Deep Learning and Applications in Medical Imaging:
Role of deep learning at various stages of quantitative image analysis (radiomics) for cancer assessment

Maryellen Giger, PhD
A. N. Pritzker Professor of Radiology / Medical Physics
The University of Chicago
m-giger@uchicago.edu

Supported in parts by various NIH grants CA 195564, CA 166945, and CA 189240; and The University of Chicago CTSA UL1 TR000430 pilot awards.

COI: M.L.G. is a stockholder in R2/Hologic, co-founder and equity holder in Quantitative Insights, shareholder in Qview, and receives royalties from Hologic, GE Medical Systems, MEDIAN Technologies, Riverain Medical, Mitsubishi, and Toshiba. It is the University of Chicago Conflict of Interest Policy that investigators disclose publicly actual or potential significant financial interest that would reasonably appear to be directly and significantly affected by the research activities

Role of deep learning at various stages of quantitative image analysis (radiomics) for disease assessment

- Applications in breast image analysis
  - Computer-aided detection (CADe)
  - Computer-aided diagnosis (CADx)
  - Risk Assessment
  - Response to neoadjuvant therapy
- Methods to handle limitations and potential pitfalls
  - Pre-processing
  - Transfer learning
  - Fine tuning
  - Data Augmentation

Deep Learning in Precision Medicine & Imaging

 → Deep Learning → Answer to Clinical Question

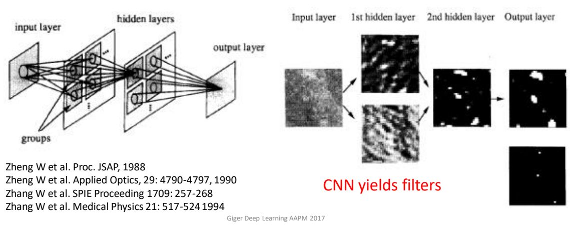
Need to consider:
- Cautious of "Garbage in, Garbage out"
  - Issue of Robustness
- There are multiple implementations of "Deep Learning" (e.g., CNNs)
  - Filtering
  - Classifier
  - Feature Extraction
  - Segmentation

Role of deep learning at various stages of quantitative image analysis (radiomics) for disease assessment

- Applications in breast image analysis
  - Computer-aided detection (CADe)
  - Computer-aided diagnosis (CADx)
  - Risk Assessment
  - Response to neoadjuvant therapy
- Methods to handle limitations and potential pitfalls
  - Pre-processing
  - Transfer learning
  - Fine tuning
  - Data Augmentation

Deep learning example in CADe

Shift-Invariant Artificial Neural Network (SIANN) for CADe in Mammography, Zhang W, Doi K, Giger ML, Wu Y, Nishikawa RM, Schmidt RA. Medical Physics 21: 517-524, 1994



Zheng W et al. Proc. JSAP, 1988
Zheng W et al. Applied Optics, 29: 4790-4797, 1990
Zhang W et al. SPIE Proceeding 1709: 257-268
Zhang W et al. Medical Physics 21: 517-524 1994

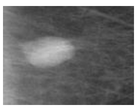
CNN yields filters

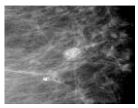
Role of deep learning at various stages of quantitative image analysis (radiomics) for disease assessment

- Applications in breast image analysis
  - Computer-aided detection (CADe)
  - Computer-aided diagnosis (CADx) - Lesions
  - Risk Assessment
  - Response to neoadjuvant therapy
- Methods to handle limitations and potential pitfalls
  - Pre-processing
  - Transfer learning
  - Fine tuning
  - Data Augmentation

Computer-aided diagnosis in the work-up of suspect lesions: malignant vs. benign lesions (on FFDMs)

