Fully Automated Postlumpectomy Breast Margin Assessment Utilizing Convolutional Neural Network Based Optical Coherence Tomography Image Classification Method

Diana Mojahed, MS, **Richard S. Ha, MD, MS**, **Peter Chang, MD**, **Yu Gan, PhD**, **Xinwen Yao, PhD**, **Brigid Angelini, MS**, **Hanina Hibshoosh, MD**, **Bret Taback, MD**, **Christine P. Hendon, PhD**

Department of Biomedical Engineering, Columbia University, New York, New York (D.M.); Department of Electrical Engineering, Columbia University, New York, New York (D.M., Y.G., X.Y., B.A., C.P.H.); Department of Radiology, Columbia University Medical Center, 622 W 168th St, PB-1-301, New York, New York 10032 (R.S.H.); Department of Radiological Sciences, University of California Irvine Medical Center, Orange, California (P.C.); Department of Pathology and Cell Biology, Columbia University Medical Center, New York, New York (H.H.); Department of Surgery, Columbia University Medical Center, New York, New York (B.T.)

Abstract

Background: The purpose of this study was to develop a deep learning classification approach to distinguish cancerous from noncancerous regions within optical coherence tomography (OCT) images of breast tissue for potential use in an intraoperative setting for margin assessment.

Methods: A custom ultrahigh-resolution OCT (UHR-OCT) system with an axial resolution of 2.7 μm and a lateral resolution of 5.5 μm was used in this study. The algorithm used an A-scan-based classification scheme and the convolutional neural network (CNN) was implemented using an 11-layer architecture consisting of serial 3×3 convolution kernels. Four tissue types were classified, including adipose, stroma, ductal carcinoma in situ, and invasive ductal carcinoma.

Results: The binary classification of cancer versus noncancer with the proposed CNN achieved 94% accuracy, 96% sensitivity, and 92% specificity. The mean five-fold validation F1 score was highest for invasive ductal carcinoma (mean standard deviation, 0.89 ± 0.09) and adipose (0.79 ± 0.17), followed by stroma (0.74 ± 0.18), and ductal carcinoma in situ (0.65 ± 0.15).

Conclusion: It is feasible to use CNN based algorithm to accurately distinguish cancerous regions in OCT images. This fully automated method can overcome limitations of manual interpretation including interobserver variability and speed of interpretation and may enable real-time intraoperative margin assessment.

Keywords

OCT; Lumpectomy; CNN

Address correspondence to: R.H.: rh2616@columbia.edu.

INTRODUCTION

Women with early-stage breast cancer undergo breast-conserving surgery, which involves the local removal of tumor and surrounding disease-free (negative) margins (1). Approximately 23% of patients require surgical re-excision (2,3), which leads to increased healthcare costs and physical and psychological stress on patients and their families (4,5). The need for re-excision could be reduced with a rapid intraoperative margin assessment tool. There is no current standard of care for routine intraoperative margin assessment. Previously intraoperative margin assessment includes touch preparation, frozen section analysis, ultrasound, and specimen radiography. Although touch preparation (6) and frozen section analysis (7) have high-resolution, they are slow and have difficulty detecting ductal carcinoma in situ (DCIS). Ultrasound (8) and specimen radiography (9) are faster than histology-based methods, but are less sensitive and have poorer performance in detecting DCIS. The MarginProbe, a handheld radiofrequency spectroscopy device, was recently developed for intraoperative margin assessment but it has been shown that there is insufficient evidence to suggest that it significantly impacts clinical outcomes (10). To our knowledge, there is no technology that has sufficiently addressed this need to gain wide use for intraoperative margin assessment, thus the motivation for our work.

Optical coherence tomography (OCT) is a high-speed, microscopic imaging modality. OCT is the optical equivalent of ultrasound, relying on the echo of near-infrared light instead of sound to produce micron-scale resolution through 1–2 mm in biological tissue (11). In contrast to high-energy X-rays and gamma rays, OCT relies on low-energy near-infrared light, which is nondestructive to tissue and can resolve microscopic structures that cannot be seen with X-rays or computed tomography (CT), and does not require contrast injection. OCT is an established medical imaging technique, and has been pioneering in ophthalmology for the past 20 years (11–13), promising in cardiology (14,15), and recently emerging in breast surgery (16–22).

