MAYO CLINIC

MAYO CLINIC PROCEEDINGS: DIGITAL HEALTH

REVIEW

Artificial Intelligence in Histologic Diagnosis of Ductal Carcinoma In Situ

Michael Co, MBBS; Yik Ching Christy Lau, MBBS; Yi Xuan Yvonne Qian, MBBS; Man Chun Ryan Chan, MBBS; Desiree Ka-ka Wong, MBBS; Ka Ho Lui, MBBS; Nicholas Yu Han So, MBBS; Stephanie Wing Sum Tso, MBBS; Yu Chee Lo, MBBS; Woo Jung Lee, MBBS; and Elaine Wong, MBBS

Abstract

A systematic review was conducted in line with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses protocol to evaluate the diagnostic accuracy of artificial intelligence (AI) in ductal carcinoma in situ. Four databases were searched for articles up to December 2022: Embase, PubMed, Scopus, and Web of Science. 23 studies were included, and a search of grey literature was not performed. The following parameters were extracted: the accuracy, sensitivity, specificity, positive predictive value, and negative predictive value of each study. Statistical analysis of the included studies revealed that AI-assisted histopathological analysis is of high accuracy (83.78%), sensitivity (83.88%), and specificity (85.49%) and has a high positive predictive value (89.43%). Our results also reported that convolutional neural network (CNN) is the most commonly used mode of machine learning—21 models used only CNN, whereas 2 models used only support vector machines (SVM). On an average, CNN reported slightly higher accuracy and sensitivity (86.71% and 85.22%, respectively) than SVM (accuracy, 85.00%; sensitivity, 70.00%). When the 2 methods were combined, a mean accuracy of 82.52% and a mean sensitivity of 83.00% were achieved. The use of AI as a diagnostic adjunct can markedly improve the accuracy and efficiency of DCIS diagnosis and can, therefore, reduce pathologists' workload.

 ■ Mayo Clin Proc Digital Health 2023;1(3):267-275

As of late 2020, 7.8 million women worldwide are diagnosed with breast cancer, making it the world's most prevalent cancer among women and the second most common cancer overall.[1] It also remains the second leading cause of cancer-related deaths worldwide.[1] Alongside correlating clinical and radiological findings, the gold standard of diagnosis is histopathologic confirmation, which not only guides treatment options but is also a strong predicting tool.

The World Health Organization (WHO) classifies ductal carcinoma in situ (DCIS) as a precursor breast lesion[2] that is characterized by the proliferation of neoplastic epithelial cells confined to the mammary ductal system and shows no evidence of invasion into the surrounding stroma.[3] Conversely, atypical ductal hyperplasia (ADH) of the breast describes a proliferation of epithelial cells that is neither qualitatively nor quantitatively abnormal enough to be labeled as carcinoma in situ but is, nonetheless, considered at risk of progression to advanced neoplasia.[4] To differentiate it from DCIS, ADH is arbitrarily defined as a proliferation of well-defined, monomorphic epithelial cells that measures <2 mm or occupies fewer than 2 separate ducts.[5]

Therefore, pathologists are faced with the problem of DCIS and ADH sharing near identical histopathological features—size and space involvement are the only differentiating factors. Manual interpretation of specimens categorized under either of the 2 is both prone to marked interobserver variability[6,7]—risking underdiagnosis and overdiagnosis—and incredibly labor intensive.

Moreover, retrospective studies have shown that most women diagnosed with ADH using biopsy do not advance to DCIS or aggressive carcinoma at the moment of

From the Department of Surgery, Center for Education and Training, Queen Mary Hospital, University of Hong Kong, Pokfulam, Hong Kong.

