Frontiers of Artificial Intelligence (AI) in Imaging

Daniel L. Rubin, MD, MS

Associate Professor of Radiology
and of Medicine (Biomedical Informatics),
and by courtesy, of Ophthalmology
Department of Radiology
Stanford University

Acknowledgements

Students, Post-docs, Residents, Staff, and Collaborators

– Assaf Hoogi
– Alfiia Galimzianova
– Imon Banerjee
– Hersh Sagreiya
– Darvin Yi
– Carson Lam
– Xuerong Xiao
– Blaine Rister
– Emel Alkim
– Sheryl John
– Niranjan Balachandar
– Selen Bozkurt
– Jocelyn Barker
– Curt Langlotz
– Matthew Lungren
– Christopher Re
– Bao Do
– Sandy Napel
– Chris Beaulieu
– Ann Leung

Funding Support
NCI QIN grants U01CA142555, 1U01CA190214, 1U01CA187947
ECOG-ACRIN
Stanford-AstraZeneca Collaboration Grant
NVIDIA Academic Hardware Grant Program

Goals of this talk

- Provide brief overview of imaging and role in medicine
- Provide examples of clinical questions where deep learning vs. feature engineering and/or statistical modeling approaches are more appropriate
- Provide perspectives on how can we integrate clinical data with imaging

Outline

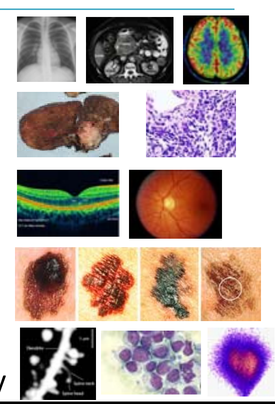
- Medical imaging and key clinical use cases motivating AI in imaging
- AI approaches and challenges
- Recent work and potential of AI in imaging

Outline

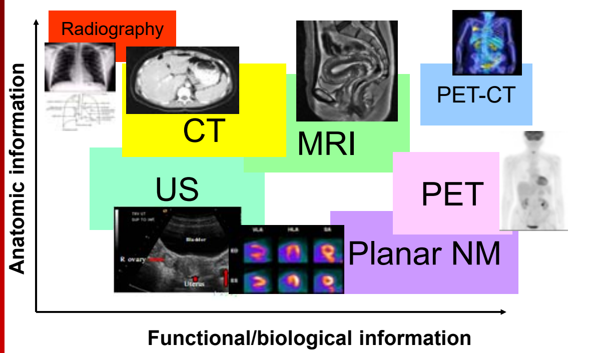
- Medical imaging and key clinical use cases motivating AI in imaging
- AI approaches and challenges
- Recent work and potential of QI in medical imaging

Imaging is key in several medical specialties

- Radiology
- Pathology
- Ophthalmology
- Dermatology
- Micro&Neurobiology



Many imaging modalities within each specialty



Key clinical uses of medical imaging (and AI)

- Disease detection
- Lesion segmentation

Current clinical practice

- Diagnosis
- Treatment selection
- Response assessment

Active research area

- Clinical prediction (of response or future disease)

Future application areas

Why do we need AI?

- Flood of image data
  - Impacts *disease detection*
- Variation in clinical practice
  - Impacts *diagnosis*
- Variation in disease among people
  - Impacts *treatment selection*

1) Flood of image data…

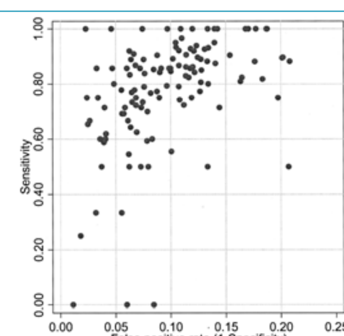
2001 2009

2) Variation in practice

- There are large variations and disparities in care (Institute of Medicine, 2001)
- *"Errors and variations in interpretation now represent the weakest aspect of clinical imaging*"*

*Robinson PJ. Radiology's Achilles' heel: error and variation in the interpretation of the Röentgen image. British Journal of Radiology. 1997 Jan 1;70(839):1085–98.

