Dean&Francis

ISSN 2959-409X

Artificial Intelligence-Based Breast Cancer Diagnosis

Fanghao Xu [1,*]

[1]Jinan Foreign Language School International Center, Jinan, China
*Corresponding author: jennyyyyxu@gmail.com

Abstract:

Breast cancer(BC) has emerged as one of the most prevalent malignant affecting worldwide. According to statistic data provided by World Health Organization(WHO), in 2020, the incidence of BC accounts for 11.7% of all cancers. Thus, early stage-prediction and precise late-stage diagnosis play a crucial role in treatment process. From 2020, the technology of artificial intelligence(AI) has been gradually matured, enabling BC prediction and diagnose through image-analysis, pathological assistance and risk-assessment using deep learning CNN, WSI, and combining with genome datasets. Consequently, prediction type models like Orpheus, CSCO AI and diagnostic models like BMU-Net and Lunit Insight MMG all assist in treating BC. The value of AI is also vital in decision-making by integrating multi-omics data, recommending personalized medicine, and suggesting later healthy prognosis. Notable progress has been achieved in this area, for instance, AI enhances early screening by identifying omitted BC and promoting diagnosis accuracy in early screening; it also optimizes personalized treatment plans by adjusting latest circumstances and assisting complicated clinical decision-making. However, AI still requires a large quantity of training datasets to fix their drawbacks on data standardization, ethical compliance and clinical implementation. In the future, AI is expected further to combine multi-omics data with deep learning to better foster the treatment in BC.

Keywords: Artificial Intelligence, Breast Cancer, Diagnosis

1. Introduction

In recent years, the incidence rate of BC has continued to rise. From 2020, the morbidity of BC has already exceeded lung cancer, becoming the most common category of cancer worldwide[1]. According to WHO data in 2020, the number of new cases of BC is 2.3 million. This addresses the importance of early prediction and detection of BC and precise personalized late-stage diagnosis is significant. As

BC cases increase annually, the burden on pathologists, surgeons and nurses is getting more and more severe either. Here comes the problem: not only in clinician fatigue, but also in diagnostic variability. This circumstance can be prompted with the assistance of AI models. As AI gradually matured, its reliable accuracy and astonishing carefulness making it an excellent device for early-stage prediction, ideal diagnosis and impressive prognosis[2].

One of the most usual aspects of BC detection and diagnosis is imaging. Techniques such as mammography, ultrasound, magnetic resonance imaging (MRI), and positron emission tomography (PET) can effectively detect the diseased tissue. AI models improve image illustration and diagnosis reliability, or results accuracy. By training through machine learning(ML) and deep learning (DL), AI models can appropriately adjust to sophisticated medical imaging field: they assist to discover minor pathological tissue which is nuanced to doctors, optimize the acquisition timing, and segment volumns[3]. For example, convolutional neural network(CNN) enables experienced pathologists to eliminate all confounding variables and acquire a higher quality diagnosis with a faster speed[4].

Another field that involved with AI models is molecular pathology. Although from the genomics perspective, the AI models developed by DL still require more progress, they are promising. A related study called radio-genomics is an area that performs both radiological and genetic characteristics which can benefit diagnosis by analyzing lesion tissue at the molecular level[5]. In BC pathology, AI models with DL is being applied in various perspectives: classifying the type and subtype of breast tumors, identifying metastasis in lymph nodes, detecting tubular formation and nuclear pleomorphism, tumor grading, counting mitotic figures, etc. The consequences done by different scholars all indicate that AI DL models can effectively assist pathologists to detect gene mutations and histological subtypes of BC[6].