OCT is being investigated as an intraoperative margin assessment tool in breast surgery. OCT has been shown to differentiate normal breast parenchyma such as lactiferous ducts, glands, adipose, and lobules, as well as pathologic conditions such as DCIS, invasive ductal carcinoma (IDC), and microcalcifications (23). OCT allows the sample to be studied in the operating room in real-time, which improves the diagnostic speed compared to histology. OCT has been investigated in a multireader clinical study and it was demonstrated that radiologists are best-suited for interpreting this modality (24). Interpretation of OCT images is typically performed by researchers and clinicians, but manual image interpretation is challenging due to its slow speed and time to train readers, high interobserver variability, and image complexity. As a result, manual interpretation is not practical in an intraoperative setting. Currently, there is no standard protocol for intraoperative margin assessment and usually takes few days at our institution for pathology processing to determine margin status. Automated image analysis has the potential to improve diagnostic accuracy with lower interobserver variability and faster speeds, which would increase the clinical impact of OCT and make it more suitable for intraoperative imaging. Deep learning could be used to automatically analyze OCT images to reduce image interpretation time.

Deep learning approaches have been developed for OCT imaging and breast cancer imaging, but very few works have investigated using deep learning for breast tissue classification using OCT images (25). Deep learning for OCT has been investigated more extensively for ophthalmology applications, including quantifying intraretinal fluid (26,27) and diagnosing retinal disease (28). In breast cancer imaging, deep learning has been used to explore several clinical problems using mammograms, MRI scans, and histology (29–35).

The purpose of this study was to develop a deep learning classification approach to distinguish cancerous from noncancerous regions within OCT images of breast tissue for potential use in an intraoperative setting for margin assessment.

MATERIALS AND METHODS

Tissue Collection

Institutional Review Board (IRB) approval was waived. Deidentified human breast specimens included both normal and non-neoplastic tissues, and were handled in accordance with Code of Federal Regulations 45CFR46. Prior to imaging, the specimens were placed in Rosewell Park Memorial Institute and imaged within 24 hours of surgical excision. The average specimen size was 1.2 cm^2 (range: $1.0–4.0 \text{ cm}^2$). 46 specimens from 23 patients were imaged with a custom ultrahigh-resolution OCT system. Seventeen specimens were normal tissue, and 29 specimens were cancerous, including 26 specimens with IDC and three specimens with DCIS confirmed with histological correlation.

Imaging Protocol

A custom in-house ultrahigh-resolution OCT system centered at 840 nm with an axial resolution of 2.7 μm and a lateral resolution of 5.5 μm measured in air was used in this study (20,36). This system was used because our previous work showed that images acquired with this system showed increased tissue classification accuracy compared to those from a commercial system with lower resolution (20). An OCT volume is composed of 800 by 800 pixels laterally covering 3 mm by 3 mm tissue surface area, and 1024 pixels in the axial direction covering 1.78 mm in depth. All specimens were imaged fresh at room temperature.

Histology

After imaging, tissue specimens were placed in 10% formalin for 24 hours and then transferred to 70% ethanol for histological processing. Specimen blocks were embedded and sliced along the OCT imaging direction. Multiple 5-μm thick slices were taken from a single specimen block, with 100 μm discarded between levels and each slide stained with Hematoxylin and Eosin. The processed slides were digitized at 40x magnification using an Aperio system. ImageScope software (v12.1.0.5029) was used to view and annotate histology images. Histology findings were evaluated by a pathologist (H.H.) with more than 20 years of experience. The dataset of specimens is listed in Table 1.

Image Labeling

Each A-line within every OCT B-scan was manually labeled into four tissue types: stroma, adipose, IDC, or DCIS. Stroma included all glandular tissue, including the parenchyma and

connective tissues. The labeling was guided by corresponding histology. These four tissue types were chosen because these are the most common features of normal and diseased breast tissue. Three OCT readers (B.A., X. Y, Y.G.) labeled the data with the guidance of histological evaluations by a pathologist (H.H).

The labeling procedure was carried out using an in-house graphical user interface that enables consecutive labeling for three-dimensional data (Fig 1). Two volumes were labeled per specimen per patient for 23 patients, corresponding to 36,800 B-scans.

Proposed Convolutional Neural Network

All software code for this study was written in Python using the TensorFlow module (1.0.0). Validation and convolutional neural network (CNN) training were done on a Linux workstation with 15 GB RAM, 946 GB disk space, Intel Xeon Silver 4110 CPU, and a NVIDIA Titan Xp GPU.

The deep learning algorithm utilized a customized hybrid 2D/1D CNN to map each 2D B-scan to a 1D label vector. Further, by using serial 2D convolutional filters, prediction of tissue type for each A-line was in part influenced by the surrounding A-lines. This design choice was important given the relatively high noise present in a single A-line taken in isolation.

The OCT image volumes were preprocessed before the CNN analysis was performed. The B-scans of each volume were down-sampled from 1024×800 pixels to 256×200 pixels. A simple z-score transformation (x - mean/S.D.) was used to normalize the histogram of each volume. The number of B-scans in each volume was maintained (800 B-scans/volume). A single tissue label class was assigned to each A-line of the down-sampled B-scans as the prediction of the CNN. Eighty percent of the B-scans within a given volume were randomly divided into the training set, and 20% of the B-scans into the validation set.