Variable Performance of Radiologists



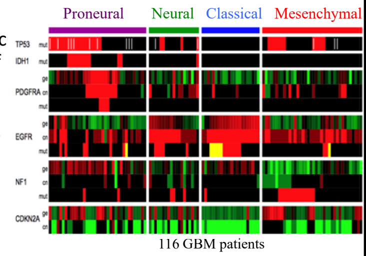
Barlow WE, Chi C, Carney PA, Taplin SH, D'Orsi C, Cutter G, et al. Accuracy of Screening Mammography Interpretation by Characteristics of Radiologists. J. Natl. Cancer Inst. 2004 Jan 15;96(24):1840–50.

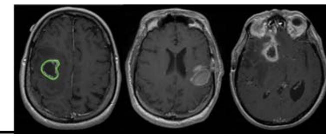
3) Variation in disease among people

People (and their diseases) differ…

Disease in different people varies

- Molecular diversity
  - Heterogeneous genomic aberration landscape of individual tumors*
- Phenotypic diversity
  - Variable appearance of lesions on images
- Clinical diversity
  - Patients have different response to treatment
- Ideally we will "profile" disease for personalized medicine

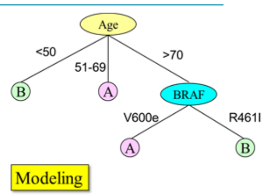


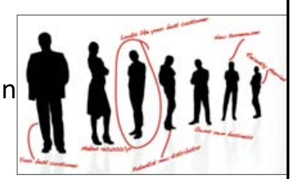


The TCGA Research Network. Cancer Cell. 2010

Data-driven, precision medicine

- Mine biological and medical data to create classifiers of disease and treatment response
- "Profile" disease in patients for personalized / precision medicine

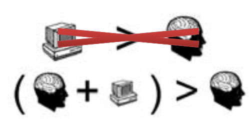




AI to the rescue?

Future jobs for AI

Future jobs for Radiologists?



Outline

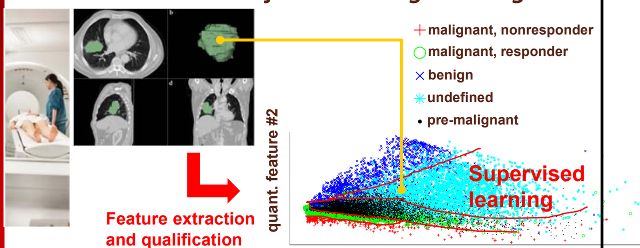
- Medical imaging and key clinical use cases motivating AI in imaging
- AI approaches and challenges
- Recent work and potential of AI in imaging

AI approaches

- Specify and process **pre-defined image features** in large volume to create clinical models
  - "radiomics"
- Process **raw image data** (unsupervised features learning) to directly create clinical models (usually classification)
  - Image patches
  - Deep learning, CNNs, etc.

"Radiomics"

Def: High-throughput extraction of quantitative image features with the intent of creating mineable databases from radiological images



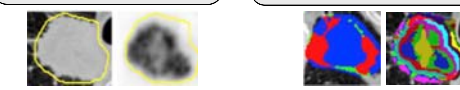
Radiomics: The process and the challenges, Magnetic Resonance Imaging, 30(9):1234-48, 2012.

Radiomics **image features** describe "imaging phenotypes"

Image phenotype: Composite of the observable characteristics of disease present in the image

Imaging phenotype =

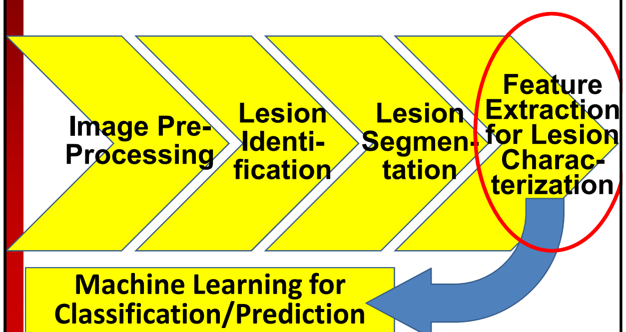
qualitative image features + *quantitative image features*



- "Mostly solid"
- "Spiculated margins"
- "Heterogeneous"

