



# An investigation of the classification accuracy of a deep learning framework-based computer-aided diagnosis system in different pathological types of breast lesions

Mengsu Xiao<sup>#</sup>, Chenyang Zhao<sup>#</sup>, Qingli Zhu, Jing Zhang, He Liu, Jianchu Li, Yuxin Jiang

Department of Ultrasound, Chinese Academy of Medical Sciences and Peking Union Medical College Hospital, Beijing 100730, China

*Contributions:* (I) Conception and design: Y Jiang; (II) Administrative support: J Li; (III) Provision of study materials or patients: Q Zhu, H Liu; (IV) Collection and assembly of data: M Xiao, C Zhao; (V) Data analysis and interpretation: J Zhang, J Li; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

<sup>#</sup>These authors contributed equally to this work.

*Correspondence to:* Yuxin Jiang, Department of Ultrasound, Chinese Academy of Medical Sciences and Peking Union Medical College Hospital, Shuaifuyuan 1st, Dongcheng District, Beijing 100730, China. Email: jiangyuxinxh@163.com.

**Background:** Deep learning-based computer-aided diagnosis (CAD) is an important method in aiding diagnosis for radiologists. We investigated the accuracy of a deep learning-based CAD in classifying breast lesions with different histological types.

**Methods:** A total of 448 breast lesions were detected on ultrasound (US) and classified by an experienced radiologist, a resident and deep learning-based CAD respectively. The pathological results of the lesions were chosen as the golden standard. The diagnostic performances of the three raters in different pathological types were analyzed.

**Results:** For the overall diagnostic performance, deep learning-based CAD presented a significantly higher specificity (76.96%) compared with the two radiologists. The area under ROC of CAD was almost equal with the experienced radiologist (0.81 *vs.* 0.81), while significantly higher than the resident (0.81 *vs.* 0.70,  $P < 0.0001$ ). In the benign lesions, deep learning-based CAD had a higher accuracy than both the two radiologists, which correctly classified as benign lesions in 119/135 of fibroadenomas (88.1%), 25/35 of adenosis (71.4%), 14/27 of intraductal papillary tumors (51.9%), 5/10 of inflammation (50%), and 4/8 of sclerosing adenosis (50%). But only the differences between CAD and the two radiologists in fibroadenomas had statistical significance ( $P = 0.0011$  and  $P = 0.0313$ ), and the differences between CAD and the resident in adenosis had statistical significance ( $P = 0.012$ ). In the malignant lesions, 151/168 of invasive ductal carcinomas (89.9%), 21/29 of ductal carcinoma in situ (DCIS) (72.4%) and 6/7 of invasive lobular carcinomas (85.7%) were diagnosed as malignancies by deep learning-based CAD, with no significant differences between CAD and the two radiologists.

**Conclusions:** In the diagnosis of these common types of breast lesions, deep learning-based CAD had a satisfying performance. Deep learning-based CAD had a better performance in the breast benign lesions, especially in fibroadenomas and adenosis. Therefore, deep learning-based CAD is a promising supplemental tool to US to increase the specificity and avoid unnecessary benign biopsies.

**Keywords:** Deep learning; diagnosis, computer-aided; breast; ultrasound (US); pathology

Submitted Sep 10, 2019. Accepted for publication Nov 27, 2019.

doi: 10.21037/jtd.2019.12.10

View this article at: <http://dx.doi.org/10.21037/jtd.2019.12.10>

## Introduction

Over the past decade, there has been a dramatic increase in breast cancer in different regions (1,2). Breast cancer has the highest incidence and mortality among all cancers in women worldwide (3,4). Consequently, various imaging methods have been designed and promoted into clinical use for early detection of breast cancer, among which breast ultrasound (US) is regarded as an essential diagnostic tool for the detection and evaluation of breast lesions (5,6). Nevertheless, despite the wide utilization of Breast Imaging Report And Data System (BI-RADS) lexicon as a standard protocol, the development of US is still hindered by its major disadvantage—low specificity and positive predictive value (PPV), which often causes a certain number of false-positive lesions and unnecessary biopsies (7-9).

