

Machine Learning-Based Image Analysis for Accelerating the Diagnosis of Complicated Preneoplastic and Neoplastic Ductal Lesions in Breast Biopsy Tissues

Shinya Sato (✉ ssato53@gancen.asahi.yokohama.jp)

Kanagawa Cancer Center Research Institute <https://orcid.org/0000-0002-2155-2379>

Satoshi Maki

Chiba University Graduate School of Medicine School of Medicine: Chiba Daigaku Daigakuin Igaku Kenkyuin Igakubu

Takashi Yamanaka

Kanagawa Cancer Center: Kanagawa Kenritsu Gan Center

Daisuke Hoshino

Kanagawa Cancer Center: Kanagawa Kenritsu Gan Center

Yukihide Ota

Kanagawa Cancer Center: Kanagawa Kenritsu Gan Center

Emi Yoshioka

Kanagawa Cancer Center: Kanagawa Kenritsu Gan Center

Kae Kawachi

Kanagawa Cancer Center: Kanagawa Kenritsu Gan Center

Kota Washimi

Kanagawa Cancer Center: Kanagawa Kenritsu Gan Center

Masaki Suzuki

Kanagawa Cancer Center: Kanagawa Kenritsu Gan Center

Yoichiro Ohkubo

Kanagawa Cancer Center: Kanagawa Kenritsu Gan Center

Toshinari Yamashita

Kanagawa Cancer Center: Kanagawa Kenritsu Gan Center

Seiji Ohtori

Chiba University Graduate School of Medicine School of Medicine: Chiba Daigaku Daigakuin Igaku Kenkyuin Igakubu

Yohei Miyagi

Kanagawa Cancer Center: Kanagawa Kenritsu Gan Center

Research Article

Keywords: Breast cancer, Machine learning, Artificial intelligence, Pathological diagnosis, Preneoplastic and neoplastic lesions, Grad-CAM

Posted Date: February 11th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-212926/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Purpose: Diagnosis of breast preneoplastic and neoplastic lesions is difficult due to their similar morphology in breast biopsy specimens. To diagnose these lesions, pathologists perform immunohistochemical analysis and consult with expert breast pathologists. These additional examinations are time-consuming and expensive. Artificial intelligence (AI)-based image analysis has recently improved, and may help in ordinal pathological diagnosis. Here, we showed the significance of machine learning-based image analysis of breast preneoplastic and neoplastic lesions for facilitating high-throughput diagnosis.

Methods: Images were obtained from normal mammary glands, hyperplastic lesions, preneoplastic lesions and neoplastic lesions, such as usual ductal hyperplasia (UDH), columnar cell lesion (CCL), ductal carcinoma in situ (DCIS), and DCIS with comedo necrosis (comedo DCIS) in breast biopsy specimens. The original enhanced convoluted neural network (CNN) system was used for analyzing the pathological images.

Results: The AI-based image analysis provided the following area under the curve values (AUC): normal lesion vs. DCIS, 0.9902; DCIS vs. comedo DCIS, 0.9942; normal lesion vs. CCL, 0.9786; and UDH vs. DCIS, 1.000. Multiple comparison analysis showed precision and recall scores similar to those of single comparison analysis. Based on the Gradient-weighted Class Activation Mapping (Grad-CAM) used to visualize the important regions reflecting the result of CNN analysis, the ratio of stromal tissue in the whole weighted area was significantly higher in UDH and CCL than that in DCIS.

Conclusions: These analyses may provide a more accurate and rapid pathological diagnosis of patients. Moreover, Grad-CAM identifies uncharted important histological characteristics for newer pathological findings and targets of research for understanding diseases.

Introduction

Breast cancer is a major cause of death in women [1,2]. Breast cancer treatments improve prognoses of patients with early stage disease, but patients with advanced stage still have poor prognosis [3-5]. Therefore, early diagnosis of non-invasive carcinoma or precancerous lesions is most important for preventing development of advanced breast cancer development. Pathological examinations are essential for detecting these lesions. The most recent classification of breast ductal lesion is based on the differences in biological behavior from basic and clinical research outcomes, such as normal mammary ducts, hyperplastic lesions, precancerous lesions, and carcinomas [6,7]. These lesions differ in their proliferative behavior and genetic background [6,8], and a precise diagnosis of these lesions contributes to good clinical outcomes in the patients. However, for the diagnosis of complicated ductal lesions, pathologists need to perform several types of immunohistochemical staining that require additional time and cost [8,9]. Moreover, despite meticulous analysis, the final diagnosis of these lesions

differs between pathologists because of the difficulty of morphological and immunohistochemical assessments [10,11].

Artificial intelligence (AI) field has been closely corroborated with pathological diagnoses in the recent decade. Along with the improvement of image processing technologies, including whole slide imaging (WSI), the AI approach has been integrated into pathological diagnosis [12-15]. Nowadays, AI-based pathological image analysis is being used in predicting cancer recurrence [16,17], therapeutic outcomes [16], and genetic mutations [18]. The U.S. Food and Drug Administration has recently approved the WSI system for digital pathology, and the use of AI-based pathological diagnosis is accelerating as a robust supporting tool for pathologists. The AI-based pathological image analyses can process very detailed and enormous information that pathologists could never assess. For example, AI can detect a slight difference in nuclear features and gland angularity that is difficult to be identified in microscopic inspection by a pathologist [19,20].

In the field of breast cancer pathology, AI-based image analysis is important for more accurate diagnosis. Recent reports have been shown that machine learning-based image analysis can distinguish invasive and non-invasive cancer [17,21]. However, these systems did not distinguish more detailed benign, preneoplastic, and neoplastic lesions, such as usual ductal hyperplasia (UDH), columnar cell lesion (CCL), and ductal carcinoma in situ (DCIS) with comedo pattern necrosis (comedo DCIS) [17,21]. Thus, to overcome problem for pathological diagnosis in breast cancer, the detailed assessment of early tumorigenic lesions and non-malignant lesions is essential for providing appropriate treatment for patients.

In this study, we performed AI-based analysis on four ductal lesions with similar morphological characteristics, but clinically different therapeutic outcomes (UDH, CCL, DCIS, and comedo DCIS). The result is promising as a novel supporting tool for the usual pathological diagnosis and may identify newer morphological characteristics for an improved diagnosis of breast lesions.

Materials And Methods

Patients and histological material

This study included 125 biopsy cases (with 15 open biopsy cases) that comprised 222 slides-77 cases of DCIS including 23 cases of comedo DCIS, 13 cases of CCL, 23 cases of UDH, and 113 cases of normal mammary glands in the cases. All cases were obtained between 2010 and 2020. Each biopsy slide included 2-6 tissues. The average age of patients with DCIS, comedo DCIS, CCL, and UDH lesions was 58.0, 55.9, 47.6, and 49.5 years, respectively. A total of 595 images, comprising 110 images for DCIS, 75 images for comedo DCIS, 150 images for CCL, 110 images for UDH, and 150 images for normal mammary glands were used for AI training or testing. The workflow is shown in Fig. 1. UDH cases were subjected to immunohistochemical examinations for ER, CK5/6, or p63 to distinguish them from DCIS. All cases were obtained from the diagnostic pathology division database at the Kanagawa Cancer Center Hospital. All images were obtained from the biopsy samples. All diagnostic criteria for DCIS, comedo

DCIS, CCL, and UDH were based on the 5th edition of the World Health Organization's blue book [7]. Six pathologists (T.Y., K.K., E.Y., Y.O., S.S., M.S., and K.W.) separately diagnosed these cases, and S.S. reviewed all cases. Institutional review boards at the Kanagawa Cancer Center and the Chiba University approved all the test set study activities.

Obtaining Images of each lesion from biopsy samples

All hematoxylin and eosin (H&E) stained slides were digitalized using Leica Aperio CS2 (Leica Biosystems Imaging, Inc., Wetzlar, Germany). Further, all images were analyzed using Qupath-0.2.0-m7 (open-source software [19]). Images were captured at 10x and 20x magnification. Each image included only one lesion and excluded any artifacts.