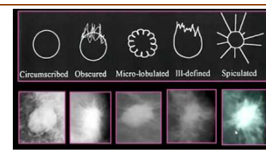
- SUVmax
- Lesion texture, shape
- Pixel histograms

Imaging processing pipeline to extract structured information from images



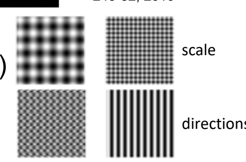
Quantitative Image Features

Shape:



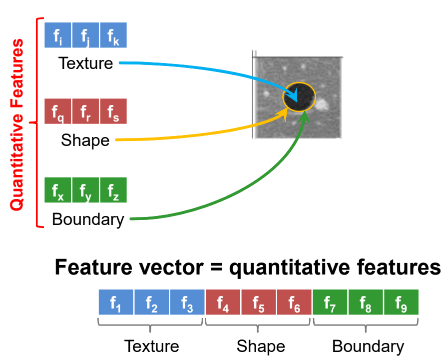
Edge:

Texture features:
(characterize lesion interior)



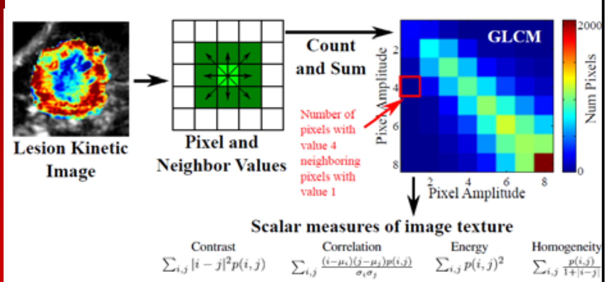
Napel S, Beaulieu CF, Rodriguez C, Cui J, Xu J, Gupta A, Korenblum D, Greenspan H, Ma Y, and Rubin DL, Radiology 256: 243-52, 2010

Structured image data ("image phenotype") represented as feature vector

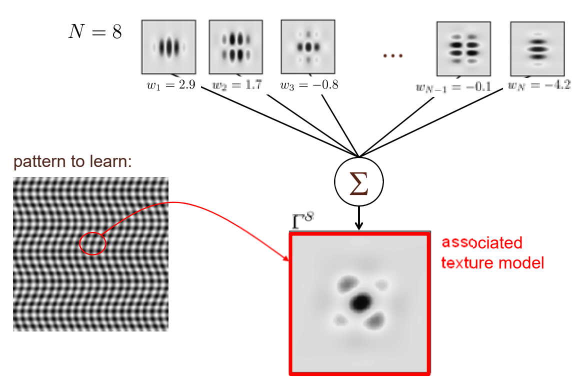


Quantifying texture: GLCM

- Gray-Level Co-Occurrence Matrix (GLCM)
- Captures heterogeneity in tissues



Describing texture as composition of of elements from Riesz filterbank



AI approaches

- Specify and process **pre-defined image features** in large volume to create clinical models
  - "radiomics"
- Process **raw image data** (unsupervised features learning) to directly create clinical models (usually classification)
  - Image patches
  - Deep learning, CNNs, etc.