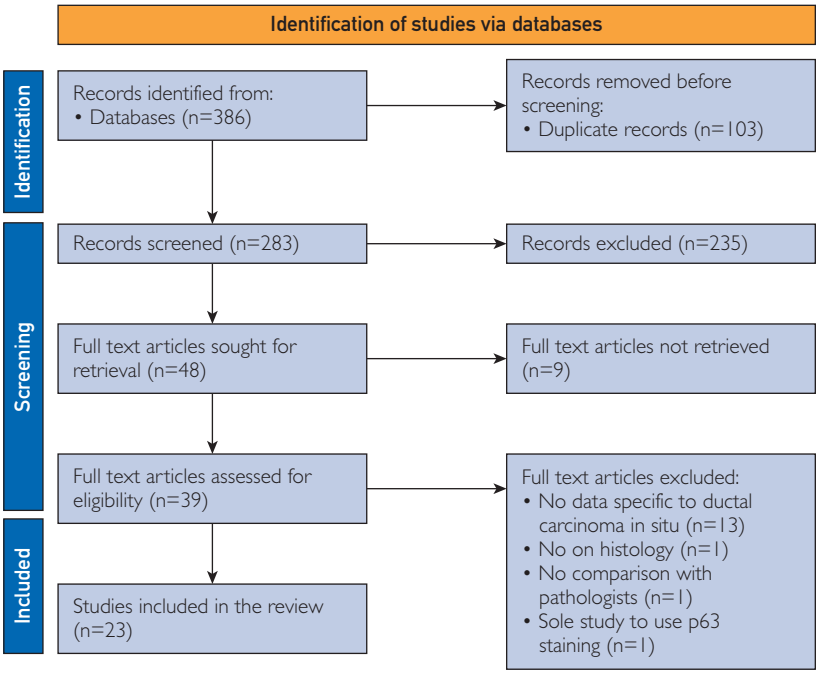

FIGURE. Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram.

surgical excision.[8] However, excision and adjuvant radiotherapy remain the standard treatment. Arguably, such an approach results in overtreatment and submitting patients to excessive morbidity.[9] In response, there has been a recent upsurge in proposals to use artificial intelligence (AI) for a more accurate histological diagnosis of breast lesions. Substantial research[10] has reported the superiority of machine learning in classifying breast histopathological images, with a large Lebanese study[11] reporting that of 128 ADH cases, surgical resection could have been avoided in 16%. With such encouraging evidence, AI has the potential to become an unparalleled adjunct for identifying patients with pure ADH, who can then be safely observed in lieu of receiving aggressive surgical intervention.

Two main modes of machine learning are currently used: convolutional neural networks (CNNs) and support vector machines (SVMs). CNNs originated in the 1980s as a type of deep learning that operates through multiple layers of a neural network to recognize patterns.[12] This detailed pattern recognition allows CNNs to efficiently classify complex characteristics of images independently of any knowledge provided by the user. Therefore, CNNs have become popular in medical imaging classification, particularly in the identification of benign and malignant pathologies.

On the contrary, SVMs discern a hyperplane to separate data points into 2 classification sets. Through this method, SVMs can identify planes in higher dimensions (depending on the input parameters), and nonlinear components (polynomial, exponential, and Gaussian) can be integrated to find the most appropriate hyperplane.[13]

Both CNNs and SVMs have distinct advantages and limitations. CNNs are superior in complexity because they can process unlimited layers, whereas SVMs are limited by the number of inputs.[14] SVMs also require manual feature extraction, whereas CNN extraction occurs automatically and, therefore, requires less user manipulation. However, CNNs require more computational power and time—an important limitation in the development of a fast and versatile diagnostic tool.[15] To circumvent this, combinations of CNN and

Mayo Clin Proc Digital Health ■ September 2023;1(3):267-275 ■ https://doi.org/10.1016/j.mcpdig.2023.05.008
www.mcpdigitalhealth.org

