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 View, and receives royalties from Hologic, GE Medical Systems, MEDIAN Technologies, Ishara 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 (CADx)
  – 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

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)

Zhang W et al. SPIE Proceeding 1709: 257-268

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

CADx: task of distinguishing between malignant and benign breast lesions

Quantitative radiomics in distinguishing between malignant and benign breast lesions

Conventional Radiomics

- 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

CNN Schematic

Giger Deep Learning AAPM 2017


Quantitative radiomics in distinguishing between malignant and benign breast lesions

Conventional Radiomics

- 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

CNN Schematic

Giger Deep Learning AAPM 2017

Due to limited size of datasets, use transfer learning

Using CNNs in feature extraction

Use of Transfer Learning in Deep learning for Feature Extraction

- 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
- Take outputs from the CNN layers and use 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.


**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.

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


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.)

Fusion Classifier

Classification on clinical question

Conventional CADx Deep Learning Convolutional Neural Networks (CNN)

Computerized, Quantitative, Tumor Features

Giger Deep Learning AAPM 2017 CNN Schematic

Classification on clinical question
ROA Analysis Evaluation: CNN vs. Analytically-extracted Features

<table>
<thead>
<tr>
<th>Model</th>
<th>AUC</th>
<th>AUC Std Dev</th>
<th>Approx. Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-trained Deep Learning CNN</td>
<td>0.81</td>
<td>0.04</td>
<td>7 minutes</td>
</tr>
<tr>
<td>&quot;Conventional CAD/Radiomics&quot;</td>
<td>0.81</td>
<td>0.03</td>
<td>5 minutes</td>
</tr>
<tr>
<td>Ensemble Classifier (Combination of both)</td>
<td>0.86</td>
<td>0.01</td>
<td>12 minutes</td>
</tr>
<tr>
<td>CNN without pre-training</td>
<td>0.71</td>
<td>0.06</td>
<td>12 hours</td>
</tr>
</tbody>
</table>

Five-fold cross validation


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 CADs

Likelihood of being cancer as determined from deep learning

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

<table>
<thead>
<tr>
<th>Imaging Modality</th>
<th>Case # of Lesions</th>
<th>FFDM # of Lesions</th>
<th>Ultrasound # of Lesions</th>
<th>DCE-MRI # of Lesions</th>
<th>Total # of ROIs</th>
<th># of Benign ROIs</th>
<th># of Malignant ROIs</th>
<th>AUC Range</th>
<th>Average Pixel Size Range</th>
<th>Form Factor Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>FFDM</td>
<td>240</td>
<td>113</td>
<td>377</td>
<td>960</td>
<td>1320</td>
<td>793</td>
<td>527</td>
<td>0.71</td>
<td>96x48 - 300x400</td>
<td>0.87</td>
</tr>
<tr>
<td>Ultrasound</td>
<td>911</td>
<td>361</td>
<td>1359</td>
<td>1265</td>
<td>2393</td>
<td>1978</td>
<td>415</td>
<td>0.66</td>
<td>100x100 - 300x400</td>
<td>0.97</td>
</tr>
<tr>
<td>DCE-MRI</td>
<td>690</td>
<td>212</td>
<td>478</td>
<td>212</td>
<td>690</td>
<td>212</td>
<td>478</td>
<td>0.69</td>
<td>48x48 - 126x126</td>
<td>0.87</td>
</tr>
</tbody>
</table>

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

Huynh et al. RSNA annual meeting 2016

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

Quantitative radiomics in distinguishing between malignant and benign breast lesions

Various Modalities

- Conventional Radiomics
  - Computerized Tumor Segmentation
  - Computer-Extracted Tumor Features
- Deep Learning
  - Convolutional Neural Networks (CNN)
  - CNN Schematic


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
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

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

4D DCE MRI images
Radiologist indicated Tumor Center
Computerized Tumor Segmentation

Computer-Extracted Image Phenotypes
Size
Shape
Morphology
Contrast Enhancement
Texture

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 (volumetric, 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

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

Image Data for input to CNN: Large ROIs

- Entire image
- Large ROI localized to tumor
- ROI mainly including only the tumor
Image Data for input to CNN: Small ROIs

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

<table>
<thead>
<tr>
<th>Task of distinguishing malignant vs. benign</th>
<th>Large ROIs</th>
<th>Small ROIs</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC</td>
<td>0.72</td>
<td>0.87</td>
</tr>
</tbody>
</table>

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

VGG19 for Feature Extraction:

Pooled Layers or Fully Connected Layer
Pooled vs. Fully-Connected

Hand-crafted vs. CNN vs. Fusion
(diagnostic task of distinguishing between cancers and non-cancers)

Hand-crafted vs. CNN vs. Fusion
(diagnostic task of distinguishing between cancers and non-cancers across breast imaging modalities)
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

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.

Patients have varying numbers of exams/scans.
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.

