

REVIEW

Open Access



Deep learning applications in breast cancer histopathological imaging: diagnosis, treatment, and prognosis

Bitao Jiang^{1,2*†}, Lingling Bao^{1,2†}, Songqin He³, Xiao Chen³, Zhihui Jin^{1,2} and Yingquan Ye^{3*}

Abstract

Breast cancer is the most common malignant tumor among women worldwide and remains one of the leading causes of death among women. Its incidence and mortality rates are continuously rising. In recent years, with the rapid advancement of deep learning (DL) technology, DL has demonstrated significant potential in breast cancer diagnosis, prognosis evaluation, and treatment response prediction. This paper reviews relevant research progress and applies DL models to image enhancement, segmentation, and classification based on large-scale datasets from TCGA and multiple centers. We employed foundational models such as ResNet50, Transformer, and Hover-net to investigate the performance of DL models in breast cancer diagnosis, treatment, and prognosis prediction. The results indicate that DL techniques have significantly improved diagnostic accuracy and efficiency, particularly in predicting breast cancer metastasis and clinical prognosis. Furthermore, the study emphasizes the crucial role of robust databases in developing highly generalizable models. Future research will focus on addressing challenges related to data management, model interpretability, and regulatory compliance, ultimately aiming to provide more precise clinical treatment and prognostic evaluation programs for breast cancer patients.

Keywords Breast cancer, Deep learning, Diagnosis, Treatment, Prognosis

Introduction

Breast cancer, a major global women's health concern, will be diagnosed in more than 2.3 million women worldwide by 2020, accounting for more than a quarter of all cancers in women [1]. Breast cancer is a highly heterogeneous disease that faces many challenges in its diagnosis, treatment and prognosis. With the development of tumor histology techniques, it not only reveals phenotypic information about the tumor, but also closely correlates with the progression and prognosis of tumor patients [2, 3]. However, assessment of histology is highly subjective and may yield different results between different observers. The uncertainty inherent in conventional pathologic analysis may lead to inappropriate treatment and complications that can adversely affect patient quality of life [4]. In patients with early-stage breast cancer,

[†]Bitao Jiang and Lingling Bao contributed equally to this work.

*Correspondence:

Bitao Jiang

jiangbitao@163.com

Yingquan Ye

dyz1989@yeah.net

¹Department of Hematology and Oncology, Beilun District People's Hospital, Ningbo 315800, China

²Department of Hematology and Oncology, Beilun Branch of the First Affiliated Hospital of Zhejiang University, Ningbo 315800, China

³Department of Oncology, The 906th Hospital of the Joint Logistics Force of the Chinese People's Liberation Army, Ningbo 315100, China



about 30% will experience a recurrence within 10 years, which necessitates appropriate and proactive treatment. However, assessing the probability that a patient will or will not recur remains a major challenge plaguing patients diagnosed with early-stage breast cancer [5, 6]. Currently, treatment stratification and prognostic prediction methods for breast cancer patients are based on molecular tests, but these tests are not readily available in many places due to cost and other barriers. Therefore, the use of artificial intelligence-based digital pathology analysis of histopathology images to determine an alternative, cost-effective method of patient stratification has great potential for patients worldwide.

Deep learning (DL) show great potential in the diagnosis, treatment and prognosis of breast cancer. In diagnosis, DL models improve image quality through data enhancement techniques and enable rapid segmentation of the lesions, greatly improving diagnostic accuracy and efficiency [7]. These models utilize convolutional neural networks (CNNs) to extract features from images such as WSI images, digital breast tomosynthesis (DBT), mammography (DM), and magnetic resonance imaging (MRI), enabling automated and accurate tumor detection and segmentation [8–10]. In terms of treatment, DL technology performs staging and subtyping of breast cancer through image-based classification techniques. These models are able to identify different tumor features and classify them into different cancer subtypes, thus providing a basis for developing personalized treatment plans [11]. DL models can also integrate multi-omics data to predict patients' responses to different treatment plans, helping physicians choose the most appropriate treatment strategy. In terms of prognosis, DL technology shows significant potential [12]. By analyzing a patient's imaging data, DL models can predict the risk of disease metastasis and other clinical outcomes. These models not only predict patient survival, but also assess the likelihood of recurrence, helping physicians develop long-term surveillance and follow-up plans [13]. In addition, DL technology, trained on large-scale databases, is able to build robust and generalizable models that improve the accuracy and reliability of prognostic predictions.

DL technology shows great potential in breast cancer applications, but it still faces challenges in practical application. Data management, model interpretability and regulatory requirements are important issues that need to be addressed. In this review, we first provide a detailed summary of the basic definitions and fundamental algorithms of DL, as well as the basic processes involved in DL. Next, we focus on reviewing the applications of DL in the early detection and diagnosis of breast cancer. Subsequently, we summarize the role of DL in the treatment and prognosis of breast cancer. We also provide a detailed review of the applications of DL in the molecular

typing and tumor microenvironment (TME) of breast cancer. Finally, we discuss the limitations of DL in the applications related to breast cancer, such as data acquisition and algorithm optimization. In the future, DL will greatly assist healthcare professionals in providing more precise and personalized treatment plans, creating a better tomorrow for breast cancer patients.

Basic definition and general process of DL

DL, as a subfield of Machine Learning (ML), mimics the learning process of the human brain through the use of multilayer neural networks to extract more advanced features from the data and enable learning of the data at a higher level. The kernel of DL lies in the ability to use multilayered neural networks for processing and extracting features from the input data in order to deal with different sources and multimodal data types.

Basic algorithms for DL

The basic algorithms of DL include Convolutional Neural Networks (CNN), Generative Adversarial Networks (GAN), Recurrent Neural Networks (RNN), and Deep Reinforcement Learning (DRL). CNNs are mainly used for image processing and analysis. GANs are used for image generation, creating high-quality image data through alternating training of a generator and a discriminator. RNNs are suitable for processing sequential data such as time series, text, and speech. DRL combines the advantages of DL and reinforcement learning for tasks that require maximizing expected rewards, such as landmark detection and lesion segmentation in medical image analysis.

CNN

CNN is the most commonly used artificial neural network (ANN) in DL, mainly used for feature extraction and analysis of image data. CNN mimics the working principles of the biological visual system, using convolutional layers, pooling layers, and fully connected layers to extract and process image features. The convolutional layer applies multiple filters to scan the input image, extracting features such as edges and textures. The pooling layer reduces the size of the feature maps through down-sampling, thereby lowering computational complexity and preventing overfitting. The fully connected layer maps the extracted features to the classification space, completing the final classification task. Common CNN architectures include ResNet, VGG, GoogleNet, and U-Net. ResNet introduces residual blocks and uses skip connections to address the training difficulties of deep networks. These skip connections allow gradients to be directly transmitted from the later layers to the earlier layers, alleviating the vanishing gradient problem. ResNet includes variants such as ResNet18, ResNet34, ResNet50,

ResNet101, and ResNet152, differing in network depth to suit various application scenarios [14, 15]. VGG employs a very deep network structure, typically having 16 or 19 layers (i.e., VGG16 and VGG19). It uses fixed-size 3×3 convolutional kernels and 2×2 pooling layers, maintaining simplicity and consistency throughout the network. VGG is primarily used for image classification tasks and is widely used as a pre-trained model in many transfer learning applications [16]. The GoogleNet family includes versions like Inception v1, v2, v3, v4, and Inception-ResNet, each improving upon the original design to enhance performance. GoogleNet excels in image classification and object detection tasks, especially demonstrating outstanding performance on large datasets [17]. U-Net, as a commonly used architecture for image segmentation tasks, has jump connections between its encoder and decoder, which are able to extract a wide range of image features, and thus efficiently accomplish the segmentation task [18, 19]. CNNs have the advantage of parameter sharing and sparse connectivity, and are able to learn hierarchical features of an image, which leads to efficient image classification and segmentation [20]. Transformer is a DL model based on the self-attention mechanism, initially designed for natural language processing (NLP) tasks [21]. The Transformer model captures global dependencies within the input sequence through the self-attention mechanism, making it highly effective in handling long-range dependencies. Recent research has primarily focused on combining CNN and Transformer models. This combination enhances the comprehensiveness of feature extraction by complementing local and global feature extraction and reduces computational complexity. Vision Transformer (ViT) is an architecture that applies the Transformer model to computer vision tasks [22]. The core idea of ViT is to divide the image into fixed-size patches, flatten these patches into vectors, and input them into the Transformer encoder for processing. ViT captures global dependencies between image patches through the self-attention mechanism, achieving efficient feature extraction.

GAN

GAN is mainly used in image generation task. It is mainly trained and learned by using different loss terms during the training process to generate images with high quality data [23]. GAN accomplishes the image generation task by alternating between generator and discriminator. The generator is mainly used to create new image data, and its results can even deceive the discriminator [24]. The discriminator's role is to determine the authenticity of the data from the false image data generated by the generator. The image data generated by the GAN through this unique alternate training is also more and more close to the real image data [25].

RNN

RNN are suitable for processing sequential data, such as 3D volumetric images, due to their internal memory's ability to memorize their inputs, allowing them to excel in solving sequential data problems [26]. RNNs are widely used in areas such as time series analysis and natural language processing (NLP) [27].

DRL

DRL combines the strengths of DL and Reinforcement Learning (RL) for tasks that require maximization of the expected reward, such as landmark detection and lesion segmentation in medical image analysis [28]. DRL has achieved significant performance gains in many areas, including computer vision, robotics, and gaming [29].