The CNN was implemented using an 11-layer architecture consisting of serial 3×3 convolutional filters applied to the down-sampled B-scans (37), with channel sizes increasing from 4 to 64 with increasing convolutional depth (Fig 2). Convolutional filters were applied with a stride of two in the superficial-to-deep dimension to collapse the image height, while a stride of one was applied in the left-to-right dimension to preserve image width. All nonlinear functions were modeled by the rectified linear unit (ReLU) (38). Batch normalization was used between the convolutional and ReLU layers to limit drift of layer activations during training (39). The feature channel sizes increased from 4 to 64 with increasing convolutional depth reflecting increasing representational complexity. Parameters were initialized using the heuristic as described (40). Training was implemented using the Adam optimizer, an algorithm for first-order gradient-based optimization of stochastic objective functions (41) and standard stochastic gradient descent technique with Nesterov momentum (42). L2 regularization was implemented to prevent over-fitting of data by limiting the squared magnitude of the kernel weights. The F1 score was used to optimize the parameters. To account for training dynamics, the learning rate was originally set to 3.0×10^{-4} and was annealed and the mini-batch size was increased from size 16 whenever

training loss plateaus. Furthermore, a normalized gradient algorithm was employed to allow for locally adaptive learning rates that adjust according to changes in the input signal (43).

The A-lines in every volume were randomly divided into training and validation sets. Two experiments were performed: first, five-fold cross-validation was used to estimate accuracy over the entire dataset. Correlation with manual annotations was calculated using a F1 score for training and validation. The F1 score is a measure of similarity between two samples and is commonly used to assess the performance of image classification methods. The F1 score is defined as:

$$\text{F1 score} = \frac{2\,\text{TP}}{2\,\text{TP} + \text{FN} + \text{FP}}$$

where TP is the number true positives, FN is the number of false negatives, and FP is the number of false positives.

Second, the entire dataset was partitioned into a single division of 80% training and 20% testing. Correlation with manual annotations was calculated using accuracy, sensitivity, and specificity for training and testing.

IRB approval was waived as this work is considered nonhuman subjects research under Code of Federal Regulations 45 CFR 46. Breast specimens were received from the Columbia University Tumor Bank, where personnel provided deidentified tissue from per their IRB approved protocol AAAB2667, therefore this research qualifies as nonhuman subject research.

RESULTS

Four different breast tissue structures were classified in this algorithm. Examples of these features of breast tissue in OCT images and corresponding Hematoxylin and Eosin histology are illustrated in Figure 3. The algorithm was used to classify these four tissue types individually, as well as perform a binary classification (cancer vs. no-cancer). For interpretation, each image (B-scan) was classified in 0.1 seconds, and each OCT 3D volume, composed of 800 B-scans, was classified in 80 seconds. The OCT image acquisition time to acquire 5 cm^3 volume specimen was 4 minutes. The mean five-fold validation F1 score was highest for IDC (mean standard deviation, 0.89×0.09) and adipose (0.79×0.17), followed by stroma (0.74×0.18), and DCIS (0.65×0.15) (Table 2). IDC and DCIS were combined as single class (cancer), and adipose and stroma were combined as the noncancer class. Using this binary classification, the mean five-fold validation F1 score for cancer was 0.88×0.04, and 0.84×0.06 for noncancer.

In the second experiment, the entire dataset was randomly divided into 80% training and 20% testing. The accuracy, sensitivity, and specificity of the proposed CNN for the training set were 95%, 96%, and 93%, respectively. The accuracy, sensitivity, and specificity for the testing set were 94%, 96%, and 92%, respectively.

DISCUSSION

In this study, we designed a CNN algorithm that achieved 94% accuracy, 96% sensitivity, and 92% specificity in a binary classification of detecting cancerous versus non-cancerous tissue in OCT images of breast specimens. Our results suggest the feasibility of using deep learning to classify cancer in OCT images of breast tissue in an intraoperative setting rather than using traditional manual image interpretation. Manually classifying OCT images of breast tissue has been investigated in a multireader study (24). Our deep learning framework had higher accuracy than the 88% accuracy of seven clinician readers combined, including radiologists, pathologists, and surgeons.

The potential to use OCT as a tool for improving clinical decision making in the intraoperative setting has been studied over the past decade and shown to be promising. OCT can differentiate normal breast parenchyma such as lactiferous ducts, glands, adipose, and lobules, as well as pathologic conditions such as DCIS, IDC, and microcalcifications (23). In the multireader study mentioned above, clinicians (radiologists, surgeons, and pathologists) were trained to distinguish suspicious from nonsuspicious areas of postlumpectomy specimens using OCT images, and the results showed that readers from different specialties could accurately read OCT images with an average training time of 3.4 hours. All clinical readers had an average accuracy of 88%. These results further validated the practical feasibility to use OCT as a real-time intraoperative margin assessment tool in breast-conserving surgery. Although clinicians can be trained to read OCT images, there remain practical concerns of high interobserver variability and slow speed, which make manual interpretation impractical for the intraoperative setting, therefore indicating a need for automated techniques to solve these problems.