The computer-aided diagnosis (CAD) system based on artificial intelligence (AI) technology is regarded as an important method in aiding diagnosis for radiologists (10,11). With the advent of deep learning approaches, the capability of interpretation of radiographic images by systems through artificial neural networks (ANNs) has been greatly enhanced after intensive training on large databases (12). Nowadays, deep learning techniques are considered to be the most advanced technology for image classification. CAD with deep learning techniques (deep learning-based CAD) exceeds conventional CAD (13). Recently, a deep learning-based CAD for the breast US (S-Detect™ for Breast in RS80A; Samsung Medison Co., Ltd., Seoul, Korea) has become commercially available (14). A good diagnostic performance of this deep learning-based CAD was shown by the previous studies (15-17).

The sonographic features of breast tumors greatly are dependent on the pathological types. The significant overlap of US features between benign and malignant tumors, especially in the rare pathological entities. In consideration that deep learning-based CAD analyzes inherent patterns from raw information data of the lesions, we hypothesized that it might possess the potential in acquiring better diagnostic efficiency in specific histologic types of breast lesions than human naked eyes. However, as far as we know, there is no published studies about deep learning-based CAD focused on the diagnostic accuracy of deep learning-based CAD in specific histologic types. Therefore, in this article, we investigated the accuracy of deep learning-based CAD in classifying a variety of lesions with different histological types.

## Methods

This research was approved by the Institutional Review Board of Peking Union Medical College Hospital. The Institutional Ethical Committee approved this prospective study (HS 1400). Informed consent was obtained from each patient.

### Patients

This study recruited 455 female patients with focal breast lesions from January 2018 to May 2018, who referred to the department of breast surgery of the hospital for further treatment of breast lesions. Eighteen patients were excluded who had no evident lesions on US, or underwent neo-adjuvant chemotherapy previously. Subsequently, 437 patients with 448 focal breast lesions were included in the study, with a mean age of 46.9 years old, and a median of 46 years old. All of the lesions were biopsied, and the pathologies were obtained. There were no masses of categories 0, 1, and 2 in our study. The category 3 masses in our study were biopsied according to the patients' choices or others reasons, including family history, nipple discharge, suspicious change at follow-up, an upgrade after additional mammography, an increase in volume of more than 20%, and patients' anxieties.

### US examinations and CAD assessments

In this study, a commercial high-end US machine (RS80A with Prestige, Samsung Medison, Co. Ltd., Seoul, Korea) equipped with a deep learning-based CAD system named S-detect was employed for breast US examination, using a 3–12 MHz linear transducer. All of the recruited patients received standard bilateral breast US examinations by an experienced radiologist (Zhu QL with 19 years of experience in breast imaging), who knew the clinical information of the patients. After US operation, the experienced radiologist evaluated and subcategorized the detected lesions based on the US descriptors regulated by the 5th edition of BI-RADS lexicon.

The lesions then were assessed by deep learning-based CAD (Samsung Healthcare, South Korea), on both the transverse and longitudinal sections reserved in the former US operation. Under the CAD mode, the selected lesion was segmented in a lined contour, which could also be adjusted by the radiologist manually. After few seconds of

**Table 1** Malignant and benign histological types of 448 focal breast lesions

Malignant lesions	Number	Benign lesions	Number
Invasive ductal carcinoma, not otherwise specified	168	Breast adenosis	35
Invasive lobular carcinoma	7	Fibroadenoma	135
Ductal carcinoma <i>in situ</i>	29	Sclerosing adenosis	8
Apocrine carcinoma	1	Intraductal papillary tumor	27
Encapsulated papillary carcinoma	3	Inflammation	10
Mucinous carcinoma	2	Benign phyllodes tumor	2
Metaplastic carcinoma	1	Adiponecrosis	3
Adenocarcinoma with spindle cell metaplasia	2	non-specific breast tissues	8
Diffuse large B-cell lymphoma	1	Dermoid cyst	1
Malignant phyllodes tumor	3	Simple cyst	1
Tubular carcinoma	1		
Total	218		230

processing, a dichotomic result (possibly benign or possibly malignant) was given by the software, also with a series of US descriptors, including shape, orientation, margins, pattern and posterior acoustic features as a reference.

After the above process, the images recorded by the radiologist for CAD evaluation were reviewed by a resident with 2-year training experience of US, who was blinded to the results of the experienced radiologist and deep learning-based CAD, as well as the medical histories of the patients.