The process underlying DL

The core of a DL model is its complex network architecture and training process that automatically learns and extracts features from raw data, capturing underlying semantic information relevant to the task. The following says the basic process of DL (Fig. 1). First is data preparation and preprocessing [30]. Data preparation and preprocessing is the first step in successfully training a DL model. This involves collecting training data from multiple sources (e.g., images, text, audio, etc.). The collected data needs to be cleaned to remove or correct noise and outliers in it. Subsequently, the data is augmented by rotating, scaling, flipping, and other operations to expand the dataset, which in turn significantly improves the generalization of the model. Finally, the data is normalized to a uniform range for better processing and learning by the model [11, 31]. The second step is model selection and design. According to the specific task, choose the appropriate DL model architecture. Currently, the most common image processing tasks use CNNs, and data such as gene sequences use RNN. The classical CNN architectures mentioned above are primarily used for various image data classification tasks [15]. U-Net, as the most commonly used primary architecture for image segmentation tasks, is favored by most current researchers due to its excellent performance [32]. The third step is model training and optimization. During the training process, the model processes the data through multiple hierarchical modules. Each module consists mainly of linear and nonlinear transformations. Although each individual module may be simple, as the number of layers increases, the network is able to model complex functions [33]. The training process updates the internal parameters or weights of the network through a backpropagation mechanism. Backpropagation is mainly used to calculate the loss function, which results in showing the error between the model output and the actual output. The network parameters are then updated using a

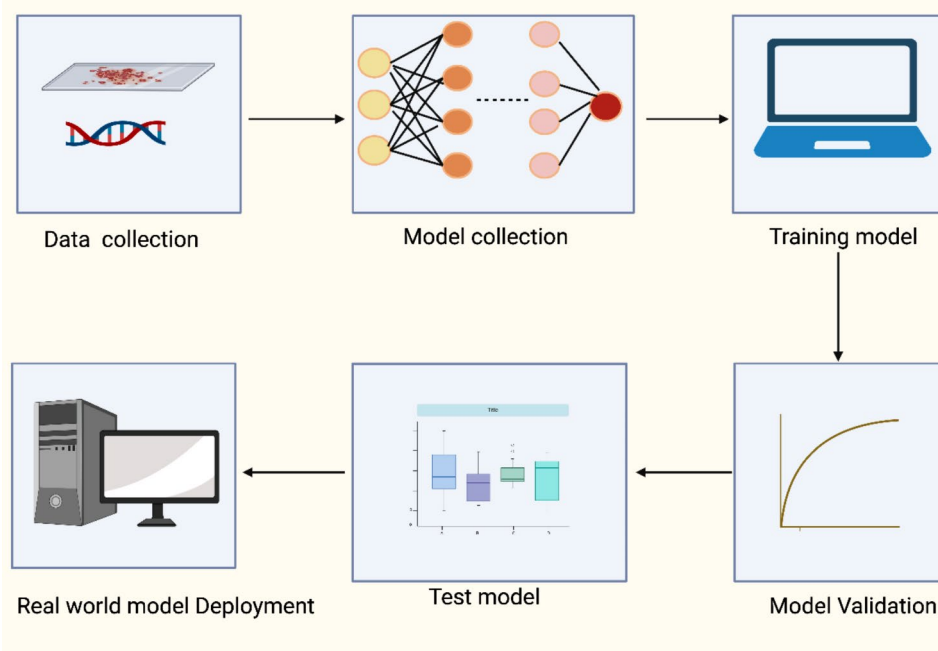


Fig. 1 Basic process of DL application in breast cancer

stochastic gradient descent algorithm to gradually reduce the error [34]. Next is the training strategy. The training strategy is crucial for utilizing the learning capability of deep neural networks. In image generation task, GAN is trained by alternating between generator and discriminator, which in turn generates high quality image data. Through this alternate training, the generated data gradually approximates the real data [35]. The fourth step is model evaluation and validation. During the training process, validation sets need to be used to evaluate the performance of the model to ensure that the model is not overfitted or underfitted [36]. Commonly used evaluation metrics include accuracy, precision, F1 score and recall. In addition, methods such as triple cross validation are usually used to evaluate and test the performance of the model more comprehensively. The final step is model deployment and application. The trained and validated models need to be deployed to real-world application environments. This may involve model optimization to improve operational efficiency and responsiveness [37]. In addition, the interpretability and security of the model need to be considered to ensure the reliability and stability of the model in real-world applications. This review first provides a detailed summary of the basic definitions and fundamental algorithms of DL, as well as the basic processes involved in DL.

DL in breast cancer diagnosis

Histopathology plays a crucial role in tumor diagnosis. Through detailed observation and analysis of tumor pathological tissue sections, pathologists can identify

and classify benign and malignant tumor characteristics. This is essential for determining the type of tumor and formulating appropriate treatment plans. Additionally, histopathology is used to assess the histological grading of tumors, providing information on tumor aggressiveness and patient prognosis. Through techniques such as immunohistochemical staining, histopathology can determine the status of key biomarkers, such as estrogen receptor (ER) and progesterone receptor (PR) expression, which are important for treatment decisions and prognosis assessment. Identifying benign and malignant tumor characteristics, histological grading, and determining biomarker status are key factors for improving the accuracy of breast cancer diagnosis. However, the main clinical issue currently is the limited number of pathologists, while the workload for pathological diagnosis is increasing. Pathologists need to meticulously examine tissue sections under a microscope to identify and classify tumor cells, a process that is both time-consuming and laborious, requiring a high level of expertise and skill. Additionally, traditional histopathological analysis is not only time-intensive but also relies on the subjective judgment of pathologists, leading to inconsistencies and errors. Variability in diagnoses among different pathologists can further exacerbate the uncertainty in diagnosis. The introduction of DL technology offers a new approach to addressing these challenges. Through automated and standardized image analysis, DL algorithms can quickly and efficiently process large volumes of tissue section images, significantly improving diagnostic efficiency and accuracy. DL can be used to identify aspects of tumor

benignity and malignancy, histological grading, and biomarker status, which in turn improves the accuracy of breast cancer diagnosis.

The role of DL in breast cancer identification and classification

The first is tumor identification and classification [38]. Cruz-Roa et al. developed a CNN to classify whether a breast cancer whole section image (WSI) plaque contains invasive ductal carcinoma. The model was trained on 400 slides and validated on 200 slides and achieved a pixel-level F1 score of 76% [39]. Han et al. on the other hand, trained a classifier to distinguish between 8 classes of benign and malignant breast tumors using the BreakHis dataset with an accuracy of 93.2% [40]. Faster Region-Based Convolutional Neural Network (FRCNN) is a DL model used for object detection. It is an improved version of the R-CNN series (including R-CNN and Fast R-CNN), significantly enhancing detection speed and accuracy [41]. Yap et al. applied the FRCNN model in breast cancer mass diagnosis and achieved excellent results. On multiple training sets, FRCNN achieved an average recall rate of 0.9236, precision rate of 0.9408, F1-score of 0.9321, and false alarm rate of 0.0621. When using RGB images, the performance of the FRCNN model was further improved, with a recall rate of 0.9572, precision rate of 0.9020, F1-score of 0.9288, and false alarm rate reduced to 0.1111 [42]. These data indicate that the FRCNN model has high accuracy and low false positive rates in breast cancer diagnosis. In the same year, Agarwal et al. also proposed a Faster Region-Based Convolutional Neural Network (Faster-RCNN) model for detecting breast cancer masses in Full-Field Digital Mammograms (FFDM). Their model achieved an accuracy of 0.93 and a false positive rate (FPI) of 0.78 on public databases. On the INbreast dataset, their model also showed superior performance, achieving an accuracy of 0.99 and an FPI of 1.69 for the diagnosis of malignant breast cancer masses, and an accuracy of 0.85 and an FPI of 1.0 for the diagnosis of benign masses [43]. RetinaNet is a DL model designed for object detection. The core idea of this model is the introduction of a loss function called Focal Loss, which addresses the issue of class imbalance between positive and negative samples, thereby enhancing detection performance [44]. One study employed a modified 3D RetinaNet model to detect breast lesions in ultrafast DCE-MRI sequences. The model leverages both spatial and temporal information to improve the detection of small lesions. In a dataset containing 572 lesions from ultrafast MRI scans, the model achieved a detection rate of 0.90, a sensitivity of 0.95, and a benign lesion detection rate of 0.81 [45]. Another study also utilized the RetinaNet model for cancer detection in mammography images. This model demonstrated outstanding

performance across multiple datasets, particularly excelling in detecting various manifestations of breast cancer, including masses, architectural distortion, and microcalcifications, achieving an AUC of 0.93 in the diagnosis of benign and malignant masses [46]. You Only Look Once (YOLO) is a DL model for real-time object detection, proposed by Joseph Redmon et al. in 2016 [47]. Unlike traditional object detection methods, YOLO transforms the object detection task into a single regression problem, achieving end-to-end object detection in an efficient manner. In the field of breast cancer diagnosis, Al-Masni et al. first proposed a Computer-Aided Diagnosis (CAD) system based on regional DL techniques, using a CNN called YOLO. This model comprises four main stages: preprocessing of mammograms, feature extraction using deep convolutional networks, mass detection with confidence, and mass classification using Fully Connected Neural Networks (FC-NNs). They utilized 600 original mammograms from the Digital Database for Screening Mammography (DDSM) and their augmented 2,400 mammograms to train and test the system. Through five-fold cross-validation tests, the results showed that the system detected mass locations with an overall accuracy of 99.7% and distinguished between benign and malignant lesions with an overall accuracy of 97% [48]. Su et al. proposed a dual model combining YOLO and LOGO (Local-Global) architecture for efficient mass detection and segmentation simultaneously. After testing on two independent mammogram datasets (CBIS-DDSM and INBreast), this model significantly outperformed previous works. On the CBIS-DDSM dataset, the true positive rate for mass detection was 95.7%, with an average precision of 65.0%; the F1-score for mass segmentation was 74.5%, and the Intersection over Union (IoU) was 64.0% [49].

The application of DL in diagnosing lymph node metastasis in breast cancer

For the diagnosis of lymph node metastasis, the seven DL algorithms developed by Bejnori et al. performed much better than the 11 pathologists in terms of diagnostic outcomes. The AUC of their best algorithm was 0.99, while the AUC of the pathologists' performance was 0.88 [50]. This suggests that DL techniques not only improve diagnostic accuracy, but also significantly reduce review time.