Machine learning has been shown to be effective for classification of OCT images. Machine learning has been used to improve OCT glaucoma detection (44) and to segment and visualize retinal structures in volumetric image sets (45). Since traditional machine learning algorithms often require several processing steps, deep learning CNNs are being developed because they require less processing and mostly use an input of images. Deep learning has been used in OCT imaging for retinal segmentation and diagnosis of retina conditions (26), coronary artery classification (46), and intracoronary segmentation (47). Machine learning and deep learning have been shown to be effective in these areas (26,46,47) and have now been applied in OCT breast imaging with success.

There are limitations of this study that will be considered in future work. The first limitation is that this was a retrospective study with a small patient sample size based on available data at the time of the study. Increasing the sample size in a future prospective study will likely improve the classification accuracy and subject variability. It will be important to increase the sample size particularly of DCIS and benign cases, because these conditions were not represented in large quantity in our study. The number of DCIS cases was the smallest, and its specificity was the lowest, so analyzing many more DCIS samples could improve the specificity of our algorithm. The second limitation is that histology can possibly include false positives and false negatives, which can impact the trained model and incorporate a bias. Only one pathologist was a reader in this study, thus introducing a limitation to this

work. A general limitation of deep learning, unlike rule-based methods, is that it is difficult to extrapolate the internal mechanisms of a CNN, which lead to its decision-making capabilities (48,49). This is an area of ongoing research to uncover the reasons why deep learning algorithms perform well.

In conclusion, our study demonstrates the feasibility of using Convolutional Neural Network (CNN) algorithms to classify cancer in OCT images of breast tissue and this study presents a unique A-line based classification scheme that can be used in real-time applications and extended beyond breast imaging to other applications. Automated processing using machine learning algorithms can overcome challenges of interobserver variability and improve speed in OCT image interpretation. Automation makes it more feasible to utilize OCT in an intraoperative setting for margin assessment and has potential to reduce the need for surgical re-excision.

ACKNOWLEDGMENTS

REFERENCES

1. Clough KB, Lewis JS, Couturaud B, et al. Oncoplastic techniques allow extensive resections for breast-conserving therapy of breast carcinomas. Ann Surg 2003; 237:26–34. doi:10.1097/00000658-200301000-00005. [PubMed: 12496527]

2. Tartter PI, Kaplan J, Bleiweiss I, et al. Lumpectomy margins, reexcision, and local recurrence of breast cancer. Am J Surg 2000; 179:81–85. doi:10.1016/S0002-9610(00)00272-5. [PubMed: 10773138]

3. McCahill LE, Single RM, Bowles EJA, et al. Variability in reexcision following breast conservation surgery. JAMA 2012; 307:467–475. doi:10.1001/jama.2012.43. [PubMed: 22298678]

4. Waljee JF, Hu ES, Newman LA, et al. Predictors of re-excision among women undergoing breast-conserving surgery for cancer. Ann Surg Oncol 2008; 15:1297–1303. doi:10.1245/s10434-007-9777-x. [PubMed: 18259820]

5. Olsen MA, Nickel KB, Margenthaler JA, et al. Increased risk of surgical site infection among breast-conserving surgery re-excisions. Ann Surg Oncol 2015; 22:2003–2009. doi:10.1245/s10434-014-4200-x. [PubMed: 25358666]

6. Simiyoshi K, Nohara T, Iwamoto M, et al. Usefulness of intraoperative touch smear cytology in breast-conserving surgery. Exp Ther Med 2010; 1:641–645. doi:10.3892/etm_00000100. [PubMed: 22993587]

7. Cendán JC, Coco D, Copeland EM. Accuracy of intraoperative frozen-section analysis of breast cancer lumpectomy-bed margins. J Am Coll Surg 2005; 201:194–198. doi:10.1016/j.jamcollsurg.2005.03.014. [PubMed: 16038815]

8. Doyle TE, Factor RE, Ellefson CL, et al. High-frequency ultrasound for intraoperative margin assessments in breast conservation surgery: a feasibility study. BMC Cancer 2011; 11:444. doi:10.1186/1471-2407-11-444. [PubMed: 21992187]

9. Goldfeder S, Davis D, Cullinan J. Breast specimen radiography. Can it predict margin status of excised breast carcinoma? Acad Radiol 2006; 13:1453–1459. doi:10.1016/j.acra.2006.08.017. [PubMed: 17138112]

10. Schnabel F, Boolbol SK, Gittleman M, et al. A randomized prospective study of lumpectomy margin assessment with use of marginprobe in patients with nonpalpable breast malignancies. Ann Surg Oncol 2014; 21:1589–1595. doi:10.1245/s10434-014-3602-0. [PubMed: 24595800]

11. Huang D, Swanson EA, Lin CP, et al. Optical coherence tomography. Science (80-) 1991;254:1178–1181.doi:10.1126/science.1957169.