Model construction and training of the convoluted neural network (CNN) for pathological image analysis

Python programming language, version 3.6.7 and Keras, version 2.2.4 with Tensorflow, version 1.14.0 at the backend were used to build the CNN architecture. In the present study, we used the InceptionV3 architectural model, that was previously trained on ImageNet. Inception-v3 is an Inception family of convolutional neural network architectures with several improvements such as label smoothing, factorized 7x7 convolution, and the use of auxiliary classifiers to propagate label information lower down the network [22]. The input images were scaled down to 299 × 299 pixels. We then fine-tuned the model with the image dataset of images of DCIS, Comedo DCIS, flat epithelial atypia (FEA), columnar cell hyperplasia, and normal glands. Weights in the first 249 layers were frozen, and weights in the other layers were retrained with our data. The network was trained for 100 epochs with a learning rate of 0.1 that was reduced if no improvement was observed. The model training convergence was monitored using cross-entropy loss. All images were randomly augmented using Image Data Generator (<https://keras.io/preprocessing/image/>) by a rotation angle range of 180°, width shift range of 0.2, height shift range of 0.2, brightness range of 0.3–1.0, and horizontal and vertical flip in 50%. The CNN was trained and validated using a computer with a GeForce RTX 2060 graphics processing unit (NVIDIA, Santa Clara, CA), a Core i7-9750 central processing unit (Intel, Santa Clara, CA), and 16 GB of random access memory. To evaluate the performance of the CNN, we plotted the receiver operating characteristic (ROC) curve and calculated the area under the curve (AUC). We also calculated the precision and recall for assessing the diagnostic accuracy of the CNN model.

Statistical analysis

For experimental data analysis, GraphPad Prism 9 software (GraphPad Software, San Diego, California, USA) was used. Normality assumption was tested using one-way ANOVA for the analysis of tumor-stromal area. Tukey-Kramer method was used for multiple comparisons in the analysis of tumor-stromal areas (Fig. 7). A $p < 0.05$ was considered statistically significant.

Results

Experiment 1: DCIS vs. normal

Distinguishing non-invasive cancer from normal mammary glands

First, we attempted to distinguish non-invasive cancer and normal mammary glands to test the AI-based diagnosis system. The differential diagnosis between the normal mammary gland and non-invasive cancer / DCIS is essential, but not difficult for pathologists. Therefore, to distinguish them is indispensable for our system. The CNN analysis successfully distinguished DCIS from normal mammary glands (Fig. 2, AUC score: 0.9902, precision: 0.935, and recall: 0.953). These results suited this system's requirement, and thus, we applied the analysis to all other benign preneoplastic and neoplastic lesions.

Experiment 2: DCIS vs. comedo DCIS

Distinguishing DCIS and comedo DCIS pattern

In case of DCIS lesions, the comedo DCIS pattern that has necrosis at the center of DCIS has a more aggressive phenotype. Comedo DCIS pattern is a risk factor for recurrence post breast-conserving surgery [23,24], and is also correlated with the expression of poor prognostic markers [25,26]. The comedo DCIS pattern is rare and usually mixed with other DCIS lesions, leading to its omission from screening. Therefore, accurately identifying the comedo DCIS pattern is vital for pathological analysis. The AUC score in the CNN analysis was very high (0.9942). Moreover, the precision and recall were 0.972 and 0.954, respectively (Fig. 3).

Experiment 3: CCLs vs. normal

Distinguishing CCLs and normal mammary glands

CCLs include columnar cell hyperplasia and FEA. FEA has a neoplastic gene alteration than that in normal mammary glands [27-29]. Additionally, CCLs, including columnar cell hyperplasia, have similar neoplastic genetic alterations [30,31]. At present, CCL cases need to be followed up closely [32-34]. Therefore, distinguishing CCLs from normal mammary glands is important for breast biopsy tissue. Thus, we distinguished CCLs from the normal mammary glands using our system. The AUC was 0.9786, and for normal vs. CCLs, the precision and recall were 0.935 and 0.953, respectively (Fig. 4).

Experiment 4: DCIS vs UDH

Distinguishing hyperplastic lesion from DCIS

Next, we focused on the hyperplastic lesions in the mammary glands. Usual ductal hyperplasia (UDH) is a common lesion in the mammary tissue. However, dense proliferation of the mammary gland epithelium resembles DCIS, and additional immunohistochemical examination is needed for the differential diagnosis between DCIS and UDH [35,36]. To simplify this differential diagnosis process, we attempted to distinguish between DCIS and UDH using the machine learning system. The morphological structure was

very similar between DCIS and UDH (Fig. 5), but the AUC score was 1.000, and the precision and recall were 1.000 and 1.000, respectively (Fig. 5).

Experiment 5: Distinguishing all 5 lesions

Finally, we attempted to simultaneously distinguish all five lesions (normal mammary glands, CCL, UDH, DCIS, and comedo DCIS) to test our system's applicability for supporting daily pathological diagnoses. The average precision and recall between the five lesions were 0.923 (0.863-0.973) and 0.927 (0.880-0.991), respectively (Fig. 6). Thus, these results are similar to those of individual comparisons performed in experiments 1-4.

Experiment 6:

Feedback from machine learning-based analysis to the practical microscopic findings

From these results, we speculated that the machine learning-based evaluation criteria/process of each lesion may differ from that of human pathologists, especially from the UDH vs. DCIS analysis. Therefore, we used Gradient-weighted Class Activation Mapping (Grad-CAM) data, to visualize the important regions reflecting results of the CNN analysis and identify newer morphological characteristics to support the pathological diagnosis. According to the Grad-CAM results, although the current pathological diagnosis only focuses on the morphology of the epithelial structure, the AI-analysis concentrated more on stromal tissue (Fig. 7). The stromal tissue ratio in the whole weighted area (red and yellow colored area) was very high in normal mammary glands than that in others (Fig. 7). The percentage of average epithelial lesions and stromal lesions were 70.5% and 29.5% in DCIS, 27.4% and 72.6% in normal, 59.3% and 40.7% in UDH, 57.8% and 42.2% in CCL, and 68.9% and 31.1% in comedo DCIS, respectively (Fig. 7). From the statistical analysis, the ratio of stromal tissue in the whole weighted area was significantly higher in the normal mammary glands than that in all other lesions (Fig. 7). Moreover, the stromal tissue ratio in the UDH and CCL was significantly higher than that in DCIS (Fig. 7).

Discussion

Pathological diagnosis of neoplastic or benign lesions with morphological similarity is always difficult, even though the patient's treatment is completely different for them. Sometimes these problematic lesions cannot be precisely diagnosed even by IHC or consulting with experts. In the field of breast pathology, some hyperplastic, preneoplastic, and neoplastic lesions such as UDH, CCL, and DCIS are considered as difficult to diagnose lesions. Our AI system could successfully distinguish normal mammary gland, UDH, CCL, DCIS, and comedo DCIS tissues. This is the first study to address the issue of the diagnosis of these indistinguishable breast lesions using machine learning-based image analysis. In previous studies, AI-based analysis for breast pathology has been used for classifying only normal, benign, *in situ*, and invasive cancer [17,21,37]. These studies combine the benign lesions and preneoplastic lesions into one criterion, such as benign hyperplastic lesions (UDH) and preneoplastic lesions (CCL). Here, we obtained high AUC scores as well as precision and recall, in both one-to-one

comparison and comparison of all 4 lesions with normal mammary glands. Therefore, our machine-learning system may be helpful for the daily pathological diagnosis of breast biopsy tissues.

In this study, we showed the importance of stroma for distinguishing breast lesions (Fig. 7). For several years, the most critical morphological characteristics for diagnosing benign and malignant tumors were those of the tumor cell itself, and findings of surrounding stromal tissue were merely supportive information. With an improved understanding of the role of stromal cells such as cancer-associated fibroblasts in cancer progression [38,39], information of the stromal structure is now getting focus for the diagnosis of cancer [40,41]. Therefore, the findings of this study support the utility of stromal structure for the diagnosis of benign and malignant breast lesions.

This study has a limitation. We used images that include only one lesion for the AI analysis. For the analysis using WSI, many lesions coexist on a slide and each lesion needs to be distinguished. The WSI-based analysis is ideal, but currently it is difficult to develop a practically usable diagnostic systems. However, our system can be applied for daily diagnosis. The pathological division of several hospitals has a light microscope with a camera system to capture images of the IHC results and for research use. Our system only needs an image of the lesion that can be captured using a camera attached to any light microscopies, or even using camera of a smartphone. It is a very high-throughput system than WSI of a single slide. In future, we plan to apply the existed system to the WSI analysis to simultaneously detect multiple lesions from WSI.