TABLE 1. Studies Included in Analysis

No.	Author, year	Title	Country
1	Elmannai et al,[40] 2021	Deep learning models combining for breast cancer histopathology image classification	Saudi Arabia
2	Mi et al,[29] 2021	Deep learning-based multi-class classification of breast digital pathology images	China
3	Narayanan et al,[41] 2021	Unmasking the immune microecology of ductal carcinoma in situ with deep learning	UK
4	Polónia et al,[31] 2021	Artificial intelligence improves the accuracy in histologic classification of breast lesions	Portugal
5	Sato et al,[42] 2021	Machine learning-based image analysis for accelerating the diagnosis of complicated preneoplastic and neoplastic ductal lesions in breast biopsy tissues	Japan
6	Carvalho et al,[23] 2020	Breast cancer diagnosis from histopathological images using textural features and CBIR	Brazil
7	Sheikh et al,[34] 2020	Histopathological classification of breast cancer images using a multi-scale input and multi-feature network	South Korea
8	Wang et al,[27] 2020	Breast cancer image classification via multi-network features and dual-network orthogonal low-rank learning	China
9	Wu et al,[36] 2020	MLCD: a unified software package for cancer diagnosis	USA
10	Yu et al,[44] 2020	A transfer learning-based novel fusion convolutional neural network for breast cancer histology classification	China
11	Mercan et al,[28] 2019	Assessment of machine learning of breast pathology structures for automated differentiation of breast cancer and high-risk proliferative lesions	USA
12	Mittal et al,[30] 2019	Digital assessment of stained breast tissue images for comprehensive tumor and microenvironment analysis	USA
13	Roy et al,[33] 2019	Path-cased system for classification of breast histology images using deep learning	India
14	Yang et al,[38] 2019	EMS-Net: ensemble of multiscale convolutional neural networks for classification of breast cancer histology images	China
15	Yao et al,[39] 2019	Parallel structure deep neural network using CNN and RNN with an attention mechanism for breast cancer histology image classification	China
16	Fondón et al,[25] 2019	Automatic classification of tissue malignancy for breast carcinoma diagnosis	Portugal
17	Gecer et al,[26] 2018	Detection and classification of cancer in whole slide breast histopathology images using deep convolutional networks	USA
18	Vo et al,[43] 2018	Classification of breast cancer histology images using incremental boosting convolution networks	South Korea

Continued on next page

TABLE 1. Continued

No.	Author, year	Title	Country
19	Araújo et al,[12] 2017	Classification of breast cancer histology images using convolutional neural networks	USA
20	Bejnordi et al,[22] 2017	Context-aware stacked convolutional neural networks for classification of breast carcinomas in whole-slide histopathology images	Netherlands
21	Radiya-Dixit et al,[32] 2017	Automated classification of benign and malignant proliferated breast lesions	UK
22	Yamada et al,[37] 2016	Quantitative nucleic features are effective for discrimination of intraductal proliferated lesions of the breast	Japan
23	Hwang et al,[27] 2005	Multi-resolution wavelet-transformed image analysis of histological sections of breast carcinomas	South Korea

CBIR, content-based image retrieval; CNN, convolutional neural network; MLCD, Machine Learning Package for Cancer Diagnosis; RNN, recurrent neural network; UK, United Kingdom; USA, United States of America.

SVM techniques—along with other machine learning modes—have been developed to optimize complexity and computing time.

Although the advent of AI-assisted histologic diagnosis is relatively recent, a preliminary model for analyzing breast images was described in as early as 1988 by the Information Technology Institute at Brighton Polytechnic.[16] By 1991, Dawson et al[17] developed a Bayesian neural network model to recognize low-grade breast lesions with an accuracy of ~70%. Although these early studies were limited to the realm of academic research, deep learning models are now capable of analyzing a wide array of characteristics important in the clinical setting: identification of the primary tumor,[18] histopathologic grading,[19] and even prognosis prediction based on morphological features.[20] Therefore, this systematic review aimed to determine the accuracy and advantages of AI-assisted diagnosis in comparison with traditional pathological evaluation in the diagnosis of breast DCIS.

MATERIALS AND METHODS

Using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses protocol,[21] 4 electronic databases were searched up to December 2022: PubMed, Embase, Scopus, and Web of Science (Figure). The search terms used were as follows: (artificial intelligence OR machine learning OR deep learning OR cognitive neural network OR convolutional neural network OR artificial neural network) AND (ductal carcinoma in situ OR noninfiltrating intraductal carcinoma OR breast carcinoma in situ) AND (pathology OR histology). This systematic review was not registered, but the protocol can be found at the Depository of the Center of Education and Training, University of Hong Kong.

Our inclusion criteria were as follows: (1) published in English, (2) retrospective and prospective studies and randomized controlled trials in peer-reviewed journals, and (3) studies on the histopathologic diagnosis of DCIS. Conversely, the following were the exclusion criteria: (1) narrative studies, conference papers, editorials, and book chapters; (2) studies not related to AI or deep learning; (3) studies on imaging or laboratory markers; (4) studies on only invasive breast cancer; (5) studies on the prognosis of DCIS; and (6) studies on neoadjuvant therapy.

