## Application and prospect of artificial intelligence in breast cancer imaging diagnosis

**Abstract**：

Breast cancer accounts for 15% of new cancers in China, ranking first, posing a serious threat to women's physical and mental health. Early diagnosis can significantly improve the survival rate of breast cancer patients. In recent years, with the development of big data and computer algorithms, the research and application of artificial intelligence (AI) such as imagomics and deep learning in the field of medical imaging have become increasingly widespread. Automatic analysis of cancer images and pathological images using artificial intelligence models can not only reduce the work burden of doctors, but also improve the accuracy of diagnostic results. Through literature search, it is found that in recent years, artificial intelligence based on medical images has been applied in imaging diagnosis to timely find lesions and judge benign and malignant tumors, and to grade breast cancer based on pathological diagnosis. This paper mainly summarizes the application of artificial intelligence in this field, in order to promote the clinical transformation of artificial intelligence technology in breast cancer diagnosis.

Ⅰ.**Introduction**：

In recent years, breast cancer has become an important factor threatening women's health. The Global Cancer Statistics 2020 report[1] shows that breast cancer is currently one of the most common causes of cancer death in women, with its incidence and mortality increasing year by year. In China, breast cancer accounts for 15% of all new cancers in women, with the highest incidence[2-3]. Early detection is the key to improve the survival rate of breast cancer patients, and the early diagnosis of breast cancer largely depends on imaging and pathology. At present, the most commonly used imaging methods for breast examination include ultrasound, mammography and MRI. In the diagnosis of breast, the information in pathological images is of great significance to the diagnosis and treatment of patients. However, the current breast cancer screening system in China is incomplete, and the overall diagnosis rate of early breast cancer is low[3]. Moreover, imaging examination results are mostly based on the evaluation of signs of lesions by radiologists, while pathology requires a lot of work by pathologists, and both of them are easily affected by doctors' subjective inference, thus easily affecting the quality of diagnosis. In recent years, with the development and application of algorithms such as machine learning (ML), Convolutional Neural Networks (CNN)[4] and graph neural networks (GNN)[5], the development of image AI and pathological AI has ushered in a new stage. The study of its application in breast imaging and pathology is also continuing to deepen. Relevant studies show that this will improve the diagnostic efficiency of early breast cancer to some extent, but artificial intelligence (AI) research on breast imaging faces many problems and challenges at the same time.

Ⅱ.**Application of artificial intelligence algorithm in breast examination and diagnosis:**

medical imaging refers to the technology and process of obtaining internal tissue images of the human body or a part of the human body in a non-invasive manner for medical treatment or medical research. With the rapid development of various science and technology, medical imaging technology has gradually formed a variety of medical imaging methods represented by X-ray (including CT), magnetic resonance imaging, ultrasound and nuclear medicine. In the field of artificial intelligence imaging, the most promising application is deep learning (DL). DL does not need to define image features in advance, and builds end-to-end models through multi-layer neural networks to automatically use value information to predict diseases. DL networks for image analysis include recurrent neural networks (RNN), CNN, deep belief networks (DBN) and their corresponding improved models. In particular, CNN has the most extensive research and application in medical imaging.

In the current analysis of pathological images of breast cancer, CNN uses pixels in the image as the unit of analysis, resulting in weak interpretability of the predicted results of the model, which is difficult to be accepted by pathologists. A new trend towards applying GNN to many computer-aided diagnostic (CAD) tasks can already be seen from work in recent years, with the help of GNN, the relationship between the spatial context of biological entities (such as cells and tissues) in pathological images and patches (which cut tissue areas into small image blocks) can be explicitly modeled. Therefore, the prediction results of the model are more easily accepted by pathologists, and the performance of the model can be further improved.

At the same time, there are also studies to explore the ability of AI-assisted ultrasound to identify breast lesions in Chinese women and the possibility of its application in breast cancer screening. It used a parallel controlled diagnostic trial and prospective follow-up study design, and included non-breast cancer women who were admitted to a specialized oncology hospital and underwent breast ultrasound. All the women received AI-assisted ultrasonography first and then routine ultrasonography to compare the differences between AI-assisted ultrasonography and conventional ultrasonography in identifying breast lesions. The incidence of breast cancer within 1 year was followed up to compare the sensitivity and specificity of the two ultrasonography methods in the diagnosis of breast cancer. It is concluded that AI-assisted ultrasound has the same ability to identify high-risk breast lesions above BI-RADS 4A and early breast cancer as conventional ultrasound, and it is an effective auxiliary diagnostic means for breast cancer, and has the potential to be applied to breast cancer screening in the population. [18]

III. **Application of AI in the differentiation of benign and malignant breast lesions**:

In imaging, benign tumors are regular in shape, clearly demarcated from surrounding tissues, and generally not enhanced significantly after enhancement. The malignant ones have irregular shape, even burrs on the edge, and the boundary between them and the surrounding tissues is not clear, and obvious enhancement can be seen after enhancement. According to the pathological characteristics, benign tumors have higher differentiation degree and less atypia. The malignant tumors showed varying degrees of dysdifferentiation or undifferentiation.

