Biomedical Informatics 260

Computational Methods for Biomedical Image Analysis and Interpretation

Spring 2019

Outline for Today

- About the course
- Expectations
- Logistics
- Resources
- Interactive teaching
- Lecture 1: Imaging modalities and need for computerized help in image interpretation

About the course

Objectives

- Understand medical imaging modalities
- Learn about how to get computers to "understand" images
- See applications for computerized image analysis
- Get hands on experience!

About the course

Audience

- Graduate students
- Medical Students
- Medical Residents / Fellows
- Undergraduates
- Auditors welcome

About the course

Daniel Rubin, MD, MS

- Professor of Biomedical Data Science, Radiology, and Medicine
- Research: Imaging AI (esp weak learning), image mining, machine learning, decision support with images

David Paik, PhD

- Adjunct Lecturer in Radiology
- Research: Visualization, image analysis, computer aided detection, image quantitation, cancer modeling

About the course

Teaching Assistants

Guhan Venkataraman **Kevin Thomas**

About the course

Expectations

Pre-requisites

- What you absolutely need to know
  - Programming ability (CS 106A)
  - Basic statistics helpful
- What would be *really* nice to know
  - Familiarity with Python 3 (initial TA sessions will do Python review, or look for online tutorials)
  - iPython/Jupyter Notebooks

Expectations

Readings

- **Articles**
  - Assigned with each lecture
  - Links posted on Canvas
- **Books**
  - Not required
  - Supplement required readings
  - (see Canvas, bmi260.stanford.edu redirects here)

Expectations

Coursework

- **Assignments (N=3)**
  - Involves programming (Python)
  - OUT on Mondays, DUE on Fridays
  - Up to groups of 2
- **Midterm** (during class, Mon, May 6)
- **Final project presentations** (during Final Exam slot, Mon, Jun 10, 3:30-6:30pm in Gates B03)
- Talking with others acceptable, all work individual
- Submissions on Canvas (one submission per group)

Expectations

Final Project

- A substantive programming project that covers element(s) of BOTH ***image quantitation*** and ***image semantics***
- Can be done in groups, up to 4 students
  - Project proposal (due Fri 4/26)
  - Milestone writeup (due Fri 5/17)
- Final Write-up (Due Monday 6/10 by 11:59 PM)
- Project Final Presentation (***Monday, June 10 from 3:30 PM - 6:30 PM, Gates B03***)

Expectations

Logistics

Schedule

- **Lectures**
  - Mon / Wed 1:30-2:50pm, Gates B03
- **Section/Office hours**
  - Fridays 10:30-11:30am, 3rd floor Conf Rm, MSOB X-393, 1265 Welch Rd
  - First one this Friday (focusing on setting up notebooks—please come!)

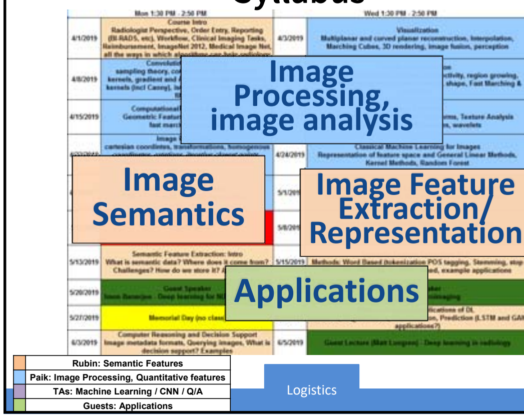
Logistics

Course Outline

- Overview of imaging modalities
- Visualization : From Machine to Screen
- Image Processing 101: (Filtering, Smoothing., Segmentation, Registration, Normalization)
- Feature Extraction (Quantitative and Semantic)
- Machine Learning (and deep learning) for Images
- Decision Support
- Clinical Applications

Logistics

Syllabus



Logistics

Grading

- **Grade Breakdown**
  - 3 Assignments 45% total (15% each)
  - Midterm exam 15%
  - Participation (piazza) 10%
  - Final project 30%
  - Total 100%
- **Class participation**: There are many different ways to participate, including but not limited to:
  - Attending Class
  - Attending TA Sections
  - Asking/answering questions on Piazza (discussion forum)

Logistics

Late Submission Policy

- Problem sets are due at or before 11:59 PM on the due date
- **You have 4 free late days total**
  - Not valid for final project assignments
- After that, **10% off your grade per day late** (late time within a day used is counted as a full late day-that is, it rounds up)