In case of early neoplastic lesions in the mammary gland, lobular neoplasms are also crucial as are DCIS or atypical ductal hyperplasia (ADH). Lobular carcinoma in situ (LCIS) lesions have totally different morphological and immunohistochemical characteristics than those of ductal lesions, and can help distinguishing benign and malignant ductal lesions based on staining of E-cadherin as a novel protein marker [42,43]. Thus, we exclude LCIS images in our system. However, benign and malignant ductal lesions are sometimes very similar, and are difficult to distinguish even using IHC methods. Therefore, we focused on the ductal early neoplastic and preneoplastic lesions in this study.

Other AI analyses have been performed for scoring or predicting the expression of biomarkers, such as HER2, Ki-67, estrogen receptor (ER) or progesterone receptor (PgR) [44-47], and assisting in identification of lymph node metastasis [48] or tumor-associated stroma [49]. However, these analyses are not directly related to the pathological diagnosis of complicated breast lesions.

The pathologists' workload on breast cancer has recently increased annually. Along with efforts for diagnosis, pathologists also need to collect data on tumor cell density and tumor area for the genetic analysis of cancer as well as evaluation of several kinds of biomarkers, such as ER, PgR, HER2, and Ki-67 positivity [7]. Moreover, some recent very effective molecular target therapies, including anti-PD-1 or -PD-L1 treatment and CDK4/6 inhibitors, require the information about the expression of particular proteins in the cancer cells or immune cells to predict therapeutic efficacy [50-52]. Therefore, this machine learning-based diagnostic support tool may shorten the duration of pathological screening of patients' biopsies

and surgical specimens, and allow pathologists to concentrate on other essential works related to therapeutic decisions.

In conclusion, this study established an AI system that can differentiate complex benign and malignant breast lesions with a very high accuracy. This system can be used for pathological diagnosis of difficult to identify breast lesions.

Abbreviations

AI, artificial intelligence; IHC, immunohistochemistry; UDH, usual ductal hyperplasia; DCIS, ductal carcinoma in situ; CCL, columnar cell lesion; Grad-CAM, Gradient-weighted Class Activation Mapping; CNN, convoluted neural network; ROC, receiver operating characteristic; AUC, area under the curve; FEA, flat epithelial atypia; WSI, whole slide imaging; ER, estrogen receptor; PgR, progesterone receptor; HER2, human epidermal growth factor receptor 2; ADH, atypical ductal hyperplasia; LCIS, lobular carcinoma in situ.

Declarations

Funding

This work was supported by JSPS Grants-in-Aid for Scientific Research (KAKENHI) Grant Number JP20K09422 to Shinya Sato, JP20K18052 to Satoshi Maki, and the Takeda Scientific Foundation to Shinya Sato.

Conflict of interest

Y. Miyagi has consigned research fund from Tosoh Corporation. T. Yamashita has received research grants and/or honoraria for lecturing from Chugai, Eisai, Novartis Pharma, Taiho, Nippon Kayaku, AstraZeneca, Kyowa Kirin, Pfizer Japan, Eli Lilly, and Daiichi Sankyo. T. Yamanaka has received honoraria for lecturing from Chugai, Eisai, Novartis Pharma, AstraZeneca, Kyowa Kirin, Pfizer Japan, Eli Lilly, and Daiichi Sankyo. No other authors have conflict of interest to report.

Authors' contributions

S. Sato and S.Maki designed this study. S. Maki established machine-learning system. S. Sato, S. Suzuki, K. Washimi, Y. Ohkubo, K. Kawachi, E. Yoshioka, T. Yokose diagnosed lesions. S. Sato prepared images for the machine-learning analysis. S. Sato carried out statistical analysis. T. Yamanaka, T. Yamashita, Y. Ota, D. Hoshino, T. Yokose, S. Ohtori, and Y. Miyagi supervised this study. S. Sato wrote the manuscript with input and editing. S. Sato and S. Maki reviewed the manuscript.

Ethics approval

Ethical approval for this study was sought and granted by the institutional review board of Kanagawa Cancer Center and Chiba University. This article does not contain any new studies with animals performed by any of the authors.

Acknowledgements and Funding Information

We thank M. Yoshihara, K. Ohrui, M. Hatayama and Y. Nakamura for help with scanning slides. This work was supported by JSPS Grants-in-Aid for Scientific Research (KAKENHI) Grant Number JP20K09422 to Shinya Sato, JP20K18052 to Satoshi Maki, and the Takeda Scientific Foundation to Shinya Sato.

References

1. Forouzanfar MH, Foreman KJ, Delossantos AM, Lozano R, Lopez AD, Murray CJ, Naghavi M (2011) Breast and cervical cancer in 187 countries between 1980 and 2010: a systematic analysis. *Lancet* 378 (9801):1461-1484. doi:10.1016/S0140-6736(11)61351-2
2. Siegel RL, Miller KD, Jemal A (2020) Cancer statistics, 2020. *CA Cancer J Clin* 70 (1):7-30. doi:10.3322/caac.21590
3. Dickler MN, Barry WT, Cirrincione CT, Ellis MJ, Moynahan ME, Innocenti F, Hurria A, Rugo HS, Lake DE, Hahn O, Schneider BP, Tripathy D, Carey LA, Winer EP, Hudis CA (2016) Phase III Trial Evaluating Letrozole As First-Line Endocrine Therapy With or Without Bevacizumab for the Treatment of Postmenopausal Women With Hormone Receptor-Positive Advanced-Stage Breast Cancer: CALGB 40503 (Alliance). *J Clin Oncol* 34 (22):2602-2609. doi:10.1200/JCO.2015.66.1595
4. Martin M, Loibl S, Hyslop T, De la Haba-Rodriguez J, Aktas B, Cirrincione CT, Mehta K, Barry WT, Morales S, Carey LA, Garcia-Saenz JA, Partridge A, Martinez-Janez N, Hahn O, Winer E, Guerrero-Zotano A, Hudis C, Casas M, Rodriguez-Martin C, Furlanetto J, Carrasco E, Dickler MN, Group GSBC, Gbg, Alliance for Clinical Trials in O (2019) Evaluating the addition of bevacizumab to endocrine therapy as first-line treatment for hormone receptor-positive metastatic breast cancer: a pooled analysis from the LEA (GEICAM/2006-11_GBG51) and CALGB 40503 (Alliance) trials. *Eur J Cancer* 117:91-98. doi:10.1016/j.ejca.2019.06.002
5. Howie LJ, Singh H, Bloomquist E, Wedam S, Amiri-Kordestani L, Tang S, Sridhara R, Sanchez J, Prowell TM, Kluetz PG, King-Kallimanis BL, Gao JJ, Ibrahim A, Goldberg KB, Theoret M, Pazdur R, Beaver JA (2019) Outcomes of Older Women With Hormone Receptor-Positive, Human Epidermal Growth Factor Receptor-Negative Metastatic Breast Cancer Treated With a CDK4/6 Inhibitor and an Aromatase Inhibitor: An FDA Pooled Analysis. *J Clin Oncol* 37 (36):3475-3483. doi:10.1200/JCO.18.02217
6. Bombonati A, Sgroi DC (2011) The molecular pathology of breast cancer progression. *J Pathol* 223 (2):307-317. doi:10.1002/path.2808
7. WHO Classification of Tumours Editorial Board., International Agency for Research on Cancer., World Health Organization. (2019) WHO classification of tumours. Breast Tumours. World Health

Organization classification of tumours, 5th edition, vol fifth, 5th edition. edn. International Agency for Research on Cancer, Lyon