### Statistical analysis

The pathological results of all of the lesion biopsies were obtained afterwards. For further statistical analysis, the BI-RADS subcategorizations of the doctors were transferred to a dichotomic form with a cutoff value of 4a for benign and malignancy. The BI-RADS 3 were deemed as benign lesions, and BI-RADS 4a, 4b, 4c and 5 were deemed as malignant lesions (18,19). The statistical procedure was conducted on the SPSS software (SPSS 19.0, IBM). The overall diagnostic performance of deep learning-based CAD was evaluated using sensitivity, specificity, PPV, negative predictive value (NPV), and area under the receiver operator characteristics curve (AUC). For each histological subtypes of breast lesions, the diagnostic accuracy of deep learning-based CAD was calculated and compared with the results of the two radiologists by chi-square test. In this study, statistical significance was considered when the P value was less than 0.05.

### Results

A total of 448 focal breast lesions were examined and classified by the experienced radiologist, deep learning-based CAD, and the in-training resident respectively. Among these lesions, 218 of the lesions were histopathologically proved as malignant masses, including invasive ductal carcinoma (not otherwise specified), invasive lobular carcinoma, ductal carcinoma in situ (DCIS), mucinous carcinoma, malignant phyllodes tumors, encapsulated papillary carcinoma, apocrine carcinoma, metaplastic carcinoma, adenocarcinoma with spindle cell metaplasia, and diffuse large B-cell lymphoma. Two hundred and thirty lesions were identified as benign ones, including adenosis, sclerosing adenosis, fibroadenoma, intraductal papillary tumor, inflammation, benign phyllodes tumor, adiponecrosis, and non-specific breast tissues (*Table 1*).

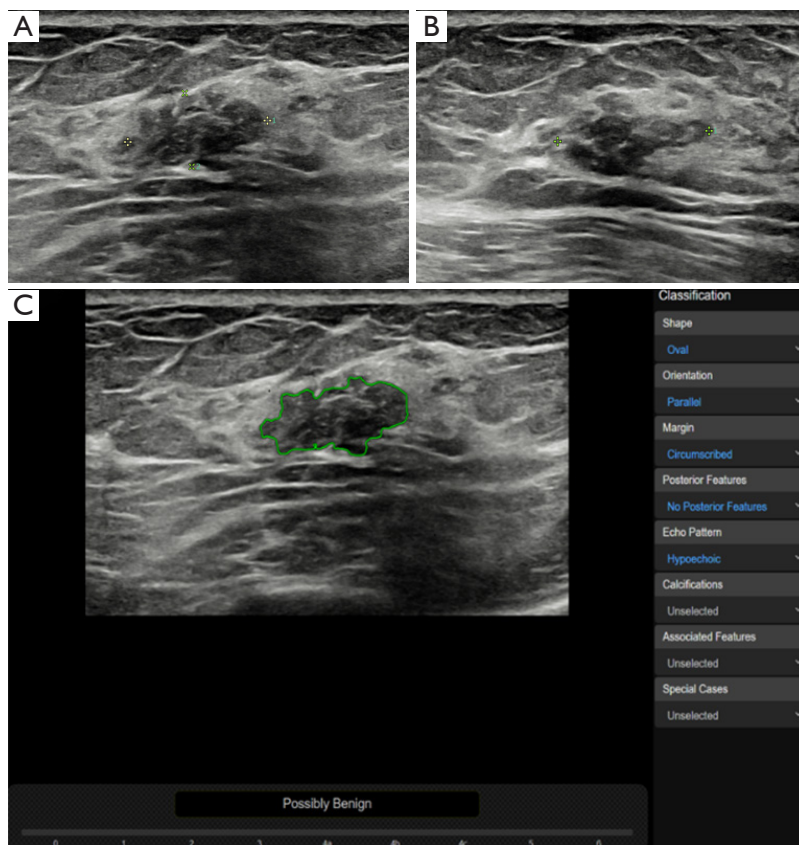
To illustrate the diagnostic performance of deep learning-based CAD, the experienced radiologist, and the resident, sensitivity, specificity, PPV, NPV, AUC were calculated and presented in *Table 2*. Deep learning-based CAD presented a significantly lower sensitivity (85.32%) and NPV (84.69%) but a significantly higher specificity (76.96%) and PPV (77.82%), compared with both the experienced radiologist and the resident ( $P < 0.05$ ). The AUC value of deep learning-based CAD was almost equal with the experienced radiologist (0.81 *vs.* 0.81), while significantly higher than the resident (0.81 *vs.* 0.70,  $P < 0.0001$ ).