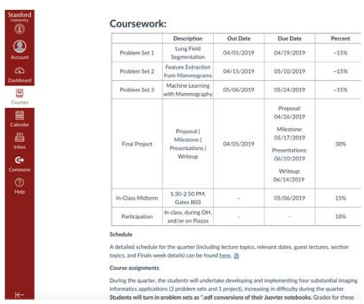
Logistics

Resources

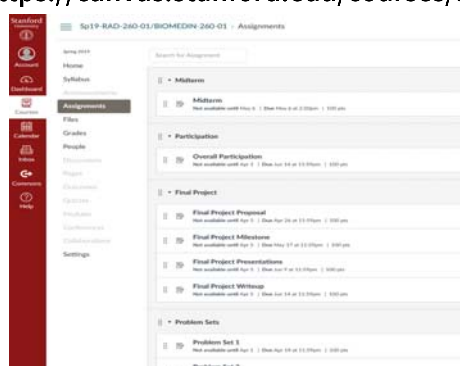
Resources

- Course Support
  - Canvas for files, assignments
  - http://bmi260.stanford.edu
  - Above link redirects to Canvas site, https://canvas.stanford.edu/courses/98045
  - Piazza for questions
  - https://piazza.com/class/jsf4uif8ocq2sb

http://bmi260.stanford.edu



https://canvas.stanford.edu/courses/98045



Interactive Teaching

Interactive Teaching

- Audience participation
  - Bring your laptop or smartphone
  - Real-time polls and quizzes
  - In class demos
- TA sessions
  - Assignment Help
  - Review of Material

Responding to Polls / Quizzes

- Open-ended or multiple choice questions
  - You send in the CODE indicating your response (with text for open-ended questions)
- Make sure you are connected on WiFi (not cellular)
- Via **phone**
  - IM the CODE to a number provided
  - e.g., 37607 (NB, your texting app must support shortcodes)
- Via **laptop**
  - Submit the CODE at http://PollEv.com
- SCPD and watching taped lectures
  - Answers provided in class, so you can play at home

Interactive Teaching

Multiple Choice Questions

What's your favorite course?

Text a **CODE** to **37607** | Submit a **CODE** to **http://PollEv.com**/imaging

BMI-217	**151954**
BMI-212	**151962**
BMI-260	**154091**
BMI-210	**154092**

0 1

Interactive Teaching

Let's try…

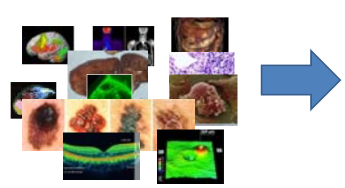
Interactive Teaching

What is your background?

Undergraduate	473219
Computer Science	473220
Biomedical Informatics	473221
Electrical Engineering	473224
Bioengineering	473242
Medical Student	473243
Postdoc	473244
Medical Trainee/Resident	473399
Professional Masters Student	473400
Other	473429

What are you hoping to learn from this class (enter 1-2 words)?

Our goal is to go from images to understanding…



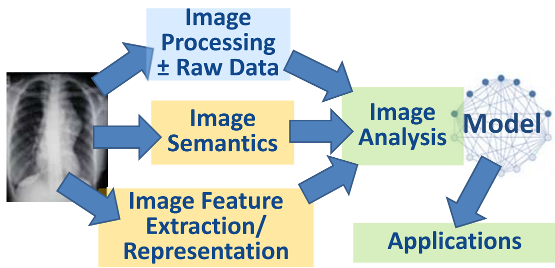
Diagnosis

Biomarker

Disease Progression

Biological Model

From Images to Understanding



Biomedical Informatics 260

Computational Methods for Image Analysis
Introduction to Imaging Modalities
Lecture 1

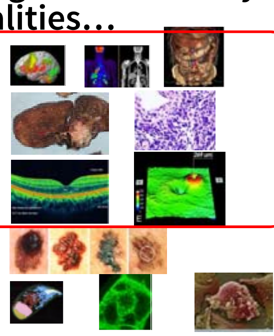
Outline of Lecture

- What are the medical imaging modalities?
- Why do we need computer help working with images?
- What are techniques for computer understanding of image data?
- What are example applications?

What are the medical imaging modalities?

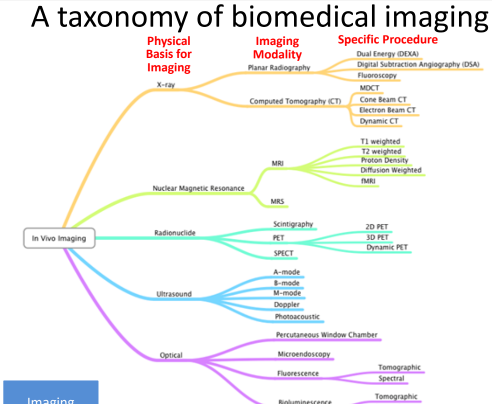
Medical imaging includes many modalities…