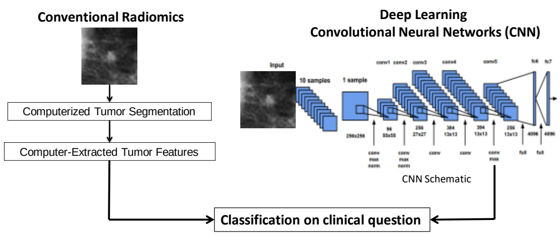
- Use computer output to help characterize (i.e., output descriptors of the lesion) and potentially indicate a computer-determined probability of malignancy of a found lesion
- The final decision on patient management is still made by the radiologist


Benign


Malignant

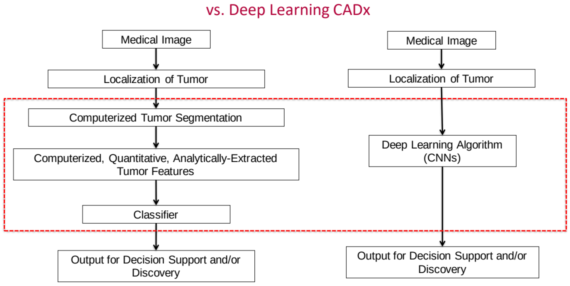

Malignant

CADx: task of distinguishing between malignant and benign breast lesions

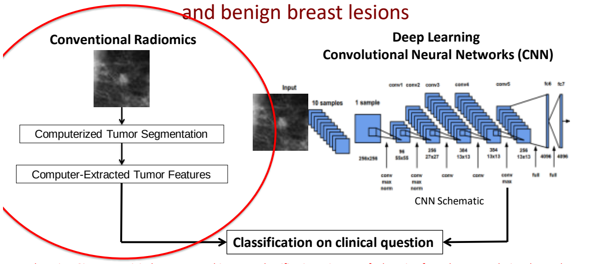


Huynh B, Li H, Giger ML: Digital mammographic tumor classification using transfer learning from deep convolutional neural networks. J Medical Imaging 3(3), 034501 (2016).

"Conventional lesion-segmentation (hand-crafted) CADx vs. Deep Learning CADx



Quantitative radiomics in distinguishing between malignant and benign breast lesions

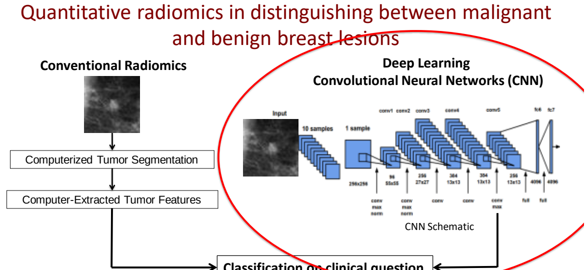


Huynh B, Li H, Giger ML: Digital mammographic tumor classification using transfer learning from deep convolutional neural networks. J Medical Imaging 3(3), 034501 (2016).

Conventional Hand-Crafted Radiomics CADx on FFDM

- Center of the lesion is indicated
- Then an automatic lesion segmentation is performed, based on a multiple transition-point, gray-level, region-growing technique.
- After the lesion is segmented, image features (i.e., mathematical descriptors) were extracted from the lesion:
  - Lesion size
  - Lesion shape
  - Intensity features (e.g., average gray level, contrast)
  - Texture within the lesion
  - Margin morphology (e.g., spiculation and sharpness) of the mass
- Features then merged by a classifier (e.g., LDA, SVM) to yield a signature indicating an estimate of the likelihood of malignancy

Quantitative radiomics in distinguishing between malignant and benign breast lesions

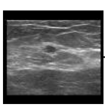


Huynh B, Li H, Giger ML: Digital mammographic tumor classification using transfer learning from deep convolutional neural networks. J Medical Imaging 3(3), 034501 (2016).