In the benign lesions, the deep learning-based CAD

**Table 2** Diagnostic performance of deep learning-based CAD, the experienced radiologist, and the resident

Diagnostic methods	SE (%) (95% CI)	SP (%) (95% CI)	PPV (%) (95% CI)	NPV (%) (95% CI)	AUC (95% CI)
Deep learning-based CAD	85.32 (79.91–89.74)	76.96 (70.97–82.24)	77.82 (72.02–82.93)	84.69 (79.08–89.29)	0.81 (0.77–0.85)
Experienced radiologist	98.62 (96.03–99.72)	63.04 (56.45–69.29)	71.67 (66.20–76.70)	97.97 (94.19–99.58)	0.81 (0.77–0.84)
Resident	92.2 (87.81–95.39)	46.96 (40.37–53.63)	62.23 (56.69–67.54)	86.4 (79.12–91.87)	0.7 (0.65–0.74)

CAD, computer-aided diagnosis; SE, sensitivity; SP, specificity; PPV, positive predictive value; NPV, negative predictive value; AUC, area under the receiver operator characteristics curve; 95% CI, 95% confidence interval.

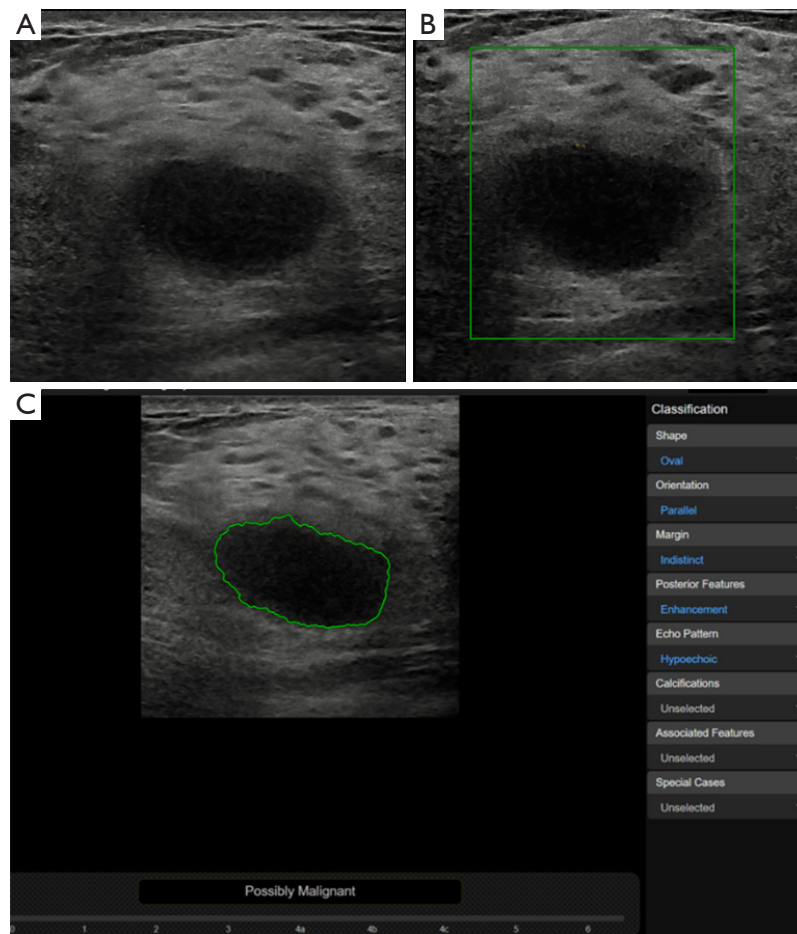


**Figure 1** A 50-year-old woman with a pathologically proved fibroadenoma in the left breast. (A) and (B), the cross section and longitude section of the irregular hypoechoic mass with ill-circumscribed margin. (C) The deep learning-based CAD correctly diagnosed it as a “possibly benign” tumor, while both the experienced radiologist and the resident diagnosed it as BI-RADS 4b. CAD, computer-aided diagnosis.

identified 119 (88.1%) lesions as benign ones in the 135 fibroadenomas (*Figure 1*). Among the 35 cases of breast adenosis, 25 of the lesions were classified accurately by deep learning-based CAD with an accuracy rate of 71.4%. As for the 27 intraductal papillary tumors, deep learning-based CAD made right classifications in 14 lesions (51.9%). Deep learning-based CAD made mistakes in 5 lesions in the 10 cases of inflammation, and 4 lesions in the 8 cases

of sclerosing adenosis. In the malignant lesions, 151/168 (89.9%) invasive ductal carcinoma were diagnosed as malignancies by deep learning-based CAD in a total of 168 lesions (*Figure 2*); 21/29 (72.4%) DCIS were classed in the right column by deep learning-based CAD. Six of the invasive lobular carcinoma were figured out by deep learning-based CAD in 7 lesions (85.7%) (*Table 3*).

