Intelligenza artificiale: lo stato dell'arte

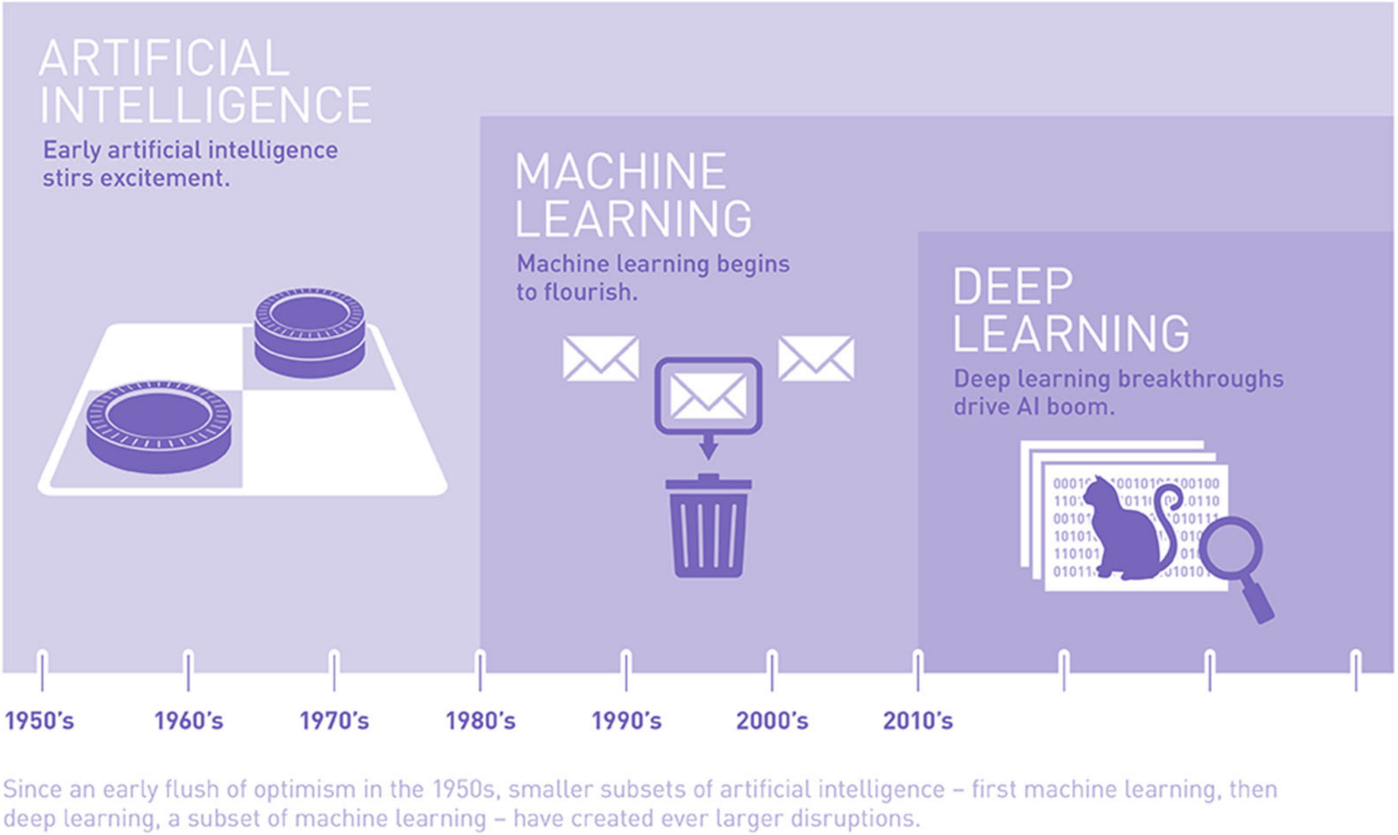
Marina Codari, PhD
(and Manuela R. Trimboli)

POLITECNICO MILANO 1863
B³ LAB Biosignals Bioimaging Bioinformatics

Programma regionale di screening per il tumore della mammella
Prevenzione serena – Workshop 2019

NOTHING TO DISCLOSE

Artificial Intelligence



https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/

1992

Computerized detection of clustered microcalcifications in digital mammograms: Applications of artificial neural networks

Yuzheng Wu, Kunio Doi, Maryellen L. Giger, and Robert M. Nishikawa
Kurt Rossmann Laboratories for Radiologic Image Research, Department of Radiology, The University of Chicago, Chicago, Illinois 60637

(Received 21 October 1991; accepted for publication 17 March 1992)

Artificial neural networks have been applied to the differentiation of actual "true" clusters from normal parenchymal patterns and also to the differentiation of actual clusters from false-positive clusters as reported by a computerized scheme for the detection of microcalcifications in digital mammograms. The differentiation was carried out in both the spatial and frequency domains. The performance of the neural networks was evaluated quantitatively by means of receiver operating characteristic (ROC) analysis. It was found that the networks could distinguish clustered microcalcifications from normal nonclustered areas in the frequency domain, and that they could eliminate approximately 50% of false-positive clusters of microcalcifications while preserving 95% of the positive clusters, when applied to the results of the automated detection scheme. A large, comprehensive training database is needed for neural networks to perform reliably in clinical situations.

Key words: microcalcification, mammography, neural network, ROC analysis, detection, classification

I. INTRODUCTION

Breast cancer is one of the leading causes of death in women. Mammography has been proven to be the primary diagnostic procedure for the early detection of breast cancer.[1] Between 30% and 50% of breast carcinomas demonstrate microcalcifications on mammograms, and between 60% and 80% of the carcinomas reveal microcalcifications upon histologic examination.[2-7] Therefore, clustered microcalcifications are an important sign in the detection of breast carcinoma. In this study, we investigated the application of artificial neural networks to the detection of microcalcifications in digital mammograms. Because an automated scheme for the detection of microcalcifications has been developed in our laboratories,[8,9] we applied the neutral networks to "positive" clusters reported by this automated detection scheme in an effort to eliminate some of the false-positive clusters, and thus to improve the overall detection efficiency.

Artificial neural networks differ from conventional algorithmic approaches to information processing in that problems are not solved by use of a predetermined algorithm, but rather by "learning" from examples presented repeatedly. The use of neural networks is thus regarded as a nonalgorithmic approach.

Neural networks have been applied to medical imaging and decision making in recent years and have been shown to be a powerful tool for pattern recognition and data classification.[10-21] Among the applications, neural networks have been employed in attempts to interpret neonatal chest radiographs,[12,13] to differentiate among patterns corresponding to various interstitial diseases in chest radiography,[14,15] and to classify mammographically evident lesions as benign or malignant.[16] These applications involved the classifications of data patterns that were subjectively extracted from images and other measurements. The neural networks have also been applied to objectively measured data patterns in applications such as radiographic signal detection,[17] x-ray spectral reconstruction from measured spectra,[18] and MRI tissue classification.[19,20] In this study, we applied the neural networks directly to images or preprocessed images in order to recognize patterns that may include microcalcifications in digital mammograms.

II. METHOD

There are two different aspects of the detection of microcalcifications in digital mammograms, namely, (1) detection of individual microcalcifications and (2) detection of clustered microcalcifications. Because only clustered microcalcifications are associated with malignancy in breasts, we focused more on the detection of clustered microcalcifications in this study. The digital mammograms used in this study were obtained by digitizing conventional screen-film (Kodak Min R/OM) mammograms on a Fuji drum scanner system with a pixel size of 0.1×0.1 mm^2. The locations of "true" microcalcifications in mammograms were identified by an expert radiologist.

We will first give an overview of our method and then describe each step in detail. We first selected ROIs of three different types, namely, positive, negative, and false-positive, from the digital mammograms, and applied a background-trend correction[21] to the selected ROIs. Two approaches were employed in the classification of microcalcifications: one approach in the spatial domain and the other in the frequency domain. For the approach in the frequency domain, power spectra were calculated by Fourier transformation of the background-corrected ROIs and scaled. The scaled power spectra were then used as input to

 0094-2405/92/30555-06$01.20

1993

Yuzheng Wu, MSc • Maryellen L. Giger, PhD • Kunio Doi, PhD • Carl J. Vyborny, MD, PhD
Robert A. Schmidt, MD • Charles E. Metz, PhD

Artificial Neural Networks in Mammography: Application to Decision Making in the Diagnosis of Breast Cancer[1]

The authors investigated the potential utility of artificial neural networks as a decision-making aid to radiologists in the analysis of mammographic data. Three-layer, feed-forward neural networks with a back-propagation algorithm were trained for the interpretation of mammograms on the basis of features extracted from mammograms by experienced radiologists. A network that used 43 image features performed well in distinguishing between benign and malignant lesions, yielding a value of 0.95 for the area under the receiver operating characteristic curve for textbook cases in a test with the round-robin method. With clinical cases, the performance of a neural network in merging 14 radiologist-extracted features of lesions to distinguish between benign and malignant lesions was found to be higher than the average performance of attending and resident radiologists alone (without the aid of a neural network). The authors conclude that such networks may provide a potentially useful tool in the mammographic decision-making task of distinguishing between benign and malignant lesions.

Index terms: Breast neoplasms, diagnosis, 00.31, 00.32 • Computers, diagnostic aid • Computers, neural network • Receiver operating characteristic curve (ROC)

Radiology 1993; 187:81–87

[1] From the Kurt Rossmann Laboratories for Radiologic Image Research, Department of Radiology, MC2026, University of Chicago, 5841 S Maryland Ave, Chicago, IL 60637. From the 1991 RSNA scientific assembly. Received June 23, 1992; revision requested September 8; revision received October 19; accepted November 23. Supported by U.S. Public Health Service grants CA48985 and CA24806 and by American Cancer Society Faculty Award FRA-390. **Address reprint requests to** M.L.G.

MAMMOGRAPHY has become a major diagnostic procedure in the early detection of breast cancer (1). However, the interpretation of mammograms for the diagnosis of breast cancer, which involves merging a large number of radiographic features of a suspicious lesion, is a difficult task. Only 15%–30% of cases that have mammographically suspicious but nonpalpable findings and are subjected to biopsy prove to be malignant (2). Automated classifiers that merge image features may be useful to radiologists in distinguishing between benign and malignant patterns in mammography and, thus, in recommending an appropriate course of action. Getty et al (3) described an expert system developed for this purpose that was based on discriminant analysis of image-feature ratings extracted by radiologists. Here, we report our investigation of an alternative approach in which an artificial neural network serves as the automated classifier.

Artificial neural networks, which constitute a nonalgorithmic approach to information processing, have been studied intensively in the field of computer science in recent years (4,5). These neural networks, which are capable of processing a large amount of information simultaneously, address problems not by means of prespecified algorithms but rather by "learning" from examples that are presented repeatedly. The popularity of neural networks is due primarily to their apparent ability to make decisions and draw conclusions when presented with complex, "noisy," or partial information and to adapt their behavior to the nature of the training data. In medical imaging, artificial neural networks have been applied to a variety of data-classification and pattern-recognition tasks, such as the differential diagnosis of interstitial diseases (6), and have been shown to provide a potentially powerful classification tool (7–12).

We employed three-layer, feed-forward neural networks with a back-propagation algorithm for the interpretation of mammograms on the basis of features that had been extracted from mammograms by experienced mammographers. Our data base consisted of features from 133 textbook cases and 60 clinical cases. Performance of the neural networks in classifying lesions as benign or malignant was evaluated with receiver operating characteristic (ROC) analysis. In addition, we evaluated the performance of attending radiologists and residents in classifying benign and malignant lesions in the same clinical cases. The performance of these observers was compared with that of neural networks that merged features that had been extracted by an experienced mammographer.