Due to limited size of datasets, use transfer learning

Using CNNs in feature extraction

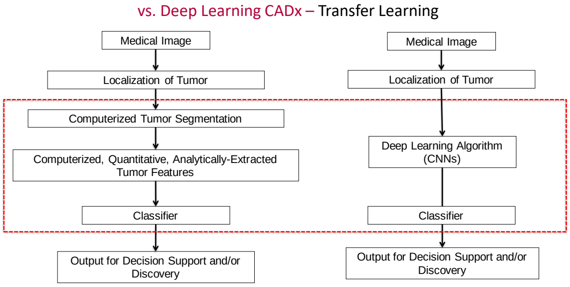
Use of Transfer Learning in Deep learning for Feature Extraction



Deep Learning Feature Extraction → Classification (LDA, SVM) → Answer to Clinical Question

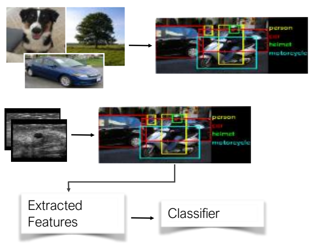
- CNNs extract features from entire ROIs without localization or segmentation of lesions.
- Advantage: No lesion segmentation is required
- Advantage: No extraction of segmentation-based features, such as size, shape, margin sharpness, texture, and kinetics
- CNNs require very large datasets -- Can we incorporate pre-trained CNNs?

"Conventional lesion-segmentation (hand-crafted) CADx vs. Deep Learning CADx – Transfer Learning



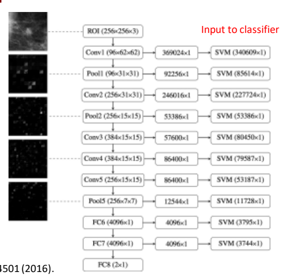
Deep learning : *Transfer Learning*

- Take a CNN trained to classify everyday objects.
- Process medical images with the pre-trained CNN.
  - Pre-trained CNN – AlexNet
  - 1.28 million training high-resolution images
  - About 1, 000 categories
  - Krizhevsky A, et al., ImageNet Classification with Deep Convolutional Neural Networks. 2012
- Take outputs from the CNN layers and use it as "features" for a classifier.



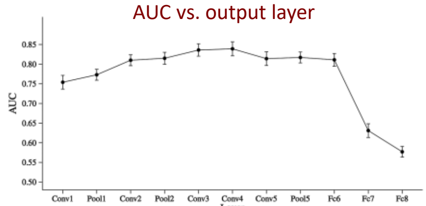
Example of Transfer Learning: Already-Trained CNN Structure (e.g., AlexNet) applied to FFDMs

- A schematic of how features are extracted using a pre-trained AlexNet.
- Each ROI is sent through the network and the outputs from each layer are preprocessed to be used as sets of features for an SVM.
- The filtered image outputs from some of the layers can be seen in the left column.
- The numbers in parentheses for the center column denote the dimensionality of the outputs from each layer.
- The numbers in parentheses for the right column denote the length of the feature vector per ROI used as an input for the SVM after zero-variance removal.
- After a feature vector has been extracted from each ROI, the SVM is then trained and evaluated by cross validation.



Huynh B, Li H, Giger ML, J Medical Imaging 3(3), 034501 (2016).

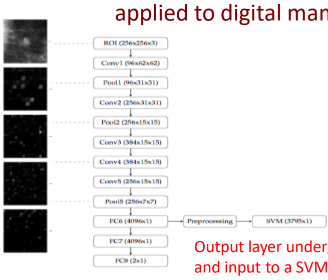
Example of Transfer Learning: AUC vs. output layer



Performance in terms of AUC for classifiers based on features from each layer of AlexNet in the task of distinguishing between malignant and benign tumors on FFDMs.

Huynh B, Li H, Giger ML: Digital mammographic tumor classification using transfer learning from deep convolutional neural networks. J Medical Imaging 3(3), 034501 (2016).

Already-Trained CNN Structure (e.g., AlexNet) applied to digital mammograms

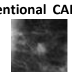


Huynh B, Li H, Giger ML: Digital mammographic tumor classification using transfer learning from deep convolutional neural networks. J Medical Imaging 3(3), 034501 (2016).