The role of DL in histological grading of breast cancer

Histologic grading is a critical step in the diagnosis of breast cancer. The Elston-Ellis revised Scarff-Bloom-Richardson grading system takes into account glandular formation, nuclear features, and mitotic activity [51]. The mitosis detection challenge presented by Veta et al. achieved a total F1 score of 0.61 using a model with a

10-layer deep convolutional neural network [52]. Tellez et al. used PHH3 stains in combination with CNN annotations, which, although not at the state-of-the-art level, demonstrated the potential of DL to improve the consistency of pathologists [53]. The above results indicate that DL is approaching or even surpassing advanced pathologists in diagnosing breast cancer. In addition, Veta and colleagues et al. proposed suppressed non-CNN algorithms, which are used for breast cancer lesion region and cell nucleus segmentation [54]. Their study focused on efficiently segmenting tumor cells directly from WSI for definitive diagnosis of breast cancer. The nuclear detection algorithm devised by Rexhepaj et al. quantified IHC staining of ER- and PR-expressed cells, with a manual and algorithmic quantitative correlation was 0.9 [55]. In addition, Couture et al. used a feature-based DL model trained on 571 H&E tissue microarray images and tested on 288 images, with a test accuracy of 84% for ER status [56]. Shamaï et al.'s DL system was able to predict 19 biomarker statuses including ER and PR, obtaining an accuracy of 92% [57].

The role of DL in diagnosing TNM staging of breast cancer

TNM staging and histopathological grading are equally important parameters in the diagnosis of breast cancer. The DL technique demonstrates significant potential in both. TNM staging assesses the risk grade and severity of the tumor by evaluating tumor size (T), lymph node involvement (N), and distant metastasis (M). The DL model of Chen et al. accurately detects metastatic cancer cells by analyzing the data from the lymph node images, which significantly improved the efficiency and accuracy of TNM staging with an AUC value of 0.99 [58].

DL utilization of thermography in the diagnosis of breast cancer

In the field of tumor detection, thermal imaging technology identifies abnormal heat distribution by capturing the infrared radiation emitted by tumor tissues. Due to the faster metabolism of tumor cells, more heat is generated, and thermal imaging can reveal these heat differences. The resulting thermal images can visually reflect the location and size of tumors [59, 60]. This technology has the advantages of being radiation-free, non-invasive, and capable of real-time detection, making it especially suitable for early screening. Using Inception V3, Inception V4, and modified Inception MV4 models in thermal-based early breast cancer detection has achieved extremely high accuracy. Different deep convolutional neural network (DCNN) models (Inception V3, Inception V4, and modified Inception MV4) have demonstrated excellent performance in thermal-based early breast cancer detection. The Inception V3 model, when using color images and Stochastic Gradient Descent with

Momentum (SGDM) optimization, achieved nearly 100% accuracy with a training time of 12.2 min and a learning rate of 1×10^{-3} . Using Adaptive Moment (ADAM) optimization, it also achieved 100% accuracy with a training time of 20.42 min at the 7th epoch and a learning rate of 1×10^{-4} . Root Mean Square Propagation (RMSPROP) optimization reached 100% accuracy with a training time of 14.87 min at 6 epochs and a learning rate of 1×10^{-3} . The Inception V4 model, when using color images and SGDM optimization, achieved 100% accuracy at 4 epochs with a learning rate of 1×10^{-4} and an area under the curve (AUC) of 1. In contrast, ADAM and RMSPROP optimizations did not reach 100% accuracy at all training stages, achieving maximum accuracies of 99.64% and 95.91%, respectively. The modified Inception MV4 model, when using color images and SGDM optimization with a learning rate of 1×10^{-4} , achieved 100% accuracy and a training speed 7% faster than Inception V4 [61]. In breast cancer self-detection, Al Husaini and his colleagues have developed a novel AI diagnostic tool. This tool, based on DL, integrates a smartphone, infrared camera, and thermal imaging of breast cancer to achieve self-detection. The process involves capturing thermal images with the infrared camera, sending these images to a cloud server via a smartphone, and then receiving the diagnostic report back on the smartphone after the cloud server processes the images. The entire process takes only 6 s. The tool they developed ensures that the transmitted thermal images do not exhibit significant differences due to image compression or transmission distance. On the contrary, the image precision remains very high, with a maximum detection accuracy deviation of only 1%. Using the DCNN model Inception MV4, they achieved highly accurate detection of breast cancer thermal images, with an accuracy rate of nearly 100%. The detection accuracy decreases by 11% only when the images are tilted during transmission [62]. The thermophysical properties of breast tissue and the application of local cooling gel play a crucial role in enhancing breast cancer detection through thermography. Researchers used COMSOL software to simulate the thermal behavior of breast tissue and applied the Pennes bioheat equation to illustrate the thermal distribution across different tissue layers. The study found that the temperature variation of the breast skin with and without tumors ranged from 2.58 °C to 0.274 °C. Larger breast sizes led to reduced temperature variations, making it difficult to observe thermal contrast, while increased tumor depth made detection more challenging. The closer the tumor was to the skin surface, the more significant the temperature change. Simulation results showed that at a tumor depth of 2 cm, the skin temperature could increase by approximately 1 °C, and at a depth of 10 cm, it was 0.98 °C. The use of cooling gel significantly enhanced the thermal contrast in breast

thermographic images. At a tumor depth of 10 cm, the use of cooling gel could achieve a temperature difference of 6 °C, which was not achievable in the simulation without cooling. After cooling, the temperature variation of the skin became more pronounced, making it easier for the detection equipment to identify abnormal thermal areas, thereby improving detection accuracy. By increasing thermal image contrast through cooling, it is possible to reduce false positives caused by normal tissue temperature variations, thereby enhancing detection reliability [63]. Recent research focuses on utilizing real-time thermography video streaming and DL models for early breast cancer detection in real-time. Equipped with a thermal imaging camera on a standard desktop computer, the study uses Inception v3, Inception v4, and a modified Inception Mv4 model to classify normal and abnormal breasts. The results demonstrate that the Inception Mv4 model, combined with real-time video streaming, can effectively detect the slightest temperature differences in breast tissue by generating thermal image sequences from different angles. This significantly enhances contrast, especially when using cooling gel, making the image acquisition process more efficient and increasing detection accuracy. For instance, when detecting a 1 cm tumor at a depth of 2 cm in breast tissue, the Inception v3 model, at 9 V, achieved a prediction accuracy range of 70–89%. After cooling with a gel for 2 min under the same conditions, the model's prediction accuracy increased to 80–94%. The Inception v4 model, at 9 V, showed a prediction accuracy range of 97–99.8%. With cooling for 2 min, the accuracy ranged from 71 to 97%. The Inception Mv4 model demonstrated the highest classification accuracy for breast cancer detection, especially with cooling. Cooling enhanced image contrast and significantly improved classification accuracy. The Inception Mv4 model performed exceptionally well in all measurements, achieving nearly 100% accuracy when cooling was used [64]. The DL technique automates and standardizes grading by automatically identifying and quantifying these features.

The application of DL technology in breast cancer diagnosis not only improves diagnostic accuracy and efficiency, but also demonstrates great potential in several aspects such as tumor identification, histological grading, and biomarker status determination. In the future, DL has an important role and a broad prospect in breast cancer diagnosis. The application of DL in breast cancer diagnosis has significantly improved the accuracy and efficiency of diagnoses. In addition to diagnosis, DL also plays an important role in predicting the prognosis of breast cancer.

DL in prognostic prediction of breast cancer

The use of DL in breast cancer prognosis prediction is also critical. Its accuracy in breast cancer prognosis prediction has been dramatically improved by evaluating the histomorphologic features of breast cancer, breast cancer grading, likelihood of recurrence, and multi-omics data.

Evaluating the histomorphological features of breast cancer to predict prognosis

Using DL to evaluate the histomorphology of breast cancer tissues can be employed to predict patient prognosis. Whitney and colleagues and others further demonstrated that the risk of recurrence in patients with ER-positive breast tumors could be independently predicted based on features such as nuclear shape and texture in pathology image data, with an even higher accuracy of 0.85 [65].

Prognostic value of tumor-infiltrating lymphocytes (TILs)

The structure and organization of TILs play an important role in the prognosis of clinical outcomes [66]. Makhoul et al.'s study found that breast cancer patients with high sTIL had significantly shorter survival times compared to those with low sTIL tumors. The hazard ratio (HR) for survival in patients with high sTIL was 1.6 (95% confidence interval [CI]=1.01–2.5, $P=0.04$) in the discovery cohort, and 2.5 (95% CI=1.3–4.5, $P=0.004$) in the validation cohort. Additionally, the presence of tumor-infiltrating lymphocytes (tTIL) within the tumor was also associated with shorter survival, with an HR of 1.7 (95% CI=1.08–2.6, $P=0.01$) in the discovery cohort and 2.0 (95% CI=1.06–3.7, $P=0.03$) in the validation cohort. Overall, the DL model effectively assessed sTIL and tTIL scores, and high levels of sTIL and tTIL were associated with poorer prognosis [67]. Choi et al. utilized a DL model to assess prognosis and treatment response in breast cancer patients. Their study results indicated that stromal TILs (sTILs) in breast cancer played a crucial role in predicting response to neoadjuvant chemotherapy. In patients with triple-negative breast cancer (TNBC) and HER2-positive breast cancer who received neoadjuvant chemotherapy, the DL model significantly improved sTIL scoring (26.8 ± 19.6 vs. 19.0 ± 16.4 , $P=0.003$), and high sTIL tumors were closely associated with better neoadjuvant chemotherapy response (odds ratio 1.28, 95% confidence interval 1.01–1.63, $P=0.039$) [68].

Evaluating histological grading to predict prognosis

Histological grading is a key prognostic factor in breast cancer. Wang and colleagues developed a new DL model for whole-slide images (WSI), named DeepGrade, aimed at improving risk stratification for breast cancer patients with Nottingham histological grade (NHG) 2. The DeepGrade model categorizes NHG 2 patients into two groups: DG2-high and DG2-low, assessing their prognostic value

with recurrence-free survival as the main outcome. The model indicated a higher risk of recurrence in the DG2-high group (hazard ratio [HR]=2.94, 95% confidence interval [CI]=1.24–6.97, $P=0.015$). The DG2-low group showed phenotypic similarities with NHG 1, while the DG2-high group was similar to NHG 3, suggesting the model can identify morphological patterns in NHG 2 associated with more aggressive tumors. Further evaluation of DeepGrade's prognostic value in external data confirmed an increased risk of recurrence in the DG2-high group (HR=1.91, CI=1.11–3.29, $P=0.019$). This model-based stratification provides prognostic information for NHG 2 tumor patients and offers a cost-effective alternative to molecular profiling [69].