12. Adhi M, Duker JS. Optical coherence tomography-current and future applications. Curr Opin Ophthalmol 2013; 24:213–221. doi:10.1097/ICU.0b013e32835f8bf8. [PubMed: 23429598]

13. Puliafito CA, Hee MR, Lin CP, et al. Imaging of macular diseases with optical coherence tomography. Ophthalmology 1995; 102:217–229. doi:10.1016/S0161-6420(95)31032-9. [PubMed: 7862410]

14. Jang I-K, Bouma BE, Kang D-H, et al. Visualization of coronary atherosclerotic plaques in patients using optical coherence tomography: comparison with intravascular ultrasound. J Am Coll Cardiol 2002; 39:604–609. doi:10.1016/S0735-1097(01)01799-5. [PubMed: 11849858]

15. Kubo T, Imanishi T, Takarada S, et al. Assessment of culprit lesion morphology in acute myocardial infarction. J Am Coll Cardiol 2007; 50:933–939. doi:10.1016/j.jacc.2007.04.082. [PubMed: 17765119]

16. Luo W, Nguyen FT, Zysk AM, et al. Optical biopsy of lymph node morphology using optical coherence tomography. Technol Cancer Res Treat 2005; 4:539–547. doi:10.1177/153303460500400507. [PubMed: 16173824]

17. Nguyen FT, Zysk AM, Chaney EJ, et al. Intraoperative evaluation of breast tumor margins with optical coherence tomography. Cancer Res 2009; 69:8790–8796. doi:10.1158/0008-5472.CAN-08-4340.Intraoperative. [PubMed: 19910294]

18. Kennedy KM, McLaughlin RA, Kennedy BF, et al. Needle optical coherence elastography for the measurement of microscale mechanical contrast deep within human breast tissues. J Biomed Opt 2013; 18:121510. doi:10.1117/1.JBO.18.12.121510. [PubMed: 24365955]

19. Scolaro L, McLaughlin RA, Kennedy BF, et al. A review of optical coherence tomography in breast cancer. Photonics Lasers Med 2014; 3. doi:10.1515/plm-2014-0013..

20. Yao X, Gan Y, Chang E, et al. Visualization and tissue classification of human breast cancer images using ultrahigh-resolution OCT. Lasers Surg Med 2017; 49:258–269. doi:10.1002/lsm.22654. [PubMed: 28264146]

21. Vakoc BJ, Fukumura D, Jain RK, et al. Cancer imaging by optical coherence tomography: preclinical progress and clinical potential. Nat Rev Cancer2012; 12:363 10.1038/nrc3235. [PubMed: 22475930]

22. Gan Y, Yao X, Chang E, et al. Comparative study of texture features in OCT images at different scales for human breast tissue classification. In:2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); 2016 p. 3926–3929.doi:10.1109/EMBC.2016.7591586.

23. Hsiung P, Phatak DR, Chen Y, et al. Benign and malignant lesions in the human breast depicted with ultrahigh resolution and three-dimensional optical coherence tomography. Radiology 2007; 244:865–874. doi:10.1148/radiol.2443061536.. [PubMed: 17630358]

24. Ha R, Friedlander LC, Hibshoosh H, et al. Optical Coherence Tomography: A Novel Imaging Method for Post-lumpectomy Breast Margin Assessment A Multi-reader Study. Acad Radiol 2018; 25:279–287. doi:10.1016/j.acra.2017.09.018.. [PubMed: 29174226]

25. Triki AR, Blaschko MB, Jung YM, et al. Intraoperative margin assessment of human breast tissue in optical coherence tomography images using deep neural networks. arXiv [statML] 2017 http://arxiv.org/abs/1703.10827%0A.

26. Lee CS, Tyring AJ, Deruyter NP, et al. Deep-learning based, automated segmentation of macular edema in optical coherence tomography.Biomed Opt Express 2017; 8. doi:10.1364/BOE.8.003440..

27. Venhuizen FG, van Ginneken B, Liefers B, et al. Deep learning approach for the detection and quantification of intraretinal cystoid fluid in multivendor optical coherence tomography. Biomed Opt Express 2018;9:1545–1569. doi:10.1364/BOE.9.001545. [PubMed: 29675301]

28. De Fauw J, Ledsam JR, Romera-Paredes B, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. Nat Med 2018. doi:10.1038/s41591-018-0107-6.