MATERIALS AND METHODS

Radiographic Features Used in the Interpretation of Mammograms

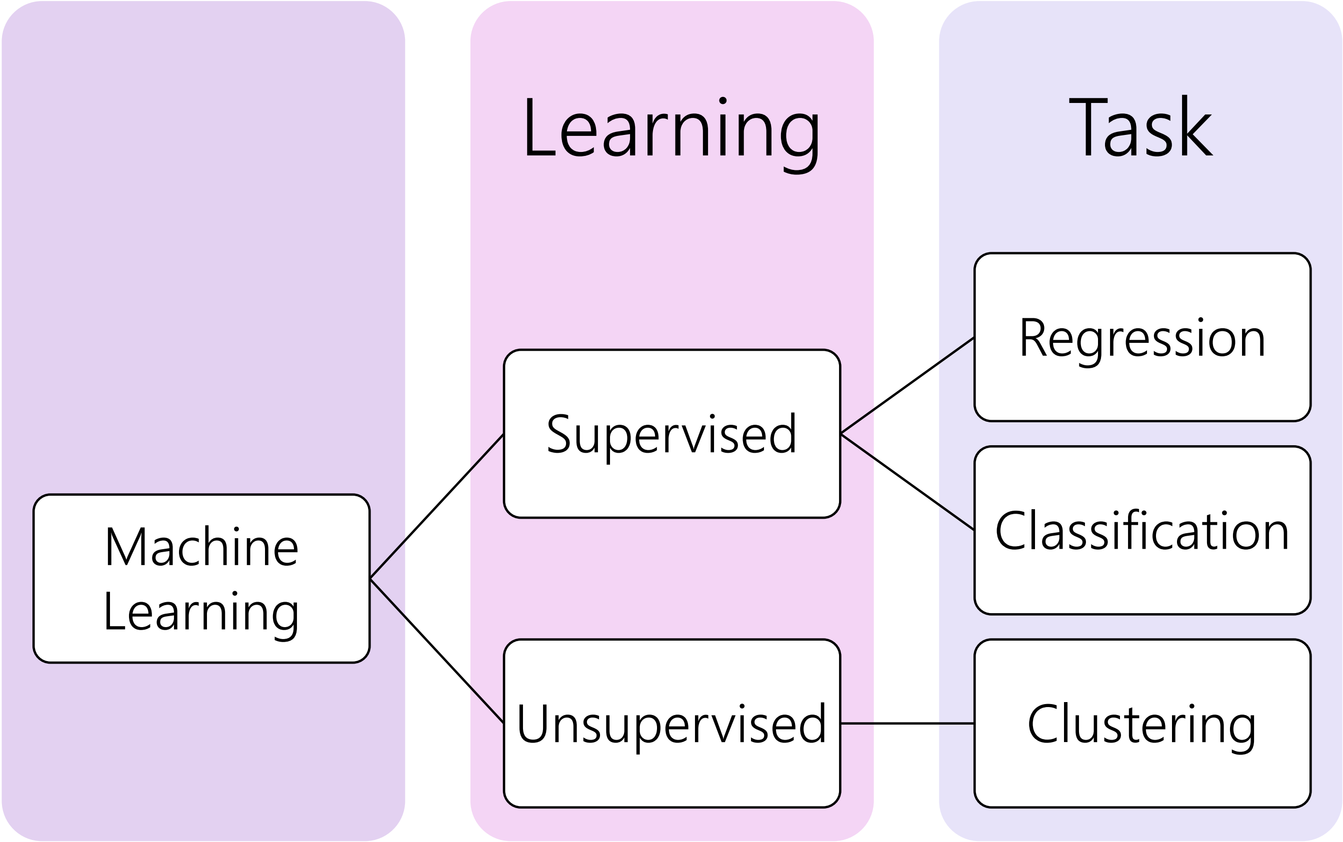
Forty-three radiographic features were selected initially as input to the neural networks. These features could be categorized into three groups: features related to masses (shape, size, margin, spiculation, pattern), features related to microcalcifications (number, shape, uniformity, distribution), and features of secondary abnormality (parenchymal distortion, skin thickening).

Figure 1 lists all of the 43 features we used for classification of benign and malignant mammographic patterns. (The term "density" was used to refer to radiographic opacity; the term "lucencies" was used to refer to areas of low opacity.) To illustrate the rating of features, we present two examples of mammograms with either a mass or a cluster of microcalcifications as the primary abnormality. Figure 2 shows an isolated abnormal density pattern (mass) of rounded shape and medium

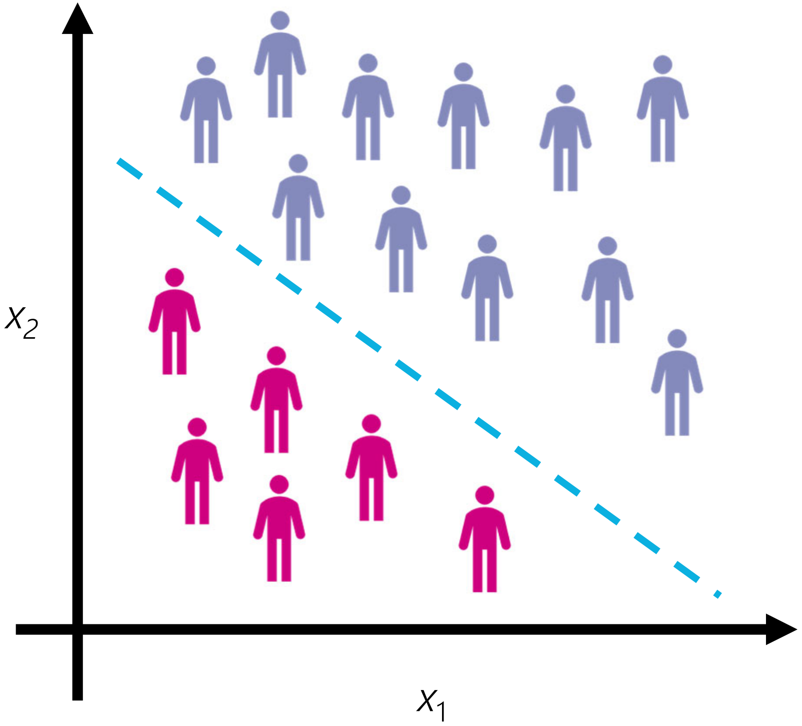
curve, ROC = receiver operating characteristic.

Marina Codari, PhD

Machine Learning



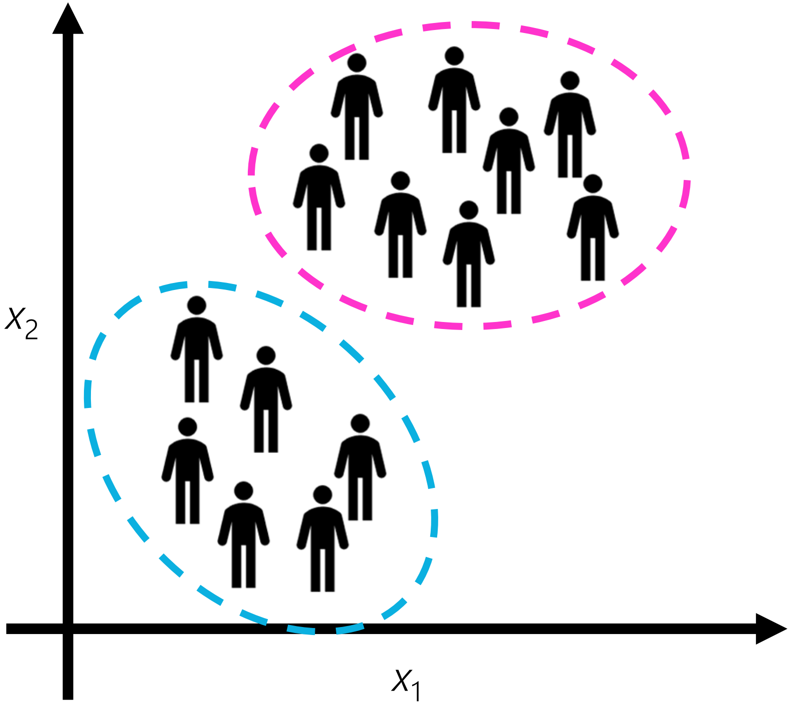
Supervised Learning



"Supervised learning involves gaining experience by using images of brain tumor examples that contain important information — specifically, "benign" and "malignant" labels — and applying the gained expertise to predict benign and malignant neoplasia on unseen new brain tumor images (test data)"

Erickson, Bradley J., et al. "Machine learning for medical imaging." Radiographics 37.2 (2017): 505-515.

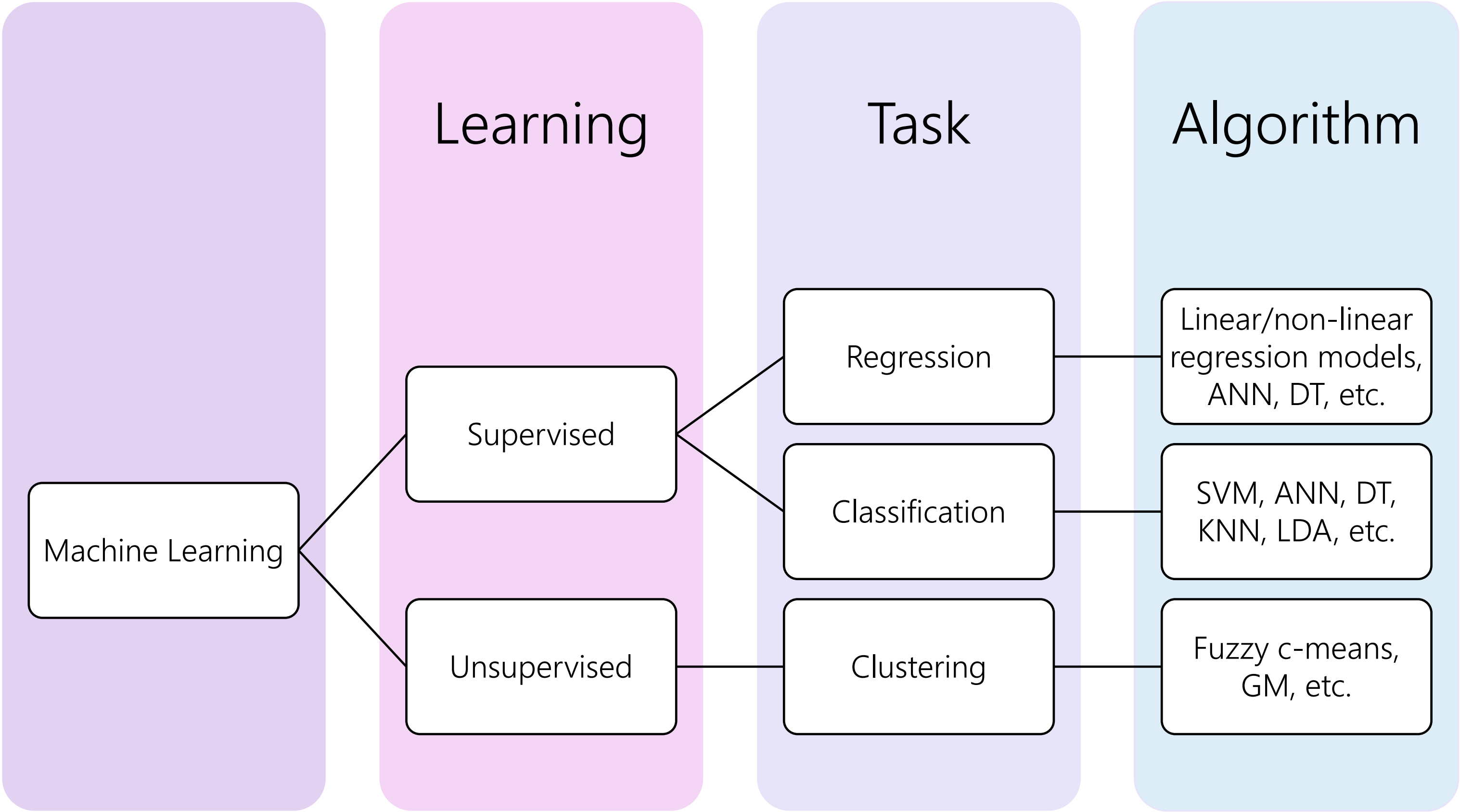
UNSUPERVISED LEARNING



"With unsupervised learning, data (eg, brain tumor images) are processed with a goal of separating the images into groups — for example, those depicting benign tumors and those depicting malignant tumors. The key difference is that this is done without the algorithm system being pro- vided with information regarding what the groups are"

Erickson, Bradley J., et al. "Machine learning for medical imaging." Radiographics 37.2 (2017): 505-515.