The titles and abstracts of each study were independently screened by 2 reviewers against the inclusion and exclusion criteria. Discrepancies between the reviewers' decisions were resolved by discussion. The full texts of the eligible studies were then sourced and their details extracted and organized on Google Sheets by the reviewers responsible for each study. Nine full texts were irretrievable. Studies with insufficient data specific to DCIS or

TABLE 2. Summary of Results

Study	Model(s)	Test images	Comparison (no. of pathologists)	Accuracy (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
Elmannai et al,[40] 2021	DCNN	80	2	95.00	99.16	99.17		
Narayanan et al,[41] 2021	CNN	178	1		77.00	98.00	76.80	98.50
Mi et al,[29] 2021	DCNN and XGBoost	33	5	90.43	84.45		90.32	
Polónia et al,[31] 2021	CNN	100	2	87.00				
Sato et al,[42] 2021	CNN	222	6		90.80	95.30	93.40	
Carvalho et al,[23] 2020	SVM	60	2	95.00				
Sheikh et al,[34] 2020	CNN	8309	1	83.00		90.00	90.00	
Wang et al,[27] 2020	DCNN	40	2	97.70	97.72	97.92		
Wu et al,[36] (2020)	CNN and SVM	240	1	63.00				
Yu et al,[44] 2020	FCNN FCNN	36			53.51 77.79			
Mercan et al,[28] 2019	CNN	240	3	70.00	85.00	45.00		
Mittal et al,[30] 2021	CNN (ResNet50) CNN (VGG19) CNN (InceptionResNetV2)	100	1	98.74 96.14 94.93				
Roy et al,[33] 2019	CNN CNN	40		87.00	85.00 100.00		80.68 90.90	
Yang et al,[38] 2019	EMS-Net	100		91.75				
Yao et al,[39] 2019	CNN and RNN	1455	2	92.00				
Fondón et al,[25] 2018	Quadratic SVM	156	1	75.00	70.00			
Gecer et al,[26] 2018	DCNN	76	3	88.00				
Vo et al,[43] 2018	DCNN	36			88.90			
Araújo et al,[12]	CNN CNN and SVM	36	2	66.70 65.00	77.75 81.15			
Bejnordi et al,[22] 2017	CNN	221	1	81.30				
Radiya-Dixit et al,[32] 2017	CNN	100		85.00				
Yamada et al,[37] 2016	SVM and LDA	200		55.00				
Hwang et al,[27] 2005	BPNN and SAS	90		96.67				

BPNN, backpropagation neural network; CNN, convolutional neural network; DCNN, deep convolutional neural network; FCNN, fusion convolutional neural network; LDA,; SAS,; SVM, support vector machine; VGG19,.

histopathology were removed (13 reported insufficient data on the former and 1 on the latter). One study did not compare AI with human pathologists and was also removed. All but 1 study used hematoxylin and eosin (H&E) staining; hence, that study was excluded to minimize discrepancies. For the final 23 included studies, the details extracted were as follows: (1) study characteristics (country, study type, AI type, sample type and size, and statistical method) and (2) accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV).

Then, the mean accuracy, sensitivity, specificity, PPV, and NPV were calculated using Google Sheets to attain our results.

RESULTS

A total of 29 AI models were evaluated in the 23 included studies (Tables 1 and 2). Of these, 21 AI models were reported by 18 studies to have a high mean accuracy of 83.78% (range, 55.00%-98.74%; 95% confidence interval (CI), 81.00%-86.56%) in the identification of DCIS in specimens stained with H&E.[12,22-39] Fifteen models were reported by 12 studies

to have a high mean sensitivity of 83.88% (range, 53.51%-100.00%; 95% CI, 80.82%-86.94%).[12,25,28,29,33,34,40-45] Five models from 5 studies reported a mean specificity of 85.49% (range, 45.00%-99.17%; 95% CI, 75.45-95.54).[28,34,40-42] Seven models from 6 studies reported a mean PPV of 89.43% (range, 76.80%-97.92%; 95% CI, 86.52%-92.34%).[23,29,33,35,41,42] Finally, 1 model reported an NPV of 98.50%.[41]

The most common AI type was CNN and its variations (25 of the 29 models), used either on its own (21 models) or in combination with other AI types (4 models). The CNN variations included deep CNN (DCNN), fusion CNN (FCNN), and EMS-Net. The second most common AI type was SVM (5 models), which was also used alone (2 models) and in combination with CNN (3 models). The remaining models (which were excluded from our pooled analysis) included recurrent neural networks, backpropagation neural networks , and XGBoost.