Predicting prognosis by assessing lymph node metastasis with DL

Lymph node metastasis significantly increases the risk of breast cancer recurrence and is closely associated with poor prognosis in patients [70]. Verghese and colleagues developed a new DL model, smuLymphNet, to analyze axillary lymph node metastases in breast cancer patients. The most significant metrics were the effective quantification of germinal centers (GC) and sinuses [51]. The quantification of these lymph node features showed a significant correlation with distant metastasis-free survival (DMFS) in triple-negative breast cancer (TNBC) patients. The results from the smuLymphNet model indicated that lymph nodes unaffected by cancer with an average of ≥ 2 GCs were associated with longer DMFS (HR=0.28, $P=0.02$), and GCs were associated with better prognosis in LN-negative TNBC patients (HR=0.39, $P=0.039$) [71]. Another similar study focused on DL radiomics (DLR) using ultrasound imaging to predict lymph node metastasis and tumor burden in patients. They employed a DL model framework to predict axillary lymph node (ALN) status in breast cancer patients. The diagnostic performance was excellent, with an area under the receiver operating characteristic curve (AUC) of 0.902. Furthermore, their model was also capable of effectively distinguishing between low and heavy tumor burdens in patients, with an AUC of 0.905. This effectively links the critical feature of ALN with patient prognosis [72].

The role of DL in predicting recurrence risk

Then there is the recurrence risk prediction. DL techniques have also achieved excellent performance in predicting the risk of breast cancer recurrence. A recent study developed a new DL model to predict the risk of ipsilateral recurrence of DCIS by selecting the WSI of 344 DCIS patients who underwent lumpectomy and combining it with clinical data. Sections were stained by H&E and then analyzed and feature extracted by the DL

model. In the DL model, classifiers were first utilized to annotate normal ducts, cancerous ducts, vascular areas, stromal cells, and lymphocytes separately. Immediately after that, among the annotated regions, eight selected structural and spatial organization features were chosen to be trained on the model to predict the risk of tumor recurrence. The results of the study showed that the prediction results were 87% accurate on a 20% validation set and were able to predict the risk of recurrence in patients after 10 years well ($P<0.0001$) [73].

The role of DL in predicting homologous recombination defects (HRD)

The prediction of HRD has been a key indicator of breast cancer prognosis in recent years. DL methods have also shown great potential in predicting HRD status in breast cancers, especially BRCA1 and BRCA2 mutation-associated cancers. A recent study developed a robust DL-based method for HRD prediction using digitized H&E-stained tumor slides. The algorithm analyzed a large number of WSIs of TNBC and ductal breast cancers and predicted HRD with high accuracy at an AUC of 0.86, while also identifying morphological features associated with HRD [74].

Predicting patient prognosis through the integration of multi-omics technologies with DL

Finally, the integration of multi-omics technology. DL technology also shows great potential in the integrated analysis of multi-omics data. By combining multi-level information from the genome, epigenome, transcriptome, proteome, and metabolome, DL methods are able to provide more accurate prognosis prediction and personalized treatment plans for breast cancer patients. Radiography's ability to provide key features of tumor imaging in combination with genotype can be of great value in accurately predicting cancer prognosis. Yu et al. collected image data from patients scanned with DCE-MRI from four centers and tallied clinical data from the patients involved. The combination of these data was used to predict ALN metastasis in a DL model with an accuracy of 0.89 [75]. The study showed that the DL method, when combined with multimodal imaging and clinical data, significantly improved prediction accuracy.

The application of DL in prognostic prediction of breast cancer continues to improve the accuracy and efficiency of prediction with the advancement of technology. Meanwhile, DL also provides a solid technical foundation for personalized medicine and precision treatment. Through continuous optimization, DL will play an increasingly important role in clinical practice and provide better prognosis and treatment strategies for breast cancer patients. Here, we summarize the application of DL in breast cancer diagnosis and prognosis (Table 1). By

Table 1 Deep learning in breast cancer diagnosis and prognosis

Aspect	Study	Details	Performance	Ref
Diagnosis				
Tumor Identification and Classification	Cruz-Roa et al.	CNN to classify WSI for invasive ductal carcinoma	F1 score: 76%	[39]
Tumor Identification and Classification	Han et al.	Classify 8 classes of breast tumors using BreakHis dataset	Accuracy: 93.2%	[40]
Diagnosis of Lymph Node Metastasis	Bejnori et al.	DL algorithms vs. pathologists	AUC: 0.99 vs. 0.88	[50]
Histologic Grading	Veta et al.	Mitosis detection challenge	F1 score: 0.61	[52]
Histologic Grading	Tellez et al.	PHH3 stains with CNN annotations	Improved consistency of pathologists	[53]
Breast Cancer Lesion and Cell Nucleus Segmentation	Veta et al.	Suppressed non-CNN algorithms	Efficient segmentation	[54]
Tumor Identification and Classification	Rexhepaj et al.	Quantified ER- and PR-expressed cells	Correlation: 0.9	[55]
Tumor Identification and Classification	Couture et al.	Feature-based DL model on H&E tissue microarray images	Accuracy: 84%	[56]
Tumor Identification and Classification	Shamai et al.	Predict 19 biomarker statuses including ER and PR	Accuracy: 92%	[57]
TNM Staging	Chen et al.	Detect metastatic cancer cells from lymph node images	AUC: 0.99	[58]
Prognosis				
Recurrence Risk Prediction	Whitney et al.	ER-positive breast tumors; nuclear shape and texture features	Accuracy: 0.85	[65]
Prognostic Value of TILs	Makhlouf et al.	High sTIL associated with shorter survival	HR: 1.6 (discovery), 2.5 (validation)	[67]
Response to Neoadjuvant Chemotherapy	Choi et al.	High sTIL tumors associated with better response	Odds ratio: 1.28	[68]
Histological Grading	Wang et al.	DeepGrade model for NHG 2 patients	HR: 2.94	[69]
Lymph Node Metastasis	Verghese et al.	smLymphNet model for axillary lymph node analysis	HR: 0.28	[71]
Lymph Node Metastasis	Zheng et al.	Ultrasound imaging for ALN status prediction	AUC: 0.902	[72]
Recurrence Risk Prediction	Klimov et al.	WSI and clinical data for recurrence prediction	Accuracy: 87%	[73]
HRD Prediction	Lazard et al.	DL method for HRD prediction using H&E slides	AUC: 0.86	[74]
Multi-Omics Integration	Yu et al.	DCE-MRI data for ALN metastasis prediction	Accuracy: 0.89	[75]

analyzing patients' imaging and pathology data, DL models can effectively predict disease metastasis and other clinical outcomes. Additionally, DL technology is widely applied in the molecular typing of breast cancer, further advancing the development of personalized treatment.

The application of DL in molecular typing of breast cancer

Molecular typing of breast cancer refers to the classification of breast cancer into different subtypes based on gene expression profiles and molecular markers to better guide treatment and predict prognosis. The main molecular typing includes hormone receptor-positive (HR+), HER2-positive, and triple-negative (TNBC) breast cancers. Each subtype is significantly different in terms of pathology, biological behavior, and response to therapy. Therefore, accurate molecular typing is essential for the development of personalized treatment strategies. In recent years, significant progress has been made in the application of DL technology in molecular typing of breast cancer [76, 77]. The following is a detailed summary of the technology's use in hormone receptor status assessment, Ki67 proliferation index scoring, HER2 status assessment, homologous recombination defect (HRD) status prediction, integration of multi-omics technologies, and spatial transcriptomics applications [78].

Hormone receptor status

Assessing hormone receptor (HR) status is a critical step in the prognosis and customization of treatment for breast cancer (BC) patients. DL technologies have achieved remarkable results in this area [79, 80]. AI-enabled digital image analysis (DIA) has been able to significantly improve pathologists' consistency in IHC HR status assessment. The largest AI study to date, which included pathologists and WSI scanners at multiple sites, showed that pathologists agreed with AI-assisted results in the majority of ER/PgR cases. In addition, deep neural network (DNN)-based analysis of IHC stained images was effective in improving the accuracy of ER and PgR scores. The study developed a Morphology-Based Molecular Profiling (MBMP) technique using logistic regression and deep neural networks, showing comparable predictive outcomes to IHC in the majority of patients, with accuracy rates for predicting various HR expressions ranging from 91–98% [57]. Similarly, a DL algorithm based on DNN multi-instance learning demonstrated efficient sensitivity and specificity for positive and negative predictive values in determining BC HR status. The AUC was 0.92, with sensitivity and specificity of 0.932 and 0.741, respectively [81]. The application of these DL techniques greatly improved the efficiency and accuracy of HR status assessment.

Ki-67 proliferation index score

The Ki-67 proliferation index is an important prognostic and predictive marker for breast cancer. The consistency of DL tools in Ki-67 scoring has improved significantly. DL models can be used to predict the Ki-67 proliferation index in breast cancer tissues [82, 83]. Through neural network-optimized image segmentation and region identification, Fulawka et al.'s study reduced errors in tumor cell recognition, enhancing classification accuracy. Using an ensemble model, for a window size of 96 pixels, the mean absolute error (MAE) of Ki-67 index prediction was reduced from 0.058 to 0.034, significantly improving the stability and accuracy of predictions. Additionally, the study found that adjusting the threshold bias effectively reduced errors, lowering the MAE of a single model from 0.058 to 0.034, demonstrating the potential of DL technology to enhance the accuracy and reliability of breast cancer Ki-67 marker assessment [84]. Lee and colleagues utilized a DL model to analyze the relationship between Ki-67 levels and recurrence-free survival (RFS) in breast cancer patients. By predicting Ki-67 expression levels, the study revealed its association with the 21-gene recurrence score (RS) and its predictive ability for treatment outcomes and risk of recurrence. Patients with high Ki-67 expression in the low genomic risk group who did not receive chemotherapy showed a significantly increased risk of recurrence ($HR=2.51$; $P=0.008$), and high Ki-67 levels were significantly associated with lower recurrence-free survival rates after three years ($P=0.003$), as well as significantly correlated with secondary endocrine resistance ($OR=2.49$; $P=0.02$) [85]. Using a DL model to calculate the Ki67 index and segmenting the relevant cells by conventional image processing methods, tumor cells can be accurately identified for efficient determination. The application of this technology not only improved the speed and accuracy of the assay, but also significantly reduced inter-observer variability.