MACHINE LEARNING



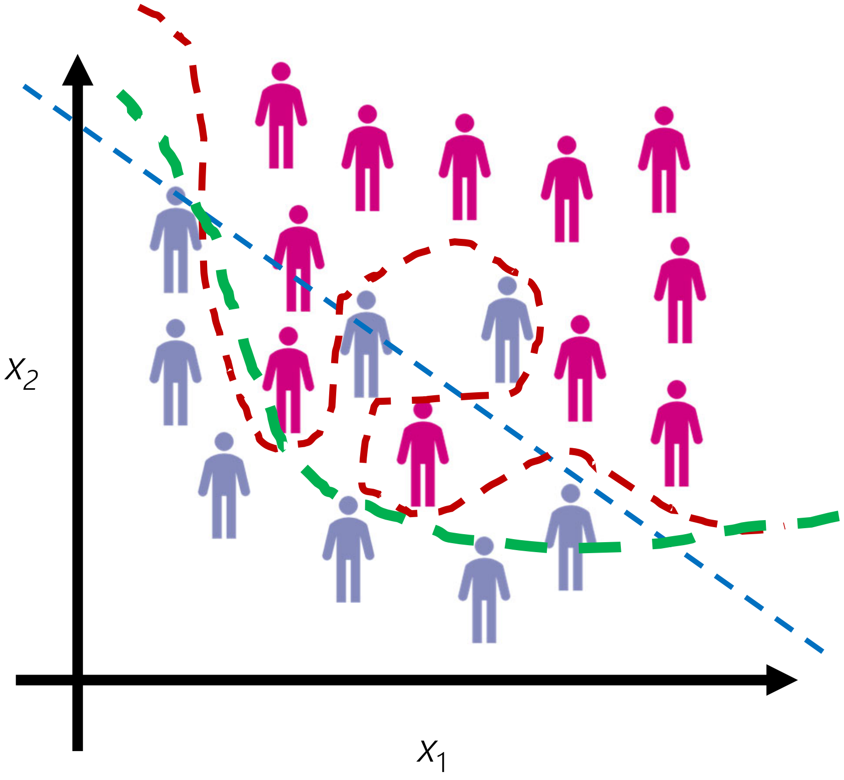
TRAINING

Inductive learning: to learn general models/concepts from specific examples.

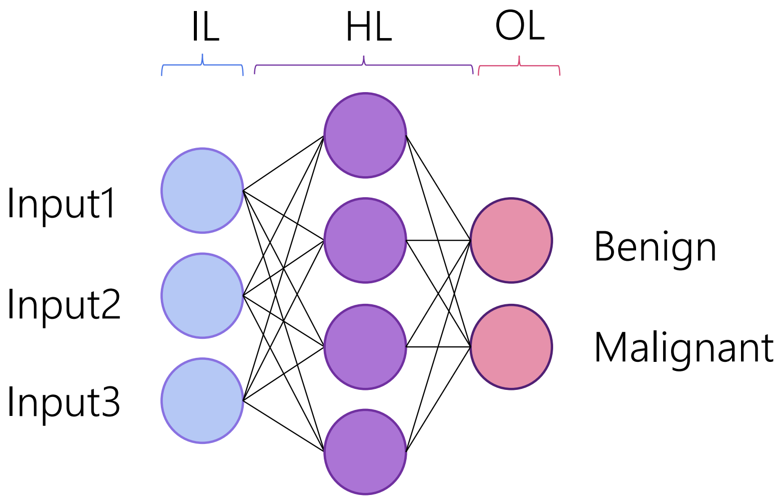
A good machine learning model must generalize well from the training data to any data from problem domain.

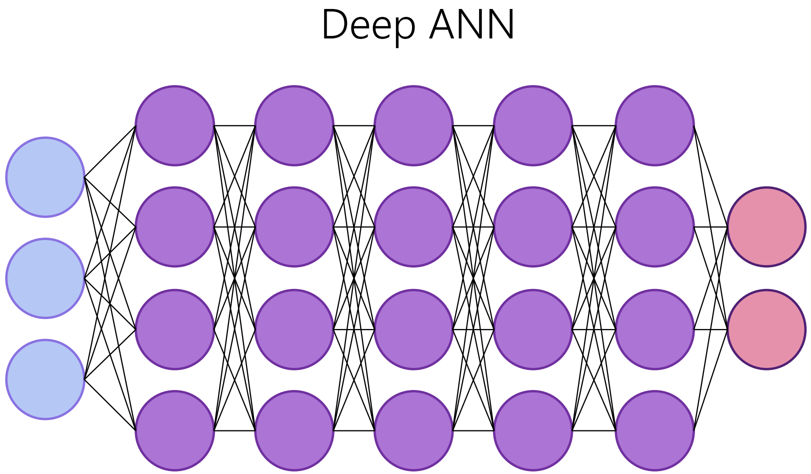
Frequent problems are:

- OVERFITTING: The model models the training data too well
- UNDERFITTING: The model can neither model the training data nor generalize to new data



Artificial Neural Network





Convolutional Neural Networks

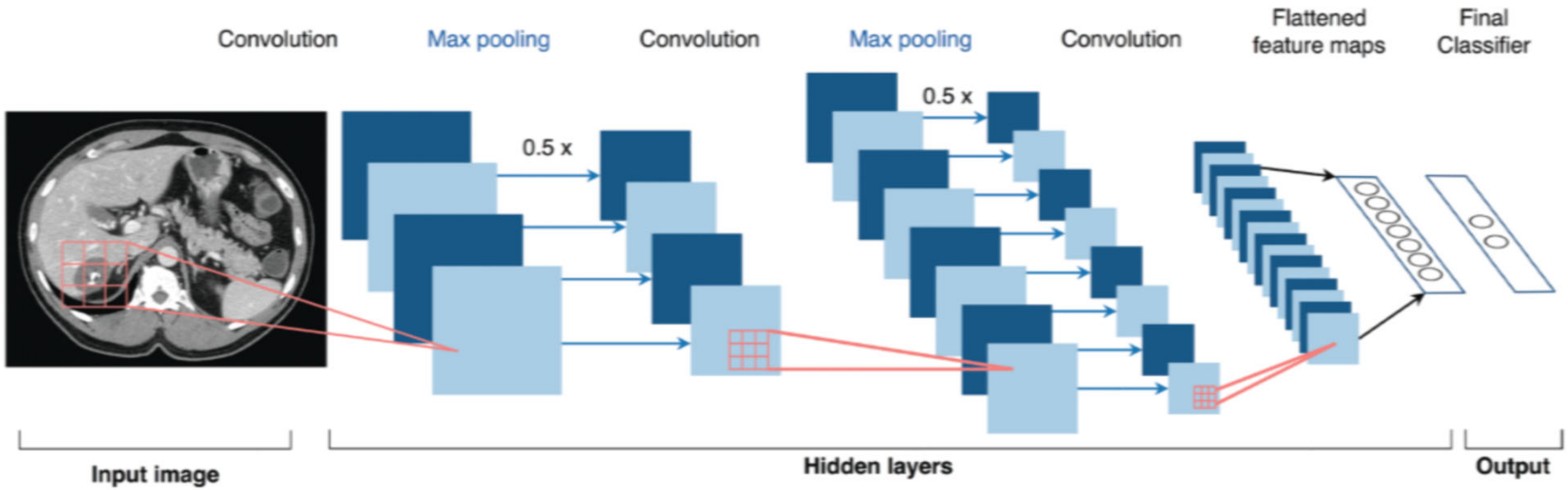
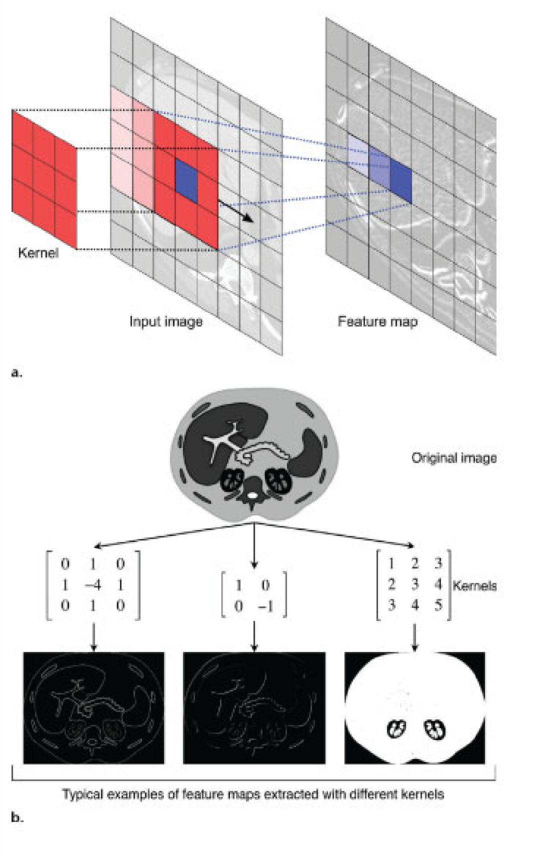
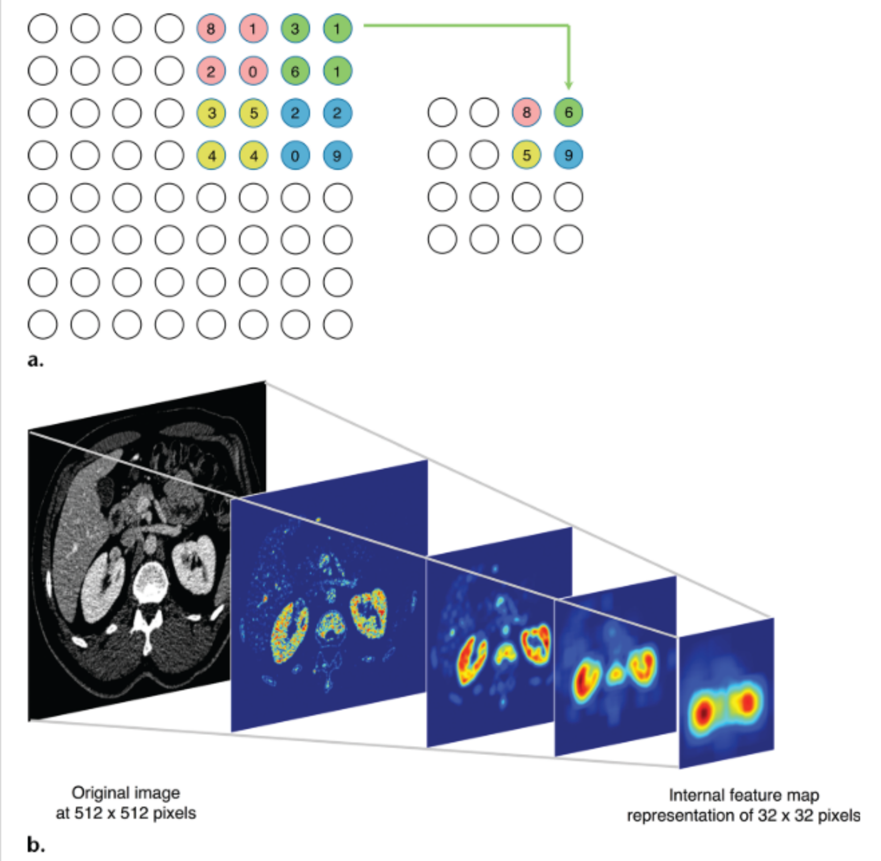


Figure 11. Integration of several concepts outlined in previous figures into a general diagram. Starting on the left, the input image is submitted to a series of convolutions with learned kernels, producing a stack of features maps containing low-level features such as edges and corners. These feature maps are then downsampled by a max pooling layer and further submitted to another set of learned convolutions, producing higher-level features such as parts of organs. Convolutions and max pooling layers can be stacked alternately until the network is deep enough to properly capture the structure of the image that is salient for the task at hand. Higher-level features are typically flattened into a single vector to perform the final classification or regression for the target task.

Chartrand, Gabriel, et al. "Deep learning: a primer for radiologists." Radiographics 37.7 (2017): 2113-2131.