In the comparisons of deep learning-based CAD with



**Figure 2** A 47-year-old female with a pathologically proved triple-negative invasive ductal carcinoma of grade III in the left breast. (A) The longitude section of the lesion, presenting a regular hypoechoic mass with well-defined margin and posterior acoustic enhancement; (B) the colour Doppler showed there was no blood flow detected in the mass; (C) the deep learning-based CAD diagnosed the mass as a “possibly malignant” tumor, while both the experienced radiologist and the resident misdiagnosed it into BI-RADS 3. CAD, computer-aided diagnosis.

the experienced radiologist and the resident by Chi-square tests, statistical differences of diagnostic accuracy were detected in some types. In the group of adenosis, the diagnostic accuracy of deep learning-based CAD was statistically higher than that of the resident ( $P=0.012$ ). Deep learning-based CAD performed better than both doctors in fibroadenomas ( $P=0.0011$  and  $P=0.0313$ ) with statistical significance. In the diagnosis of these common types of breast lesions, deep learning-based CAD had a satisfying performance (*Table 3, Figure 3*).

## Discussion

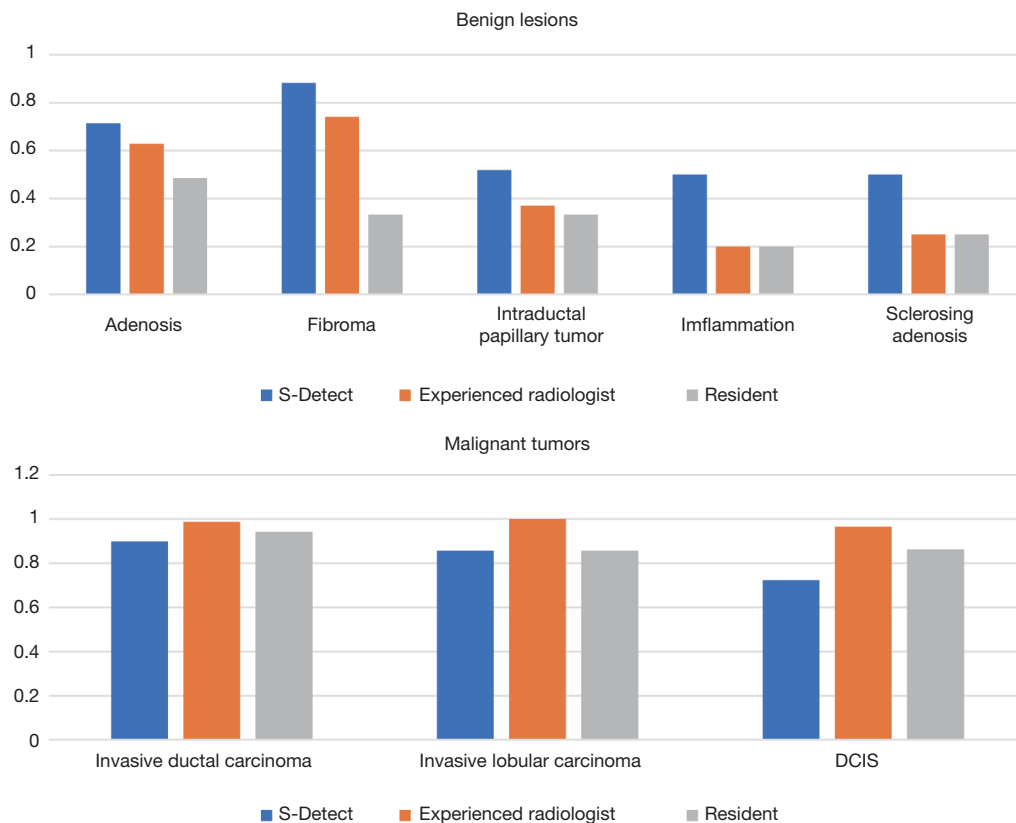
The emergence of deep learning methods has brought

unprecedented changes to the field of AI, which also has implementing a profound influence on the medical society. Deep learning is a type of machine learning that was inspired by the structure and function of the human brain. It uses ANNs that contain multiple hidden processing layers to learn inherent patterns from raw information through a complex hierarchical framework. After training on large databases and consecutive iterative procedures, the machine can make progresses itself by improving the algorithms linking of the input data and output data. ANNs are believed to have great potential in interpreting medical images (20,21). The deep learning-based CAD used in our study is typically developed on ANNs. According to the previous studies and the result of this study, the deep learning-based