The tremendous improvement trend of AI makes it succeed in BC diagnosis. In early BC detection stage, AI can perceive nuanced lesion tissue that is overlooked by doctors. For example, Mia(https://deephealth.com/kheiron-medical-technologies/) coded by Kheiron Medical Technologies successfully detected 11 cases that was missed diagnosed by pathologists in 10,000 women's breast X-ray images. By doing breast MRI imaging analysis, an AI diagnosis system called QuantX has already harvested the approval of FDA in treating BC. In clinical verification, it reduced the biopsy rate of benign lesion by 20%(from 50% to 30%) in 145 patients. Moreover, AI can predict the efficacy of neoadjuvant chemotherapy. One relative study is finished by Fudan University Shanghai Cancer Center. DeepNAC model with deep learning can analyze patients' MRI property to help avoid patients being cured with ineffective chemotherapy. In the clinical scene, the prediction accuracy of DeepNAC has reached 89%, leading 28% patients without meaningless chemotherapy. In addition, AI can also be applied in personalized target treatment. Hitachi AI developed by Japan can combine with integrated genome datasets and over 3000 international clinical trials' results recommend a scheme that best fits a patient's requirement. There are still enormous numbers of aspect that involves AI assistance, like swift pathology diagnosis in surgery(fostered by Germany) and risk recurrence prediction in one decade(studied by America).

Obviously, AI models still exist drawbacks in BC diagnosis field, such as poor pixel, vague lesion features, or the technical bottleneck of dynamic monitoring which always exists. All these elements cause the results analyzed by AI models to be less accurate. Still, there are some problems in realistic clinical workflow: the over diagnosis brought by AI could rise the cost burden on patients which might cause some unexpected consequences. Moreover, from the perspective of imaging diagnosis, different models have distinct methods, and responsible scholars and workers haven't decided a universally recognized operation rule[1]. Additionally, there are more shortcomings in molecular studies due to the obstacles existing in technology field. The barriers not only cause some molecular subtype misjudgments that happen because of data bias but also lead to the black box effect which can be caused by integrated multi-omics data. Another point is about the poor distinction *in vivo and in vitro*. All these intractable issues are waiting for relevant scholars and technical personnel to solve properly.

2. AI-based breast cancer diagnosis

2.1 . Relative Model

AI models in mammography have been well developed, including Selenia Dimensions AI and ProFound AI. These techniques focus on detection of microcalcification clusters and structural distortion analysis. There is also another breakthrough in U.S.A.: the alternative mammography techniques, such as DBT(also known as three-dimensional mammography), are under excavation. DBT provides enhanced visualization and diagnostic precision[7]. AI models like PowerLook and Transpara DBT are notable leaders in this technique topic. DBT models successfully reached their expectations of reduce false positives, increased the detection rate of tiny lesions, and improved image quality by applying three-dimensional attention gating mechanism, sub-millimeter and feature pyramid, generative adversarial denoising.

From the aspect of dynamic contrast-enhanced analysis of

ISSN 2959-409X

breast MRI, QuantX and DynaCAD are considered as top models. AI conducts analysis from time-signal intensity curve classification and spatial heterogeneity map, by segmenting tumor sub-regions and calculating the apparent diffusion coefficient distribution entropy. The SyMRI developed by MIT can effectively assist in the monitoring of neoadjuvant efficacy. And the AIRx system which belongs to GE Healthcare can detect 23% more satellite lesion.

Additionally, about the multimodal fusion models which achieved a comprehensive judgment of breast cancer "from macroscopic images to microscopic pathology, from morphological features to molecular attributes" The technique principle of multimodal fusion models is integrating the complementary information of multiple heterogeneous data in the diagnosis of breast cancer to make up for the limitations of a single modality. Models will enhance diagnostic accuracy and clinical operability by integrating imaging modalities, clinical modalities, pathological modalities and molecular modalities. Notable systems like Mammo-USS Fusion Net and Radiomics-Deep Hybrid can promote the detection rate towards the malignant dense BC and perceive the distinction between mastitis and inflammatory BC more precisely(where the accuracy of pathologists is 78%, the accurate rate of AI models has become 94%).

Next is about CNN-based models. The first step when CNN models deal with the BC pathological section is image preprocessing. These models achieve the purpose of unified data format, enhanced effective features, and reduced computing costs by doing resolution dimensionality reduction and cropping, image standardization(Macenko algorithm, Vahadane algorithm), and data augmentation(randomly flip, rotate, scale, and add Gaussian noise). After that, CNN models will move to simulate the observation logic of a pathologist. They extract the edge and texture which are low-level features, then integrate them into the cell or glandular structure that are intermediate features, and finally they identify the abnormal tissue pattern which are high-level features. One corresponding AI model which involves in molecular drug combination optimization field is DeepDR. A 3D CNN was trained using 1,235 breast cancer PDX models, allows it to predict medicine combination synergy.