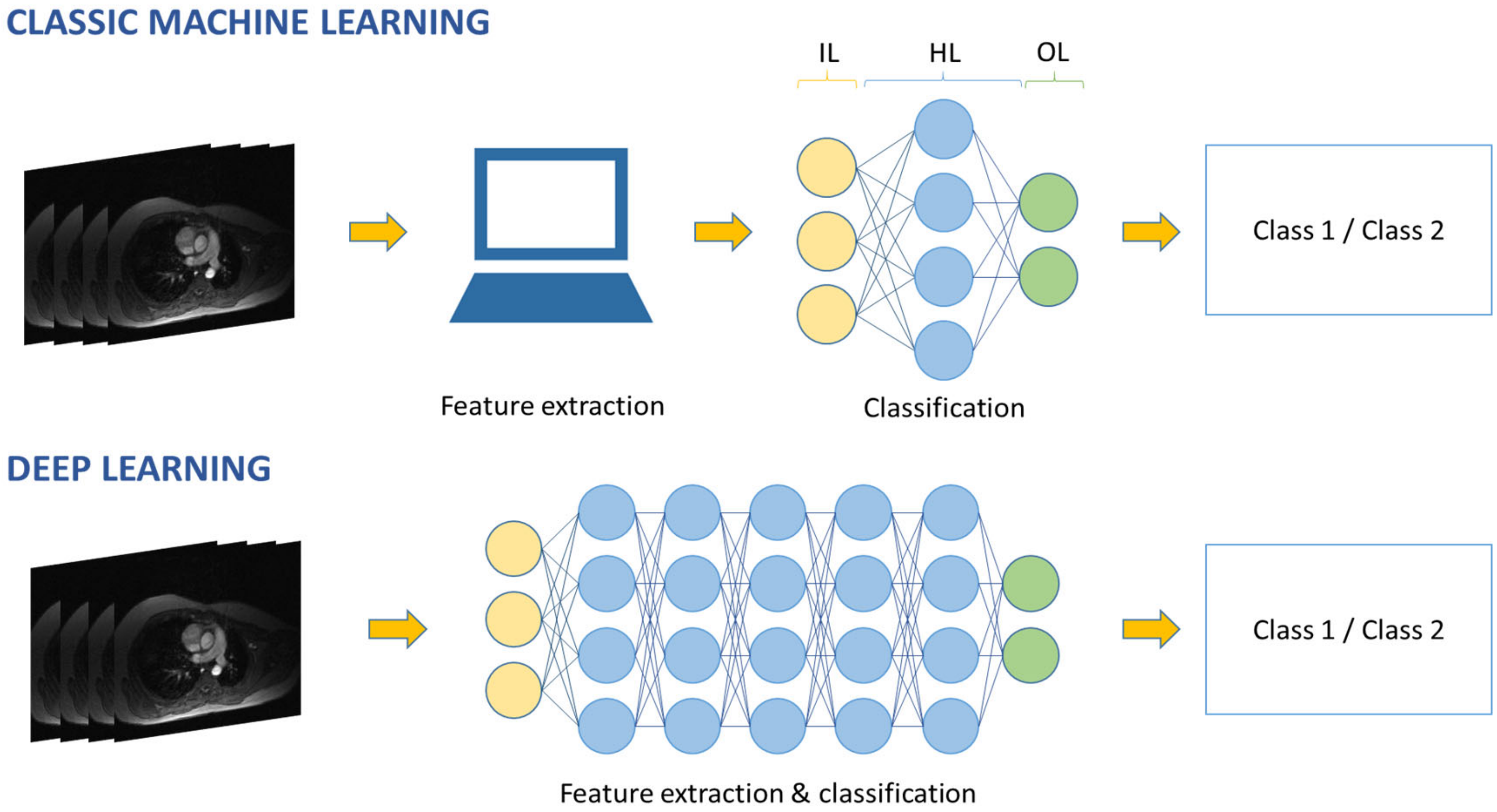
Convolutional Neural Network





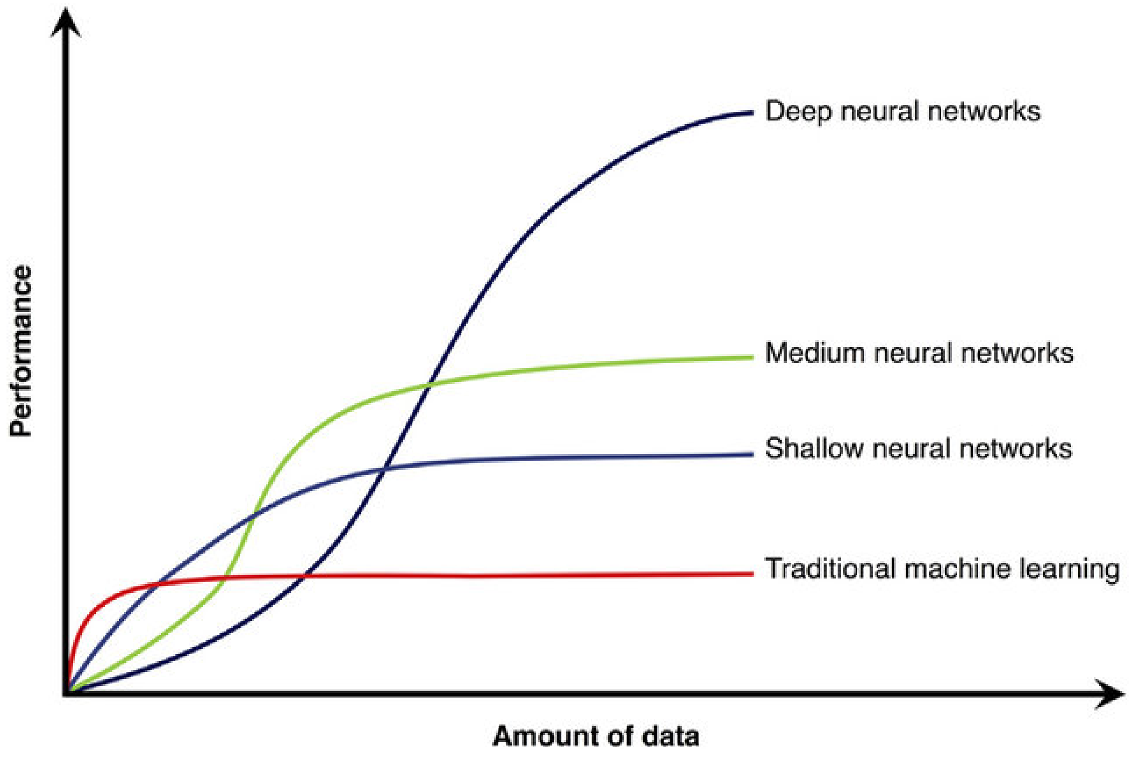
Images from: Chartrand, Gabriel, et al. "Deep learning: a primer for radiologists." Radio graphics 37.7 (2017): 2113-2131.

Deep Learning



Pesapane F, Codari M., and Sardanelli F. "Artificial intelligence in medical imaging: threat or opportunity? Radiologists again at the forefront of innovation in medicine." European radiology experimental 2.1 (2018): 35.

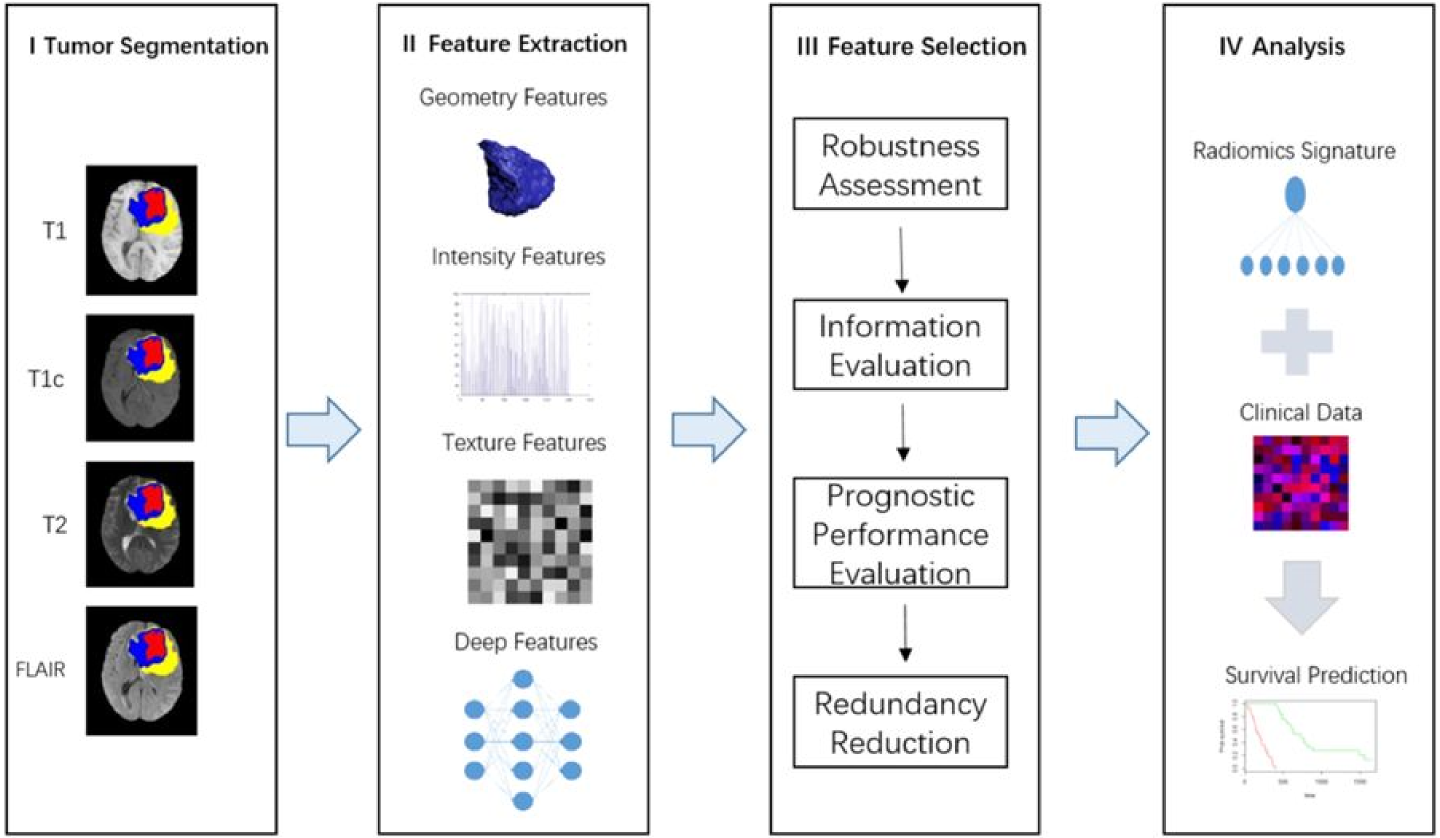
ARTIFICIAL INTELLIGENCE



Impact of sample size on performance of traditional machine learning algorithms (hand-crafted features) and neural networks with few (shallow), moderate (medium), or large (deep) numbers of layers.

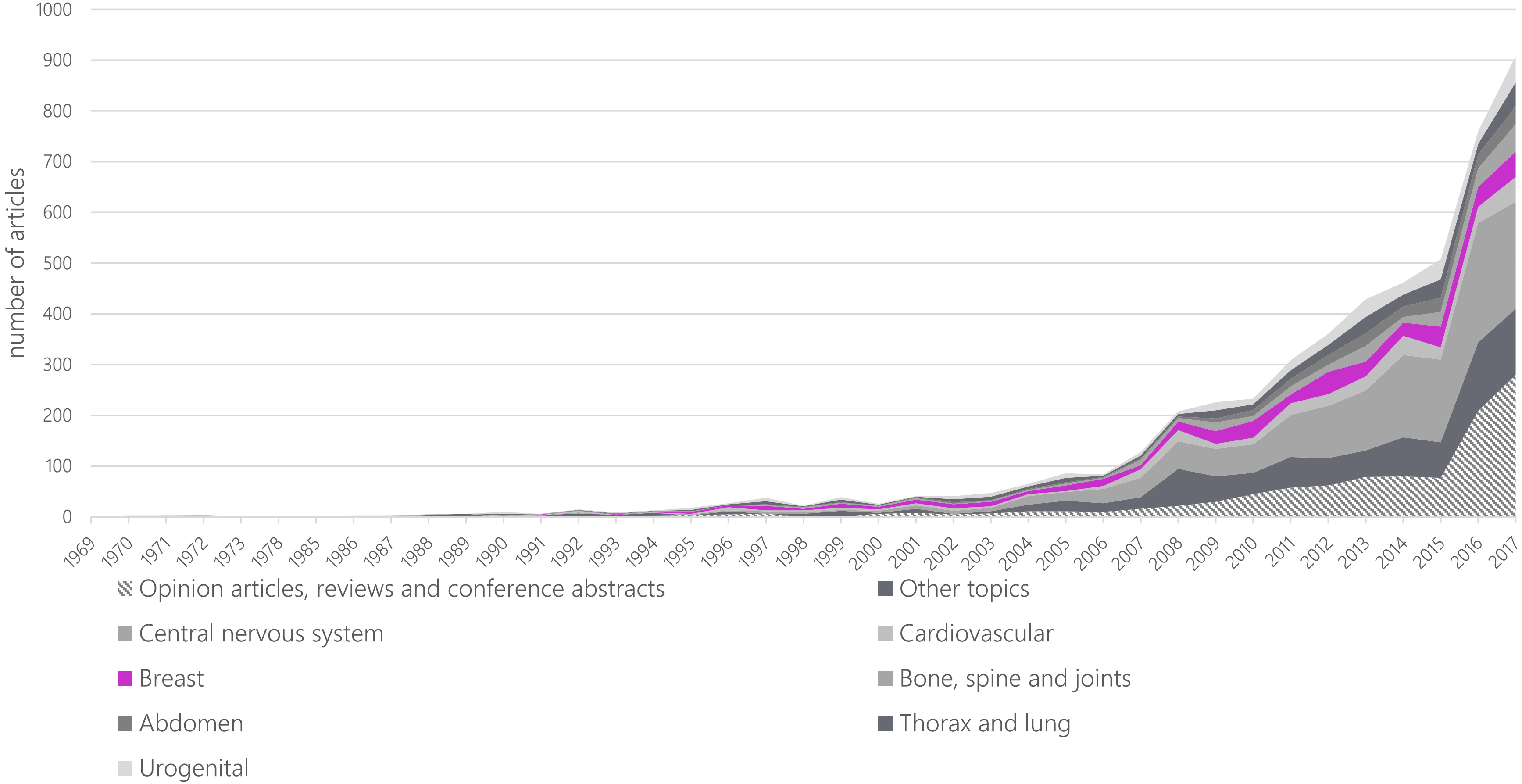
Tang, An, et al. "Canadian Association of Radiologists white paper on artificial intelligence in radiology." Canadian Association of Radiologists Journal (2018).