Mammography is one of the most commonly used breast cancer screening methods and is particularly good at detecting calcification. However, its diagnostic performance is easily affected by the type of gland, and it is easy to miss the non-calcifying lesions of dense glands[6]. However, there are obvious limitations in the application of X-ray examination in Asian women represented by China. X-ray examination has low screening sensitivity (Se) for young Asian women and those with high breast density. Ultrasound has the advantages of portability, simple operation and non-invasive, and is one of the common methods for breast cancer screening. Compared with mammography, ultrasound is more suitable for dense glands in Asian women. Recently on the depth of neural network model of breast cancer group development SonoBreast (open. Baai. Ac. Cn/SonoBreast) trial in some hospital, The results showed that it assessed the likelihood of malignancy more accurately than the breast imaging reporting and data system (BI-RADS)[7]. MRI is of great value in the diagnosis of breast cancer due to its high resolution of soft tissue and multi-sequence imaging. In view of the fact that dynamic contrast enhanced MRI (DCE-MRI) can reflect functional information such as hemodynamics of cancer foci, AI studies are mostly based on this[8].

Ⅳ. **The role of AI technology in pathological evaluation:**

computer-aided diagnosis (CAD) technology is part of the application of artificial intelligence in the medical field. In 1959, Ledley[9] in the United States established the mathematical model of clinical medicine for the first time, applied Bayes theorem and Boolean algebra to the computer-aided diagnosis of the model, and successfully used the model to diagnose lung cancer. With the continuous development of pathological image data processing technology, computers can continuously accumulate and deeply learn pathological data, so that they have the ability to process a vast amount of medical information and have good reproducibility of pathological pictures, so that computer-aided diagnosis technology can be further applied in pathology. At present, it has been reported in the fields of cytopathology diagnosis, histopathology diagnosis and immunohistochemical diagnosis. [10] The goal of computer-aided diagnosis is to use a combination of qualitative and quantitative methods to provide reliable pathological diagnosis for individualized precision therapy. Whether it is cytopathology, histopathology, immunohistochemical detection and diagnosis, it has contributed. [11]

Ⅴ**. Conclusion and prospect：**

Looking forward to the potential future development trend of artificial intelligence in breast cancer diagnosis, new technologies such as deep learning and graph neural networks will improve the accuracy of breast cancer diagnosis and assist doctors to make faster and more accurate decisions. Machine learning is primarily based on learning patterns from data sets to make classifications and predictions that are ultimately used to support the decision-making process. The development of machine learning has enabled the analysis of massive amounts of data, and digital methods have enabled the acquisition, integration, and analysis of data far beyond the capabilities of traditional technology, and can do this efficiently and at scale. This has the potential to reshape what can be achieved in terms of yield, accuracy and reliability, allowing large sample data to be summarized and compared [12]. Common machine learning methods include regression analysis, neural network analysis, support vector machine analysis, decision tree method and random forest method [13]. At present, the application of machine learning in breast cancer diagnosis mainly involves mammograms, magnetic resonance, ultrasound imaging data, histopathological data and genomic data. Machine learning is a great way to use existing data to build models to help predict and detect breast cancer early. [14] Of course, AI technology still faces many challenges in transforming clinical practice. (1) Accessibility of large databases: the premise of AI to achieve "intelligence" requires "artificial" to provide a large number of high-quality data for training, the accuracy of intelligence depends on artificial accuracy, and there is no standard specification and expert consensus on the acquisition, management and application of data. At the same time, subject to the impact of patient privacy and research protection policies, it is difficult for medical centers to achieve data sharing, and the phenomenon of data island is a key problem plaguing the clinical application of AI. (2) Interpretability of results: Decisions and predictions made by machine learning algorithms (especially deep learning) are not properly explained. In order to demonstrate the reliability of these decisions, they need to be properly explained. Therefore, some scholars focus on using explainable models to make high-risk decisions [15]. The ultimate responsibility of the research community is to make AI algorithms fully interpretable, which will contribute to the widespread clinical application of AI techniques [16]. (3) Generalization and robustness of the model: generalization refers to the accuracy of the prediction of the model in medical center B obtained from the data of medical center A, and robustness refers to the anti-interference ability of the model. The model generalization and robustness are based on high-quality multi-center data that has been standardized, and the consistency of the data will be affected by different detection equipment, lighting conditions, and the subjectivity of data label annotation. Liu et al. [17] proposed that multi-source networks and shape-sensing meta-learning algorithms could solve the heterogeneity problem of multi-center data, thereby improving the model's learning performance on multi-center data.