8. Tomasino RM, Morello V, Gullo A, Pompei G, Agnese V, Russo A, Rinaldi G (2009) Assessment of "grading" with Ki-67 and c-kit immunohistochemical expressions may be a helpful tool in management of patients with flat epithelial atypia (FEA) and columnar cell lesions (CCLs) on core breast biopsy. *J Cell Physiol* 221 (2):343-349. doi:10.1002/jcp.21858
9. Boecker W, Stenman G, Schroeder T, Schumacher U, Loening T, Stahnke L, Lohnert C, Siering RM, Kuper A, SamoiloVA V, Tiemann M, Korsching E, Buchwalow I (2017) Multicolor immunofluorescence reveals that p63- and/or K5-positive progenitor cells contribute to normal breast epithelium and usual ductal hyperplasia but not to low-grade intraepithelial neoplasia of the breast. *Virchows Arch* 470 (5):493-504. doi:10.1007/s00428-017-2073-7
10. Ghofrani M, Tapia B, Tavassoli FA (2006) Discrepancies in the diagnosis of intraductal proliferative lesions of the breast and its management implications: results of a multinational survey. *Virchows Arch* 449 (6):609-616. doi:10.1007/s00428-006-0245-y
11. Sinn HP, ElSawaf Z, Helmchen B, Aulmann S (2010) Early Breast Cancer Precursor Lesions: Lessons Learned from Molecular and Clinical Studies. *Breast Care (Basel)* 5 (4):218-226. doi:10.1159/000319624
12. Leong FJ, McGee JO (2001) Automated complete slide digitization: a medium for simultaneous viewing by multiple pathologists. *J Pathol* 195 (4):508-514. doi:10.1002/path.972
13. Ho J, Parwani AV, Jukic DM, Yagi Y, Anthony L, Gilbertson JR (2006) Use of whole slide imaging in surgical pathology quality assurance: design and pilot validation studies. *Hum Pathol* 37 (3):322-331. doi:10.1016/j.humpath.2005.11.005
14. Gilbertson JR, Ho J, Anthony L, Jukic DM, Yagi Y, Parwani AV (2006) Primary histologic diagnosis using automated whole slide imaging: a validation study. *BMC Clin Pathol* 6:4. doi:10.1186/1472-6890-6-4
15. Kayser K, Gortler J, Bogovac M, Bogovac A, Goldmann T, Vollmer E, Kayser G (2009) AI (artificial intelligence) in histopathology—from image analysis to automated diagnosis. *Folia Histochem Cytobiol* 47 (3):355-361. doi:10.2478/v10042-009-0087-y
16. Bychkov D, Linder N, Turkki R, Nordling S, Kovanen PE, Verrill C, Walliander M, Lundin M, Haglund C, Lundin J (2018) Deep learning based tissue analysis predicts outcome in colorectal cancer. *Sci Rep* 8 (1):3395. doi:10.1038/s41598-018-21758-3
17. Klimov S, Miligy IM, Gertych A, Jiang Y, Toss MS, Rida P, Ellis IO, Green A, Krishnamurti U, Rakha EA, Aneja R (2019) A whole slide image-based machine learning approach to predict ductal carcinoma in situ (DCIS) recurrence risk. *Breast Cancer Res* 21 (1):83. doi:10.1186/s13058-019-1165-5
18. Coudray N, Ocampo PS, Sakellaropoulos T, Narula N, Snuderl M, Fenyo D, Moreira AL, Razavian N, Tsirigos A (2018) Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat Med* 24 (10):1559-1567. doi:10.1038/s41591-018-0177-5

19. Veta M, Kornegoor R, Huisman A, Verschuur-Maes AH, Viergever MA, Pluim JP, van Diest PJ (2012) Prognostic value of automatically extracted nuclear morphometric features in whole slide images of male breast cancer. *Mod Pathol* 25 (12):1559-1565. doi:10.1038/modpathol.2012.126
20. Lee G, Sparks R, Ali S, Shih NN, Feldman MD, Spangler E, Rebbeck T, Tomaszewski JE, Madabhushi A (2014) Co-occurring gland angularity in localized subgraphs: predicting biochemical recurrence in intermediate-risk prostate cancer patients. *PLoS One* 9 (5):e97954. doi:10.1371/journal.pone.0097954
21. Mercan E, Mehta S, Bartlett J, Shapiro LG, Weaver DL, Elmore JG (2019) Assessment of Machine Learning of Breast Pathology Structures for Automated Differentiation of Breast Cancer and High-Risk Proliferative Lesions. *JAMA Netw Open* 2 (8):e198777. doi:10.1001/jamanetworkopen.2019.8777
22. Szegedy CV, V.; Ioffe, J.; Shlens, J.; Wojna Z. (2015) Rethinking the inception architecture for computer vision. *arXiv preprint arXiv 1512:00567*
23. Group EBCC, Group ER, Bijker N, Meijnen P, Peterse JL, Bogaerts J, Van Hoorebeeck I, Julien JP, Gennaro M, Rouanet P, Avril A, Fentiman IS, Bartelink H, Rutgers EJ (2006) Breast-conserving treatment with or without radiotherapy in ductal carcinoma-in-situ: ten-year results of European Organisation for Research and Treatment of Cancer randomized phase III trial 10853—a study by the EORTC Breast Cancer Cooperative Group and EORTC Radiotherapy Group. *J Clin Oncol* 24 (21):3381-3387. doi:10.1200/JCO.2006.06.1366
24. Cutuli B (2000) Pathologic findings from the National Surgical Adjuvant Breast Project (NSABP) eight-year update of protocol B-17: intraductal carcinoma. *Cancer* 88 (8):1976-1977
25. Toss MS, Miligy IM, Goringe KL, AlKawaz A, Khout H, Ellis IO, Green AR, Rakha EA (2018) Prolyl-4-hydroxylase Alpha subunit 2 (P4HA2) expression is a predictor of poor outcome in breast ductal carcinoma in situ (DCIS). *Br J Cancer* 119 (12):1518-1526. doi:10.1038/s41416-018-0337-x
26. Aljohani AI, Toss MS, Kurozumi S, Joseph C, Aleskandarany MA, Miligy IM, Ansari RE, Mongan NP, Ellis IO, Green AR, Rakha EA (2020) The prognostic significance of wild-type isocitrate dehydrogenase 2 (IDH2) in breast cancer. *Breast Cancer Res Treat* 179 (1):79-90. doi:10.1007/s10549-019-05459-7
27. Moinfar F, Man YG, Bratthauer GL, Ratschek M, Tavassoli FA (2000) Genetic abnormalities in mammary ductal intraepithelial neoplasia-flat type ("clinging ductal carcinoma in situ"): a simulator of normal mammary epithelium. *Cancer* 88 (9):2072-2081
28. Lamb LR, Bahl M, Gadd MA, Lehman CD (2017) Flat Epithelial Atypia: Upgrade Rates and Risk-Stratification Approach to Support Informed Decision Making. *J Am Coll Surg* 225 (6):696-701. doi:10.1016/j.jamcollsurg.2017.08.022
29. Alvarado-Cabrero I, Valencia-Cedillo R, Estevez-Castro R (2020) Preneoplasia of the Breast and Molecular Landscape. *Arch Med Res* 51 (8):845-850. doi:10.1016/j.arcmed.2020.09.011
30. Dabbs DJ, Carter G, Fudge M, Peng Y, Swalsky P, Finkelstein S (2006) Molecular alterations in columnar cell lesions of the breast. *Mod Pathol* 19 (3):344-349. doi:10.1038/modpathol.3800538

31. Troxell ML, Brunner AL, Neff T, Warrick A, Beadling C, Montgomery K, Zhu S, Corless CL, West RB (2012) Phosphatidylinositol-3-kinase pathway mutations are common in breast columnar cell lesions. *Mod Pathol* 25 (7):930-937. doi:10.1038/modpathol.2012.55
32. Seo M, Chang JM, Kim WH, Park IA, Lee SH, Cho N, Moon WK (2013) Columnar cell lesions without atypia initially diagnosed on breast needle biopsies: is imaging follow-up enough? *AJR Am J Roentgenol* 201 (4):928-934. doi:10.2214/AJR.12.9906
33. Grabenstetter A, Brennan S, Salagean ED, Morrow M, Brogi E (2020) Flat Epithelial Atypia in Breast Core Needle Biopsies With Radiologic-Pathologic Concordance: Is Excision Necessary? *Am J Surg Pathol* 44 (2):182-190. doi:10.1097/PAS.0000000000001385
34. Miller-Ocuin JL, Fowler BB, Coldren DL, Chiba A, Levine EA, Howard-McNatt M (2020) Is Excisional Biopsy Needed for Pure FEA Diagnosed on a Core Biopsy? *Am Surg* 86 (9):1088-1090. doi:10.1177/0003134820943546
35. Otterbach F, Bankfalvi A, Bergner S, Decker T, Krech R, Boecker W (2000) Cytokeratin 5/6 immunohistochemistry assists the differential diagnosis of atypical proliferations of the breast. *Histopathology* 37 (3):232-240. doi:10.1046/j.1365-2559.2000.00882.x
36. MacGrogan G, Arnould L, de Mascarel I, Vincent-Salomon A, Penault-Llorca F, Lacroix-Triki M, Bibeau F, Baranzelli MC, Fridman V, Antoine M, Becette V, Brouste V, Jacquemier J, Mathoulin-Pelissier S, group G (2008) Impact of immunohistochemical markers, CK5/6 and E-cadherin on diagnostic agreement in non-invasive proliferative breast lesions. *Histopathology* 52 (6):689-697. doi:10.1111/j.1365-2559.2008.03016.x
37. Alom MZ, Yakopcic C, Nasrin MS, Taha TM, Asari VK (2019) Breast Cancer Classification from Histopathological Images with Inception Recurrent Residual Convolutional Neural Network. *J Digit Imaging* 32 (4):605-617. doi:10.1007/s10278-019-00182-7
38. Houthuijzen JM, Jonkers J (2018) Cancer-associated fibroblasts as key regulators of the breast cancer tumor microenvironment. *Cancer Metastasis Rev* 37 (4):577-597. doi:10.1007/s10555-018-9768-3
39. Becker LM, O'Connell JT, Vo AP, Cain MP, Tampe D, Bizarro L, Sugimoto H, McGow AK, Asara JM, Lovisa S, McAndrews KM, Zielinski R, Lorenzi PL, Zeisberg M, Raza S, LeBleu VS, Kalluri R (2020) Epigenetic Reprogramming of Cancer-Associated Fibroblasts Deregulates Glucose Metabolism and Facilitates Progression of Breast Cancer. *Cell Rep* 31 (9):107701. doi:10.1016/j.celrep.2020.107701
40. Tomas D, Kruslin B (2004) The potential value of (Myo)fibroblastic stromal reaction in the diagnosis of prostatic adenocarcinoma. *Prostate* 61 (4):324-331. doi:10.1002/pros.20109
41. Mai KT, Hogan K (2016) Stromal tissue as an adjunct tool in the diagnosis of follicular thyroid lesions by fine-needle aspiration biopsy. *Cytojournal* 13:20. doi:10.4103/1742-6413.189639
42. Goldstein NS, Bassi D, Watts JC, Layfield LJ, Yaziji H, Gown AM (2001) E-cadherin reactivity of 95 noninvasive ductal and lobular lesions of the breast. Implications for the interpretation of problematic lesions. *Am J Clin Pathol* 115 (4):534-542. doi:10.1309/B0DD-4M7H-GJG1-7KCW