RADIOMICS



Lao, Jiangwei, et al. "A deep learning-based radiomics model for prediction of survival in glioblastoma multiforme." Scientific reports 7.1 (2017): 10353.

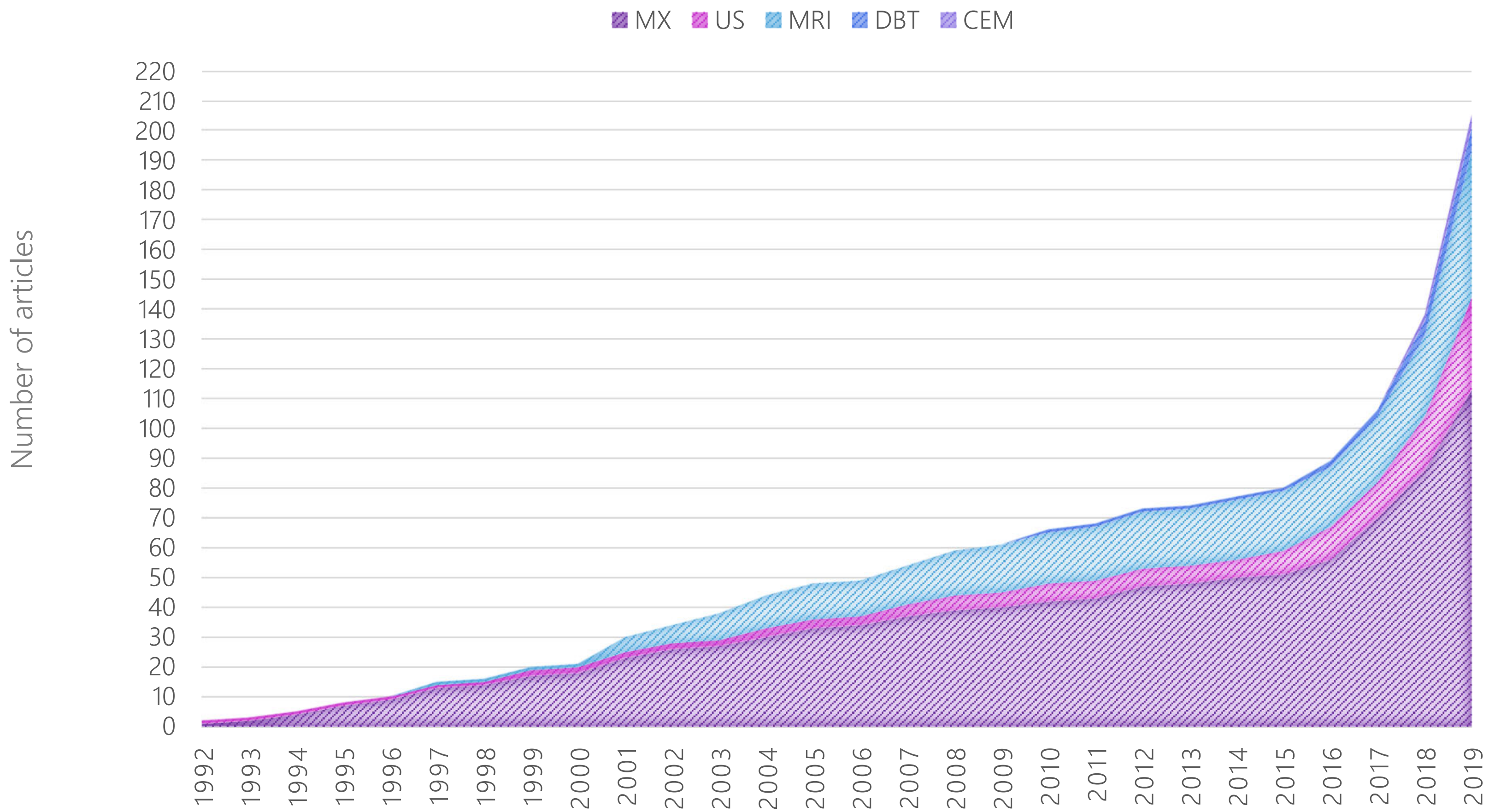
Machine Learning and Breast Imaging



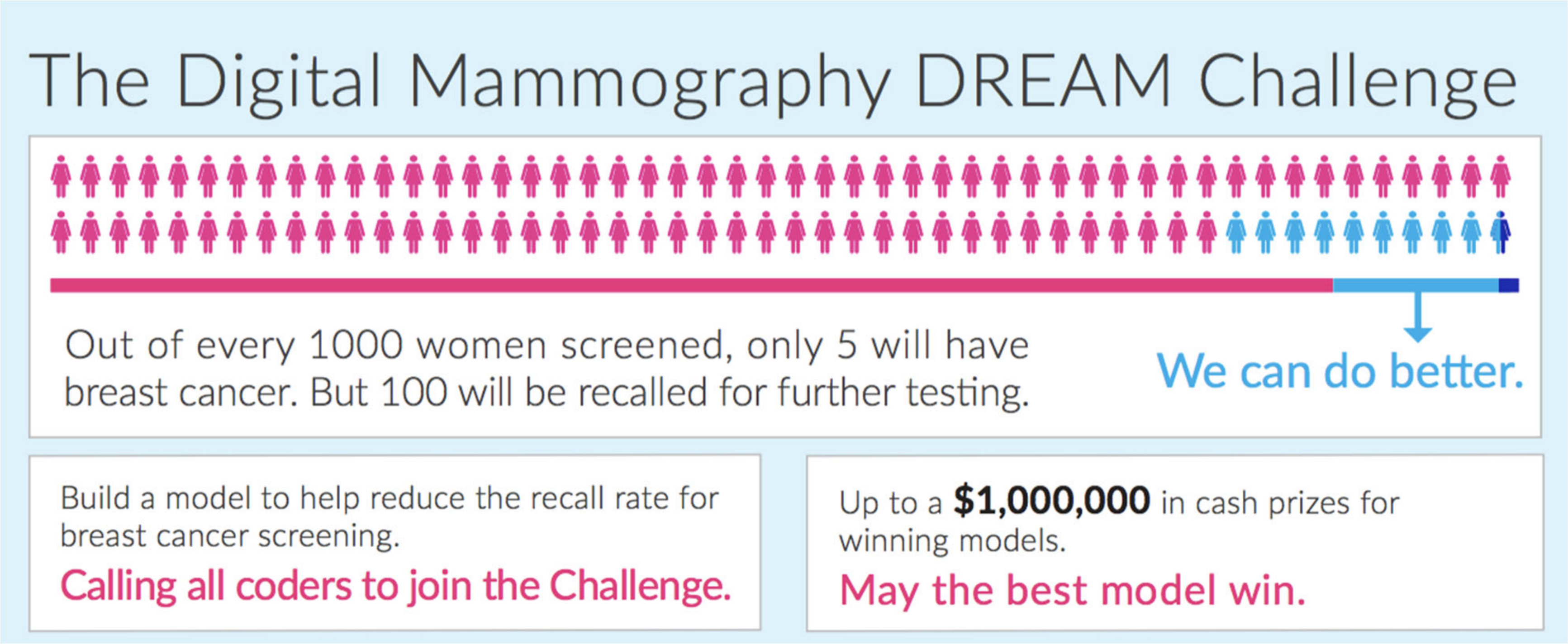
Adapted from: Pesapane F., Codari M. and Sardanelli F. Artificial intelligence in medical imaging: threat or opportunity? Radiologists again on the wavefront of innovation in medicine Eur Rad Exp, In pressc

Marina Codari, PhD

MACHINE LEARNING AND BREAST IMAGING



AI AND MAMMOGRAPHY



Challenge will include a Community Phase after the Competitive Phase, where top-performing teams will work together to further refine prediction algorithms that can ultimately be used in routine clinical practice.

AI and Mammography

640,000 de-identified digital mammography images (146,000 mammography exams, 86,000 women) + demographic, clinical and longitudinal data (Kaiser Permanente Washington)

+ Independent dataset with **15,000** images (3,200 exams and 1,400 women from Icahn School of Medicine at Mount Sinai)

1,150 coders

To determine the **cancer status of each breast of a subject (positive/negative)**

1. given only a screening digital mammography exam
2. given a screening exam + clinical/demographic information + previous screening exam(s).

Therapixel
(France)

First task: predictive accuracy of 80.3%, which was 5% percent more accurate than the runner up.

Second task: Tied first place
- Therapixel (Acc: 80.4%)
- Yuanfang Guan (Acc: 77.5%)

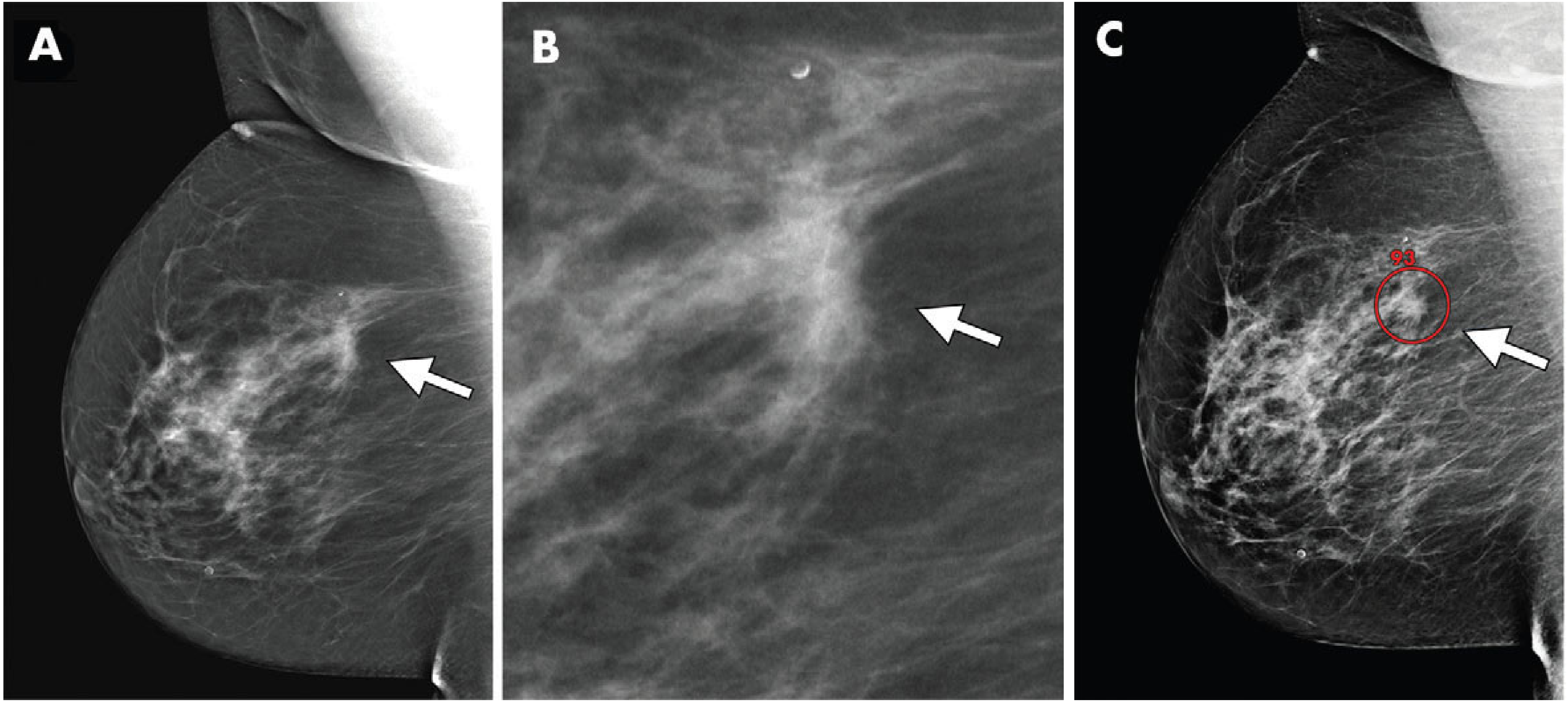
Both winning teams used **Deep Learning** approaches

https://www.ibm.com/blogs/research/2017/06/dream-challenge-results/

AI AND MAMMOGRAPHY

1. DEALING WITH AN **IMBALANCED DATASET** (Only 0.35% cancers)
2. PROCESSING **HIGH-RESOLUTION IMAGES**
3. **LARGE DIFFERENCE** IN APPEARANCE OF NORMAL BREAST TISSUE AMONG DIFFERENT VENDORS
4. **SEPARATE NN** FOR DETECTING MASSES AND CALCIFICATIONS