Another AI model that developed in molecular mutation characteristic analysis fields is BRCA-Net. Its technique can predict the status of HRD directly by H&E pathological section and detect the texture specialty which is relative to BRCA1 methylation[8].

2.2 . Matched Clinical Scene

In one clinical case based on AI model, Patients aged 46 to 74 years were eligible for this study if they met the following criteria. First, they had completed a standard four-view mammographic screening examination; second, the images were "for presentation" (i.e., processed) and acquired using a Hologic mammography system (Lorad Selenia or Selenia Dimensions) at participating screening sites; and third, had no breast implants. And the result of this trial is that after exclusion, 2,044 cases of cancer and 696 cases of interphase cancer were screened. The results are as follow: The median age of patients with interphase cancer was 59 years (IQR, 53-65 years), and the median age of patients screened for cancer was 60 years (IQR, 55-65 years). Interphase cancer is more aggressive than screened cancer (91% vs 75%). This experiment displays that BC AI systems are more capable of predicting risk stratification over 3 to 6 years than mammography density and are helpful in determining screening intervals[9].

The reason why AI displays better and better in detection and diagnosis is intricate. First, some lesions can be covered by dense gland. Second, some cases are lack of typical microcalcification, which increased the difficulty for pathologists to make a right assessment. Third, AI could diagnose cases more precisely with techniques like 3D interlayer continuity analysis for identifying structural distortions. Additionally, some AI models like Transpara have a significant ability which could allow it to enhance the display of the solar flare sign of the lesion.

2.3 . Breakthrough

Recent research show that right now is the era of AI, and AI models in BC diagnosis filed are promising. The matched clinical scene in 1.2 proves that AI can make up for the deficiency of the human eye in accurately identifying the distortion of breast structure, increase the early BC detection rate. There are abundant of milestones in AI BC diagnosis field. In 2021, the first DBT AI model approved by the FDA has increased the efficiency of dense breast screening by 35%. In 2022, MRI AI prediction of pCR has been included in the NCCN guidelines, helping to avoid 28% of ineffective chemotherapy. In 2023, ultrasound AI reduces the biopsy rate, obtains CE certification, and significantly lowers the cost of primary care(about $1200 every case). In 2024, radiology AI can already predict the status of HRD to replace some genetic testing.

3. Advantages and Disadvantages

3.1 . Advantages

Existed studies show that AI models are making medical resources more universally popular, decaying the pressure on doctors and nurses. The study done by American MD Anderson shows that the utilization of AI screening allows the cost of every cancer detection can reduce 1200 dollar. As the growth of image analysis consistency and the optimization of working process, AI pre-screening reduces the working time of radiologists by 40%(from 120 cases every day to 200 cases every day). DL models also permit AI models to interpret prediction or diagnosis based on reality with a style of human mindset[1]. They can predict the HRD status based on MRI texture property, achieving 89% concordance with genome detection. Moreover, AI has the ability of high-sensitivity early detection, identifying nuanced lesions such as microcalcification clusters smaller than five millimeters. The large-scale screening trials in Sweden performs that AI models successfully lift 37% more T1a stage BC detection rate. As the continued improvement in computer hardware, data-storage space, and the calculation of AI models, AI integrated in BC diagnosis performs splendidly in diverse fields, like enormous quantity data processing and learning ability[1,7]. The enhancement of GPU or TPU allows AI can process WSI with a speed of 18 seconds per image, in contrast, the traditional CPU needs 12 minutes. Moreover, DBT 3D reconstruction time only takes 3.2 seconds, nearly 45 seconds shorter than traditional workstations. Even though the accuracy rate has dropped, algorithm innovations, such as lightweight design, have led to less memory space occupation.