Output layer undergoes post processing and input to a SVM classifier

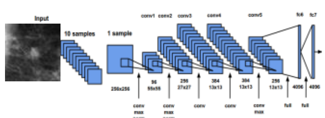
Conventional CADx vs. CNN CADx in distinguishing between malignant and benign breast lesions (Huynh et al.)

Conventional CADx



Computerized, Quantitative, Tumor Features

Deep Learning
Convolutional Neural Networks (CNN)



CNN Schematic

Classification on clinical question

Conventional CADx vs. CNN CADx in distinguishing between malignant and benign breast lesions (Huynh et al.)

Conventional CADx

Computerized, Quantitative, Tumor Features

Deep Learning
Convolutional Neural Networks (CNN)

CNN Schematic

Fusion Classifier

Classification on clinical question

ROC Analysis Evaluation: CNN vs. Analytically-extracted Features

Model	AUC	AUC Std Dev	Approx. Time
Pre-trained Deep Learning CNN	0.81	0.04	7 minutes
"Conventional CAD/Radiomics"	0.81	0.03	5 minutes
Ensemble Classifier (Combination of both)	0.86	0.01	12 minutes
CNN without pre-training	0.71	0.06	12 hours

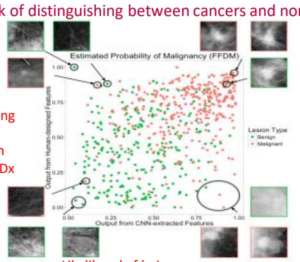
Five-fold cross validation

Huynh B, Li H, Giger ML: Digital mammographic tumor classification using transfer learning from deep convolutional neural networks. J Medical Imaging 3(3), 034501 (2016).

Giger Deep Learning AAPM 2017

Conventional CAD/Radiomics & Deep Learning CAD/Radiomics

(task of distinguishing between cancers and non cancers)



RED = CANCER

GREEN = Non-CANCER

Likelihood of being cancer as determined from conventional CADx

Likelihood of being cancer as determined from deep learning

Huynh et al. RSNA annual meeting 2016

Giger Deep Learning AAPM 2017

Deep Learning Applied across Multiple Modalities: FFDM, Ultrasound, MRI

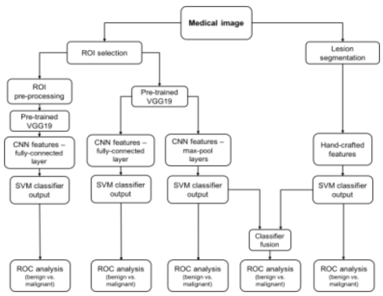
Imaging Modality	Total # of Lesions	# of Benign Lesions	# of Malignant Lesions	Total # of ROIs	# of Benign ROIs	# of Malignant ROIs	ROI size Range	Average Pixel Size
FFDM	245	113	132	739	328	411	512x512	0.10 mm
Ultrasound	1125	967	158	2393	1978	415	100x100-300x400	0.10 mm
DCE-MRI	690	212	478	690	212	478	48x48 - 126x126	0.69 mm

Can train with multiple ROIs of a lesion, however when testing all ROIs of a case need to be in either "training" or "testing".

Giger Deep Learning AAPM 2017

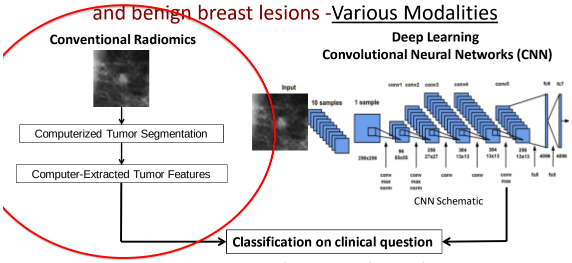
Various Parameters Investigated

- Pre-Processing
- Transfer learning
  - Pooled features
  - Fully-connected features
- Data augmentation
  - Images from multiple time points or views
- Classifier fusion



Giger Deep Learning AAPM 2017

Quantitative radiomics in distinguishing between malignant and benign breast lesions -Various Modalities