[1] SUNG H, FERLAY J, SIEGEL R L, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide

for 36 cancers in 185 countries[J]. CA Cancer J Clin, 2021, 71(3): 209-249. DOI: 10.3322/caac.21660.

[2] CHEN W Q, ZHENG R S, BAADE P D, et al. Cancer statistics in China, 2015[J]. CA Cancer J Clin, 2016, 66(2): 115-132. DOI: 10.3322/

caac.21338.

[3]WANG F, YU Z G. Current status of breast cancer prevention in China[J].

Chronic Dis Transl Med, 2015, 1(1): 2-8. DOI: 10.1016/j.cdtm.2015.02.003

[4]周飞燕, 金林鹏 & 董军. 卷积神经网络研究综述. 计算机学报 40, 1229–1251 (2017).

[5]郑光远, 刘峡壁 & 韩光辉. 医学影像计算机辅助检测与诊断系统综述. 软件学报 29, 1471–1514 (2018).

[6] SECHOPOULOS I, TEUWEN J, MANN R. Artificial intelligence for breast cancer detection in mammography and digital breast tomosynthesis:state of the art[J/OL]. Semin Cancer Biol, 2021, 72: 214-225 [2023-03-23]. https://pubmed.ncbi.nlm.nih.gov/32531273/.DOI:10.1016/j.semcancer.

2020.06.002

[7]QIAO M Y, FANG Z, GUO Y, et al. Breast calcification detection based on multichannel radiofrequency signals via a unified deep learning framework. Expert Syst Appl,2021,168:114218.

[8]王浩天, 于韬 & 徐姝. MRI影像组学在乳腺癌新辅助化疗中应用的研究进展. 中国临床医学影像杂志 34, 892–896 (2023).

[9] Gómez MMY, Escalante HJ, Segura A, et al.Advances in artificial intelligence:IBERAMIA 2016[M].Springer International Publishing, 2016.Proceedings.

[10]于鹤,赵稳兴.计算机辅助诊断技术在病理学中的应用进展[J].诊断病理学杂志,2018,25(03):223-226.

[11]郭蕾,石峰,郑闪等.乳腺癌Ki-67计算机全自动分析可行性探讨[J].中华肿瘤防治杂志,2015,22(09):678-681.DOI:10.16073/j.cnki.cjcpt.2015.09.008.

[12] Bini SA.Artificial Intelligence,machine Learning,deep learning, and cognitive computing: what do these terms mean and ho w will they impact health care?[J].J Arthroplasty,2018,33(8):2358-2361.

[13] Greener JG,Kandathil SM, Moffat L,et al.A guide to machine learning for biologists[J]. Nat Rev Mol Cell Biol,2022,23(1):40-55.

[14]卢峰,张开炯,吴立春等.机器学习对乳腺癌诊断应用现状研究进展[J].社区医学杂志,2023,21(24):1315-1322.DOI:10.19790/j.cnki.JCM.2023.24.11.

[15] Rudin C.Stop explaining black box machine learning models for high stakes decisions and use interpretable models inst ead[J]. Nat Mach Intell,2019, 1(5):206-215.

[16] Shah SM, Khan RA,Arif S,et al. Artificial intelligence for breast cancer analysis:trends & directions[J]. Comput Biol Med,20

22,142:105221.

[17] Liu Q,Dou Q, Yu L,et al. MS-net:multi-site network for improving prostate segmentation with heterogeneous MRI data|].I EEE Trans Med Imaging, 2020,39(9):2713-2724.

[18]沈洁,刘雅静,莫淼等.人工智能辅助超声对中国女性乳腺病灶识别的有效性研究[J].中国癌症杂志,2023,33(11):1002-1008.DOI:10.19401/j.cnki.1007-3639.2023.11.005.