29. Wang J, Yang X, Cai H, et al. Discrimination of breast cancer with microcalcifications on mammography by deep learning. Sci Rep 2016; 6. doi:10.1038/srep27327. [PubMed: 28442741]

30. Ha R, Chang P, Karcich J, et al. Convolutional neural network based breast cancer risk stratification using a mammographic dataset. Acad Radiol 2018 10.1016/j.acra.2018.06.020.

31. Ha R, Chang P, Mutasa S, et al. Convolutional neural network using a breast MRI tumor dataset can predict oncotype dx recurrence score. J Magn Reson Imaging 2018; 0. doi:10.1002/jmri.26244.

32. Ha R, Chang P, Karcich J, et al. Predicting post neoadjuvant axillary response using a novel convolutional neural network algorithm. Ann Surg Oncol 2018. doi:10.1245/s10434-018-6613-4.

33. Ha R, Chang P, Karcich J, et al. Axillary lymph node evaluation utilizing convolutional neural networks using MRI dataset. J Digit Imaging 2018. doi:10.1007/s10278-018-0086-7.

34. Bejnordi BE, Veta M, Van Diest PJ, et al. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. JAMA 2017; 318:2199–2210. doi:10.1001/jama.2017.14585. [PubMed: 29234806]

35. Ciresan DC, Giusti A, Gambardella LM, et al. Mitosis detection in breast cancer histology images with deep neural networks In: Mori K, Sakuma I, Sato Y, Barillot C, Navab N, eds. Medical image computing and computer-assisted intervention – MICCAI 2013, Berlin, Heidelberg: Springer Berlin Heidelberg; 2013:411–418.

36. Yao X, Gan Y, Marboe CC, et al. Myocardial imaging using ultrahigh-resolution spectral domain optical coherence tomography. J Biomed Opt 2016; 21:061006. doi:10.1117/1.JBO.21.6.061006.

37. Springenberg JT, Dosovitskiy A, Brox T, et al. Striving for simplicity: the all convolutional net. arXiv [csLG] 2014 http://arxiv.org/abs/1412.6806 Accessed September 26, 2018.

38. Nair V, Hinton GE. Rectified linear units improve restricted Boltzmann machines. In: Proc 27th Int Conf Mach Learn; 2010 p. 807–814. doi:10.1.1.165.6419.

39. Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. arXiv 2015; 150203167:1–11. doi:10.1007/s13398-014-0173-7.2.

40. He K, Zhang X, Ren S, et al. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: Proceedings of the IEEE International Conference on Computer Vision, 2015; 2015 p. 1026–1034. doi:10.1109/ICCV.2015.123.

41. Kingma DP, Ba J. Adam: a method for stochastic optimization. IEEE Signal Process Lett 2014.

42. Bengio Y, Boulanger-Lewandowski N, Pascanu R. Advances in optimizing recurrent networks. In: ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings; 2013:8624–8628. doi:10.1109/ICASSP.2013.6639349.

43. Mandic DP. A generalized normalized gradient descent algorithm. IEEE Signal Process Lett 2004; 11:115–118. doi:10.1109/LSP.2003.821649.

44. Burgansky-Eliash Z, Wollstein G, Chu T, et al. Optical coherence tomography machine learning classifiers for glaucoma detection: a preliminary study. Invest Ophthalmol Vis Sci 2005; 46:4147–4152. doi:10.1167/iovs.05-0366. [PubMed: 16249492]

45. Zawadzki RJ, Fuller AR, Wiley DF, et al. Adaptation of a support vector machine algorithm for segmentation and visualization of retinal structures in volumetric optical coherence tomography data sets. J Biomed Opt 2007; 12:041206. doi:10.1117/1.2772658. [PubMed: 17867795]

46. Abdolmanafi A, Duong L, Dahdah N, et al. Deep feature learning for automatic tissue classification of coronary artery using optical coherence tomography. Biomed Opt Express 2017; 8:1203. doi:10.1364/BOE.8.001203. [PubMed: 28271012]

47. Zahnd G, Karanasos A, Soest Gijs Van, et al. Quantification of fibrous cap thickness in intracoronary optical coherence tomography with a contour segmentation method based on dynamic programming. Int J CARS 2015; 10:1383–1394. doi:10.1007/s11548-015-1164-7.

48. Coates A, Arbor A, Ng AY. An analysis of single-layer networks in unsupervised feature learning. Aistats 2011; 2011:215–223. doi:10.1109/ICDAR.2011.95.