1. **INCREASE CANCER DETECTION** RATE AND REDUCE THE RECALL RATE
2. QUANTITATIVE AND REPRODUCIBLE **ASSESSMENT OF BREAST DENSITY** TO STRATIFY RISK FOR BREAST CANCER
3. RADIOMICS **TO IMPROVE TREATMENT AND PROGNOSIS**



Geras KJ, Mann RM, Moy L. Artificial Intelligence for Mammography and Digital Breast Tomosynthesis. Current Concept and Future Perspectives. Radiology (2019)

AI AND MAMMOGRAPHY (BEYOND BREAST CANCER)

"CARDIOVASCULAR DISEASE, OFTEN THOUGHT TO BE A "MALE" PROBLEM, IS **THE MAIN KILLER OF OLDER PEOPLE OF BOTH SEXES** ALMOST EVERYWHERE IN THE WORLD." WHO, 2018

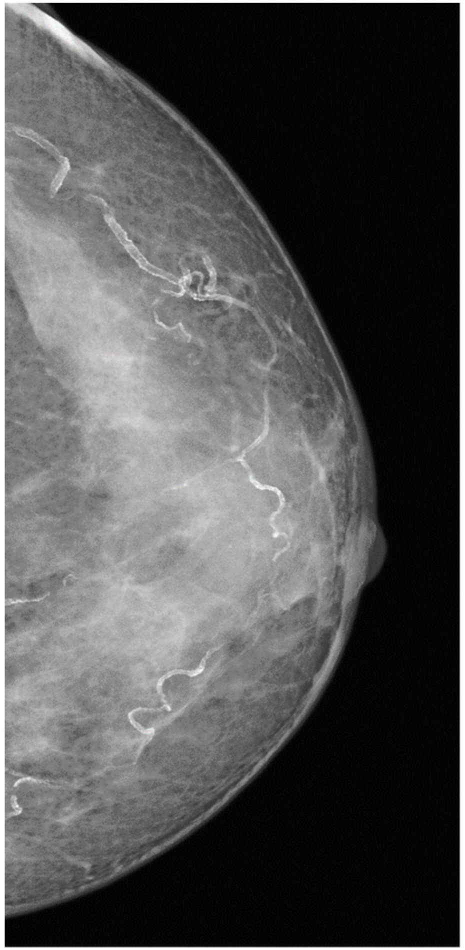
BREAST ARTERIAL CALCIFICATIONS, ARE ASSOCIATED WITH AN INCREASED RISK OF CARDIOVASCULAR DISEASE EVENTS.

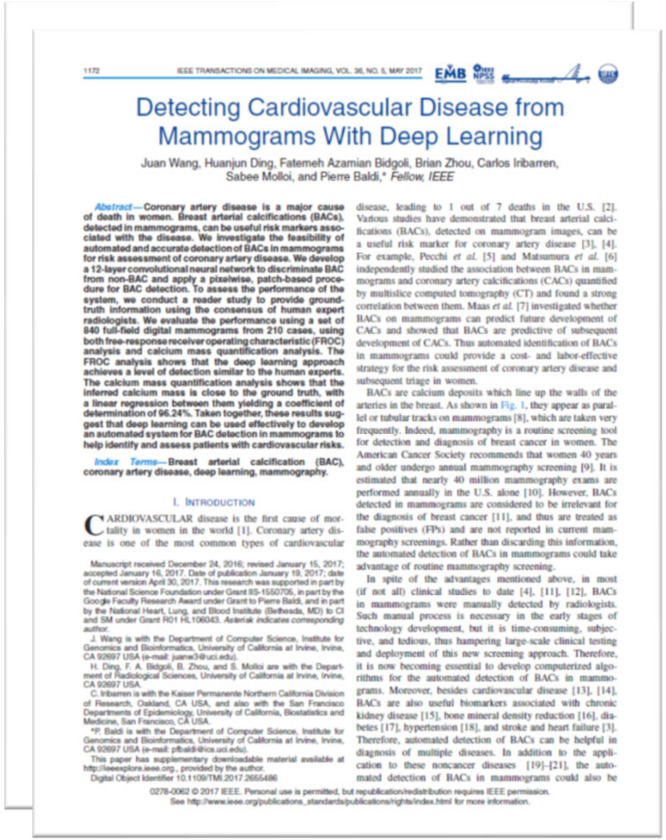
64 MILIONS UNDERGOING MAMMOGRAPHY

8 MILIONS HAVING BAC

2 DISEASES

1 SCREENING





1172 IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 36, NO. 5, MAY 2017

Detecting Cardiovascular Disease from Mammograms With Deep Learning

Juan Wang, Huanjun Ding, Fatemeh Azamian Bidgoli, Brian Zhou, Carlos Iribarren, Sabee Molloi, and Pierre Baldi,* *Fellow, IEEE*

***Abstract*—Coronary artery disease is a major cause of death in women. Breast arterial calcifications (BACs), detected in mammograms, can be useful risk markers associated with the disease. We investigate the feasibility of automated and accurate detection of BACs in mammograms for risk assessment of coronary artery disease. We develop a 12-layer convolutional neural network to discriminate BAC from non-BAC and apply a pixelwise, patch-based procedure for BAC detection. To assess the performance of the system, we conduct a reader study to provide ground-truth information using the consensus of human expert radiologists. We evaluate the performance using a set of 840 full-field digital mammograms from 210 cases, using both free-response receiver operating characteristic (FROC) analysis and calcium mass quantification analysis. The FROC analysis shows that the deep learning approach achieves a level of detection similar to the human experts. The calcium mass quantification analysis shows that the inferred calcium mass is close to the ground truth, with a linear regression between them yielding a coefficient of determination of 96.24%. Taken together, these results suggest that deep learning can be used effectively to develop an automated system for BAC detection in mammograms to help identify and assess patients with cardiovascular risks.**

***Index Terms*—Breast arterial calcification (BAC), coronary artery disease, deep learning, mammography.**

I. INTRODUCTION

CARDIOVASCULAR disease is the first cause of mortality in women in the world [1]. Coronary artery disease is one of the most common types of cardiovascular disease, leading to 1 out of 7 deaths in the U.S. [2]. Various studies have demonstrated that breast arterial calcifications (BACs), detected on mammogram images, can be a useful risk marker for coronary artery disease [3], [4]. For example, Pecchi *et al.* [5] and Matsumura *et al.* [6] independently studied the association between BACs in mammograms and coronary artery calcifications (CACs) quantified by multislice computed tomography (CT) and found a strong correlation between them. Maas *et al.* [7] investigated whether BACs on mammograms can predict future development of CACs and showed that BACs are predictive of subsequent development of CACs. Thus automated identification of BACs in mammograms could provide a cost- and labor-effective strategy for the risk assessment of coronary artery disease and subsequent triage in women.

BACs are calcium deposits which line up the walls of the arteries in the breast. As shown in Fig. 1, they appear as parallel or tubular tracks on mammograms [8], which are taken very frequently. Indeed, mammography is a routine screening tool for detection and diagnosis of breast cancer in women. The American Cancer Society recommends that women 40 years and older undergo annual mammography screening [9]. It is estimated that nearly 40 million mammography exams are performed annually in the U.S. alone [10]. However, BACs detected in mammograms are considered to be irrelevant for the diagnosis of breast cancer [11], and thus are treated as false positives (FPs) and are not reported in current mammography screenings. Rather than discarding this information, the automated detection of BACs in mammograms could take advantage of routine mammography screening.

In spite of the advantages mentioned above, in most (if not all) clinical studies to date [4], [11], [12], BACs in mammograms were manually detected by radiologists. Such manual process is necessary in the early stages of technology development, but it is time-consuming, subjective, and tedious, thus hampering large-scale clinical testing and deployment of this new screening approach. Therefore, it is now becoming essential to develop computerized algorithms for the automated detection of BACs in mammograms. Moreover, besides cardiovascular disease [13], [14], BACs are also useful biomarkers associated with chronic kidney disease [15], bone mineral density reduction [16], diabetes [17], hypertension [18], and stroke and heart failure [3]. Therefore, automated detection of BACs can be helpful in diagnosis of multiple diseases. In addition to the application to these noncancer diseases [19]–[21], the automated detection of BACs in mammograms could also be

Manuscript received December 24, 2016; revised January 15, 2017; accepted January 16, 2017. Date of publication January 19, 2017; date of current version April 30, 2017. This research was supported in part by the National Science Foundation under Grant IIS-1550705, in part by the Google Faculty Research Award under Grant to Pierre Baldi, and in part by the National Heart, Lung, and Blood Institute (Bethesda, MD) to CI and SM under Grant R01 HL106043. *Asterisk indicates corresponding author.*

J. Wang is with the Department of Computer Science, Institute for Genomics and Bioinformatics, University of California at Irvine, Irvine, CA 92697 USA (e-mail: juanw3@uci.edu).

H. Ding, F. A. Bidgoli, B. Zhou, and S. Molloi are with the Department of Radiological Sciences, University of California at Irvine, Irvine, CA 92697 USA.

C. Iribarren is with the Kaiser Permanente Northern California Division of Research, Oakland, CA USA, and also with the San Francisco Departments of Epidemiology, University of California, Biostatistics and Medicine, San Francisco, CA USA.

*P. Baldi is with the Department of Computer Science, Institute for Genomics and Bioinformatics, University of California at Irvine, Irvine, CA 92697 USA (e-mail: pfbaldi@ics.uci.edu).

This paper has supplementary downloadable material available at http://ieeexplore.ieee.org., provided by the author.

Digital Object Identifier 10.1109/TMI.2017.2655486

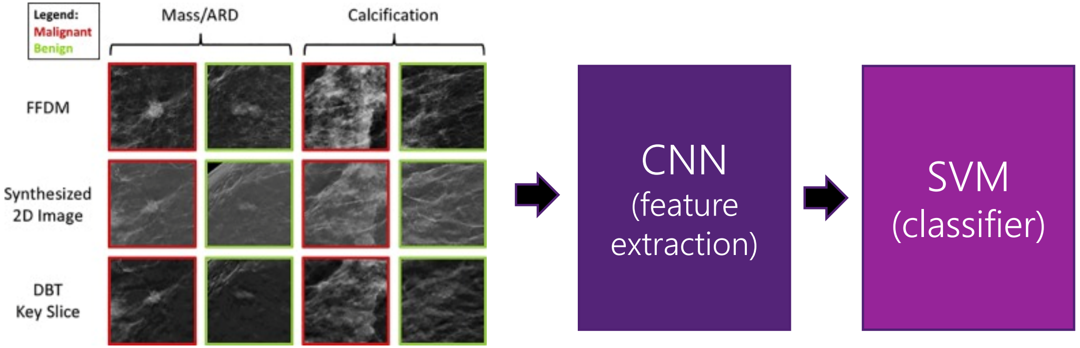
AI and Digital Breast Tomosynthesis (DBT)

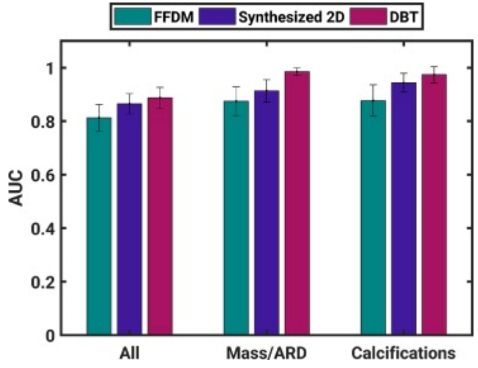
- Helps in the detection of masses and in women with increased breast density

- Increases reading times (50% - 200%)
- Limited resolution
- Limited amount of data
- Difference among vendors

Lesion detection using DBT images with conventional radiomic methods yielded promising results.

