, Stand-Alone Artificial Intelligence for Breast Cancer Detection in Mammography: Comparison With 101 Radiologists, JNCI: Journal of the National Cancer Institute, Volume 111, Issue 9, September 2019, Pages 916– 922, <https://doi.org/10.1093/jnci/djy222>
- 2 Hyo-Eun Kim, Hak Hee Kim, Boo-Kyung Han, Ki Hwan Kim, Kyunghwa Han, Hyeonseob Nam, Eun Hye Lee, Eun-Kyung Kim. Changes in cancer detection and false-positive recall in mammography using artificial intelligence: a retrospective, multireader study. Lancet 2020; 138–148
- 3 McKinney, S.M., Siemiek, M., Godbole, V. et al. International evaluation of an AI system for breast cancer screening. Nature 577, 89–94 (2020). <https://doi.org/10.1038/s41586-019-1799-6>
- 4 Watanabe AT, Lim V, Vu HX, et al. Improved Cancer Detection Using Artificial Intelligence: a Retrospective Evaluation of Missed Cancers on Mammography. J Digit Imaging. 2019;32(4):625-637. doi:10.1007/s10278-019-00192-5
- 5 Rodriguez-Ruiz, Alejandro & Krupinski, Elizabeth & Mordang, Jan-Jurre & Schilling, Kathy & Heywang-Köbrunner, Sylvia & Sechopoulos, Ioannis & Mann, Ritse. (2018). Detection of Breast Cancer with Mammography: Effect of an Artificial Intelligence Support System. Radiology. 290. 181371. 10.1148/radiol.2018181371
- 6 Adam Yala, Tal Schuster, Randy Miles, Regina Barzilay, and Constance Lehman. A Deep Learning Model to Triage Screening Mammograms: A Simulation Study Radiology 2019;293:1, 38-46
- 7 Schaffter T, Buist DSM, Lee CI, et al. Evaluation of Combined Artificial Intelligence and Radiologist Assessment to Interpret Screening Mammograms. JAMA Netw Open. 2020;3(3):e200265. doi:10.1001/jamanetworkopen.2020.0265