Huynh B, Li H, Giger ML: Digital mammographic tumor classification using transfer learning from deep convolutional neural networks. J Medical Imaging 3(3), 034501 (2016). Giger Deep Learning AAPM 2017

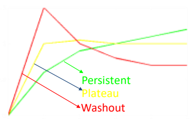
FFDM & Ultrasound Hand-Crafted Features

- After the lesion is segmented, image features (i.e., mathematical descriptors) were extracted from the lesion:
  – Lesion size
  – Lesion shape
  – Intensity features (e.g., average gray level, contrast)
  – Texture within the lesion
  – Margin morphology (e.g., spiculation and sharpness) of the mass

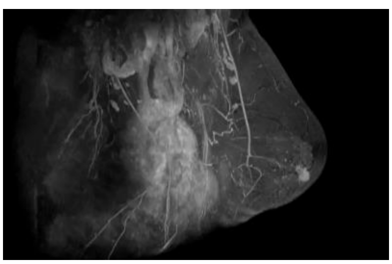
Giger Deep Learning AAPM 2017

Dynamic Contrast-Enhanced Magnetic Resonance Imaging: Additional hand-crafted features related to dynamic imaging

- Tumors have increased blood vessels and differ in micro-vascular density and vessel permeability
- Dynamic-Contrast MRI [DCE-MRI]
  - Contrast agent (Gd-DTPA) shortens T1 relaxation time which leads to increase of signal in T1-weighted images
  - Pre-contrast and a series of post-contrast images are obtained to provide functional information regarding lesions



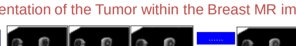
Clinical 3D Breast MRI image



Giger Deep Learning AAPM 2017

University of Chicago High-Throughput MRI Phenotyping System: **Hand-Crafted** (Segmentation of the Tumor within the Breast MR image)

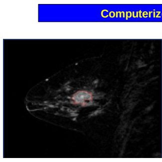
4D DCE MRI images



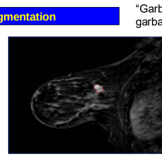
Radiologist-indicated Tumor Center

Computerized Tumor Segmentation

"Garbage in, garbage out"



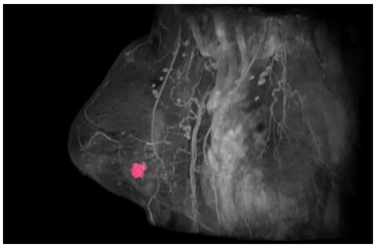
ER-negative



ER-positive

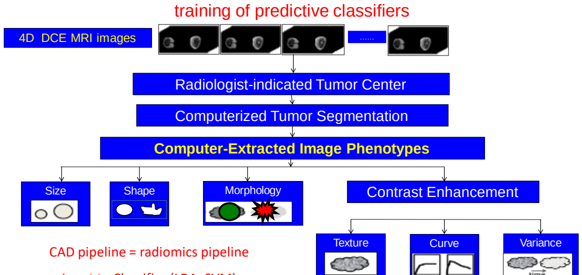
Giger Deep Learning AAPM 2017

Computer-extracted Breast Cancer on MRI (can analyze as a "virtual" digital biopsy of the tumor)



- non-invasive
- covers complete tumor
- repeatable

Computer-extraction of hand-crafted, lesion-based features followed by training of predictive classifiers

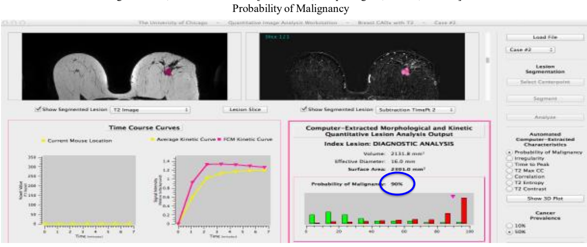


CAD pipeline = radiomics pipeline

Input to Classifier (LDA, SVM)