Image patch analysis

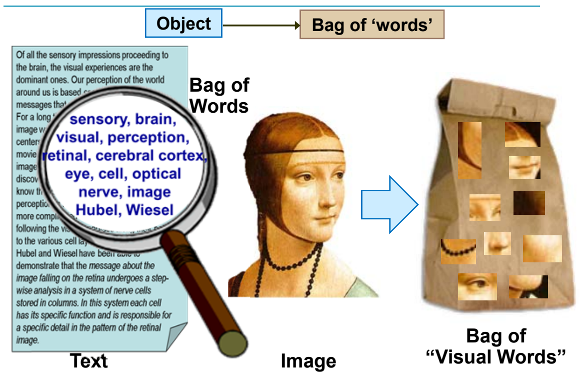


Image credit: Fei-Fei Li

Image patch analysis: Feature vectors of visual words

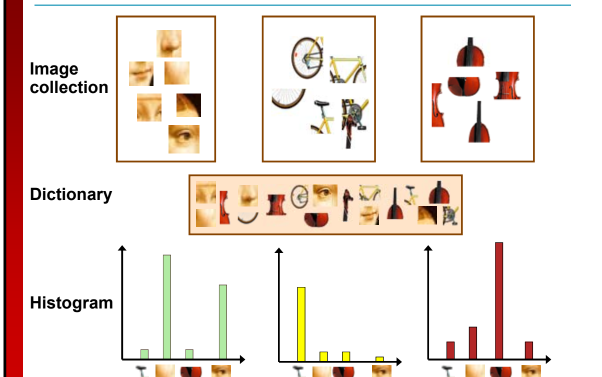
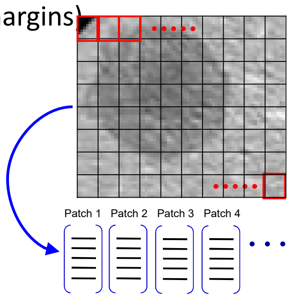


Image credit: Fei-Fei Li

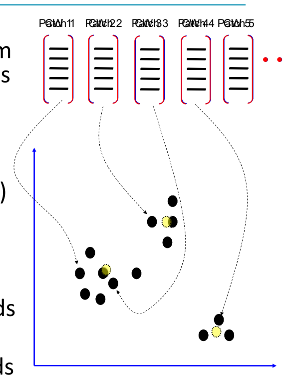
1. Patch Feature Extraction

- Uniform-sized patches extracted (including entire lesion and its margins)
- Sliding grid
- Normalize patches
  - Subtract mean
  - Divide by SD
- Final result = set of patch feature vectors



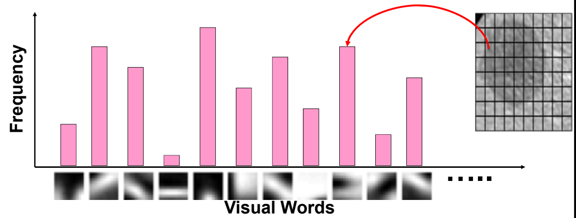
2. Dictionary Generation

- Represents all the unique "visual words" derived from all raw patches in all images
- Collect all raw patches
- *Dimensionality reduction*: PCA on patches to select components ("codewords") with highest variance
- Cluster patches using K-means
- **Visual Word** = The centroids of each clusters
- **Dictionary** = all visual words

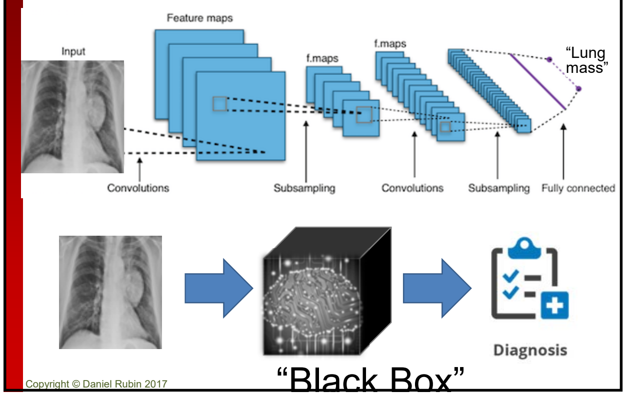


Feature Vector Generation

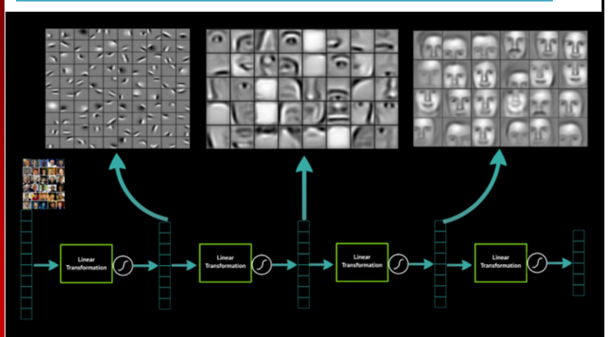
- Each image is a feature vector based on histogram of visual words
  - Dimensions = visual words in dictionary
  - Value of each dimension = count of patches closest to that visual word
- Normalize by dividing by number of patches
- Get separate feature vectors using interior and boundary dictionaries