43. Sokolova A, Lakhani SR (2020) Lobular carcinoma in situ: diagnostic criteria and molecular correlates. *Mod Pathol*. doi:10.1038/s41379-020-00689-3
44. Vandenberghe ME, Scott ML, Scorer PW, Soderberg M, Balcerzak D, Barker C (2017) Relevance of deep learning to facilitate the diagnosis of HER2 status in breast cancer. *Sci Rep* 7:45938. doi:10.1038/srep45938
45. Shamaï G, Binenbaum Y, Slossberg R, Duek I, Gil Z, Kimmel R (2019) Artificial Intelligence Algorithms to Assess Hormonal Status From Tissue Microarrays in Patients With Breast Cancer. *JAMA Netw Open* 2 (7):e197700. doi:10.1001/jamanetworkopen.2019.7700
46. Anand D, Kurian NC, Dhage S, Kumar N, Rane S, Gann PH, Sethi A (2020) Deep Learning to Estimate Human Epidermal Growth Factor Receptor 2 Status from Hematoxylin and Eosin-Stained Breast Tissue Images. *J Pathol Inform* 11:19. doi:10.4103/jpi.jpi_10_20
47. Feng M, Deng Y, Yang L, Jing Q, Zhang Z, Xu L, Wei X, Zhou Y, Wu D, Xiang F, Wang Y, Bao J, Bu H (2020) Automated quantitative analysis of Ki-67 staining and HE images recognition and registration based on whole tissue sections in breast carcinoma. *Diagn Pathol* 15 (1):65. doi:10.1186/s13000-020-00957-5
48. Steiner DF, MacDonald R, Liu Y, Truszkowski P, Hipp JD, Gammage C, Thng F, Peng L, Stumpe MC (2018) Impact of Deep Learning Assistance on the Histopathologic Review of Lymph Nodes for Metastatic Breast Cancer. *Am J Surg Pathol* 42 (12):1636-1646. doi:10.1097/PAS.0000000000001151
49. Ehteshami Bejnordi B, Mullooly M, Pfeiffer RM, Fan S, Vacek PM, Weaver DL, Herschorn S, Brinton LA, van Ginneken B, Karssemeijer N, Beck AH, Gierach GL, van der Laak J, Sherman ME (2018) Using deep convolutional neural networks to identify and classify tumor-associated stroma in diagnostic breast biopsies. *Mod Pathol* 31 (10):1502-1512. doi:10.1038/s41379-018-0073-z
50. Mills AM, Dill EA, Moskaluk CA, Dziegielewski J, Bullock TN, Dillon PM (2018) The Relationship Between Mismatch Repair Deficiency and PD-L1 Expression in Breast Carcinoma. *Am J Surg Pathol* 42 (2):183-191. doi:10.1097/PAS.0000000000000949
51. Nanda R, Chow LQ, Dees EC, Berger R, Gupta S, Geva R, Pusztai L, Pathiraja K, Aktan G, Cheng JD, Karantza V, Buisseret L (2016) Pembrolizumab in Patients With Advanced Triple-Negative Breast Cancer: Phase Ib KEYNOTE-012 Study. *J Clin Oncol* 34 (21):2460-2467. doi:10.1200/JCO.2015.64.8931
52. Johnston S, Puhalla S, Wheatley D, Ring A, Barry P, Holcombe C, Boileau JF, Provencher L, Robidoux A, Rimawi M, McIntosh SA, Shalaby I, Stein RC, Thirlwell M, Dolling D, Morden J, Snowdon C, Perry S, Cornman C, Batten LM, Jeffs LK, Dodson A, Martins V, Modi A, Osborne CK, Pogue-Geile KL, Cheang MCU, Wolmark N, Julian TB, Fisher K, MacKenzie M, Wilcox M, Huang Bartlett C, Koehler M, Dowsett M, Bliss JM, Jacobs SA (2019) Randomized Phase II Study Evaluating Palbociclib in Addition to Letrozole as Neoadjuvant Therapy in Estrogen Receptor-Positive Early Breast Cancer: PALLET Trial. *J Clin Oncol* 37 (3):178-189. doi:10.1200/JCO.18.01624