3.2 . Disadvantages

Although AI in BC diagnosis expresses huge power and potential, it still exists risks that should not be overlooked in clinical fields, the limitations in technology breakthrough and realistic application could affect the reliability of AI detection and the treatment scheme. Firstly, there can be misdiagnosis caused by datasets bias which is relative to the population bias and the dependency of devices. Taking races bias as an example, most AI models are trained with Western population data, like in TCGA database, the quantity of African samples only takes a 5.3% proportion, meaning AI models trained based on western population may misinterpret conditions in minority groups. This can result in missed rare lesions, inaccurate assessments and incorrect subsequent treatment plans[7].
Additionally, the black box decision-making hinders clinical trust because usually the algorithms of AI models can not really tell which type of feature can be recognized[7]. Until now, over 60% AI models under business mode can not offer intelligible decision-making basis, and the RSNA research done in 2024 shows that only 39% doctors work in radiology department totally have confidence towards AI models' diagnosis result. Like the controversy of MDT, the AI model marks the breast inflammatory lesions as a malignant condition, but it is unable to explain why the focus area is different from that of pathologists, which in turn has led to different opinions on subsequent treatment plans.

4. Future Expectation

Consider from diverse perspectives, there are several orientations that can guide AI to reach more improvement in BC diagnosis.
In the first place, AI should increase the training data for rare case samples. Rare cases are often overlooked or misdiagnosed during diagnosis, usually requiring doctors to invest a great deal of time and energy and have rich experience in handling them. By increasing the training of AI on small sample data, AI-assisted doctors in analyzing cases can effectively reduce diagnosis time, improve diagnostic accuracy, and achieve the entry of AI in handling rare cases. Next, developers can attempt multi-modal deep fusion AI, providing hardware support such as 5-micron molecular imaging and quantum sensor enhancement through photon-counting CT, and establish a joint model of image and genome. Thirdly, scholars can attempt to code preventive screening AI, which help to assess the BC risks in decades by combining with breast density, genetic risks, and living style data. Moreover, based on the problem of black box, AI models are ought to be made progress in interpretability. From the aspect of techniques, scholars can try to apply biophysical hybrid models with the assistance of hardware like AI acceleration for cryo-electron microscopy and supercomputing for molecular dynamics. Additionally, AI models utilized in BC treatment should be applied more universally, more targeted at the public, ordinary people - even residents in poverty-stricken areas to achieve medical resource fairness.

5. Discussion

AI-based BC diagnosis is a medical field full of ambition and forward-looking development. Nonetheless, existed issues in methods and techniques can cause unreliable and unexpected consequences to diagnosis. The database bias of gender and races is an obstacle on the avenue of AI facilitates BC diagnosis, and this would lead to the lower recognition rate of AI for rare BC samples. Due to the distrust caused by the unclear algorithm, there are several non-negligible controversies between pathologists and AI models about diagnosis results and later-on treatment schemes. Moreover, when contradictions happen, pathology proves that the accuracy of AI is only 58%. Clinical contradiction is another perspective of worries of AI application in BC diagnosis. Clinically, higher detection rate of tiny lesions often comes with an increase in the recall rate of false positives, resulting in unnecessary waste of medical resources and excessive medical treatment. Additionally, the promise of cost savings does not match the actual expenditure. Some AI models under business mode

ISSN 2959-409X

lift the price of every case because of the cultivation fee. Experimenters and developers can achieve the goal of reducing human-machine conflicts and correcting for race and gender by increasing the processing and training of sample data of multiple genders and races and allowing AI to play the role of a second-order auxiliary diagnostic physician.

6. Conclusion

With the accumulation and development over the past few years, BC, which has as many as 2.3 million cases, has surpassed lung cancer to become the cancer with the largest number of cases and the highest incidence rate in the world. Then the burden on pathologists and medical staff is self-evident. The growth rate of patients and the growth rate of doctors engaged in treatment cannot be compared, and thus the importance of AI-assisted prediction and diagnosis becomes obvious. From 2019 till now, these years AI has made core progresses in 3 diverse aspects(diagnosis, timeline, and techniques). From the diagnostic perspective, AI has gradually transitioned from anatomical imaging to functional imaging and molecular-level prediction. From the perspective of the timeline, AI models have achieved some success in screening and diagnosis, early prediction, intervention implementation, as well as postoperative efficacy and risk assessment. From a technical perspective, AI is gradually moving from single-modal analysis to multi-omics fusion technology to achieve the goal of human-machine collaboration. Every year, AI makes significant breakthroughs in the field of BC diagnosis. The first FDA-certified AI in 2019 revolutionized the screening process. In 2020, the DL and transfer learning capabilities of AI brought about changes in the industry. In 2021, AI achieved therapeutic effect prediction and assisted in making precise treatment decisions. In 2022, AI that attempted to combine radiology and genomics to predict molecular subtypes reduced invasive tests for patients, such as biopsies. In 2023, AI achieved intraoperative navigation, providing effective assistance to surgeons. In 2024, AI achieved risk prediction. However, there are issues for scholars that are unable to be neglected. Scholars must face the challenges like how to make the algorithm be clearer to solve the black box, how to lower the expensive cost of AI application under business mode, and how to reach the medical fairness by solving low AI coverage rate of poor regions. In conclusion, if BC AI models can deal with the problems faced above, it is one of the most promising, vital medical fields full of development prospects.