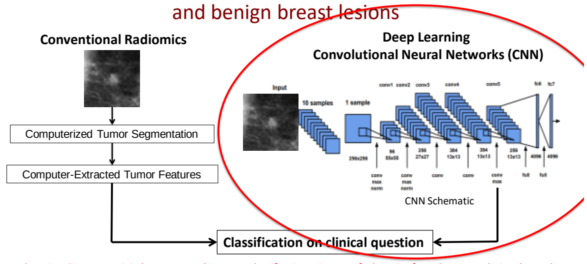
Quantitative Image Analysis Workstation for the High Throughput MRI Phenotyping of Breast Lesions – **DIAGNOSTIC TASKS**

Automated Lesion Segmentation, Feature Extraction [volumetrics, morphological, texture, kinetics] and Estimation of the Probability of Malignancy



Giger et al., RSNA 2010

Quantitative radiomics in distinguishing between malignant and benign breast lesions



Huynh B, Li H, Giger ML: Digital mammographic tumor classification using transfer learning from deep convolutional neural networks. J Medical Imaging 3(3), 034501 (2016).

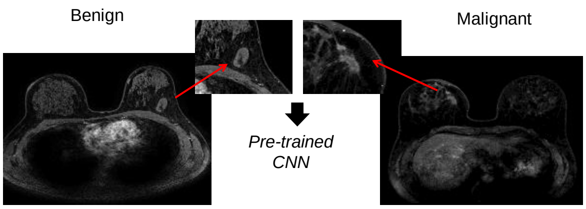
Giger Deep Learning AAPM 2017

Various Parameters Investigated

- Pre-Processing
  - Since ROIs of different sizes
  - Add frame or mirror padding to obtain equal input ROI sizes
- Transfer learning
  - Pooled features
  - Fully-connected features
- Data augmentation
  - Images from multiple time points or views
- Classifier fusion

Giger Deep Learning AAPM 2017

Image Data for input to CNN: Large ROIs

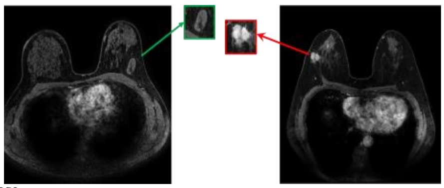


- Entire image
- Large ROI localized to tumor
- ROI mainly including only the tumor

Antropova N et al. SPIE Proc. Med Imag 2017

Giger Deep Learning AAPM 2017

Image Data for input to CNN: Small ROIs

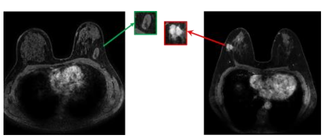


- Entire image
- Large ROI localized to tumor
- ROI mainly including only the tumor

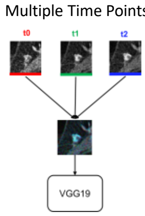
Task of distinguishing malignant vs. benign	Large ROIs	Small ROIs
AUC	0.72	0.87

Antropova N et al. SPIE Proc. Med Imag 2017
Antropova N, et al Medical Physics (in press), 2017 Giger Deep Learning AAPM 2017

Data Augmentation: Use images from multiple time points to incorporate the dynamic characteristics



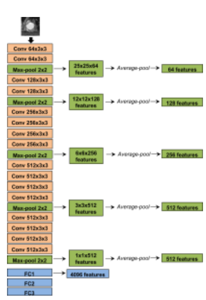
Use Images from Multiple Time Points



Giger Deep Learning AAPM 2017

VGG19 for Feature Extraction:

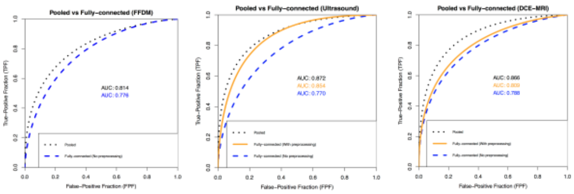
Pooled Layers
or
Fully Connected Layer



Giger Deep Learning AAPM 2017

Pooled vs. Fully-Connected

FFDM | Ultrasound | MRI

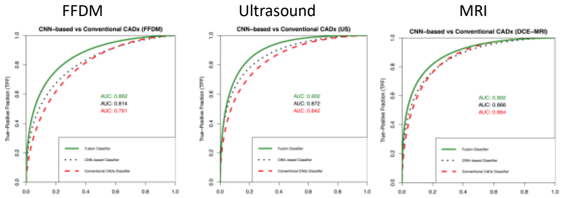


Antropova N, Huynh BQ, Giger ML: A deep fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets. Medical Physics (in press), 2017.