49. Marcus G Deep Learning: A Critical Appraisal. arXiv Prepr arXiv 180100631. 2018; 180100631:1–27. http://arxiv.org/abs/1801.00631.
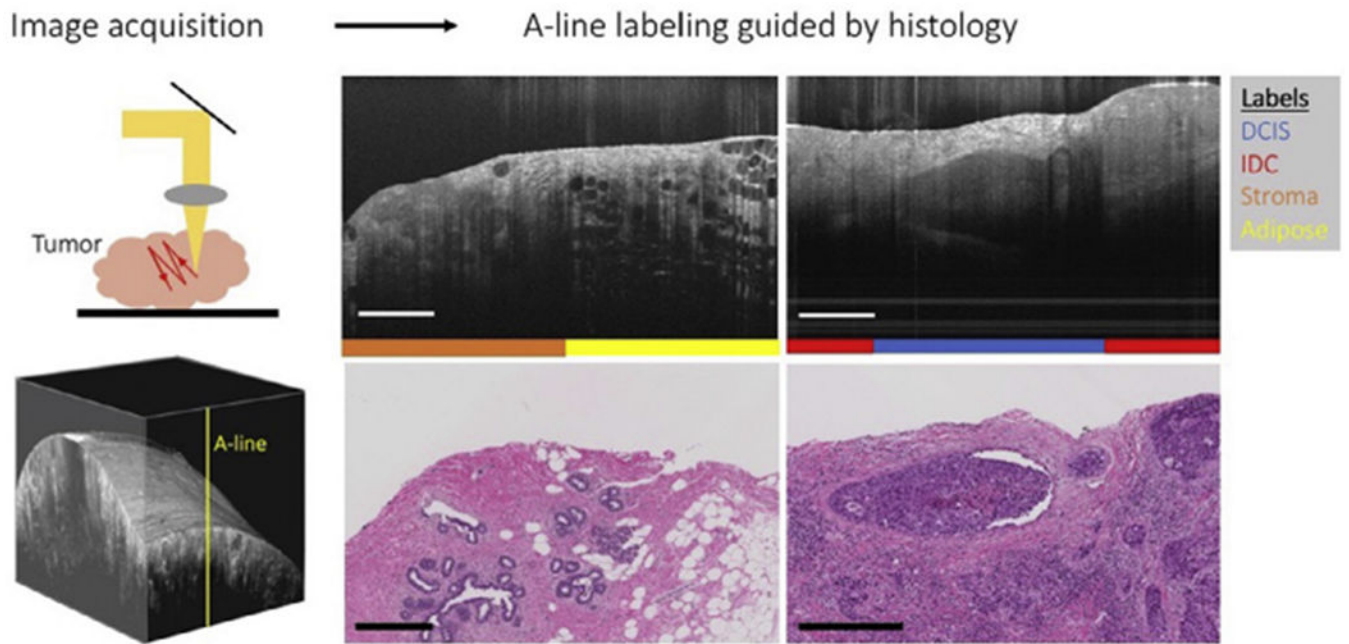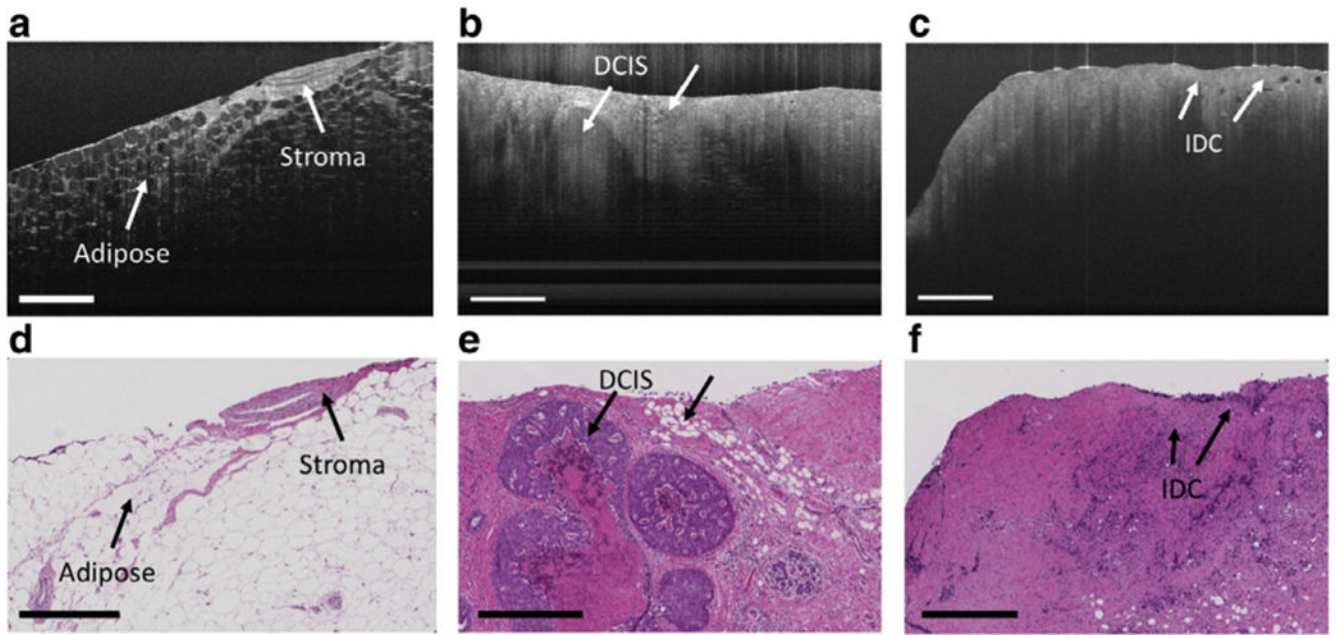
Figure 1.
Work-flow of OCT images acquisition and labeling. OCT system scans excised breast specimen on bench-top. Each 3D image set contains subsurface information of tissue structure. After imaging, specimen is sent for histology processing. Readers label each A-line of each 3D image volume as noncancer vs cancer using histology as a guide. Scale bar = 500 μm. OCT, optical coherence tomography. (Color version of figure is available online.)
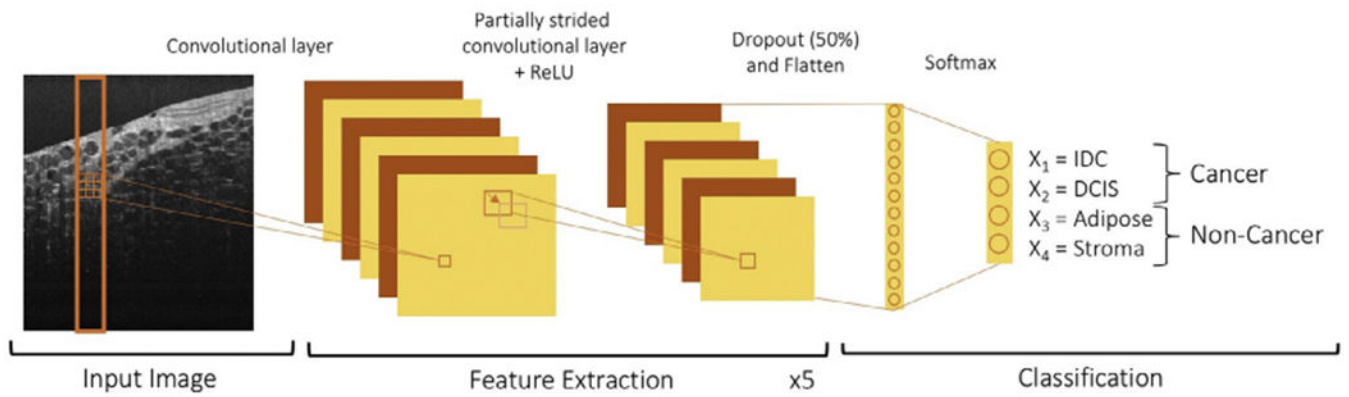
Figure 2.
The CNN architecture is composed of 3×3 convolutions, five partially strided convolutions and 50% dropout in the second to last convolutional layer. The output is flattened to produce a dimensional column-wise feature vector, upon which a softmax classifier is used to produce final network outputs. (Color version of figure is available online.)