Deep learning



Deep learning learns *feature hierarchies*

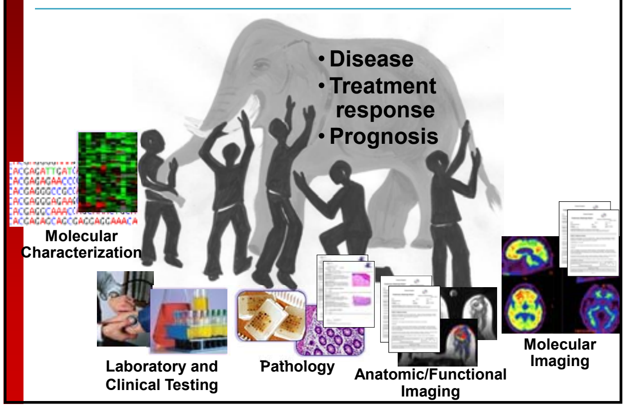


Source: http://www.datarobot.com/blog/a-primer-on-deep-learning/

Why not do everything with deep learning?

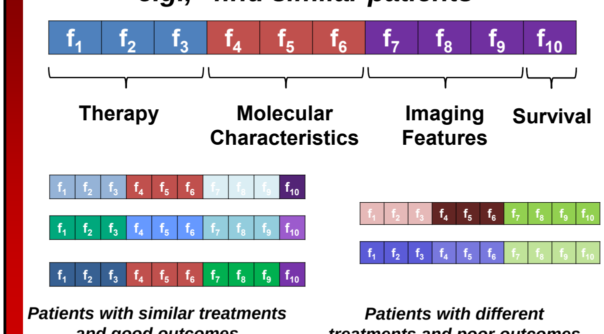
- Need ***lots of data*** to train models
- Need ***powerful hardware***
- Large amounts of ***tagged training data*** is in short supply and expensive to produce
- Many ***parameters*** need to be tuned, requires expertise and labor intensive
- Main applications limited to only ***classification*** and ***segmentation***

Integrating clinical data



This infrastructure enables learning from data

e.g., "find similar patients"



Outline

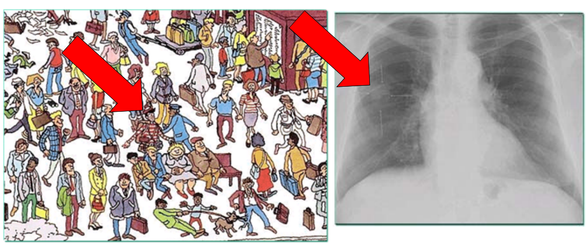
- Medical imaging and key clinical use cases motivating AI in imaging
- AI approaches and challenges
- Recent work and potential of AI in imaging

Key clinical uses of medical imaging (and AI)

1. Disease detection
2. Lesion segmentation
3. Diagnosis
4. Treatment selection
5. Clinical prediction (of response or future disease)

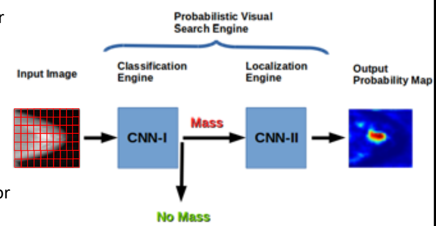
1) **Detection** of image abnormalities

AKA "where's Waldo?"



Detection of breast masses with deep learning

- Digital Database for Screening Mammography (DDSM)
- 2420 mass ROIs
- 80%/10%/10% training/test evaluation sets
- 256x256 patches, labeled as "mass" or "non-mass"
- Data augmentation: cropping, translation, rotation, flipping and scaling of image tiles
- Probability classification map of location (fully connected CNN)

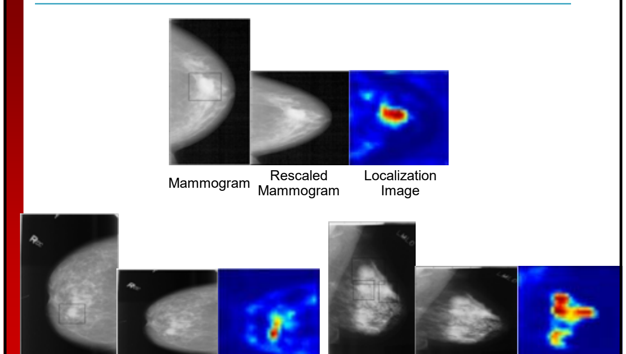


Performance:

	Number of Layers	Number of Parameters	Validation Accuracy
AlexNeT	8	60M	84%
VGG Net-16	16	140M	82%
GoogLeNeT	22	4M	85%

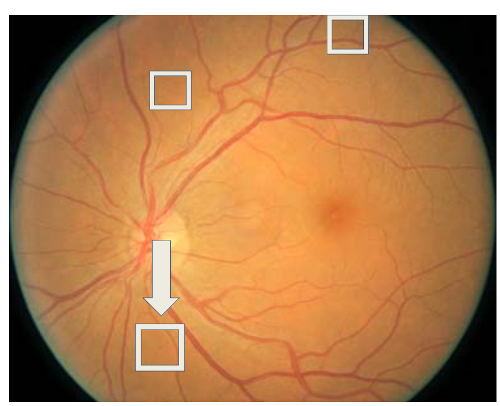
IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 1310-1315, 2015

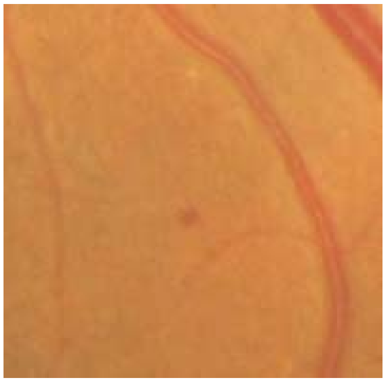
Examples



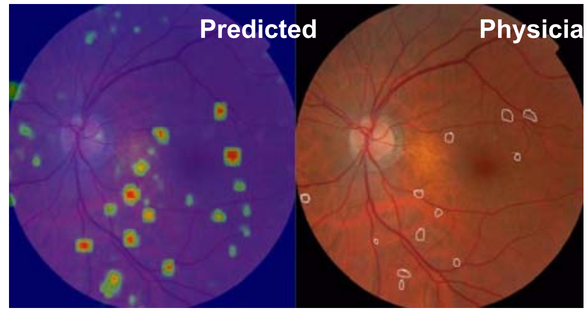
IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 1310-1315, 2015

Detecting retinal hemorrhages





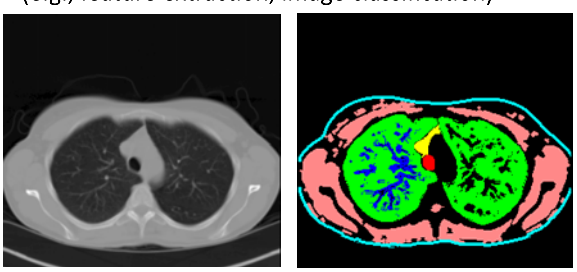
Sliding window detection of small features compared to physician manual detection



Single 224x224x3 input CNN sliding window: Detecting any abnormal feature
Red = P~.99 Green = P ~ 0.5 Blue = ~ 0.01

2) **Segmentation** of image regions

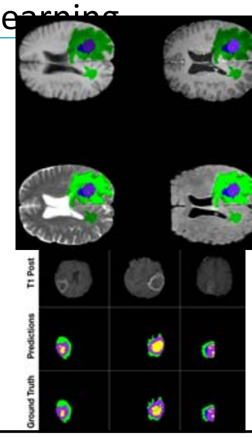
- Division of image into non-overlapping, homogeneous regions
- Segmented regions often input to other processing (e.g., feature extraction, image classification)



Segmentation of brain tumors using deep learning

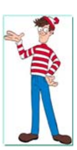
- **BRA**in **T**umor **S**egmentation (BRATS)
  - Glioblastoma Segmentation
  - 257 Patients
  - 4 Modalities of Co-registered MR Data
  - Expert Segmentations
- Algorithm: 3-Dimensional 4-Channel Fully Convolutional Neural Network (AlexNet)
- Dice Score Accuracy: 0.89
  Inter-radiologist Dice Score: 0.89

Yi D, et al., in press



3) Diagnosis: Classification of images

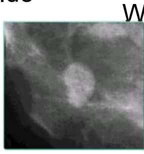
AKA "is it Waldo?"