**Table 3** Comparisons of deep learning-based CAD, the experienced radiologist, and the resident in the accuracy of classifying breast lesions of different histological types

Histologic types	Total	Deep learning-based CAD		Experienced radiologist		P <sup>a</sup>	Resident		P <sup>b</sup>
		Accurate	Ratio	Accurate	Ratio		Accurate	Ratio	
<b>Benign breast lesions</b>									
Adenosis	35	25	0.714	22	0.628	0.543	17	0.486	0.012
Fibroadenoma	135	119	0.881	100	0.741	0.0011	72	0.533	0.0313
Intraductal papillary tumor	27	14	0.519	10	0.370	0.0648	9	0.333	0.892
Inflammation	10	5	0.5	2	0.2	–	2	0.2	–
Sclerosing adenosis	8	4	0.5	2	0.25	–	2	0.25	–
<b>Malignant lesions</b>									
Invasive ductal carcinoma	168	151	0.899	166	0.988	0.483	158	0.940	0.127
Ductal carcinoma <i>in situ</i>	29	21	0.724	28	0.966	0.61	25	0.862	0.633
Invasive lobular carcinoma	7	6	0.857	7	1	–	6	0.857	–

<sup>a</sup>, P value between deep learning-based CAD and the experienced radiologist; <sup>b</sup>, P value between deep learning-based CAD and the resident. CAD, computer-aided diagnosis.

**Figure 3** The classification accuracy of deep learning-based CAD, the experienced radiologist, and the resident in different pathological types of breast lesions. CAD, computer-aided diagnosis.

CAD had an excellent diagnostic performance in classifying focal breast lesions (15-17). In general, the deep learning-based CAD can act as a second reader for aiding diagnosis. On US, some benign and malignant lesions look so similar that it is easy to misdiagnose each other, which affects the performance of US, especially increases the false positive rate and causes unnecessary biopsies. Since deep learning-based CAD underwent an iterative process by raw data, we supposed that it may have the better capability in identifying atypical breast lesions than human naked eyes.

In the benign lesions of our study, deep learning-based CAD had a higher accuracy than both the experienced radiologist and the resident, including fibroadenomas (88.1% *vs.* 74.1% and 53.3%), adenosis (71.4% *vs.* 62.8% and 48.6%), intraductal papillary tumors (51.9% *vs.* 37.0% and 33.3%), inflammation (50% *vs.* 20% and 20%), and sclerosing adenosis (50% *vs.* 20% and 20%). These results indicated that deep learning-based CAD had a better performance than conventional US in the benign lesions since the diagnostic information utilized by deep learning-based CAD is largely different from conventional US imaging analysis. Therefore, deep learning-based CAD may be potentially a helpful supplemental tool for the US to increase the specificity, reduce the false positive rate, and avoid unnecessary biopsies. In a recent study, Choi *et al.* (22) used the same deep learning-based CAD as ours, and the CAD decreased 10.1% false-positive biopsies by correcting the management decision of the radiologists.

Fibroadenoma and adenosis are considered to be the two most common types of benign breast lesions. Fibroadenoma is usually characterized by round or oval shape, clear boundary. However, sometimes the complex presentations that overlap malignant tumors are also detectable, including lobulated shape, uncircumscribed margin, posterior acoustic shadowing, heterogenicity, and microcalcification. The previous study reported uncircumscribed fibroadenomas were in 21.7% of the cases, lobulated shapes in 28.3%, intratumoral calcification in 9.8%, and heterogenicity in 2.2%, respectively (23). Therefore, sometimes we may diagnose fibroadenomas as BI-RADS 4 category which need biopsy. Adenosis lesions, including sclerosing adenosis and adenosis tumors (24), are considered as important entities because they often have irregular shapes and unclear boundaries mimicking the features of malignancy on imaging (24). Thus, a radiologist cannot reliably distinguish adenosis from cancers based on conventional US or mammography. In our study, CAD had a significantly higher accuracy than both the experienced radiologist ( $P=0.0011$ )