Figures

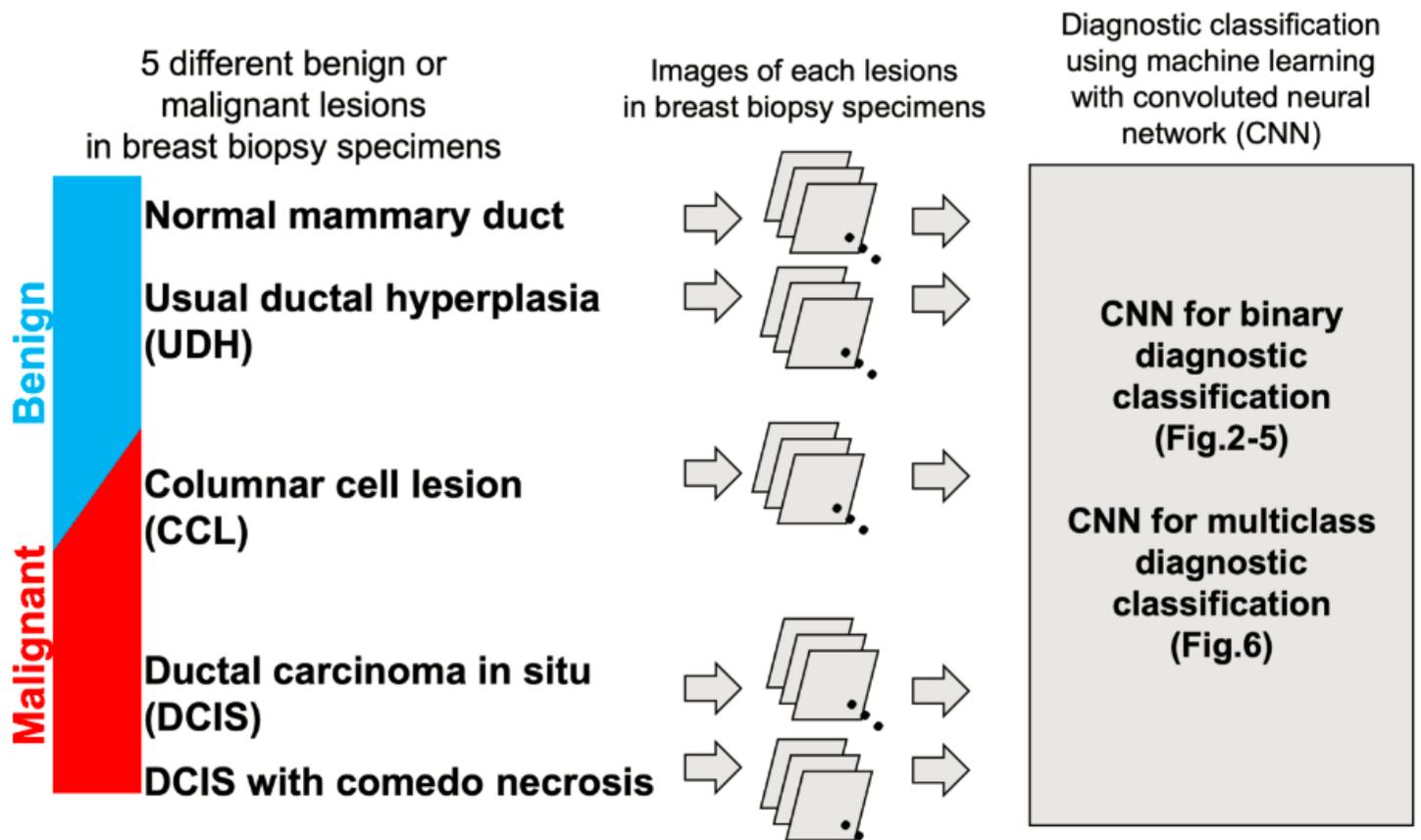


Figure 1

Workflow for the image analysis of benign and malignant lesions in breast biopsy specimens Five different benign, preneoplastic and malignant lesions are selected from breast biopsy specimens, and images from these lesions are pooled and used for machine learning analysis. We performed two type of analysis-binary comparison of lesions, and simultaneous comparison of all five lesions.

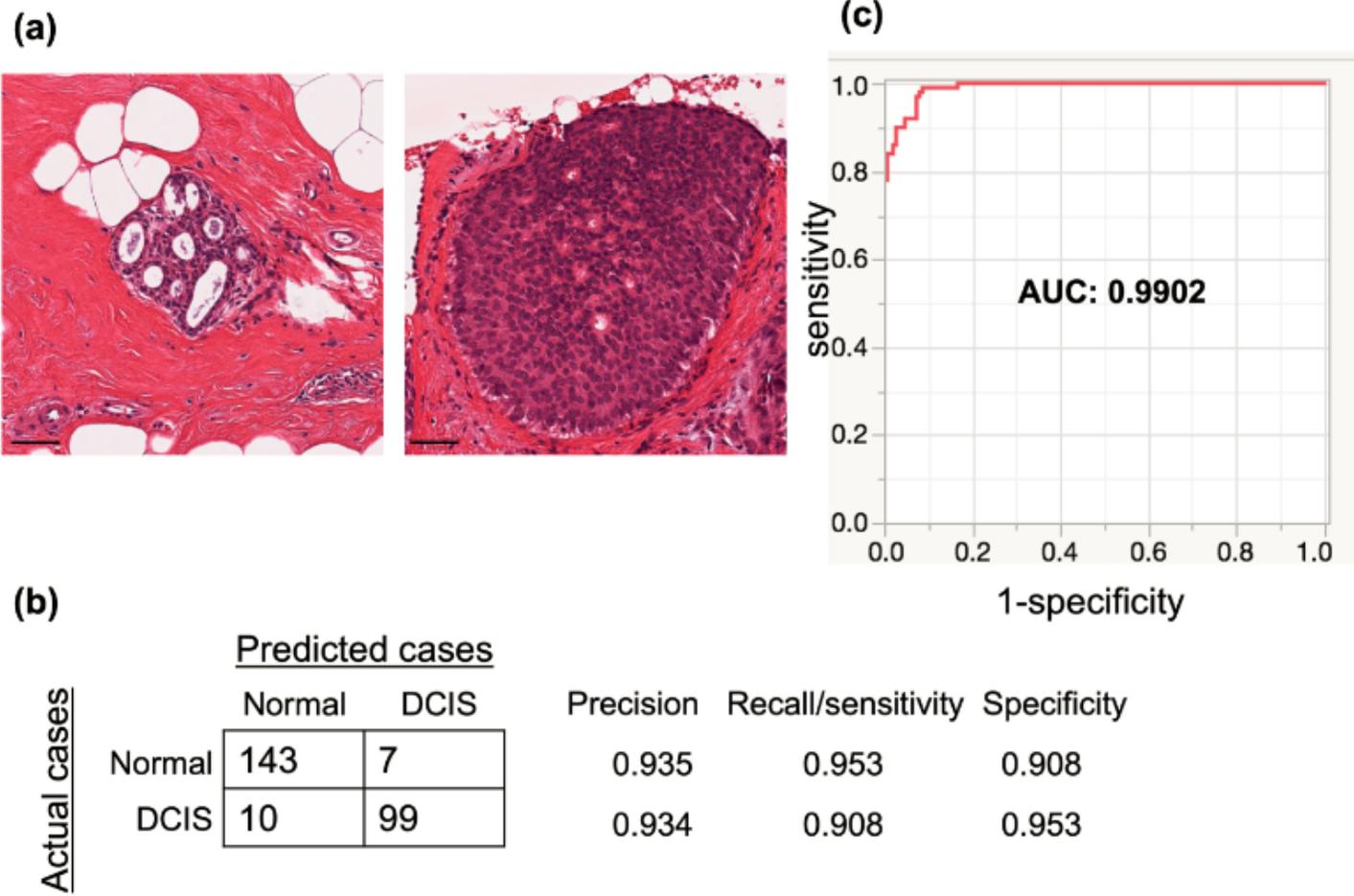


Figure 2

(a) Representative images of normal mammary gland (left) and ductal carcinoma in situ (DCIS, right). Scale bar: 50 μ m. (b) Precision, recall/sensitivity and specificity of analysis results for predicting normal mammary gland or DCIS. (c) Receiver operating characteristic (ROC) for the deep learning model to predict the normal mammary gland or DCIS in total 259 histopathological images.