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

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 augmentation
  - Pre and post processing to handle images of differing sizes
Thank you &
Acknowledgements

Recent & Current Graduate Students
Wei-Jie Chen, PhD
Joel Wilkie, PhD
Martin King, PhD
Nick Goutsmakis, PhD
Yuling Yuan, PhD
Robert Tomock, MS
Naha Bhushan, PhD
Andrew Janisse, PhD
Hazen-Chi Kuo, PhD
Hui Li, PhD
Mike Chinander, PhD
Li Lan, MS
John Popinnoor, MS
Sandra (Alexandra) Edwards
Chau Wai Chan, MS
Benjamin Hayash
Summer medical students, undergraduates, and high school students

Research Lab
Karen Drukker, PhD
Hui Li, PhD
Mike Chinander, PhD
Li Lan, MS
John Popinnoor, MS
Sandra (Alexandra) Edwards
Chau Wai Chan, MS
Benjamin Hayash

Recent & Current Graduate Students
Weijie Chen, PhD
Joel Wilkie, PhD
Martin King, PhD
Nick Goutsmakis, PhD
Yuling Yuan, PhD
Robert Tomock, MS
Naha Bhushan, PhD
Andrew Janisse, PhD
Hazen-Chi Kuo, PhD
Hui Li, PhD
Mike Chinander, PhD
Li Lan, MS
John Popinnoor, MS
Sandra (Alexandra) Edwards
Chau Wai Chan, MS
Benjamin Hayash
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
Neha Bhooshan, PhD
Andrew Jamieson, PhD
Hsien-Chi Kuo, PhD
Martin Andrews, PhD
William Weisz, Ph.D.
Charlene Sennett, MD*
Charles E. Metz, PhD*
Robert Nishikawa, PhD
Funmi Olopade, MD
Neha Bhooshan, PhD
Andrew Jamieson, PhD
Hsien-Chi Kuo, PhD
Martin Andrews, PhD
William Weisz, Ph.D.
Charlene Sennett, MD*
Charles E. Metz, PhD*
Robert Nishikawa, PhD
Funmi Olopade, MD
Neha Bhooshan, PhD
Andrew Jamieson, PhD
Hsien-Chi Kuo, PhD
Martin Andrews, PhD
William Weisz, Ph.D.
Charlene Sennett, MD*
Charles E. Metz, PhD*
Robert Nishikawa, PhD
Funmi Olopade, MD
Neha Bhooshan, PhD
Andrew Jamieson, PhD
Hsien-Chi Kuo, PhD
Martin Andrews, PhD
William Weisz, Ph.D.
Charlene Sennett, MD*
Charles E. Metz, PhD*
Robert Nishikawa, PhD
Funmi Olopade, MD
Neha Bhooshan, PhD
Andrew Jamieson, PhD
Hsien-Chi Kuo, PhD
Martin Andrews, PhD
William Weisz, Ph.D.
Charlene Sennett, MD*
Charles E. Metz, PhD*
Robert Nishikawa, PhD
Funmi Olopade, MD
Neha Bhooshan, PhD
Andrew Jamieson, PhD
Hsien-Chi Kuo, PhD
Martin Andrews, PhD
William Weisz, Ph.D.
Charlene Sennett, MD*
Charles E. Metz, PhD*
Robert Nishikawa, PhD
Funmi Olopade, MD
Neha Bhooshan, PhD
Andrew Jamieson, PhD
Hsien-Chi Kuo, PhD
Martin Andrews, PhD
William Weisz, Ph.D.
Charlene Sennett, MD*
Charles E. Metz, PhD*
Robert Nishikawa, PhD
Funmi Olopade, MD
Neha Bhooshan, PhD
Andrew Jamieson, PhD
Hsien-Chi Kuo, PhD
Martin Andrews, PhD
William Weisz, Ph.D.
Charlene Sennett, MD*
Charles E. Metz, PhD*
Robert Nishikawa, PhD
Funmi Olopade, MD
Neha Bhooshan, PhD
Andrew Jamieson, PhD
Hsien-Chi Kuo, PhD
Martin Andrews, PhD
William Weisz, Ph.D.
Charlene Sennett, MD*
Charles E. Metz, PhD*
Robert Nishikawa, PhD
Funmi Olopade, MD
Neha Bhooshan, PhD
Andrew Jamieson, PhD
Hsien-Chi Kuo, PhD
Martin Andrews, PhD
William Weisz, Ph.D.
Charlene Sennett, MD*
Charles E. Metz, PhD*
Robert Nishikawa, PhD
Funmi Olopade, MD
Neha Bhooshan, PhD
Andrew Jamieson, PhD
Hsien-Chi Kuo, PhD
Martin Andrews, PhD
William Weisz, Ph.D.
Charlene Sennett, MD*
Charles E. Metz, PhD*
Robert Nishikawa, PhD
Funmi Olopade, MD
Neha Bhooshan, PhD
Andrew Jamieson, PhD
Hsien-Chi Kuo, PhD
Martin Andrews, PhD
William Weisz, Ph.D.