HER2 status assessment

HER2 expression is critical for effective breast cancer treatment. The latest DL algorithms have shown great accuracy in differentiating between HER2 positive and negative patients [86]. When Che et al. conducted DL analysis on HER2 images of breast cancer, the DL model they developed demonstrated extremely high accuracy in identifying HER2 expression in IHC 0 and 3+ images, with an overall accuracy of 97.9% [87]. In addition, Meng and her colleagues developed a DL Radiomics (DLR) model that combines ultrasound images and clinical data from breast cancer patients to predict the HER2 expression status. Utilizing Extreme Gradient Boosting (XGBoost) and Logistic Regression (LR) methods combined with clinical parameters, this DL model demonstrated excellent diagnostic performance on the test set,

with a specificity reaching 0.917 and an AUC of 0.810 [88]. For patients with HER2 scores of 1+ or 2+ or ISH-negative, the DL model also accurately identifies these tumors with very low HER2 expression. This has important implications for new therapies for treating BC with low HER2 (e.g., trastuzumab-democicam for BC with low HER2). Farahmand and colleagues developed a CNN-based classifier for accurately predicting HER2 status and trastuzumab treatment response. On an independent test set, the AUC for predicting HER2 status using WSI was 0.81. Additionally, this DL model achieved an AUC of 0.80 in predicting trastuzumab treatment response after a 5-fold cross-validation [89]. Interestingly, automated quantification of specific protein overexpression due to gene mutations is currently an important direction in DL as digital pathology images of H&E staining are intensively studied. This line of thought emphasizes that a multi-stage DL pipeline should be employed, which in turn improves the efficiency, accuracy and interpretability of the analysis. A DL architecture utilizing a spatial transformer network (STN) in combination with a visual transformer was used to detect HER2 expression in the absence of IHC staining with excellent results.

HRD status prediction

HRD is a loss of genetic repair mechanisms, often associated with mutations in the BRCA1 or BRCA2 genes, which are crucial for maintaining the repair of DNA double-strand breaks [90]. In breast cancer, the presence of HRD increases the sensitivity of breast cancer cells to certain drugs, such as PARP inhibitors, leading to drug resistance in tumor cells [91]. Therefore, predicting HRD status is particularly important for personalized treatment of breast cancer, especially in terms of selecting treatment strategies and predicting therapeutic efficacy. Lazard and colleagues used DL to predict HRD from H&E-stained WSI of breast cancer patients. Initially trained and tested on the public TCGA database, achieving an AUC of 0.83. Furthermore, their research was also able to identify specific morphological patterns associated with HRD, such as necrosis, high density of TIL, and nuclear heterogeneity, with an average AUC of 0.76 [74]. DL methods have shown significant potential in predicting HRD status. DL models using the Resnet architecture were able to successfully predict the status of tumors with high or low genomic instability (CIN), with results correlating with survival. In addition, other studies have directly inferred HRD by BRCA mutation status, with results showing significant predictive power. For example, the DeepSmile system combined self-supervised learning and multi-instance learning networks to achieve efficient HRD status prediction on the TCGA dataset, with an AUC of 0.81 [92]. The application of these DL methods helps to predict HRD status more accurately,

thus providing a more reliable reference for breast cancer treatment.

Integration of multi-omics techniques

DL technologies have also made significant progress in integrating multi-omics data. By combining genomic, epigenomic, transcriptomic, proteomic and metabolomic information at multiple levels, DL methods can provide more accurate prognosis prediction and personalized treatment plans for breast cancer patients [93, 94]. In a study involving a multi-omics analysis of breast cancer, researchers evaluated 1404 cases of invasive breast cancer, focusing on 11 biomarkers. By integrating DL and multiplex fluorescence immunohistochemistry, the study identified five biomarkers—progesterone receptor (PR), estrogen receptor (ER), androgen receptor (AR), GATA3, and PD-L1—that were significantly associated with overall survival in patients [95]. These markers emerged as independent risk factors in multivariate analysis, with p-values for each biomarker being less than or equal to 0.0095, demonstrating strong prognostic relevance with an overall p-value of less than 0.0001 [96]. An AI-based radiogenomics approach was able to significantly improve the accuracy of ALN metastasis prediction by analyzing DCE-MRI images and clinical data. DL methods have also been used to predict RNA expression in breast cancer patients, and the results show the effectiveness of these methods in providing spatially relevant genomic insights [97]. Spatial transcriptomics techniques are capable of mapping expression data to cells in histopathologic images, and despite their limited cost and availability, some studies have yielded preliminary results. For example, the ST-net algorithm was developed to capture the expression heterogeneity of multiple genes and successfully predicted the expression of multiple genes, with results closely correlating with RNA seq data from the TCGA dataset. In addition, other studies have successfully predicted TP53 and PIK3CA genomic alterations in breast cancer by utilizing convolutional neural networks and attentional mechanisms to associate the spatial content of WSI with genomic mutations [98, 99]. These studies demonstrate that DL technology can effectively combine spatial transcriptomics data with histopathology images to provide more in-depth gene expression and mutation information.

The application of DL in molecular typing of breast cancer has significantly improved the accuracy and consistency of the assessment of hormone receptors, Ki67 proliferation index and HER2 status. Meanwhile, the DL approach also showed great potential in predicting homologous recombination defect status and integrating multi-omics data. These findings suggest that DL will be a major breakthrough in molecular typing of breast cancer in the future. Here, we summarize the specific

applications of the DL model in breast cancer (Fig. 2). Through DL technology, the molecular typing of breast cancer has been greatly improved, aiding in the development of more precise treatment plans. Despite the significant potential of DL technology in breast cancer applications, there are still many challenges in its practical implementation.

Challenges in the application of DL in breast cancer

Although DL shows great potential in breast cancer research and diagnosis, it still faces many challenges in practical applications. These challenges focus on algorithms, data scarcity and heterogeneity, model performance and interpretability, ethical issues, and related software and hardware. The following is a detailed summary of these challenges.

Algorithmic challenges

Lack of annotated data for algorithm training

Training effective AI algorithms requires a large sample of high-quality, annotated images. The task of annotating these images is typically done by professional organizational pathologists, a time-consuming and expensive process. Additionally, pathologists may suffer from issues such as low image resolution, feature ambiguity, and slow network connectivity when annotating. While crowdsourcing labeled data is a viable solution, it may introduce inter-observer variability, leading to inconsistent data quality. Small sample sizes also limit the management of labeled data [100, 101]. To cope with these issues, methods such as data augmentation and active learning can be employed to maximize the utility of existing samples.

Integration into real-world algorithm validation

Extensive integration of AI algorithms into digital pathology workflows requires extensive validation of ML and DL algorithms against multi-institutional data to ensure the reproducibility and accuracy of their predictions. Differences in slide preparation methods, scanner models, and image formats and analysis methods underscore the importance of standardization schemes and data normalization techniques [102]. These techniques not only improve the generalizability of algorithms, but also allow for the creation of merged, common datasets that can be used to retrain models to account for natural variation across populations. Currently, there is a lack of official guidelines on the number of annotations, images, and laboratories needed to capture natural variation, although there are some guiding protocols (e.g., TRI-POD-AI and PROBAST-AI) [103].

Lack of transparency and interpretability

In digital pathology, most AI algorithms categorize samples in a binary way and lack transparency into the

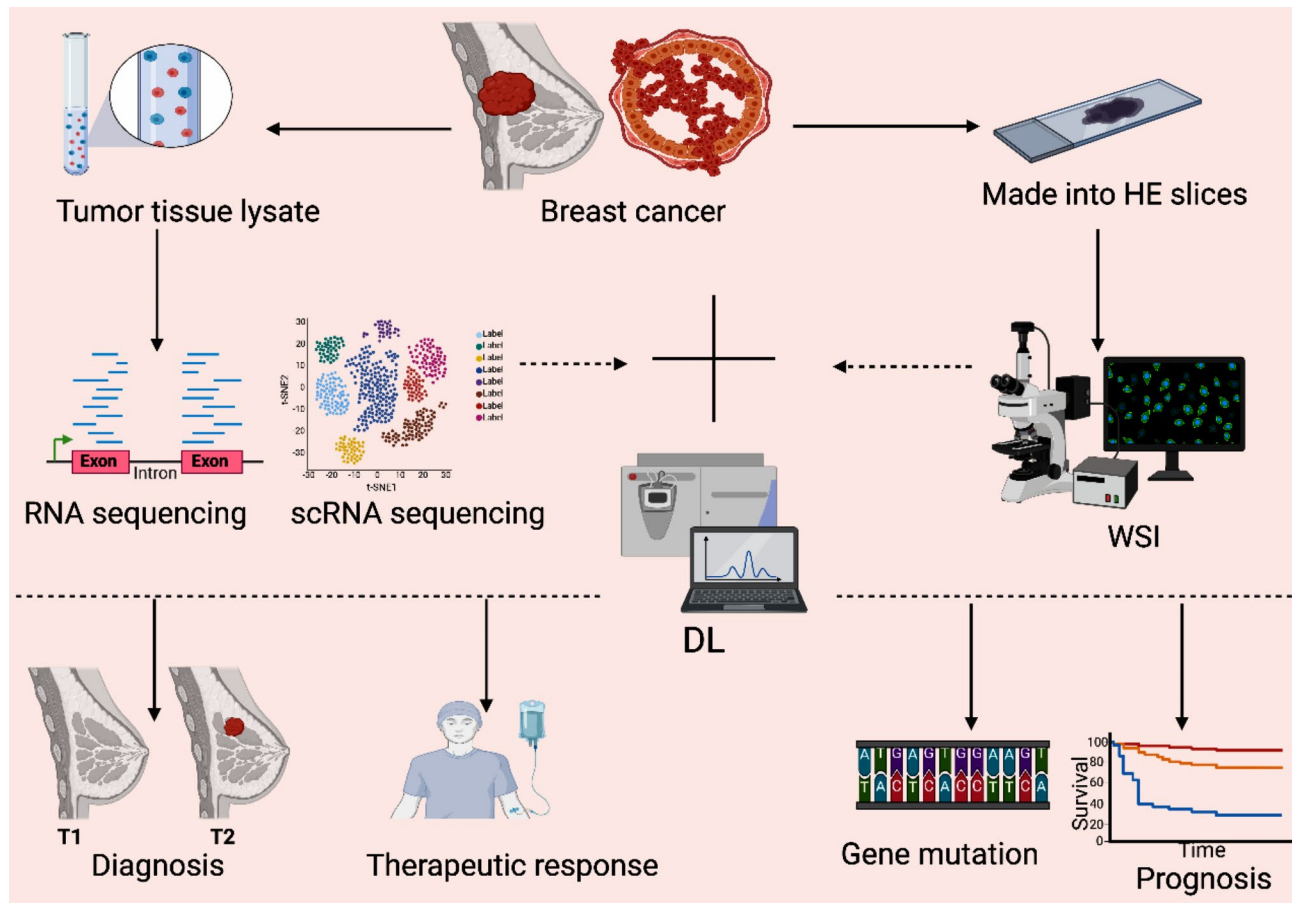


Fig. 2 Specific application of DL in breast cancer

model’s decision-making process, a problem known as the “black box” problem. In contrast, pathologists’ decisions combine experience, cognition, clinical context, and descriptive language to better label and discuss complex or rare cases [104]. The opacity of AI models may affect their adoption in clinical settings, as clinicians need to explain the rationale for their decisions to patients in order to gain their trust.