When used alone, CNN and its modifications reported a mean accuracy of 86.71% (13 data points; range, 66.70%-98.74%; 95% CI, 84.02%-89.40%), a mean sensitivity of 85.22% (12 data points; range, 53.51%-100.00%; 95% CI, 81.58%-88.86%), a mean specificity of 85.49% (5 data points; range, 45.00%-99.17%; 95% CI, 75.45%-95.54%), a mean PPV of 87.94% (5 data points; range, 76.80%-97.92%; 95% CI, 84.05%-91.83%), and an NPV of 98.50% (1 data point).[12,22,26,28,30-35,38-43]

In comparison, SVM alone reported a slightly lower mean accuracy of 85.00% (2 data points; range, 75.00%-95.00%), a lower sensitivity of 70.00% (1 data point), and a higher PPV of 96.00% (1 data point).

Six models combined 2 different AI types and reported the lowest mean accuracy of 82.52% (6 data points; range, 63.00%-96.67%; 95% CI, 76.66%-88.37%), a mean sensitivity of 83.00% (2 data points; range, 81.15%-84.85%), and a PPV of 90.32% (1 data point).[12,28,29,36-38]

DISCUSSION

Methodology

Although reviews on the use of AI in breast histopathology have been published previously, no systematic review particularly on DCIS and ADH has been conducted. Therefore, this review aimed to provide fresh insight into this clinically pertinent area. We used stringent inclusion and exclusion criteria and adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses protocol. Four databases were searched, which provided a broad range of studies that allowed for a comprehensive analysis of AI-assisted DCIS diagnosis. Of the included studies, most were published in the past 5 years, making our review highly relevant to the most recent advancements in AI biotechnology.

However, a limitation is the inclusion of only studies written in or translated to English; therefore, relevant studies from institutions publishing in other languages may have been excluded. In addition, given the rapid development of AI technologies, by the time of publication, there may have been important new studies that were not able to be included.

Evaluation of the Included Studies

Of the 23 included studies, 16 had pathologists' diagnoses as the gold standard against which the accuracy, sensitivity, specificity, PPV, and NPV of the AI models were determined. Among these 16 studies, the numbers of pathologists ranged considerably from 1 to 6; 7 studies had only 1 pathologist. Given the significant interobserver variability in the histopathological diagnosis of DCIS (a study[24] found an agreement rate of only 75.3% between 115 pathologists), this may limit the validity of those studies. One included study[28] calculated the accuracy of an additional group of pathologists and compared their diagnoses with the reference data, allowing for a more direct comparison between the performance of pathologists and that of the AI models.

There was also considerable variability in the validation and testing processes of each AI model—each study had an arbitrary number of samples against which the model was verified, and a wide range of statistical models were used. Moreover, most of our included studies used the same data sets to test and train their AI systems, the images of which may not be sufficiently representative of the full scope that pathologists may encounter in the clinical setting.

Furthermore, CNN and SVM are merely umbrella categories; each study developed its own variation to diagnose DCIS—they may differ markedly in their image processing methods, their mechanisms for identifying and extracting the region of interest, and the specific algorithm used to classify the images. Therefore, our pooled results should only be referred to as an estimation of the general performance of the evaluated AI models. Beyond CNN and SVM, there were insufficient studies to conduct a pooled analysis of numerous other AI techniques.

All 23 studies also diagnosed histology slides in isolation from their clinical and radiological contexts. Therefore, their applicability in the clinical setting—in which triple assessment remains the standard diagnostic process—is yet to be confirmed. Among the screened studies, tissue staining and storage methods also differed—all but 1 used H&E staining. Hence, we excluded the single study that used p63 staining to minimize discrepancies—whether the type of stain affects AI-assisted diagnosis may be explored in future studies. Regarding tissue storage, most of the screened studies used samples stored in paraffin; similarly, we excluded studies that used other storage methods. Finally, because all included studies reported favorable results and because many developed the AI models evaluated in the studies, publication bias may be a concern.