and the resident ( $P=0.0313$ ) in the fibroadenomas, and also had a significantly higher accuracy than the resident in the adenosis ( $P=0.012$ ). These results indicated that deep learning-based CAD was better at diagnosing the benign lesions since it can extract contain hidden information from raw imaging data and recognize the characteristics of the atypical fibroadenomas and adenosis, which are indistinguishable to the radiologist's naked eyes (25). So that deep learning-based CAD could increase the specificity of US and narrow down the number of patients who undergo unnecessary pathologic sampling. Thus, we deem the deep learning-based CAD to be an adjunctive tool to the US, just like elastography, which has been used as a complementary implement to decrease the number of benign biopsies, with improving the specificity of US from 47.6–61.1% to 55.6–78.5% without loss of sensitivity (18,26,27).

Invasive ductal carcinoma is the most common type of breast malignant tumors. In the previous study, US has been validated to have a better performance in the detection and diagnosis of invasive ductal carcinomas compared with mammography (83.6% *vs.* 54.3%,  $P<0.001$ ) (5). In our study, deep learning-based CAD had a satisfying diagnostic accuracy in the invasive ductal carcinomas (89.9%), with no significant difference from the experienced radiologist ( $P=0.483$ ) and the resident ( $P=0.127$ ). We consider that deep learning-based CAD had learned and been trained in abundant cases of invasive ductal carcinomas, establishing a compact capability in identifying the typical breast lesions. On the other hand, CAD needs more supervised mode learning and training in DCIS and some rare pathological types in the future.

There were some limitations in our study. First, there were not enough cases of some rare types of breast lesions for further statistical analysis, such as mucinous carcinomas, phyllodes tumors, etc. The diagnostic accuracy of deep learning-based CAD in these rare types need further analyses. Second, in our study, the static images analyzed by deep learning-based CAD were obtained by a radiologist with 19 years of experience in breast US. The quality of the images may be better in this study, and the diagnostic performance of deep learning-based CAD needs further verification.

## Conclusions

Deep learning-based CAD is a powerful tool for aiding US diagnosis of breast lesions with a high-level of diagnostic performance. It did well in the most common pathological types of breast lesions, especially in benign lesions. The

deep learning-based CAD may play an essential role in avoiding unnecessary biopsies in benign lesions.

## Acknowledgments

*Funding:* This work was supported by the National Natural Science Foundation of China [81771855], CAMS Innovation Fund for Medical Sciences [2017-I2M-1-006], Fundamental Research Funds for the Central Universities [2017320002]; and Beijing Natural Science Foundation [7192177].

## Footnote

*Conflicts of Interest:* The authors have no conflicts of interest to declare.

*Ethical Statement:* The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. This research was approved by the Institutional Review Board of Peking Union Medical College Hospital. The Institutional Ethical Committee approved this prospective study (HS 1400). Informed consent was obtained from each patient.