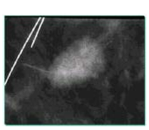
Waldo

Atypical Waldo

Pseudo-Waldo

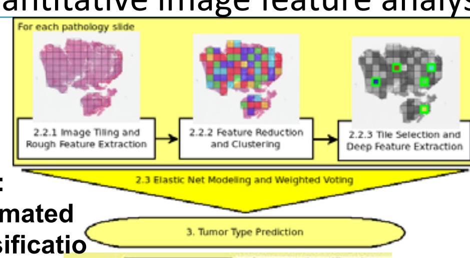


Benign Lymph Node



Infiltrating Cancer

Pathology classification using quantitative image feature analysis



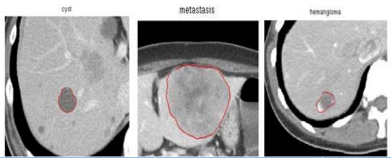
Goal: Automated classification of high- and low-grade glioma

		Actual	
		GBM	LGG
Predicted	GBM	23	1
	LGG	0	21

- Correct classification in 44 of the total 45 tissue slices
- Accuracy of 97.78% with 95% CI 88.23%-99.94%
- NIR of 51.11% gives p-value = 3.366x10^-12

Barker, J, Hoogi, A, Depeursinge, A, Rubin, D, Medical Image Analysis 30:60-71,, 2016.

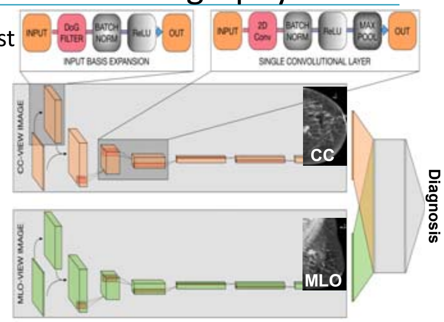
Diagnosis of liver lesions with image patches



Type of Dictionary	Only Boundary Dictionary				Only Interior Dictionary				Both Boundary and Interior Dictionary			
# of Errors	16				17				**3**			
Confusion Matrix	Pre-dicted → Actual ↓	Cyst	Met	Hem	Pre-dicted → Actual ↓	Cyst	Met	Hem	Pre-dicted → Actual ↓	Cyst	Met	Hem
	Cyst	23	**2**	0	Cyst	25	0	0	Cyst	25	0	0
	Met	0	24	0	Met	**8**	16	0	Met	0	24	0
	Hem	**2**	**12**	10	Hem	0	**9**	15	Hem	0	**3**	21

Deep learning: Diagnosis of masses on mammography

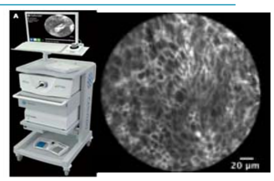
- Classify breast masses as benign vs. malignant
- Branching structure of CNN to account for two views of breast
- Predictive accuracy ~ 0.8

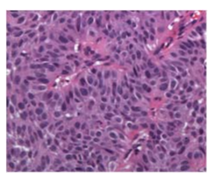


Yi D and Rubin DL, in preparation

Confocal Endomicroscopy

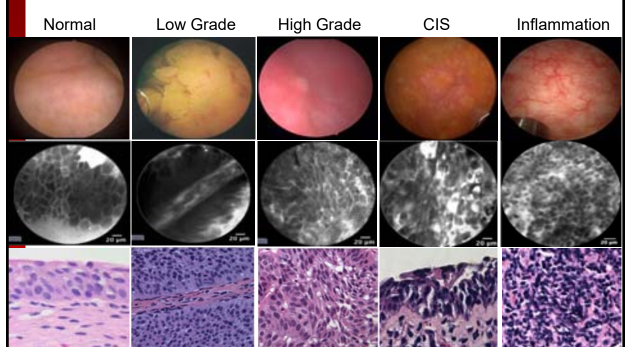
- Enables clinicians to obtain real time microscopic images
- “Optical biopsy”
  - based on confocal microscopy
- High resolution, dynamic, sub-surface imaging
- Used in GI and pulmonary applications
  - Barrett’s esophagus, colonic dysplasia



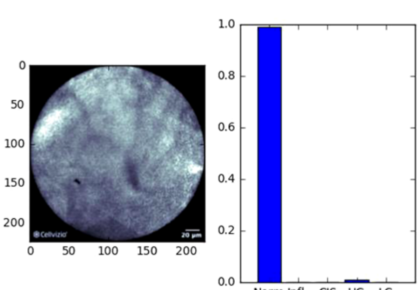


Sonn et al. J Urol. 2009. 182(4):1299-305.