Diverse and scarce data sources

Multi-center and high-quality data are the basis for developing relevant DL models. However, there are issues with data sharing across different centers, scanners used in different centers vary significantly due to manufacturer and operator skill level. The pre-processing process after image acquisition can also lead to significant heterogeneity in the quality of images due to different algorithms [105]. Although DL can increase the quantity of data through data augmentation and improve the quality of data using methods such as unsupervised learning, there is still a lack of standardized algorithms and processes [106]. The problem of data sources is a key constraint

to maximize the generalization ability, repeatability, and generalizability of DL models.

Model performance and interpretability

Most AI algorithms lack prospective validation, limiting their application to different patient populations. Prospective validation trials should test different patient populations to realistically assess the clinical utility of AI. In addition, the “black box” nature of AI models prevents clinicians from trusting their results [107, 108]. Although data visualization tools can support the understanding of algorithmic decisions, more research is needed to improve the interpretability and transparency of models.

Ethical issues

The use of AI systems in clinical decision-making raises a number of ethical issues, including attribution of responsibility for wrong decisions, public perception of AI decision-making tools, and data security and privacy concerns. Standard regulations should be developed alongside the establishment of new technologies to address the transparency and accountability of AI decision-making [109]. There is also a need to avoid over-reliance on

automation and to ensure that AI systems do not ignore common sense and ethical considerations in assisting the decision-making process.

Software and hardware challenges

Affordability of computing costs

Widespread adoption of AI-based digital pathology in clinical settings requires high hardware, software, and storage costs. DL algorithm training relies on dedicated Graphics Processing Units (GPUs), which typically require dedicated computers and high-cost GPU server clusters. Long-term data storage costs are also a major challenge, especially for programs that need to store hundreds of gigabytes of images [110]. Despite the availability of cloud-based storage solutions, many hospitals rely on on-premise data storage, further exacerbating the cost issue.

Software scalability and reliability

Although commercial DL-based software has gained some acceptance, it still has limitations in covering all clinical workflows. In actual clinical practice, DL systems need to overcome data scarcity, heterogeneity, and ethical issues to achieve widespread adoption [111]. In addition, more efficient unsupervised learning models need to be developed to reduce the reliance on large amounts of labeled data and to address label noise and data imbalance in medical image analysis.

Medical data privacy and security

Protecting medical data privacy poses a serious challenge to ML and DL algorithms. These algorithms require large amounts of training data, but privacy regulations limit data sharing. Collaborative and decentralized training methods allow model development without sharing patient data. In addition, the robustness of data acquisition methods needs to be addressed to improve the accuracy and consistency of image analysis [112].

Multimodal data integration

Only a few DL models currently use non-imaging features in conjunction with imaging data. More models that combine radiomics features with image data need to be developed in the future to improve predictive performance and clinical relevance. The development of these models will provide new avenues for personalized treatment and prognosis prediction [113].

Although DL technology shows great potential in breast cancer research and diagnosis, it still faces many challenges in practical applications. These challenges include algorithm transparency and interpretability, data scarcity and heterogeneity, computational costs, ethical issues, and medical data privacy protection. Addressing these challenges requires interdisciplinary collaboration

and innovation to ensure that DL technologies maximize their potential in clinical practice and provide better diagnostic and treatment strategies for breast cancer patients.

Perspectives and conclusion

With the continuous development of AI and deep DL technologies, the future application of these technologies in breast cancer diagnosis and treatment is promising. First, AI and DL are expected to significantly improve the diagnostic performance of breast cancer and its metastasis through the integration of multiple imaging modalities, such as MRI and ultrasound, especially in areas with limited medical resources. This combination of multimodal imaging will compensate for the shortcomings of a single imaging modality and improve diagnostic accuracy and early detection. In addition, AI demonstrates great potential in integrating multi-omics data streams to provide personalized diagnostic, staging, treatment and prognostic recommendations by generating large-scale language models (LLMs). These models can combine tumor characteristics, individual information, and social network data to optimize the output of current visual models and present information in a natural language format, enhancing the accuracy and efficiency of clinical decision-making. However, data sharing limitations remain a major challenge. Methods such as federated learning can enhance the performance and generalization of DL models by protecting patient privacy for model training, facilitating cross-institutional data sharing, and helping to build large-scale and diverse datasets. Another key development direction is to enhance the transparency and interpretability of algorithms to increase clinicians' trust in AI systems and to ensure the operability of these systems in real clinical settings. There is also a need to address legal and ethical issues, such as decision-making responsibility and data privacy, to ensure the safety and reliability of AI systems. By overcoming these challenges, AI and DL will play a greater role in the full-stack analysis of breast cancer, optimizing clinical workflows and improving patient outcomes. Here, we summarize relevant clinical trials of DL in breast cancer diagnosis, treatment, and prognosis (Table 2).

Although AI and deep DL show great potential in breast cancer diagnosis, prognosis and molecular typing, their practical application still faces many challenges. These challenges include the lack of large-scale high-quality annotated datasets, algorithmic transparency and interpretability issues, and high computational and storage costs. In addition, legal and ethical issues, such as decision responsibility and patient privacy, need to be fully discussed and resolved before clinical implementation. In order to achieve effective application of DL in clinical practice, large-scale prospective trials must

Table 2 Clinical trials of deep learning in breast cancer diagnosis, treatment, and prognosis

NCT Number	Title	Status	Study Results	Conditions	Interventions	Characteristics
NCT06016790	Exploring a Breast Cancer Early Screening Model Based on cfDNA	Recruiting	No Results Available	•Breast Cancer •Breast Fibroadenoma	•Genetic: cfDNA sequencing	Phase: Not Applicable
NCT05443672	Multi-center Study of Deep Learning AI in Breast Mass	Unknown status	No Results Available	•Breast Neoplasms	•Device: Yizhun BUSMS	Phase: Not Applicable
NCT05243121	Artificial Intelligence Analysis for Magnetic Resonance Imaging in Screening and Diagnosis of Breast Cancer	Recruiting	No Results Available	•Breast Neoplasms •Magnetic Resonance Imaging	•Diagnostic Test: MRI	Phase: Not Applicable
NCT05021055	Multicentric Study for External Validation of a Deep Learning Model for Mammographic Breast Density Categorization	Unknown status	No Results Available	•Breast Cancer		Phase: Not Applicable
NCT04996615	Artificial Intelligence Analysis for Magnetic Resonance Imaging in Screening Breast Cancer in High-risk Women	Recruiting	No Results Available	•Breast Cancer •Magnetic Resonance Imaging	•Other: no intervention	Phase: Not Applicable
NCT04270032	Using Deep Learning Methods to Analyze Automated Breast Ultrasound and Hand-held Ultrasound Images, to Establish a Diagnosis, Therapy Assessment and Prognosis Prediction Model of Breast Cancer.	Recruiting	No Results Available	•Breast Cancer	•Diagnostic Test: ABUS and HHUS	Phase: Not Applicable
NCT04156880	Artificial Intelligence in Mammography-Based Breast Cancer Screening	Withdrawn	No Results Available	•Breast Cancer	•Other: mammography	Phase: Not Applicable
NCT04003558	Deep Learning Algorithms for Prediction of Lymph Node Metastasis and Prognosis in Breast Cancer MRI Radiomics (RBC-01)	Recruiting	No Results Available	•Breast Neoplasm Female •Early-stage Breast Cancer •Radiomics •Axillary Lymph Node •Survival, Prosthesis	•Other: No interventions	Phase: Not Applicable
NCT03851497	Application of Deep-learning and Ultrasound Elastography in Opportunistic Screening of Breast Cancer	Completed	No Results Available	•Breast Cancer		Phase: Not Applicable
NCT03706534	Breast Ultrasound Image Reviewed With Assistance of Deep Learning Algorithms	Unknown status	No Results Available	•Breast Cancer •Breast Lesions •Breast Mass	•Device: Ultrasound Image review with CADe •Device: Ultrasound Image review with CADx •Device: Ultrasound Image manual review •Procedure: Biopsy	Phase: Not Applicable

be conducted to validate the performance and practical effects of DL systems, and standardized data sharing and validation mechanisms must be established. Through continuous optimization and innovation, AI and DL will play an increasingly important role in the diagnosis and treatment of breast cancer, ultimately improving patient outcomes and quality of life.

Acknowledgements
Not applicable.

Author contributions
B.J., Y.Y. and L.B. conceived the study. B.J., L.B. and S.H. drafted the manuscript. Z.J., X.C. and Y.Y. revised the manuscript. All authors read and approved the final manuscript.

Funding
Not applicable.

Data availability
No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
All authors consent to publication.

Competing interests
The authors declare no competing interests.