Hand-crafted vs. CNN vs. Fusion

(diagnostic task of distinguishing between cancers and non cancers)

FFDM | Ultrasound | MRI



Antropova N, Huynh BQ, Giger ML: A deep fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets. Medical Physics (in press), 2017.

Hand-crafted vs. CNN vs. Fusion

(diagnostic task of distinguishing between cancers and non cancers across breast imaging modalities)

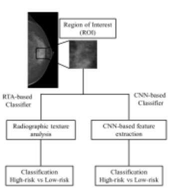
	Conventional CADx Classifier	CNN-based Classifier	Fusion Classifier
FFDM	0.79 (se = 0.01)	0.81 (se = 0.01)	0.86 (se = 0.01)
		p = 0.0043	
		p = 2.114e-09	
Ultrasound	0.84 (se = 0.01)	0.87 (se = 0.01)	0.90 (se = 0.01)
		p = 7.491e-07	
		p =1.467e-08	
DCE-MRI	0.86 (se = 0.01)	0.87 (se = 0.01)	0.89 (se = 0.01)
		p = 0.026	
		p = 3.676e-05	

Antropova N, Huynh BQ, Giger ML: A deep fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets. Medical Physics (in press), 2017.

Role of deep learning at various stages of quantitative image analysis (radiomics) for disease assessment

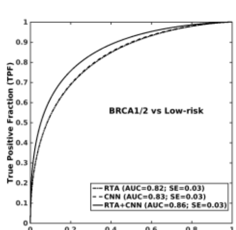
- Applications in breast image analysis
  - Computer-aided detection (CADe)
  - Computer-aided diagnosis (CADx)
  - Risk Assessment – Breast Parenchyma (microenvironment)
  - Response to neoadjuvant therapy
- Methods to handle limitations and potential pitfalls
  - Pre-processing
  - Transfer learning
  - Fine tuning
  - Data Augmentation

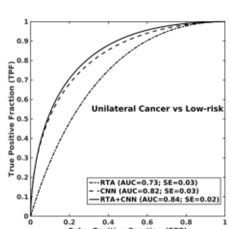
Deep Learning in Breast Cancer Risk Assessment: Evaluation of Convolutional Neural Networks on a Clinical Dataset of FFDMs



	BRCA1/2 gene-mutation carriers	Unilateral cancer women	Low-risk women
N	53	75	328
Mean age (SD)	40.2 (11.8)	55.8 (15.0)	58.4 (11.9)
Race			
White, non-Hispanic	49	28	107
Black, non-Hispanic	3	34	194
Asian	0	3	7
American Indian or Alaskan Native	0	0	1
Hispanic	1	1	8
Other/mixed	0	9	11
BI-RADS density			
A	4 (7.5%)	5 (6.7%)	34 (10.4%)
B	18 (34.0%)	43 (57.3%)	200 (61.0%)
C	25 (47.2%)	22 (29.3%)	90 (27.4%)
D	6 (11.3%)	5 (6.7%)	4 (1.2%)
Breast mean percent density (%)	27.2 (18.2)	22.5 (18.4)	18.7 (17.4)

Hand-crafted RTA vs. Deep Learning CNN



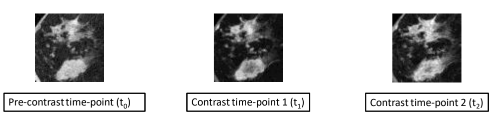


Role of deep learning at various stages of quantitative image analysis (radiomics) for disease assessment