Figure 3.
OCT images of breast tissue structures acquired with ultrahigh-resolution OCT (UHR-OCT) system. (a) Stroma and adipose (b) Ductal carcinoma in situ (DCIS) (c) Invasive ductal carcinoma (IDC) (d-f) Images of corresponding histology (Hematoxylin-eosin stain; Original magnification, x40.) Scale bar: 500 μm. OCT, optical coherence tomography. (Color version of figure is available online.)

TABLE 1.

Distribution of Tissue Types in Dataset. Forty-six Specimens From 23 Patients Were Imaged With a Custom Ultrahigh-Resolution OCT (UHR-OCT) System. Seventeen Specimens Were Normal Tissue, and 29 Specimens Were Cancerous

Characteristic	Value (n)
Number of patients	23
Number of specimens	46
Specimen histological confirmations	
Normal	17
Cancer	29
IDC	26
DCIS	3

TABLE 2.

Distribution of Tissue Types in Dataset and Corresponding Five-Fold Cross-Validation F1 Scores

Tissue Type	Five-Fold Cross-Validation F1 Scores	Binary Classification Five-Fold Cross-Validation F1 Scores
IDC	0.82–0.95	0.84–0.94
DCIS	0.54–0.75	
Adipose	0.67–0.91	0.81–0.93
Stroma	0.61 –0.86	