Received: 5 July 2024 / Accepted: 16 September 2024

Published online: 20 September 2024

References

- Nolan E, Lindeman GJ, Visvader JE. Deciphering breast cancer: from biology to the clinic. *Cell*. 2023;186:1708–28.
- Onkar SS, Carleton NM, Lucas PC, et al. The great Immune escape: understanding the Divergent Immune response in breast Cancer subtypes. *Cancer Discov*. 2023;13:23–40.
- Cardoso MJ, Poortmans P, Senkus E, et al. Breast cancer highlights from 2023: knowledge to guide practice and future research. *Breast*. 2024;74:103674.
- Laws A, Punglia RS. Endocrine therapy for primary and secondary Prevention after diagnosis of high-risk breast lesions or preinvasive breast Cancer. *J Clin Oncol*. 2023;41:3092–9.
- Loibl S, André F, Bachelot T, et al. Early breast cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. *Ann Oncol*. 2024;35:159–82.
- Nicholson WK, Silverstein M, Wong JB, et al. Screening for breast Cancer: US Preventive Services Task Force Recommendation Statement. *JAMA*. 2024;331:1918–30.
- Unger M, Kather JN. Deep learning in cancer genomics and histopathology. *Genome Med*. 2024;16:44.
- Koetzier LR, Mastrodicasa D, Szczykutowicz TP, et al. Deep Learning Image Reconstruction for CT: technical principles and clinical prospects. *Radiology*. 2023;306:e221257.
- Zhang J, Wu J, Zhou XS, et al. Recent advancements in artificial intelligence for breast cancer: image augmentation, segmentation, diagnosis, and prognosis approaches. *Semin Cancer Biol*. 2023;96:11–25.
- Zhao X, Bai JW, Guo Q, et al. Clinical applications of deep learning in breast MRI. *Biochim Biophys Acta Rev Cancer*. 2023;1878:188864.
- Zhang C, Xu J, Tang R, et al. Novel research and future prospects of artificial intelligence in cancer diagnosis and treatment. *J Hematol Oncol*. 2023;16:114.
- Cooper M, Ji Z, Krishnan RG. Machine learning in computational histopathology: challenges and opportunities. *Genes Chromosomes Cancer*. 2023;62:540–56.
- Amorim JP, Abreu PH, Fernandez A, et al. Interpreting Deep Machine Learning models: an Easy Guide for oncologists. *IEEE Rev Biomed Eng*. 2023;16:192–207.
- Anwar SM, Majid M, Qayyum A, et al. Medical Image Analysis using Convolutional neural networks: a review. *J Med Syst*. 2018;42:226.
- He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition. 2016; 770–778.
- Simonyan K, Zisserman AJ. Very deep convolutional networks for large-scale image recognition. 2014.
- Szegedy C, Liu W, Jia Y et al. Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition. 2015; 1–9.
- Zhang J, Zhang Y, Jin Y, et al. MDU-Net: multi-scale densely connected U-Net for biomedical image segmentation. *Health Inf Sci Syst*. 2023;11:13.
- Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5–9, 2015, proceedings, part III 18. Springer. 2015; 234–241.
- Xuan X, Zhang X, Kwon OH, Ma KL. VAC-CNN: a visual Analytics System for comparative studies of deep convolutional neural networks. *IEEE Trans Vis Comput Graph*. 2022;28:2326–37.
- Vaswani A, Shazeer N, Parmar N et al. Attention is all you need. 2017; 30.
- Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16x16 words. Transformers for image recognition at scale; 2020.
- Iqbal A, Sharif M, Yasmin M, et al. Generative adversarial networks and its applications in the biomedical image segmentation: a comprehensive survey. *Int J Multimed Inf Retr*. 2022;11:333–68.
- Kazemina S, Baur C, Kuijper A, et al. GANs for medical image analysis. *Artif Intell Med*. 2020;109:101938.
- Shokraei Fard A, Reutens DC, Vegh V. From CNNs to GANs for cross-modality medical image estimation. *Comput Biol Med*. 2022;146:105556.
- Chen L, Pan X, Zhang YH, et al. Classification of widely and rarely expressed genes with recurrent neural network. *Comput Struct Biotechnol J*. 2019;17:49–60.
- Wen A, Fu S, Moon S, et al. Desiderata for delivering NLP to accelerate health-care AI advancement and a Mayo Clinic NLP-as-a-service implementation. *NPJ Digit Med*. 2019;2:130.
- Matsuo Y, LeCun Y, Sahani M, et al. Deep learning, reinforcement learning, and world models. *Neural Netw*. 2022;152:267–75.
- Zhou SK, Le HN, Luu K, et al. Deep reinforcement learning in medical imaging: a literature review. *Med Image Anal*. 2021;73:102193.
- Mandair D, Reis-Filho JS, Ashworth A. Biological insights and novel biomarker discovery through deep learning approaches in breast cancer histopathology. *NPJ Breast Cancer*. 2023;9:21.
- Rai HM, Yoo J. A comprehensive analysis of recent advancements in cancer detection using machine learning and deep learning models for improved diagnostics. *J Cancer Res Clin Oncol*. 2023;149:14365–408.
- Xiao B, Xu B, Bi X, Li W. Global-feature encoding U-Net (GEU-Net) for Multi-focus Image Fusion. *IEEE Trans Image Process*. 2021;30:163–75.
- Jiang X, Hu Z, Wang S, Zhang Y. Deep learning for medical image-based Cancer diagnosis. *Cancers (Basel)*. 2023; 15.
- Petinrin OO, Saeed F, Toseef M, et al. Machine learning in metastatic cancer research: potentials, possibilities, and prospects. *Comput Struct Biotechnol J*. 2023;21:2454–70.
- Yang H, Chen R, Li D, Wang Z. Subtype-GAN: a deep learning approach for integrative cancer subtyping of multi-omics data. *Bioinformatics*. 2021;37:2231–7.
- Daneshjou R, He B, Ouyang D, Zou JY. How to evaluate deep learning for cancer diagnostics - factors and recommendations. *Biochim Biophys Acta Rev Cancer*. 2021;1875:188515.
- Huang S, Yang J, Fong S, Zhao Q. Artificial intelligence in cancer diagnosis and prognosis: opportunities and challenges. *Cancer Lett*. 2020;471:61–71.
- Duggento A, Conti A, Mauriello A, et al. Deep computational pathology in breast cancer. *Semin Cancer Biol*. 2021;72:226–37.
- Cruz-Roa A, Gilmore H, Basavanahally A, et al. High-throughput adaptive sampling for whole-slide histopathology image analysis (HASHI) via convolutional neural networks: application to invasive breast cancer detection. *PLoS ONE*. 2018;13:e0196828.
- Han Z, Wei B, Zheng Y, et al. Breast Cancer multi-classification from histopathological images with structured Deep Learning Model. *Sci Rep*. 2017;7:4172.
- Ren S, He K, Girshick R, Sun, J. Faster r-cnn: towards real-time object detection with region proposal networks. 2015; 28.
- Yap MH, Goyal M, Osman F, et al. Breast ultrasound region of interest detection and lesion localisation. *Artif Intell Med*. 2020;107:101880.
- Agarwal R, Diaz O, Yap MH, et al. Deep learning for mass detection in full field Digital mammograms. *Comput Biol Med*. 2020;121:103774.
- Lin T-Y, Goyal P, Girshick R et al. Focal loss for dense object detection. In Proceedings of the IEEE international conference on computer vision. 2017; 2980–2988.
- Ayatollahi F, Shokouhi SB, Mann RM, Teuwen J. Automatic breast lesion detection in ultrafast DCE-MRI using deep learning. *Med Phys*. 2021;48:5897–907.
- Ueda D, Yamamoto A, Onoda N, et al. Development and validation of a deep learning model for detection of breast cancers in mammography from multi-institutional datasets. *PLoS ONE*. 2022;17:e0265751.
- Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition. 2016; 779–788.
- Al-Masni MA, Al-Antari MA, Park JM, et al. Simultaneous detection and classification of breast masses in digital mammograms via a deep learning YOLO-based CAD system. *Comput Methods Programs Biomed*. 2018;157:85–94.
- Su Y, Liu Q, Xie W, Hu P. YOLO-LOGO: a transformer-based YOLO segmentation model for breast mass detection and segmentation in digital mammograms. *Comput Methods Programs Biomed*. 2022;221:106903.
- Ehteshami Bejnordi B, Veta M, van Johannes P, et al. Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in women with breast Cancer. *JAMA*. 2017;318:2199–210.
- Zhu X, Wolfgruber TK, Leong L, et al. Deep learning predicts interval and screening-detected Cancer from Screening mammograms: a case-control study in 6369 women. *Radiology*. 2021;301:550–8.
- Veta M, van Diest PJ, Willems SM, et al. Assessment of algorithms for mitosis detection in breast cancer histopathology images. *Med Image Anal*. 2015;20:237–48.
- Tellez D, Balkenhol M, Otte-Holler I, et al. Whole-slide mitosis detection in H&E breast histology using PHH3 as a reference to Train Distilled Stain-Invariant Convolutional Networks. *IEEE Trans Med Imaging*. 2018;37:2126–36.