Factors Contributing to the Higher Accuracy in the Studies Using CNNs

In analyzing the studies that produced the highest diagnostic accuracies, we identified 2 main contributing factors. First, attention mechanisms, which combine image features extracted from different types of AI models, improve AI performance when added to the AI architecture.[39] Although our review found combined models to have a lower mean accuracy than CNN alone, Yao et al[39] reported that attention mechanisms with a more extensive array of AI methods (the study included CNNs, recurrent neural networks, switchable normalization, and targeted dropouts) can achieve an excellent overall accuracy of 98.30%.

The second factor was the use of image patches of different scales, which mimics pathologists' inspection of histology slides under different levels of magnification, allowing for a better analysis of structural and textural information.

Accuracy of SVMs

Of importance, although CNNs reported a slightly higher mean accuracy, Carvalho et al[23] developed an SVM with an exceptional accuracy of 95%, which may be attributed to its addition of a phylogenetic diversity index for extracting histological features. The excluded study that used p63 staining mentioned earlier[45] also operated an SVM—it highlighted the importance of detecting subtle morphological changes in the myoepithelial cells for the accurate diagnosis of DCIS, suggesting that, by selecting such pertinent features and with the aid of extraction algorithms, SVMs can achieve accuracies comparable with those of CNNs but using far less computational time and power.

Implications

Our review has shown that the use of AI in diagnosing DCIS has high accuracy and can, therefore, assist pathologists in considerably quickening the process of pathological evaluation. However, as in most research on new and evolving technology, publication bias is a concern. In addition, AI is a rapidly evolving field, and the diagnostic accuracy of AI changes rapidly with time. Nevertheless, our systematic review based on the best available evidence from the literature has suggested promising diagnostic accuracy with AI for histopathologic diagnosis of DCIS. In fact, the 2021 study by Polónia et al[31] found that when pathologists were aided by AI, the diagnostic accuracy for DCIS was higher than when pathologists or AI alone examined breast lesions. Therefore, AI as an adjunct can be a powerful tool for overcoming interobserver variability.

However, published studies have only examined the pathological applications of AI in isolation from the clinical and radiologic aspects of triple assessment. They have also not considered the cost-effectiveness of implementing AI. For AI to be clinically useful, it must prove that its costs would not be inordinately higher than current practices.

Although a presiding reason for the use of AI is its greater diagnostic accuracy, problems may still arise from diagnostic errors, which can lead to undertreatment or overtreatment. Therefore, it would be important for pathologists to act as gatekeepers and for multidisciplinary teams managing patients with breast cancer to carefully synthesize the available histopathologic, clinical, and radiological data.

POTENTIAL COMPETING INTERESTS

The authors report no competing interests.

Abbreviations and Acronyms: **AI,** artificial intelligence; **PPV,** positive predictive value; **NPV,** negative predictive value; **CNN,** convolutional neural network; **SVM,** support vector machine; **DCIS,** diagnosis of ductal carcinoma in situ; **ADH,** atypical ductal hyperplasia; **H&E,** hematoxylin and eosin

Correspondence: Address to Michael Co, MBBS, K1401, Queen Mary Hospital, Pokfulam, Hong Kong (mcth@hku.hk).