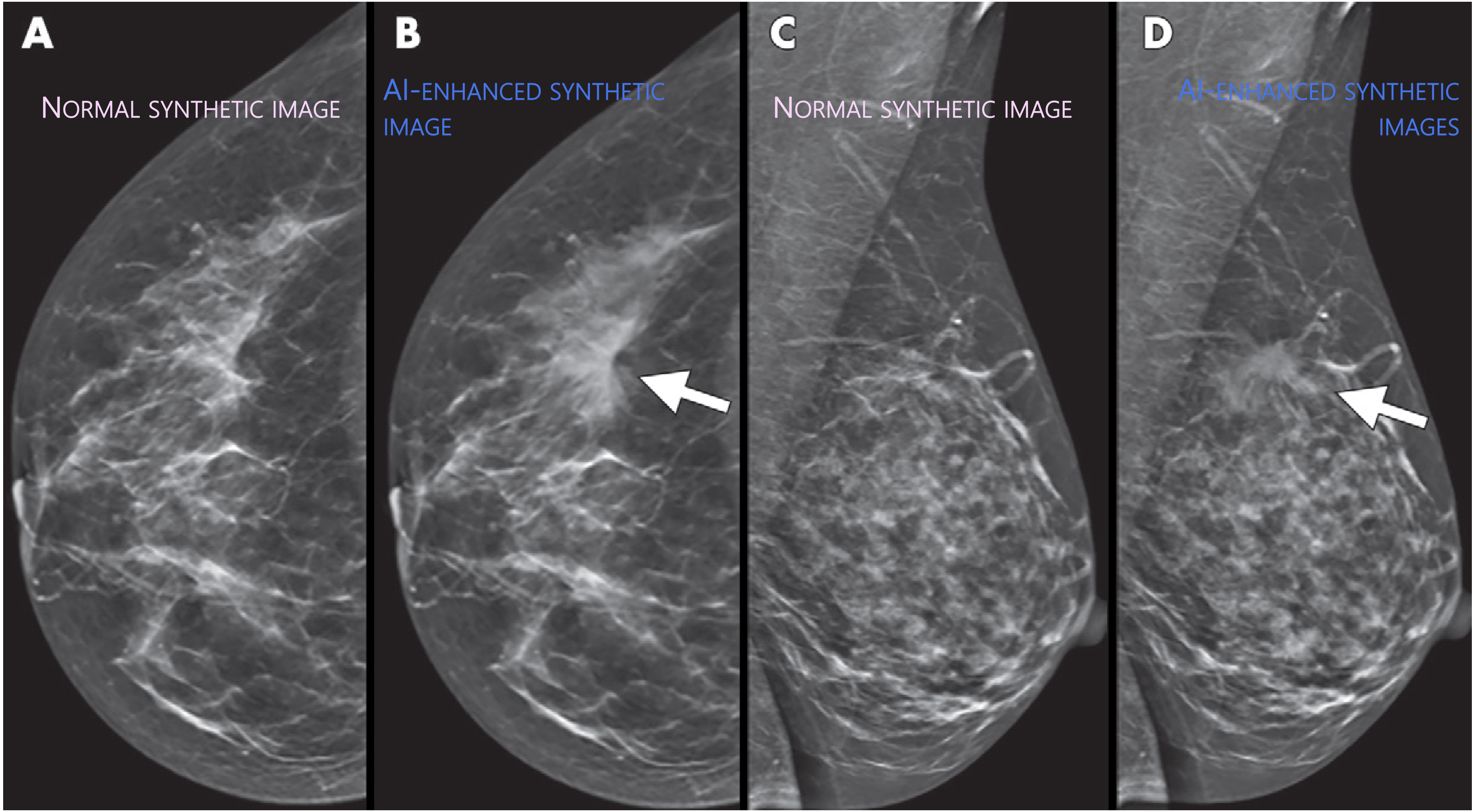
- Reducing radiation dose
- Enhance of synthetic images
- Enhance the conspicuity of calcifications
- Improve lesion classification (3D)
- Remove normal fibroglandular tissue





Classification performance of the merged-view classifier on each subset of lesions considered in this study. AUC is plotted with error bars showing one standard error.

Adapted from: Mendel, Kayla, et al. "Transfer learning from convolutional neural networks for computer-aided diagnosis: a comparison of digital breast tomosynthesis and full-field digital mammography." Academic radiology 26.6 (2019): 735-743.



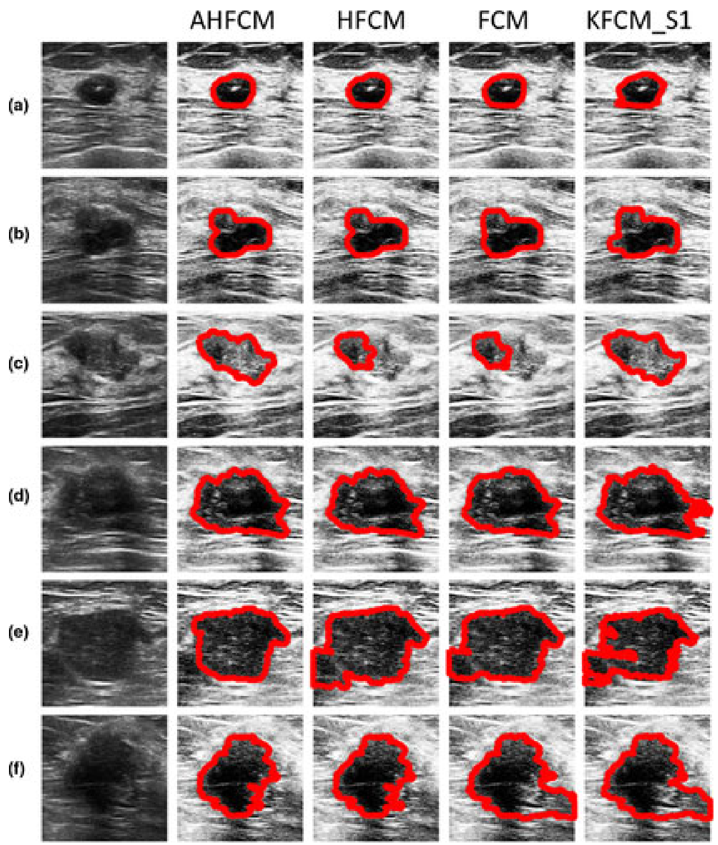
Geras KJ, Mann RM, Moy L. Artificial Intelligence for Mammography and Digital Breast Tomosynthesis. Current Concept and Future Perspectives. Radiology (2019)

AI and Breast Ultrasounds

✓ High availability, cost-effectiveness, acceptable diagnostic performance, non invasive, real-time

✗ Operator dependency, speckle noise, low contrast, and blurred boundaries

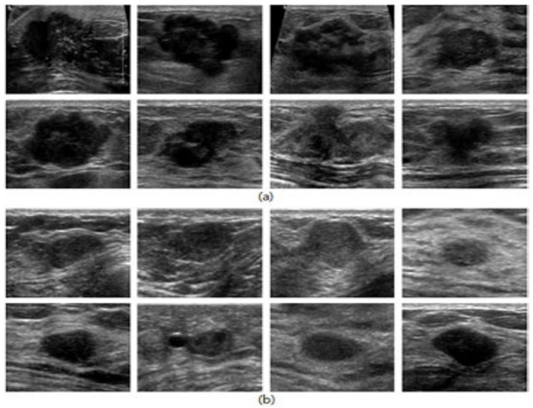
- AI has been increasingly applied in breast US and proved to be a powerful tool to provide a reliable diagnosis with higher accuracy and efficiency and reduce the workload of physicians.
- Great progress has been made in processing and segmentation of US breast images



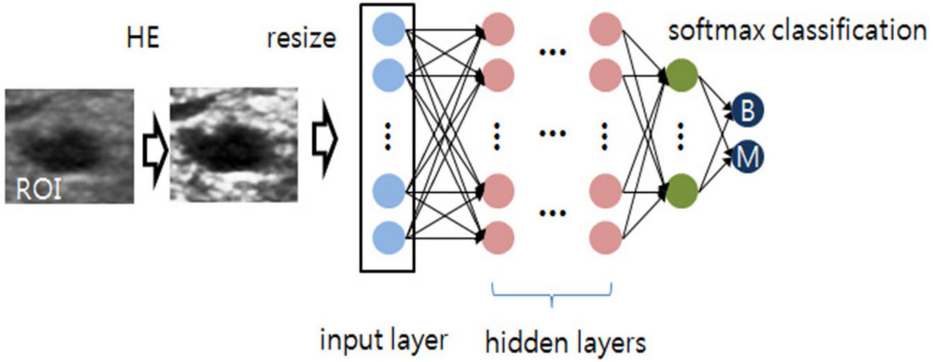
Feng, Yuan, et al. "An adaptive fuzzy C-means method utilizing neighboring information for breast tumor segmentation in ultrasound images." Medical physics 44.7 (2017): 3752-3760.

AI and Breast Ultrasounds

- Great progress has been made in Lesion classification in US breast images
- Zhang et al. established a DL architecture that could automatically classify benign and malignant breast tumors from shear-wave elastography (accuracy of 93.4%, a sensitivity of 88.6%, specificity of 97.1%, and AUC of 0.947.
- Han et al. used CNN DL framework to classify benign and malignant lesions on breast images acquired by ultrasound (accuracy 0.91, sensitivity of 0.86, specificity of 0.96 and AUC of 0,90)



BIRADS distribution of (a) training data (7408 images) and (b) test data (829 images).

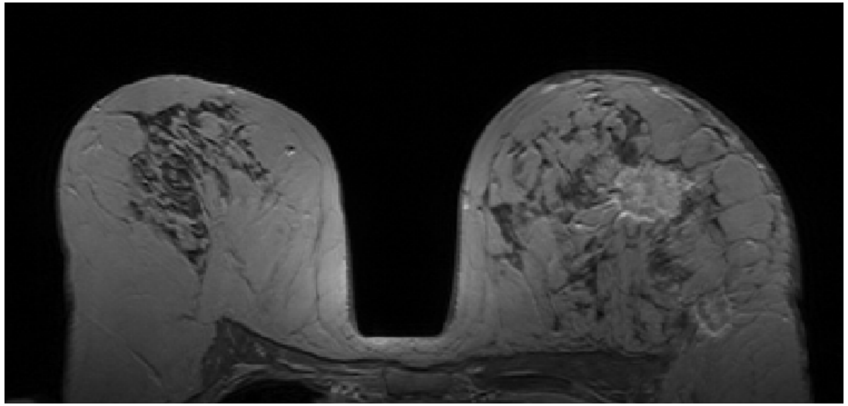


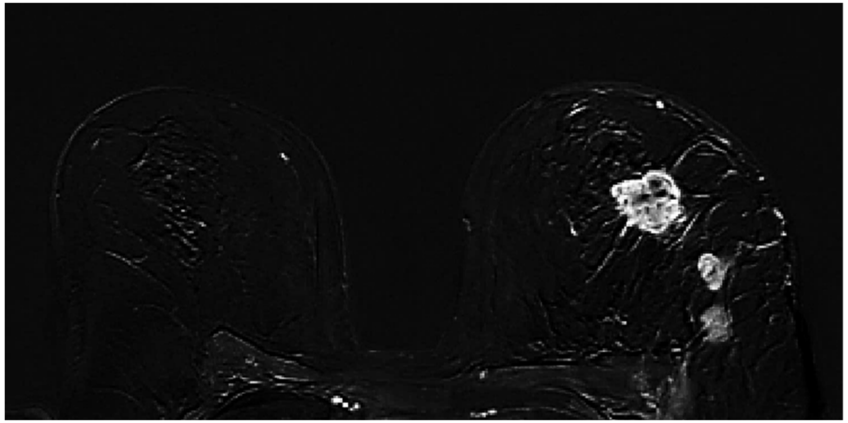
Han, Seokmin, et al. "A deep learning framework for supporting the classification of breast lesions in ultrasound images." Physics in Medicine & Biology 62.19 (2017): 7714.