54. Veta M, Kornegoor R, Huisman A, et al. Prognostic value of automatically extracted nuclear morphometric features in whole slide images of male breast cancer. *Mod Pathol*. 2012;25:1559–65.
55. Rexhepaj E, Brennan DJ, Holloway P, et al. Novel image analysis approach for quantifying expression of nuclear proteins assessed by immunohistochemistry: application to measurement of oestrogen and progesterone receptor levels in breast cancer. *Breast Cancer Res*. 2008;10:R89.
56. Couture HD, Williams LA, Geradts J, et al. Image analysis with deep learning to predict breast cancer grade, ER status, histologic subtype, and intrinsic subtype. *NPJ Breast Cancer*. 2018;4:30.
57. Shamai G, Binenbaum Y, Slossberg R, et al. Artificial Intelligence Algorithms to assess Hormonal Status from tissue microarrays in patients with breast Cancer. *JAMA Netw Open*. 2019;2:e197700.
58. Chen J, Yang Y, Luo B, et al. Further predictive value of lymphovascular invasion explored via supervised deep learning for lymph node metastases in breast cancer. *Hum Pathol*. 2023;131:26–37.
59. Jalloul R, Chethan HK, Alkhatib R. A review of machine learning techniques for the classification and detection of breast Cancer from medical images. *Diagnostics (Basel)* 2023; 13.
60. Singh D, Singh AK. Role of image thermography in early breast cancer detection- Past, present and future. *Comput Methods Programs Biomed*. 2020;183:105074.
61. Al Husaini MAS, Habaebi MH, Gunawan TS, et al. Thermal-based early breast cancer detection using inception V3, inception V4 and modified inception MV4. *Neural Comput Appl*. 2022;34:333–48.
62. Al Husaini MAS, Hadi Habaebi M, Gunawan TS, Islam MR. Self-detection of early breast Cancer Application with Infrared Camera and Deep Learning. *Electronics*. 2021;10:2538.
63. Al Husaini MAS, Habaebi MH, Suliman FM, et al. Influence of tissue thermo-physical characteristics and situ-cooling on the detection of breast Cancer. *Appl Sci*. 2023;13:8752.
64. Al Husaini MAS, Habaebi MH, Islam MR. Utilizing Deep Learning for the Real-Time Detection of Breast Cancer through Thermography. In 2023 9th International Conference on Computer and Communication Engineering (ICCCCE). IEEE. 2023; 270–273.
65. Whitney J, Corredor G, Janowczyk A, et al. Quantitative nuclear histomorphometry predicts oncotype DX risk categories for early stage ER+ breast cancer. *BMC Cancer*. 2018;18:610.
66. Saltz J, Gupta R, Hou L, et al. Spatial Organization and molecular correlation of Tumor-infiltrating lymphocytes using Deep Learning on Pathology images. *Cell Rep*. 2018;23:181–e193187.
67. Makhlof S, Wahab N, Toss M, et al. Evaluation of tumour infiltrating lymphocytes in luminal breast cancer using artificial intelligence. *Br J Cancer*. 2023;129:1747–58.
68. Choi S, Cho SI, Jung W, et al. Deep learning model improves tumor-infiltrating lymphocyte evaluation and therapeutic response prediction in breast cancer. *NPJ Breast Cancer*. 2023;9:71.
69. Wang Y, Acs B, Robertson S, et al. Improved breast cancer histological grading using deep learning. *Ann Oncol*. 2022;33:89–98.
70. Vrdoljak J, Krešo A, Kumrić M et al. The role of AI in breast Cancer lymph node classification: a Comprehensive Review. *Cancers (Basel)* 2023; 15.
71. Verghese G, Li M, Liu F, et al. Multiscale deep learning framework captures systemic immune features in lymph nodes predictive of triple negative breast cancer outcome in large-scale studies. *J Pathol*. 2023;260:376–89.
72. Zheng X, Yao Z, Huang Y, et al. Deep learning radiomics can predict axillary lymph node status in early-stage breast cancer. *Nat Commun*. 2020;11:1236.
73. Klimov S, Miligy IM, Gertych A, et al. A whole slide image-based machine learning approach to predict ductal carcinoma in situ (DCIS) recurrence risk. *Breast Cancer Res*. 2019;21:83.
74. Lazard T, Bataillon G, Naylor P, et al. Deep learning identifies morphological patterns of homologous recombination deficiency in luminal breast cancers from whole slide images. *Cell Rep Med*. 2022;3:100872.
75. Yu Y, Tan Y, Xie C, et al. Development and validation of a Preoperative Magnetic Resonance Imaging Radiomics-Based Signature to Predict Axillary Lymph Node Metastasis and Disease-Free Survival in patients with early-stage breast Cancer. *JAMA Netw Open*. 2020;3:e2028086.
76. Furtney I, Bradley R, Kabuka MR. Patient Graph Deep Learning to predict breast Cancer Molecular Subtype. *IEEE/ACM Trans Comput Biol Bioinform*. 2023;20:3117–27.
77. Malik V, Kalakoti Y, Sundar D. Deep learning assisted multi-omics integration for survival and drug-response prediction in breast cancer. *BMC Genomics*. 2021;22:214.
78. Liu J, Zhao H, Zheng Y, et al. DrABC: deep learning accurately predicts germline pathogenic mutation status in breast cancer patients based on phenotypic data. *Genome Med*. 2022;14:21.
79. Hanker AB, Sudhan DR, Arteaga CL. Overcoming endocrine resistance in breast Cancer. *Cancer Cell*. 2020;37:496–513.
80. Moisan A, Madéry M, Boyer T et al. Hormone receptor signaling and breast Cancer Resistance to Anti-tumor Immunity. *Int J Mol Sci* 2023; 24.
81. Naik N, Madani A, Esteva A, et al. Deep learning-enabled breast cancer hormonal receptor status determination from base-level H&E stains. *Nat Commun*. 2020;11:5727.
82. Kreipe H, Harbeck N, Christgen M. Clinical validity and clinical utility of Ki67 in early breast cancer. *Ther Adv Med Oncol*. 2022;14:17588359221122725.
83. Zhang A, Wang X, Fan C, Mao X. The role of Ki67 in evaluating neoadjuvant endocrine therapy of hormone receptor-positive breast Cancer. *Front Endocrinol (Lausanne)*. 2021;12:687244.
84. Fulawka L, Blaszczyk J, Tabakov M, Halon A. Assessment of Ki-67 proliferation index with deep learning in DCIS (ductal carcinoma in situ). *Sci Rep*. 2022;12:3166.
85. Lee J, Lee YJ, Bae SJ, et al. Ki-67, 21-Gene recurrence score, Endocrine Resistance, and survival in patients with breast Cancer. *JAMA Netw Open*. 2023;6:e2330961.
86. Tarantino P, Viale G, Press MF, et al. ESMO expert consensus statements (ECS) on the definition, diagnosis, and management of HER2-low breast cancer. *Ann Oncol*. 2023;34:645–59.
87. Che Y, Ren F, Zhang X et al. Immunohistochemical HER2 recognition and analysis of breast Cancer based on deep learning. *Diagnostics (Basel)* 2023; 13.
88. Quan MY, Huang YX, Wang CY, et al. Deep learning radiomics model based on breast ultrasound video to predict HER2 expression status. *Front Endocrinol (Lausanne)*. 2023;14:1144812.
89. Farahmand S, Fernandez AI, Ahmed FS, et al. Deep learning trained on hematoxylin and eosin tumor region of interest predicts HER2 status and trastuzumab treatment response in HER2+ breast cancer. *Mod Pathol*. 2022;35:44–51.
90. Doig KD, Fellowes AP, Fox SB. Homologous recombination Repair Deficiency: an overview for pathologists. *Mod Pathol*. 2023;36:100049.
91. Daly GR, Naidoo S, Alabulrahman M, et al. Screening and testing for homologous recombination Repair Deficiency (HRD) in breast Cancer: an overview of the Current Global Landscape. *Curr Oncol Rep*; 2024.
92. Schirris Y, Gavves E, Nederlof I, et al. DeepSMILE: contrastive self-supervised pre-training benefits MSI and HRD classification directly from H&E whole-slide images in colorectal and breast cancer. *Med Image Anal*. 2022;79:102464.
93. Ortiz MMO, Andrechek ER. Molecular characterization and Landscape of breast cancer models from a multi-omics perspective. *J Mammary Gland Biol Neoplasia*. 2023;28:12.
94. Neagu AN, Whitham D, Bruno P et al. Omics-Based Investigations Breast Cancer Molecules 2023; 28.
95. Shamai G, Livne A, Polónia A, et al. Deep learning-based image analysis predicts PD-L1 status from H&E-stained histopathology images in breast cancer. *Nat Commun*. 2022;13:6753.
96. Ogunleye AZ, Piyawajanusorn C, Gonçalves A, et al. Interpretable Machine Learning Models to predict the resistance of breast Cancer patients to Doxorubicin from their microRNA profiles. *Adv Sci (Weinh)*. 2022;9:e2201501.
97. He B, Bergensträhle L, Stenbeck L, et al. Integrating spatial gene expression and breast tumour morphology via deep learning. *Nat Biomed Eng*. 2020;4:827–34.
98. Grimm LJ, Mazurowski MA. Breast Cancer Radiogenomics: current status and future directions. *Acad Radiol*. 2020;27:39–46.
99. Low SK, Zembutsu H, Nakamura Y. Breast cancer: the translation of big genomic data to cancer precision medicine. *Cancer Sci*. 2018;109:497–506.
100. Bera K, Braman N, Gupta A, et al. Predicting cancer outcomes with radiomics and artificial intelligence in radiology. *Nat Rev Clin Oncol*. 2022;19:132–46.
101. Huang S, Yang J, Shen N, et al. Artificial intelligence in lung cancer diagnosis and prognosis: current application and future perspective. *Semin Cancer Biol*. 2023;89:30–7.
102. Shmatko A, Ghaffari Laleh N, Gerstung M, Kather JN. Artificial intelligence in histopathology: enhancing cancer research and clinical oncology. *Nat Cancer*. 2022;3:1026–38.
103. He X, Liu X, Zuo F, et al. Artificial intelligence-based multi-omics analysis fuels cancer precision medicine. *Semin Cancer Biol*. 2023;88:187–200.

104. Yan S, Li J, Wu W. Artificial intelligence in breast cancer: application and future perspectives. *J Cancer Res Clin Oncol*. 2023;149:16179–90.
105. Wu X, Li W, Tu H. Big data and artificial intelligence in cancer research. *Trends Cancer*. 2024;10:147–60.
106. Saleh H, Abd-El Ghany SF, Alyami H, Alosaimi W. Predicting Breast Cancer Based on Optimized Deep Learning Approach. *Comput Intell Neurosci*. 2022; 2022: 1820777.
107. Zhang YP, Zhang XY, Cheng YT, et al. Artificial intelligence-driven radiomics study in cancer: the role of feature engineering and modeling. *Mil Med Res*. 2023;10:22.
108. Hussain S, Lafarga-Osuna Y, Ali M, et al. Deep learning, radiomics and radiogenomics applications in the digital breast tomosynthesis: a systematic review. *BMC Bioinformatics*. 2023;24:401.
109. Shah SM, Khan RA, Arif S, Sajid U. Artificial intelligence for breast cancer analysis: Trends & directions. *Comput Biol Med*. 2022;142:105221.
110. Zhang Y, He G, Ma L, et al. A GPU-based computational framework that bridges neuron simulation and artificial intelligence. *Nat Commun*. 2023;14:5798.
111. Nassif AB, Talib MA, Nasir Q, et al. Breast cancer detection using artificial intelligence techniques: a systematic literature review. *Artif Intell Med*. 2022;127:102276.
112. Corti C, Cobanaj M, Dee EC, et al. Artificial intelligence in cancer research and precision medicine: applications, limitations and priorities to drive transformation in the delivery of equitable and unbiased care. *Cancer Treat Rev*. 2023;112:102498.
113. La Porta CA, Zapperi S. Artificial intelligence in breast cancer diagnostics. *Cell Rep Med*. 2022;3:100851.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.