ORCID

Michael Co: https://orcid.org/0000-0002-1705-051X

REFERENCES

1. Breast cancer. World Health Organization. https://www.who.int/news-room/fact-sheets/detail/breast-cancer. Accessed December 12, 2021.
2. Sinn HP, Kreipe H. A brief overview of the WHO classification of breast tumours, 4th edition, focusing on issues and updates from the 3rd edition. *Breast Care*. 2013;8(2):149-154.
3. Collins LC, Laronga C, Wong JS. Breast ductal carcinoma in situ: epidemiology, clinical manifestations, and diagnosis. UpToDate. https://www.uptodate.com/contents/breast-ductal-carcinoma-in-situ-epidemiology-clinical-manifestations-and-diagnosis?search=Breast%20ductal%20carcinoma%20in%20situ:%20Epidemiology,%20clinical%20manifestations,%20and%20diagnosis&source=search_result&selectedTitle=1~56&usage_type=default&display_rank=1. Accessed December 12, 2021.
4. Myers DJ, Walls AL. Atypical breast hyperplasia. StatPearls. https://www.ncbi.nlm.nih.gov/books/NBK562244/. Accessed December 12, 2021.
5. Sabel MS, Collins LC. Atypia and lobular carcinoma in situ: high-risk lesions of the breast. UpToDate [about 14 screens]. https://www.uptodate.com/contents/atypia-and-lobular-carcinoma-in-situ-high-risk-lesions-of-the-breast?search=atypia%20and%20lobular%20carcinoma%20in%20situ&source=search_result&selectedTitle=1~150&usage_type=default&display_rank=1. Accessed December 12, 2021.
6. Ibrahim A, Gamble P, Jareoensri R, et al. Artificial intelligence in digital breast pathology: techniques and applications. *Breast*. 2020;49:267-273.
7. Moxley-Wyles B, Colling R, Verrill C. Artificial intelligence in pathology: an overview. *Diagn Histopathol*. 2020;26(11):513-520.
8. Ha R, Mutasa S, Sant E, et al. Accuracy of distinguishing atypical ductal hyperplasia from ductal carcinoma in situ with convolutional neural network-based machine learning approach using mammographic image data. *AJR Am J Roentgenol*. 2019;212(5):1166-1171.
9. van Seijen M, Lips EH, Thompson AM, et al. Ductal carcinoma in situ: to treat or not to treat, that is the question. *Br J Cancer*. 2019;121(4):285-292.
10. Schiaffino S, Calabrese M, Melani EF, et al. Upgrade rate of percutaneously diagnosed pure atypical ductal hyperplasia: systematic review and meta-analysis of 6458 lesions. *Radiology*. 2020;294(1):76-86.
11. Harrington L, diFlorio-Alexander R, Trinh K, MacKenzie T, Suriawinata A, Hassanpour S. Prediction of atypical ductal hyperplasia upgrades through a machine learning approach to reduce unnecessary surgical excisions. *JCO Clin Cancer Inform*. 2018;2:1-11.
12. Araújo T, Aresta G, Castro E, et al. Classification of breast cancer histology images using convolutional neural networks. *PLoS One*. 2017;12(6):e0177544.
13. Gandhi R. Support vector machine—introduction to machine learning algorithms. Medium. https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47. Accessed December 12, 2021.
14. Zhu C, Song F, Wang Y, Dong H, Guo Y, Liu J. Breast cancer histopathology image classification through assembling multiple compact CNNs. *BMC Med Inform Decis Mak*. 2019;19(1):198.
15. De Luca G. SVM vs neural network. Baeldung. https://www.baeldung.com/cs/svm-vs-neural-network. Accessed December 12, 2021.
16. Heathfield HA, Winstanley G, Kirkham N. Computer-assisted breast cancer grading. *J Biomed Eng*. 1988;10(5):379-386.
17. Dawson AE, Austin RE, Weinberg DS. Nuclear grading of breast carcinoma by image analysis. Classification by multivariate and neural network analysis. *Am J Clin Pathol*. 1991;95(4 suppl 1):S29-S37.
18. Osareh A, Shadgar B. Machine learning techniques to diagnose breast cancer. In: *Proceedings of the 2010 5th International Symposium on Health Informatics and Bioinformatics*. IEEE; 2010:114-120.
19. 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(1):30.
20. 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(1):610.
21. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;10(1):89.
22. Bejnordi BE, Zuidhof G, Balkenhol M, et al. Context-aware stacked convolutional neural networks for classification of breast carcinomas in whole-slide histopathology images. *J Med Imaging*. 2017;4(4):044504.
23. Carvalho ED, Filho AO, Silva RR, et al. Breast cancer diagnosis from histopathological images using textural features and CBIR. *Artif Intell Med*. 2020;105:101845.
24. Elmore JG, Longton GM, Carney PA, et al. Diagnostic concordance among pathologists interpreting breast biopsy specimens. *JAMA*. 2015;313(11):1122-1132.
25. Fondón I, Sarmiento A, García AI, et al. Automatic classification of tissue malignancy for breast carcinoma diagnosis. *Comput Biol Med*. 2018;96:41-51.
26. Gecer B, Aksoy S, Mercan E, Shapiro LG, Weaver DL, Elmore JG. Detection and classification of cancer in whole slide breast histopathology images using deep convolutional networks. *Pattern Recognit*. 2018;84:345-356.
27. Hwang HG, Choi HJ, Lee BI, Yoon HK, Nam SH, Choi HK. Multi-resolution wavelet-transformed image analysis of histological sections of breast carcinomas. *Cell Oncol*. 2005;27(4):237-244.