## References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. *CA Cancer J Clin* 2015;65:5-29.
2. Fan L, Strasser-Weippl K, Li JJ, et al. Breast cancer in China. *Lancet Oncol* 2014;15:e279-89.
3. Lynge E, Napolitano G, Vejborg I, et al. Overdiagnosis in breast cancer screening. *Transl Cancer Res* 2018;7:1313-8.
4. Barton H, Shatti D, Jones CA, et al. Review of radiological screening programmes for breast, lung and pancreatic malignancy. *Quant Imaging Med Surg* 2018;8:525-34.
5. Pan B, Yao R, Zhu QL, et al. Clinicopathological characteristics and long-term prognosis of screening detected non-palpable breast cancer by ultrasound in hospital-based Chinese population (2001-2014). *Oncotarget* 2016;7:76840-51.
6. Yan F, Song Z, Du M, et al. Ultrasound molecular imaging for differentiation of benign and malignant tumors in patients. *Quant Imaging Med Surg* 2018;8:1078-83.
7. Corsetti V, Houssami N, Ferrari A, et al. Breast screening with ultrasound in women with mammography-negative dense breasts: evidence on incremental cancer detection and false positives, and associated cost. *Eur J Cancer* 2008;44:539-44.
8. Nothacker M, Duda V, Hahn M, et al. Early detection of breast cancer: benefits and risks of supplemental breast ultrasound in asymptomatic women with mammographically dense breast tissue. A systematic review. *BMC Cancer* 2009;9:335.
9. Sprague BL, Stout NK, Schechter C, et al. Benefits, harms, and cost-effectiveness of supplemental ultrasonography screening for women with dense breasts. *Ann Intern Med* 2015;162:157-66.
10. Kohli M, Geis R. Ethics, Artificial Intelligence, and Radiology. *J Am Coll Radiol* 2018;15:1317-9.
11. Hou Z, Yang Y, Li S, et al. Radiomic analysis using contrast-enhanced CT: predict treatment response to pulsed low dose rate radiotherapy in gastric carcinoma with abdominal cavity metastasis. *Quant Imaging Med Surg*. 2018;8:410-20.
12. Dromain C, Boyer B, Ferre R, et al. Computed-aided diagnosis (CAD) in the detection of breast cancer. *Eur J Radiol* 2013;82:417-23.
13. Cheng JZ, Ni D, Chou YH, et al. Computer-aided diagnosis with deep learning architecture: applications to breast lesions in US images and pulmonary nodules in CT scans. *Sci Rep* 2016;6:24454.
14. Han S, Kang HK, Jeong JY, et al. A deep learning framework for supporting the classification of breast lesions in ultrasound images. *Phys Med Biol* 2017;62:7714-28.
15. Di Segni M, de Soccio V, Cantisani V, et al. Automated classification of focal breast lesions according to S-detect: validation and role as a clinical and teaching tool. *J Ultrasound* 2018;21:105-18.
16. Cho E, Kim EK, Song MK, et al. Application of Computer-Aided Diagnosis on Breast Ultrasonography: Evaluation of Diagnostic Performances and Agreement of Radiologists According to Different Levels of Experience. *J Ultrasound Med* 2018;37:209-16.
17. Zhao C, Xiao M, Jiang Y, et al. Feasibility of computer-assisted diagnosis for breast ultrasound: the results of the diagnostic performance of S-detect from a single center in China. *Cancer Manag Res* 2019;11:921-30.
18. Lee SH, Cho N, Chang JM, et al. Two-view versus single-view shear-wave elastography: comparison of observer performance in differentiating benign from malignant breast masses. *Radiology* 2014;270:344-53.
19. Zhi W, Gu X, Qin J, et al. Solid breast lesions: clinical experience with US-guided diffuse optical



- tomography combined with conventional US. *Radiology* 2012;265:371-8.
20. Chougrad H, Zouaki H, Alheyane O. Deep Convolutional Neural Networks for breast cancer screening. *Comput Methods Programs Biomed* 2018;157:19-30.
  21. Danaee P, Ghaeini R, Hendrix DA. A Deep Learning Approach For Cancer Detection and Relevant Gene Identification. *Pac Symp Biocomput* 2017;22:219-29.
  22. Choi JS, Han BK, Ko ES, et al. Effect of a Deep Learning Framework-Based Computer-Aided Diagnosis System on the Diagnostic Performance of Radiologists in Differentiating between Malignant and Benign Masses on Breast Ultrasonography. *Korean J Radiol* 2019;20:749-58.
  23. Namazi A, Adibi A, Haghighi M, et al. An Evaluation of Ultrasound Features of Breast Fibroadenoma. *Adv Biomed Res* 2017;6:153.
  24. Gity M, Arabkheradmand A, Taheri E, et al. Magnetic Resonance Imaging Features of Adenosis in the Breast. *J Breast Cancer* 2015;18:187-94.
  25. Aerts HJ, Velazquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun* 2014;5:4006.
  26. Berg WA, Cosgrove DO, Doré CJ, et al. Shear-wave elastography improves the specificity of breast US: the BE1 multinational study of 939 masses. *Radiology* 2012;262:435-49.
  27. Suvannarerg V, Chitchumnong P, Apiwat W, et al. Diagnostic performance of qualitative and quantitative shear wave elastography in differentiating malignant from benign breast masses, and association with the histological prognostic factors. *Quant Imaging Med Surg* 2019;9:386-98.

**Cite this article as:** Xiao M, Zhao C, Zhu Q, Zhang J, Liu H, Li J, Jiang Y. An investigation of the classification accuracy of a deep learning framework-based computer-aided diagnosis system in different pathological types of breast lesions. *J Thorac Dis* 2019;11(12):5023-5031.doi: 10.21037/jtd.2019.12.10