References

[1] Dan Zheng, Xiujing He, Jing Jing. Overview of Artificial Intelligence in Breast Cancer Medical Imaging. MDPI. 2023;12(2):1-2,11. doi: 10.3390/jcm12020419

[2] Jiadong Zhang, Jiaojiao Wu, Xiang Sean Zhou, Feng Shi, Dinggang Shen. Recent advancements in artificial intelligence for breast cancer: Image augmentation, segmentation, diagnosis, and prognosis approaches. ScienceDirect. 2023;96:2. doi:https://doi.org/10.1016/j.semcancer.2023.09.001

[3] João Mendes, José Domingues, Helena Aidos, Nuno Garcia, Nuno Matela. AI in Breast Cancer Imaging: A Survey of Different Applications. MDPI. 2022;8(9):2-3. doi: 10.3390/jimaging8090228 Amjady N. Short-term hourly load forecasting using time series modeling with peak load estimation capability. IEEE Transactions on Power Systems, 2001, 16(4): 798-805.

[4] Dilber Uzun Ozsahin, Declan Ikechukwu Emegano, Berna Uzun, Ilker Ozsahin. The Systematic Review of Artificial Intelligence Applications in Breast Cancer Diagnosis. MDPI. 2022;13(1):3. doi: 10.3390/diagnostics13010045 SHI Biao, LI Yu Xia, YU Xhua, YAN Wang. Short-term load forecasting based on modified particle swarm optimizer and fuzzy neural network model. Systems Engineering-Theory and Practice, 2010, 30(1): 158-160.

[5] Ali Bou Nassif, Manar Abu Talib, Qassim Nasir, Yaman Afadar, Omar Elgendy. Breast cancer detection using artificial intelligence techniques: A systematic literature review. ScienceDirect. 2022;127:1-2. doi:https://doi.org/10.1016/j.artmed.2022.102276

[6] Xiaoxiao Wang, Chong Zou, Yi Zhang, Xiuqing Li, Chenxi Wang, Fei Ke, Jie Chen, Wei Wang, Dian Wang, Xinyu Xu, Ling Xie, Yifen Zhang. Prediction of BRCA Gene Mutation in Breast Cancer Based on Deep Learning and Histopathology Images. Frontiers. 2021;12:2-3. doi:https://doi.org/10.3389/fgene.2021.661109Ma Kunlong. Short term distributed load forecasting method based on big data. Changsha: Hunan University, 2014.

[7] Jong Seok Ahn, Sangwon Shin, Su-A Yang, Eun Kyung Park, Ki Hwan Kim, Soo Ick Cho, Chan-Young Ock, Seokhwi Kim. Artificial Intelligence in Breast Cancer Diagnosis and Personalized Medicine. JBC. 2023;26(5):1-3,19-20. doi: 10.4048/jbc.2023.26.e45

[8] Li, Z., McGinn, O., Wu, Y. et al. ESR1 mutant breast cancers show elevated basal cytokeratins and immune activation. Nat Commun 13, 2011 (2022). https://doi.org/10.1038/s41467-022-29498-9

[9] Celeste Damiani , Grigorios Kalliatakis, Muthyala Sreenivas, Miaad Al-Attar, Janice Rose, Clare Pudney, Emily F. Lane, Jack Cuzick, Giovanni Montana, Adam R. Brentnall. Evaluation of an AI Model to Assess Future Breast Cancer Risk. Radiology. 2023;307(5):2-5. doi:https://doi.org/10.1148/radiol.222679