28. Mercan E, Mehta S, Bartlett J, Shapiro LG, Weaver DL, Elmore JG. Assessment of machine learning of breast pathology structures for automated differentiation of breast cancer and high-risk proliferative lesions. *JAMA Netw Open*. 2019;2(8):e198777.
29. Mi W, Li J, Guo Y, et al. Deep learning-based multi-class classification of breast digital pathology images. *Cancer Manag Res*. 2021;13:4605-4617.
30. Mittal S, Stoean C, Kajdacsy-Balla A, Bhargava R. Digital assessment of stained breast tissue images for comprehensive tumor and microenvironment analysis. *Front Bioeng Biotechnol*. 2019;7:246.
31. Polónia A, Campelos S, Ribeiro A, et al. Artificial intelligence improves the accuracy in histological classification of breast lesions. *Am J Clin Pathol*. 2021;155(4):527-536.
32. Radiya-Dixit E, Zhu D, Beck AH. Automated classification of benign and malignant proliferative breast lesions. *Sci Rep*. 2017;7(1):9900.
33. Roy K, Banik D, Bhattacharjee D, Nasipuri M. Patch-based system for classification of breast histology images using deep learning. *Comput Med Imaging Graph*. 2019;71:90-103.
34. Sheikh TS, Lee Y, Cho M. Histopathological classification of breast cancer images using a multi-scale input and multi-feature network. *Cancers*. 2020;12(8):2031.
35. Wang Y, Lei B, Elazab A, et al. Breast cancer image classification via multi-network features and dual-network orthogonal low-rank learning. *IEEE Access*. 2020;8:27779-27792.
36. Wu W, Li B, Mercan E, et al. MLCD: a unified software package for cancer diagnosis. *JCO Clin Cancer Inform*. 2020;4:290-298.
37. Yamada M, Saito A, Yamamoto Y, et al. Quantitative nucleic features are effective for discrimination of intraductal proliferative lesions of the breast. *J Pathol Inform*. 2016;7(1):1.
38. Yang Z, Ran L, Zhang S, Xia Y, Zhang Y. EMS-Net: ensemble of multiscale convolutional neural networks for classification of breast cancer histology images. *Neurocomputing*. 2019;366:46-53.
39. Yao H, Zhang X, Zhou X, Liu S. Parallel structure deep neural network using CNN and RNN with an attention mechanism for breast cancer histology image classification. *Cancers*. 2019;11(12):1901.
40. Elmannai H, Hamdi M, AlGarni A. Deep learning models combining for breast cancer histopathology image classification. *Int J Comput Intell Syst*. 2021;14(1):1003-1013.
41. Narayanan PL, Raza SE, Hall AH, et al. Unmasking the immune microecology of ductal carcinoma in situ with deep learning. *NPJ Breast Cancer*. 2021;7(1):19.
42. Sato S, Maki S, Yamanaka T, et al. Machine learning-based image analysis for accelerating the diagnosis of complicated preneoplastic and neoplastic ductal lesions in breast biopsy tissues. *Breast Cancer Res Treat*. 2021;188(3):649-659.
43. Vo DM, Nguyen N, Lee SW. Classification of breast cancer histology images using incremental boosting convolution networks. *Inf Sci*. 2019;482:123-138.
44. Yu X, Chen H, Liang M, Xu Q, He L. A transfer learning-based novel fusion convolutional neural network for breast cancer histology classification. *Multimed Tools Appl*. 2020;81(5):1-15.
45. Yamamoto Y, Saito A, Tateishi A, et al. Quantitative diagnosis of breast tumors by morphometric classification of microenvironmental myoepithelial cells using a machine learning approach. *Sci Rep*. 2017;7(1):1-12.