Bladder pathology

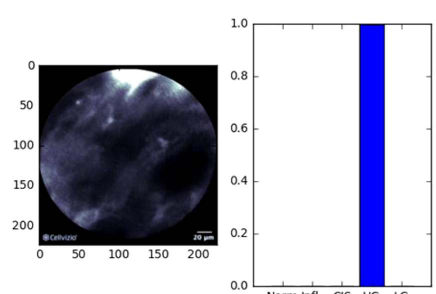


Normal Bladder Tissue



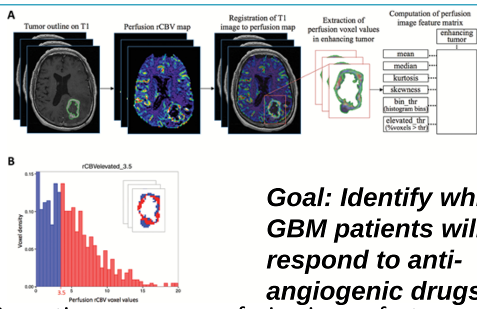
Work of Darvin Yi

Bladder Cancer



Work of Darvin Yi

4) Treatment selection



Goal: Identify which GBM patients will respond to anti-angiogenic drugs

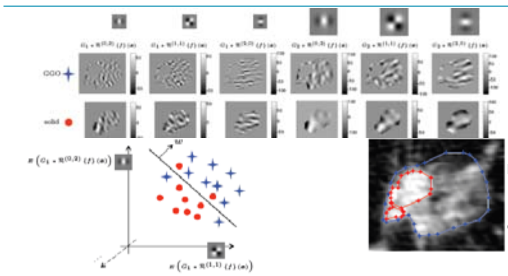
Magnetic resonance perfusion image features uncover a **subgroup of GBM patients** with **poor survival** and **better response to drug treatment**

Work of Tiffany Liu Neuro Oncol. 2016;19(7):997-1007

5) Clinical prediction

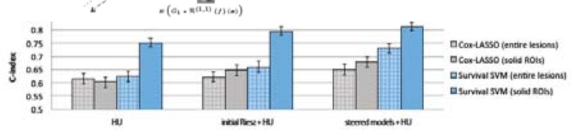
- Will disease respond to the treatment?
- Will the disease progress?
- Will disease recur?

Predicting lung cancer recurrence using quantitative analysis of lesion texture



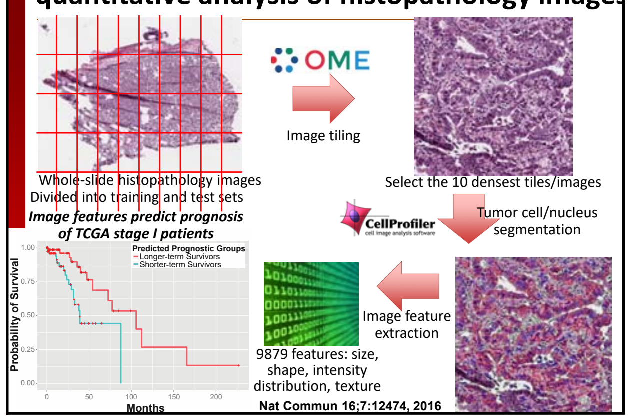
Riesz wavelets for texture

Lung nodule with solid and ground glass components



Depeursinge, et. al, and Rubin, Med Phys 42(4):2054-63, 2015

Prediction: Predicting survival from quantitative analysis of histopathology images



Nat Commun 16;7:12474, 2016

Summary: Key points

- Medical imaging is key to many important clinical use cases
- Clinicians who interpret images need assistance to reduce variations in care
- AI methods are promising for decision support and for reducing variations in care
- Deep learning methods are promising, but there are challenges, and best machine learning approach depends on the clinical problem

Thank you.

Contact info:
dlrubin@stanford.edu