SIIM 2017 Scientific Session
Analytics & Deep Learning Part 2

Friday, June 2 | 8:00 am – 9:30 am

Mammographic Breast Density Classification by a Deep Learning Approach

Mohamed A. Aly, PhD, University of Pittsburgh; Robert Nishikawa, PhD

Hypothesis

We hypothesis that a deep learning based computerized tool can improve the clinical assessment of breast density from digital mammograms, thereby, potentially providing more accurate and consistent breast density notifications to women who receive a screening mammogram.

Introduction

Mammography is the standard screening examination for breast cancer. Breast density is a measure used to describe the proportion of fibroglandular tissue in a woman's breast depicted on a digital mammogram. According to the American College of Radiology's Breast Imaging Reporting and Data System (BI-RADS) criteria [1], currently, mammographic breast density is clinically evaluated by four qualitative categories, i.e., fatty, scattered density, heterogeneously dense, or extremely dense. Several large clinical studies have established that mammographic breast density is an imaging-based risk factor for breast cancer [2], independent of age and menopausal status. Women with extremely dense breasts have a 4-6 fold higher risk compared to women with fatty breasts [3]. When comparing a woman with heterogeneously or extremely dense breasts to women with average breast density, her risk is about 1.2 or 2.1 times more likely to develop breast cancer. In terms of the BI-RADS categories, a clinical assessment of either the "heterogeneously dense" or "extremely dense" category defines dense breasts, and physicians may execute additional clinical workup (e.g., supplementary screening) for women identified as having dense breasts.

In the U.S., 27 states have enacted breast density notification legislation. This legislation requires notifying women with some level of information regarding their breast density after a screening mammogram. Depending on their clinical BI-RADS assessment, different courses of action may be taken for patients based on the patient's perception of risk and potential supplementary screening. Therefore, an accurate assessment of breast density is critical in order to better inform patients about their risks and subsequent screening options.

Based on radiologists' visual observations, current qualitative assessment of breast density by BI-RADS categories is subjective, with noticeable reader variability [3], but it is currently the clinical standard. In order to improve the accuracy of the density assessment, one major concern lies in improving the consistency of the assessment. While it is relatively easy to consistently distinguish fatty breasts (BI-RADS category 1) from extremely dense breasts (category 4), there is particular difficulty for radiologists to distinguish the categories of scattered density (category 2) from heterogeneously dense (category 3). However, distinguishing between these two categories represents a very critical division, because of the different indications caused by assigning breast density as either "scattered density" or "heterogeneously dense". The latter case will triage a woman to the "dense breasts" group and, therefore, trigger the suggestion for additional screening workup, as well as the anxiety and stress resulting from the indicated higher risk of developing breast cancer.

The aim of this work is to provide a machine learning based breast density assessment tool to aid radiologists in breast density categorization, and potentially to support a more accurate and consistent lawful breast density notification. More specifically, our goal is to develop a deep learning based approach using Convolutional Neural Networks (CNNs) to more accurately classify between the two BI-RADS categories: scattered density versus heterogeneously dense, which is particularly difficult to distinguish by radiologists' visual assessments.

Methods

Materials and Methods

3.1 Image data

We are working towards collecting a large data set of digital mammogram images, including both the mediolateral oblique (MLO) and cranial caudal (CC) views of the bilateral breasts, and their corresponding BI-RADS breast density categories, routinely annotated in standard clinical workflow by radiologists. The mammograms are the post-processed (i.e., "FOR PRESENTATION") images acquired from the mammography platforms. In order to challenge the classification algorithms, all these images have either a "scattered dense" or "heterogeneously dense" assessment. At this stage, we used 12,000 of the images for training the breast density classifier and a separate unseen set of 3909 images for testing the classification performance.

3.2 *Deep learning by using CNN*

A novel component of our work is employing a deep learning based approach to develop the computer-aided breast density assessment tool. The accuracy of most conventional classification systems is based on appropriate data representation and strong feature engineering [4]. Feature engineering is a difficult and time consuming process that needs prior expert domain knowledge of the data in order to build good features. In this work, we applied a deep learning based approach to automatically learn and organize essential image features from training data without a hand-crafted process of defining image features. We implemented the deep learning model by CNN, a very promising technique in image classification and pattern recognition [4]. We constructed a 2-class CNN model to classify the two BI-RADS breast density categories: scattered density vs heterogeneously dense. We used a large set of images, i.e., 12,000 digital mammogram images, to train the CNN model. We used the Caffe platform and an improved version of the AlexNet model [5]. In addition, the implementation of our approach is also supported by the use of Graphic Processing Units (GPUs). Our CNN model was trained on an Intel® Core™ i7-2670QM CPU@2.20GHz with 8GB RAM and a GeForce GT 525M/PCIe/SSE2 GPU. In addition to the GPU acceleration, we used rectified linear units (ReLU) in place of the traditional tangent function and the sigmoid function as the activation function [4] to further speed up the training.

Results

We used the trained CNN model to make predictions of the breast density categories on 3909 unseen testing mammogram images with known ground truth: 1909 with a "scattered density" assessment and 2000 with a "heterogeneously dense" assessment. The classification performance of testing our CNN model was 90% and 85% for the scattered density and heterogeneously dense categories, respectively. The overall accuracy of our CNN model was 87.3% based on the metric of image recognition rate.

Discussion

In breast cancer screening, assessing breast density is a routine clinical need. The BI-RADS based assessment is often challenging for radiologists due to the difficulty of discerning the image features of dense breasts in coming up with a consistent classification to one of the four BI-RADS categories, particularly between the scattered density (category 2) and heterogeneously dense categories (category

3). Our ultimate goal is to develop a computerized tool to help radiologists make a clinical assessment of breast density, potentially integrating the tool as a second quantitative reader to aid the decision-making. We created a novel deep learning based model that avoids hand-crafted image features. This method, therefore, will substantially improve the accuracy and consistency of current clinical breast density assessment so that patients will benefit from receiving more accurate density assessments of their breasts.

Conclusion

Our preliminary results show encouraging classification performance by our CNN-based deep learning approach to distinguish scattered density and heterogeneously dense categories. We are performing more experiments with an increasing number of images and by incorporating quantitative breast density features. We believe that our approach will provide a promising computer tool to enhance current clinical assessment of breast density in breast cancer screening.

References

1. American College of Radiology. Breast Imaging Reporting and Data System (BI-RADS). 4th ed. Reston, VA: American College of Radiology; 2003.
2. Boyd NF, Guo H, Martin LJ, Sun L, Stone J, Fishell E, Jong RA, Hislop G, Chiarelli A, Minkin S, Yaffe MJ. Mammographic density and the risk and detection of breast cancer. N Engl J Med. 356(3):227-36. 2007.
3. Huo CW, Chew GL, Britt KL, et al. Mammographic density-A review on the current understanding of its association with breast cancer. Breast Cancer Res Treat 144:479-502, 2014.
4. Y. LeCun, Y. Bengio, and G. Hinton, Deep learning, Nature, vol. 521, pp. 436–444, 2015.
5. Https://github.com/BVLC/caffe/tree/master/models

Keywords

mammography, breast density, BI-RADS, lawful density notification, deep leaning, convolutional neural networks