AI and Breast MRI

Breast MRI data well fits to DL application

- Morphologic / Spatial
- Dynamic / Temporal
- Huge amount of data for a single patient





Study Design

87% have retrospective design

B_0 Strength

56% using 1.5 T, 18% using 3 T, 20% both

MRI protocol

74% DCE only

12% DCE and (DWI or MRS or T2-w)

Codari M., Schiaffino S, Sardanelli F and Trimboli RM.
Artificial Intelligence for Breast MRI 2008-2018: A Systematic Mapping Review. AJR (2019)

STUDY DESIGN

STUDY DESIGN

87% HAVE RETROSPECTIVE DESIGN

B_0 STRENGTH

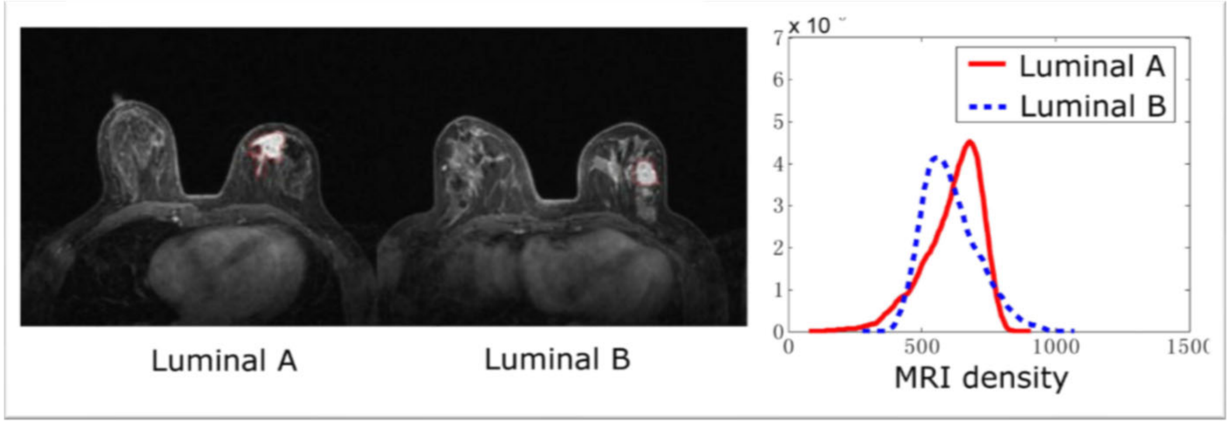
56% USING 1.5 T, 18% USING 3 T AND 20% USING BOTH

MRI PROTOCOL

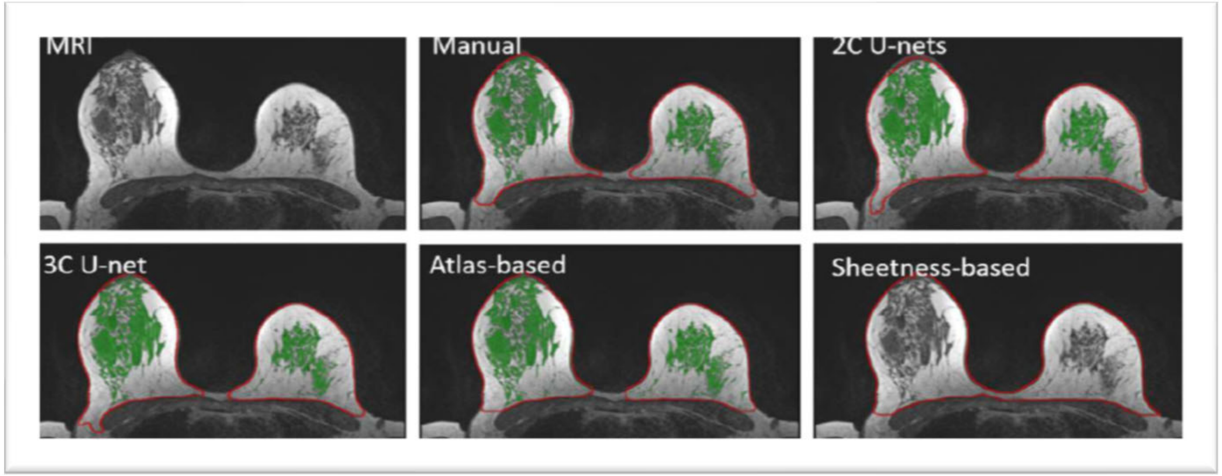
- 74% DCE ONLY
- 12% DCE AND (DWI OR MRS OR T2-W)
- 5% DIXON'S METHOD
- 5% DWI

Addressed Aim

- Image processing
- Lesion classification
- Prognostic imaging
- Response to NAT

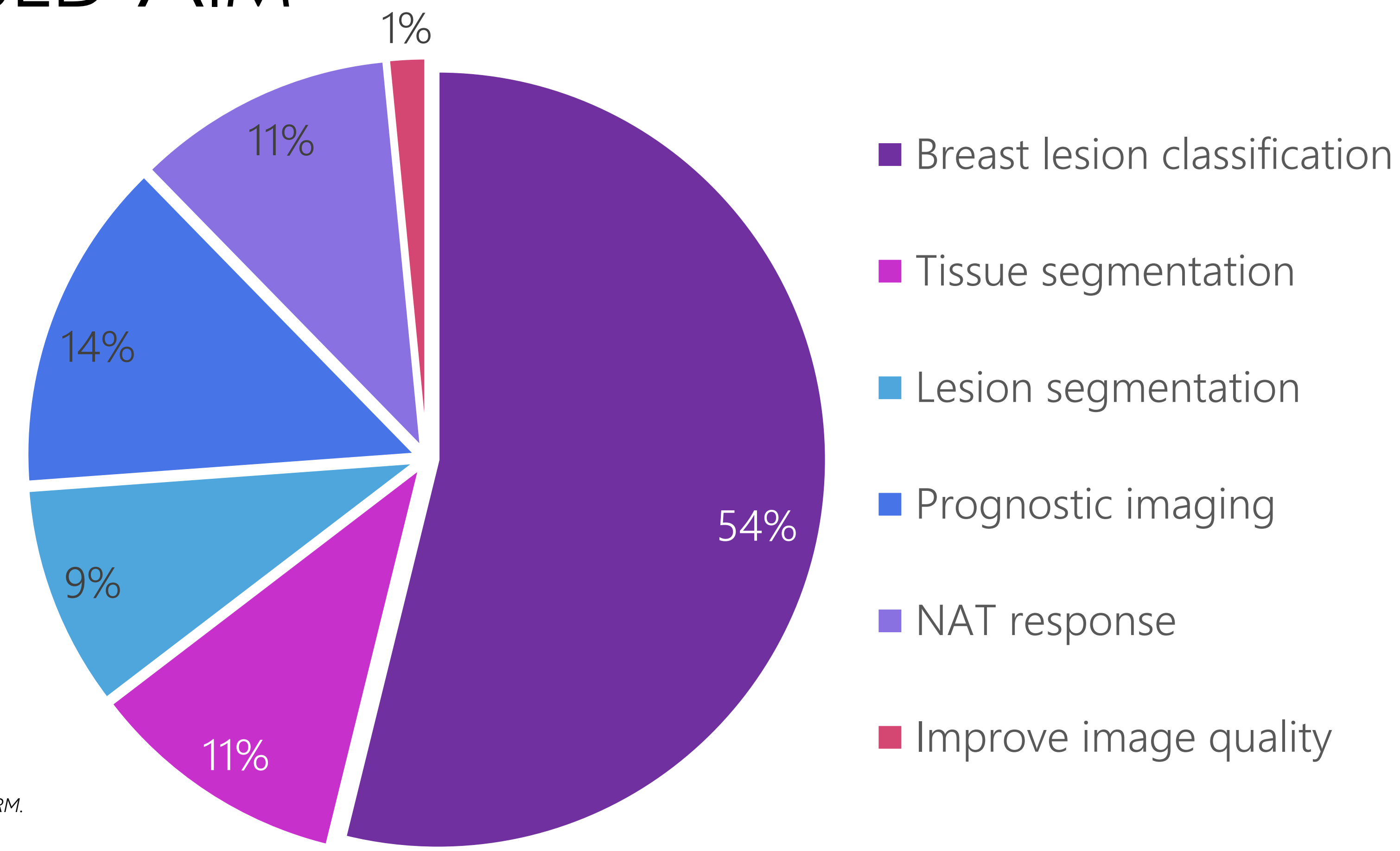


Fan, et al, PlosONE (2017)



Dalmış MU, et al., Medical physics (2017)

Addressed Aim

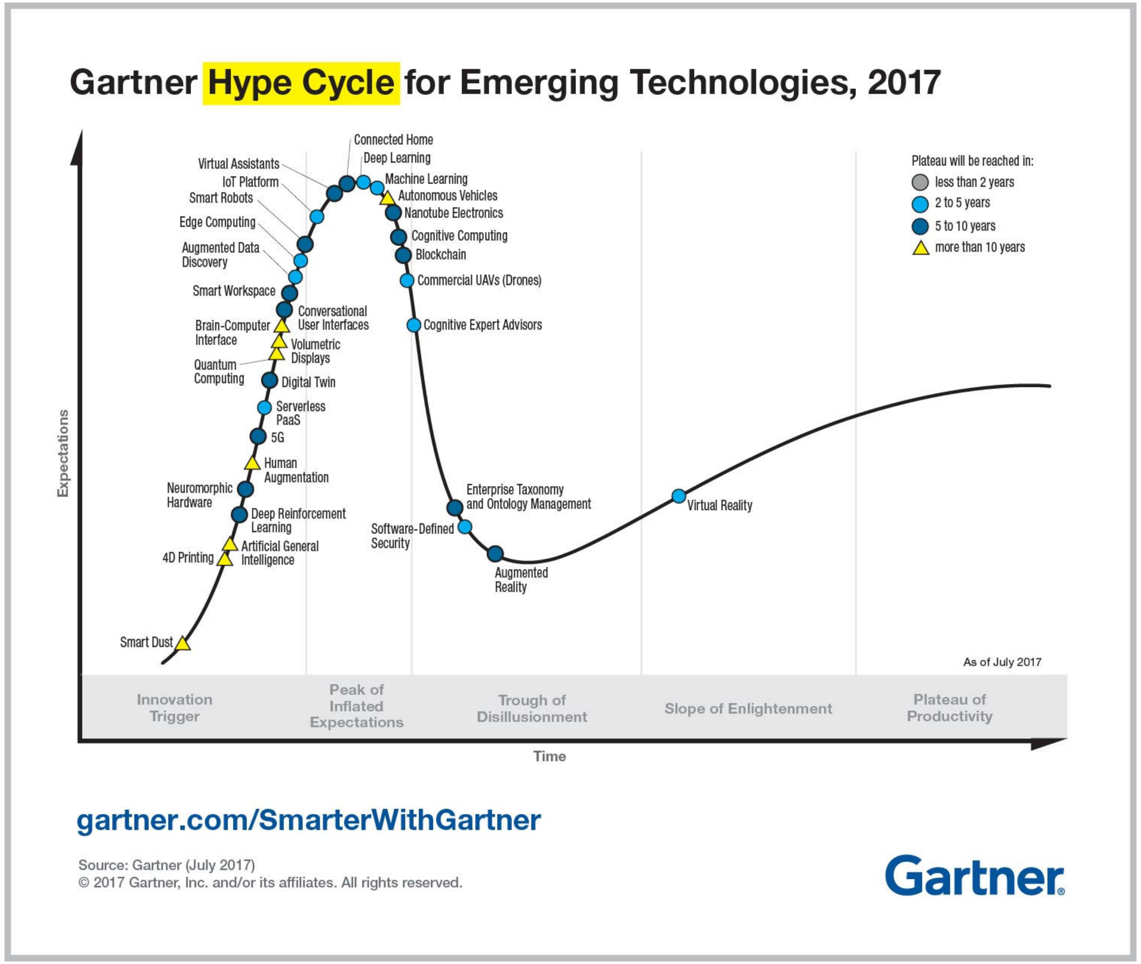


Codari M., Schiaffino S, Sardanelli F and Trimboli RM. Artificial Intelligence for Breast MRI 2008-2018: A Systematic Mapping Review. AJR, In press

Take home message

- Breast imaging represents a fertile ground for AI application
- AI has the **potential to improve it**
- **Lesion Classification** and **Image Processing** are the current focus
- **Close Partnership** between Clinical and Data scientists will be the key of success

What to expect



marina.codari@polimi.it