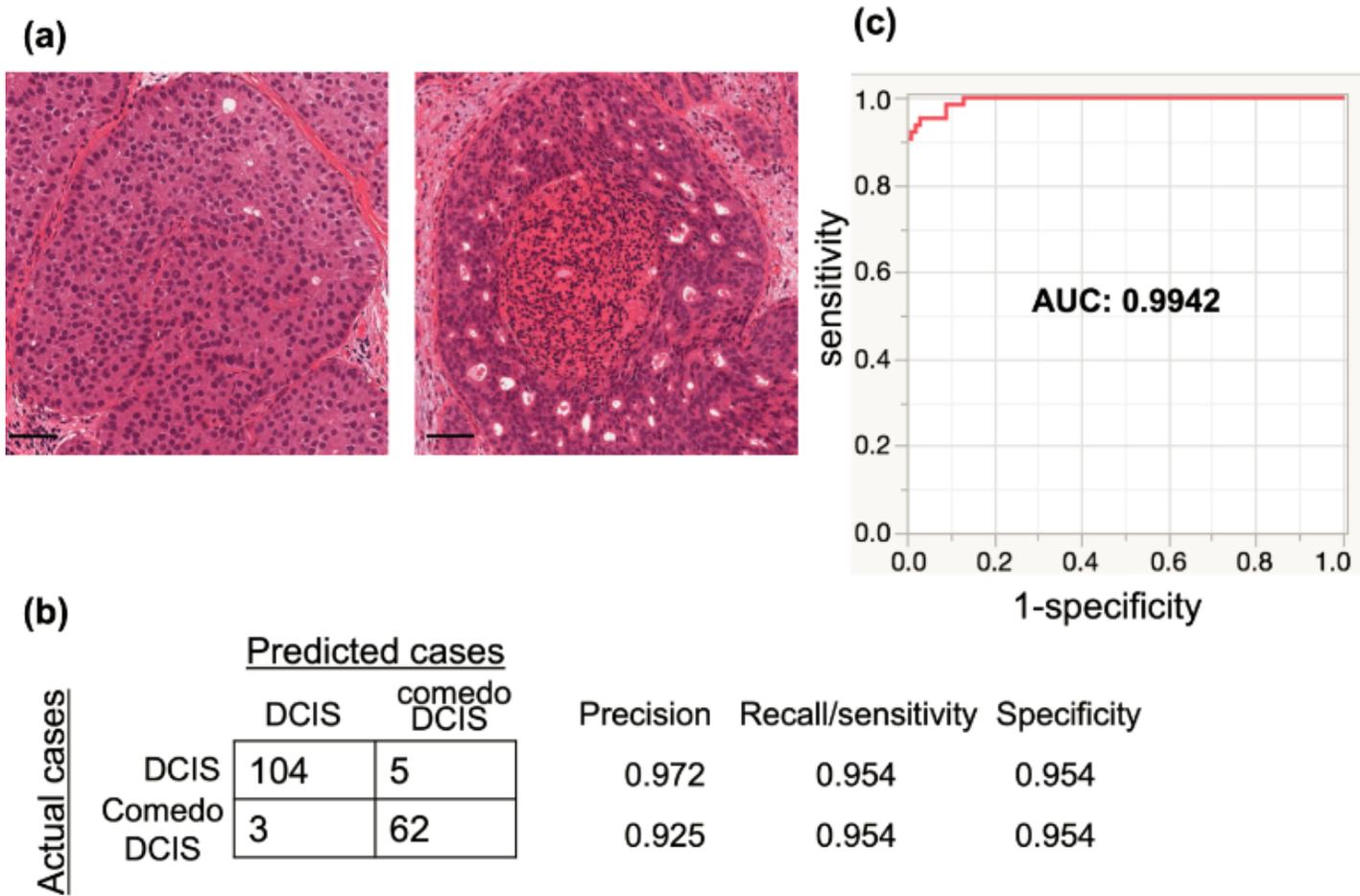


Figure 3

(a) Representative images of DCIS (left) and comedo DCIS (right). Scale bar: 50 μ m. (b) Precision, recall/sensitivity and specificity of analysis results for predicting DCIS or comedo DCIS. (c) Receiver operating characteristic (ROC) for the deep learning model to predict the DCIS or comedo DCIS on total 174 histopathological images.

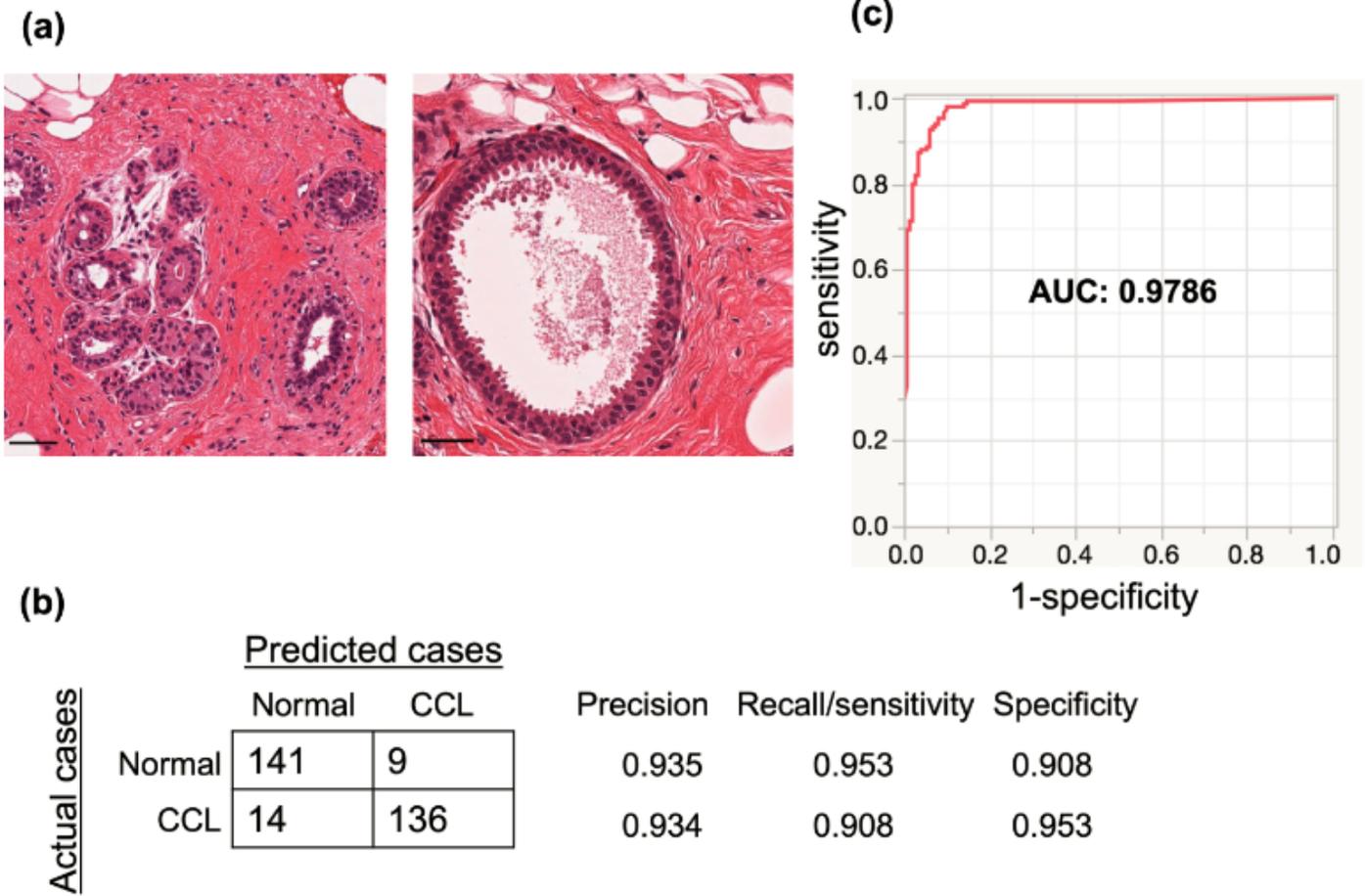


Figure 4

(a) Representative images of normal mammary gland (left) and columnar cell lesion (CCL, right). Scale bar: 50 μ m. (b) Precision, recall/sensitivity and specificity of analysis results for predicting normal mammary glands or CCL. (c) Receiver operating characteristic (ROC) for the deep learning model to predict normal mammary glands or CCL in total 300 histopathological images.