- Applications in breast image analysis
  - Computer-aided detection (CADe)
  - Computer-aided diagnosis (CADx)
  - Risk Assessment – Breast Parenchyma (microenvironment)
  - Response to neoadjuvant therapy
- Methods to handle limitations and potential pitfalls
  - Pre-processing
  - Transfer learning
  - Fine tuning
  - Data Augmentation

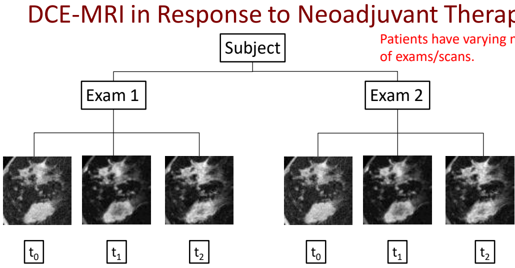
DCE-MRI in Response to Neoadjuvant Therapy

- Incorporate the dynamic (temporal) aspect of DCE-MRI.
- Multiple scans per exam.
- Multiple exams per subject.
- Each contrast time-point provides different physiological information.



DCE-MRI in Response to Neoadjuvant Therapy

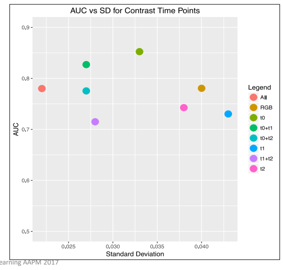
Patients have varying numbers of exams/scans.



Huynh B et al. SPIE Proc. Med Imag 2017

Deep Learning & DCE-MRI in Response to Therapy

- All subsets performed well (AUC ~0.70-0.85).
- Using only the pre-contrast time-point worked the best.
- Incorporating more time-points decreases the variance.



Huynh B et al. SPIE Proc. Med Imag 2017

Giger Deep Learning AAPM 2017

Role of deep learning at various stages of quantitative image analysis (radiomics) for disease assessment

- Applications in breast image analysis
  - Computer-aided detection (CADe)
  - Computer-aided diagnosis (CADx)
  - Risk Assessment – Breast Parenchyma (microenvironment)
  - Response to neoadjuvant therapy
- Methods to handle limitations and potential pitfalls
  - Pre-processing
  - Transfer learning
  - Fine tuning
  - Data Augmentation

Giger Deep Learning AAPM 2017

Summary

- Image analysis tasks are continuing to be developed using both hand-crafted methods and deep learning methods
- Understanding the CNN is important in optimizing and in using in interpretations (don't just say "black box")
- Methods available to handle limited data sets
  - Transfer learning, data augmentaton
  - Pre and post processing to handle images of differing sizes

Giger Deep Learning AAPM 2017

Thank you
&
Acknowledgements

Recent & Current Graduate Students
Weijie Chen, PhD
Joel Wilkie, PhD
Martin King, PhD
Nick Gruszauskas, PhD
Yading Yuan, PhD
Robert Tomek, MS
Neha Bhooshan, PhD
Andrew Jamieson, PhD
Hsien-Chi Kuo, PhD
Martin Andrews, PhD
William Weiss, Ph.D.
Chris Haddad, Ph.D.
Adam Sibley
Natasha Antropova
Kayla Mendel

Research Lab
Karen Drukker, PhD
Hui Li, PhD
Mike Chinander, PhD
Li Lan, MS
John Papaioannou, MS
Sasha (Alexandra) Edwards
Chun Wai Chan, MS
Benjamin Huynh
Summer medical students, undergraduates, and high school students

Collaborators
Gillian Newstead, MD
Charlene Sennett, MD*
Charles E. Metz, PhD*
Robert Nishikawa, PhD
Funmi Olopade, MD
Greg Karczmar, PhD
Milica Medved, PhD
Yulei Jiang, PhD
Anna Di Rienzo, PhD
Hiro Abe, MD
Yuan Ji, PhD
Yitan Zhu, PhD