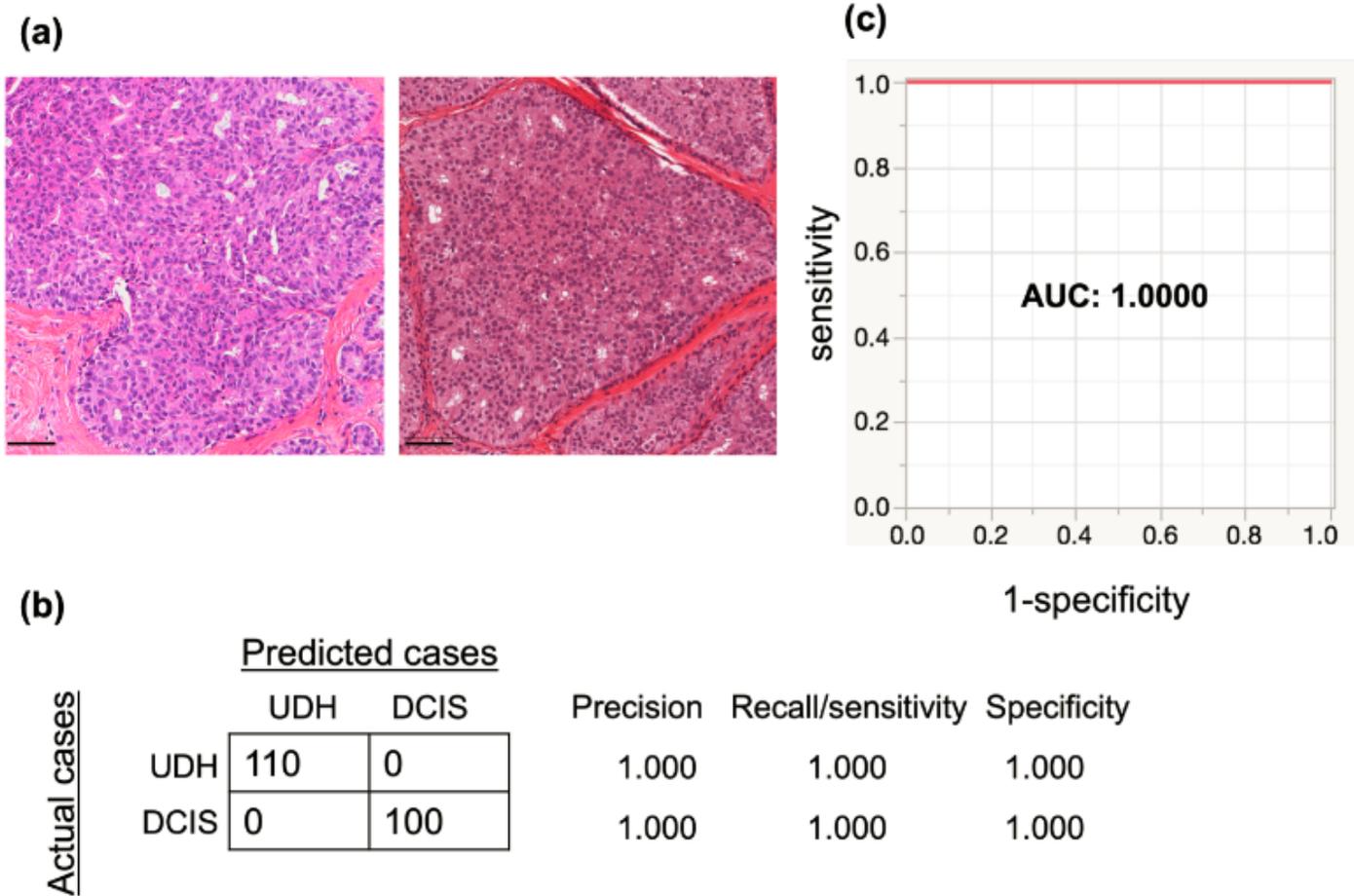
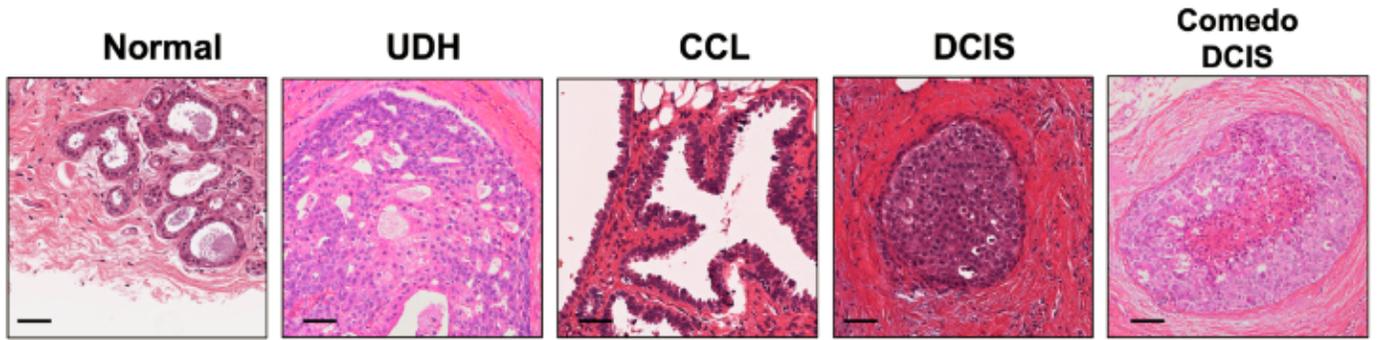


Figure 5

(a) Representative images of usual ductal hyperplasia (UDH, left) and DCIS (right). Scale bar: 50 μ m. (b) Precision, recall/sensitivity and specificity of analysis results for predicting UDH or DCIS. (c) Receiver operating characteristic (ROC) for the deep learning model to predict UDH or DCIS in total 210 histopathological images.

5 category classification

(a)



(b)

Predicted cases

		Normal	UDH	CCL	DCIS	DCIS with comedo	Precision	Recall
Actual cases	Normal	134	0	9	4	3	0.893	0.893
	UDH	0	109	0	0	1	0.973	0.991
	CCL	9	0	139	0	2	0.939	0.927
	DCIS	7	0	0	88	5	0.946	0.880
	Comedo DCIS	0	3	0	1	69	0.863	0.945

Figure 6

(a) Representative images of five lesions (from left: normal mammary gland, UDH, CCL, DCIS, and comedo DCIS). Scale bar: 50 μ m. (b) Precision and recall of analysis results for predicting all five lesions on total 574 images.

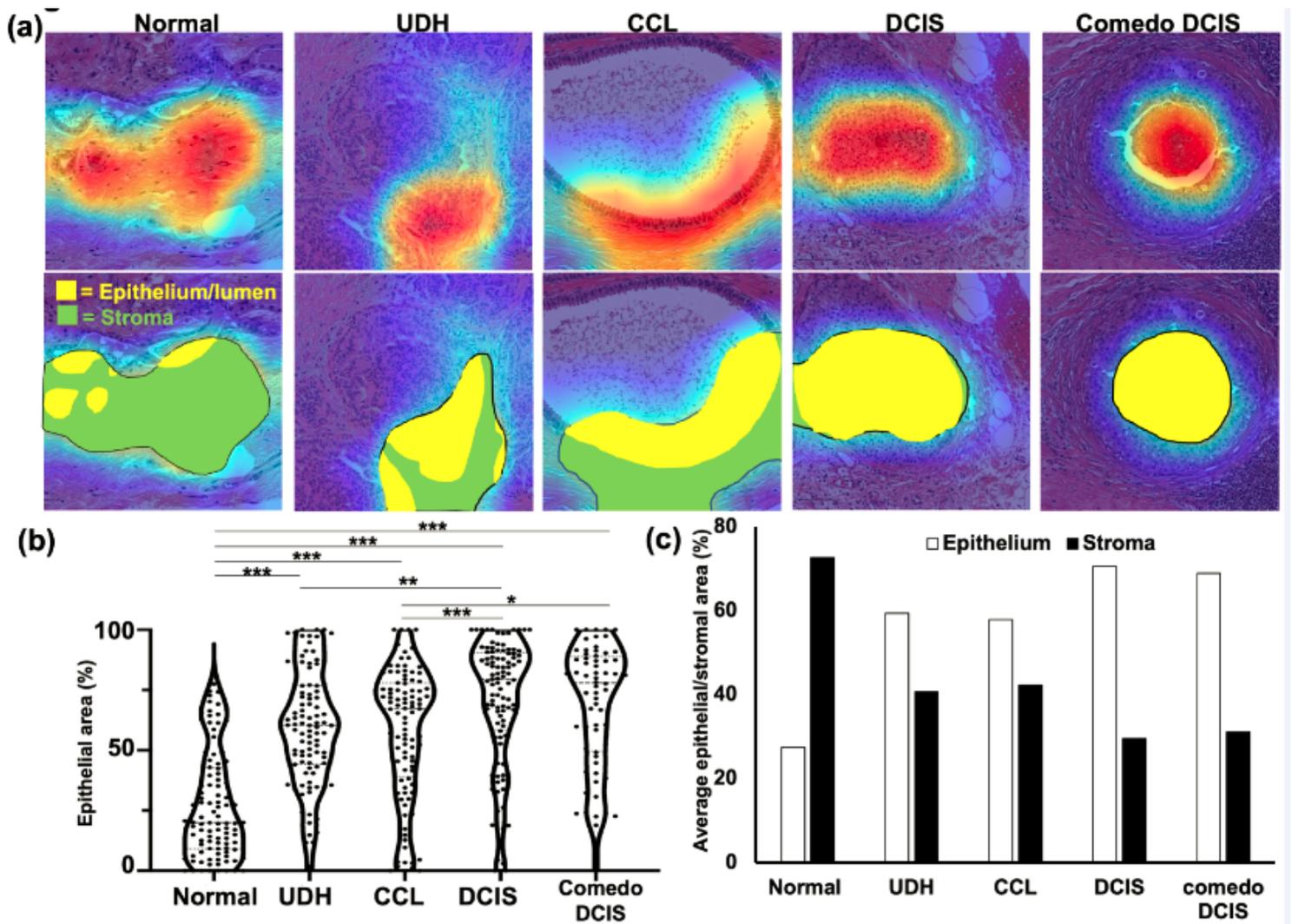


Figure 7

(a) Representative images of Grad-CAM in five lesions. Upper images: red is the most focused area in Grad-CAM. Lower images: Epithelial/lumen area (yellow) and stromal area (green) in Grad-CAM focused area. Scale bar: 50 μ m. (b) Analysis of the percentage of epithelial area in Grad-CAM focused area in each lesion. Tukey-Kramer method was used for statistical analysis. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$ (c) Average of epithelial or stromal area in Grad-CAM focused area in each lesion.