- Radiology
- Pathology
- Ophthalmology
- Dermatology
- Microscopy



Imaging Modalities

A taxonomy of biomedical imaging

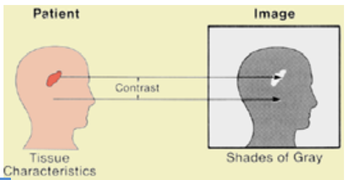


Slide courtesy David Paik

Imaging Modalities

Image contrast

- Medical imaging is the process of converting tissue characteristics into a visual image
- **Image contrast**: Difference in the image pixel values between closely adjacent regions on the image (seen by human or computer)



Imaging Modalities

Contrast resolution & spatial resolution in images

- **Contrast resolution**
  - The ability of the imaging modality to distinguish between differences in image intensity
  - Differs according to the physical principles governing image generation
- **Spatial resolution**
  - The ability of the imaging modality to visualize small objects
  - Differs according to the amount of signal generated by the modality (e.g., photon flux)

Imaging Modalities

Basis for contrast resolution depends on the imaging modality

- Radiography, CT
  - X-ray attenuation, absorption
- Nuclear medicine/molecular imaging
  - Uptake of targeted agents; attenuation/absorption
- Ultrasound
  - Sound transmission, reflection
- MRI
  - Proton relaxation (generates RF signal)

Imaging Modalities

Basis for spatial resolution depends on the imaging modality

- High signal flux → high resolution
  - Radiography
  - CT
  - MRI
- Low signal flux → low resolution
  - Nuclear medicine/molecular imaging
  - Ultrasound

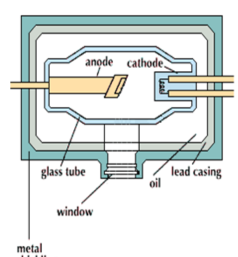
Imaging Modalities

Endogenous and exogenous factors in image contrast

- **Endogenous** differences in tissue properties
  - Xray, CT: Atomic number, density of tissue
  - MRI: Differences in T1/T2 relaxation
  - PET/NM: none without contrast agent
- **Exogenous** (pharmaceutical agents)
  - Intravenous or oral
  - Xray: Iodine, Barium
  - CT: Iodine
  - MRI: Gadolinium
  - PET/NM: Various radionuclide tagged compounds

Imaging Modalities

Radiography

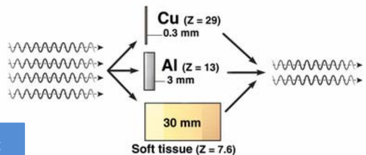


- High energy electromagnetic radiation
- Produced in a "cathode-ray tube"
- Pass through tissues or are absorbed based on the tissue composition

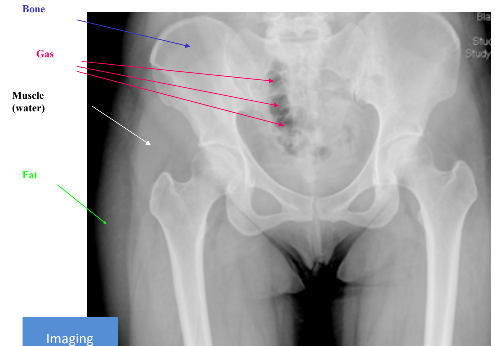
Imaging Modalities

Contrast resolution in X-ray imaging

- Contrast resolution based on differences in X-ray attenuation
  - Atomic number (calcium, iodine, lead)
  - Tissue density/thickness
- Bright on image: Bone, contrast agent, markers, foreign bodies, very dense tissue
- Dark on image: Air/gas



A Pelvic "X-ray"

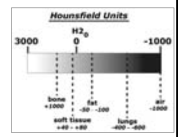


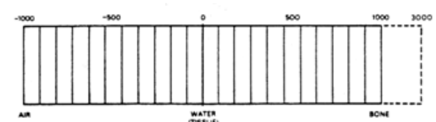
Computed Tomography

- Row of x-ray detectors
- X-ray cross-sections
- Tube-detector apparatus rotates 360º
- Computer reconstruction of xray attenuation in patient

Hounsfield units

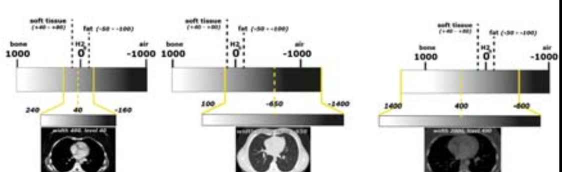
- Each pixel in CT images is assigned a number (Hounsfield unit) that is related to the linear attenuation coefficient (μ) of tissue within each voxel
- Water defined HU = 0
- CT Number = $1000 \cdot \frac{\mu_{object} - \mu_{water}}{\mu_{water}}$
- Large range of HU in CT images:





Limitation in viewing all the CT numbers

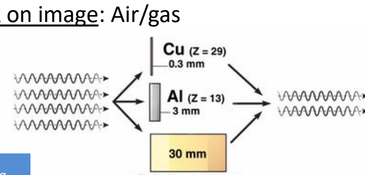
- Human eye cannot appreciate full range of CT numbers
- Visualized by setting window width and level



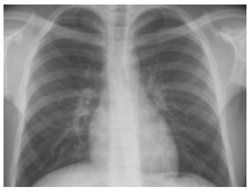
"Soft tissue window" "Lung window" "Bone window"

Contrast resolution in CT imaging

- As with radiography, contrast resolution based on differences in X-ray attenuation
  - Atomic number (calcium, iodine, lead)
  - Tissue density
- Bright on image: Bone, contrast agent, markers, foreign bodies
- Dark on image: Air/gas



Contrast resolution much higher in CT



Chest X-Ray

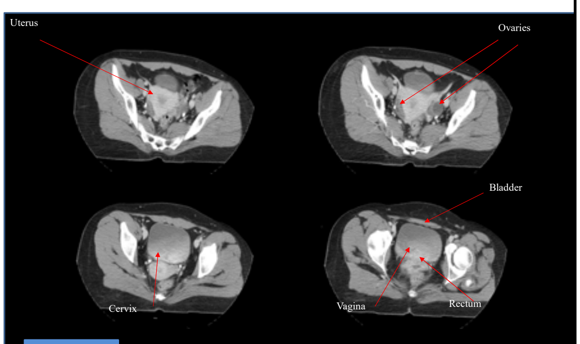
CT

Imaging Modalities

Superiority in contrast resolution on CT

Why need computer help?

CT studies have lots of slices



Imaging Modalities

Ultrasound

- High frequency (>1 MHz) sound
- Echoes of tissue interfaces
- Image in real time/motion
- Blood flow imaging without contrast agents
- No ionizing radiation
- Any scan plane

Imaging Modalities

Contrast resolution in ultrasound imaging

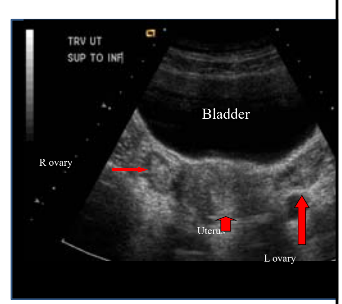
- Contrast resolution based on differences in sound impendence among tissues
  - Homogeneous tissue transmits sound
  - Heterogeneous tissue or gas reflects sound
- Bright on image: Tissue interfaces, stones/calcification, gas, heterogeneous tissue
- Dark on image: Fluid filled structures, homogenous solid organs

Imaging Modalities

Ultrasound image

Grayscale reflects amplitude of echoes generated at tissue interfaces

- Fluid filled structures: black
- Solid tissue: gray
- Fat: white
- Calcifications, air: white with "shadow"



Imaging Modalities

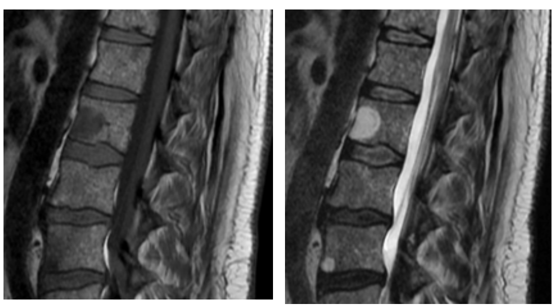
Magnetic Resonance Imaging (MRI)

- Basis for imaging: radio waves emitted from patient
- Superb and varied soft tissue contrast
- Any scan plane
- Superb depiction of musculoskeletal system and bone marrow
- Used to characterize pathology seen on ultrasound or CT

Contrast resolution in MRI

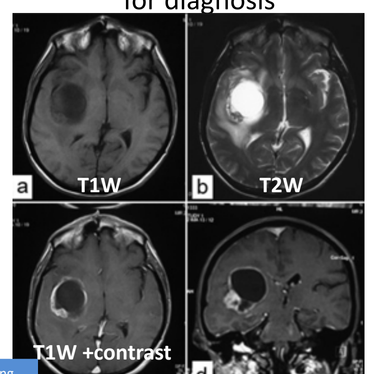
- Contrast resolution based on differences in T1, T2 relaxation and proton density among tissues, as well as the pulse sequence used for imaging
  - "T1 weighted" emphasizes T1 contrast
  - "T2 weighted" emphasizes T2 contrast
- Bright on T1 image: Iron deposits, contrast agents, fat
- Dark on T1 image: Fluid filled structures, iron deposits
- Bright on T2 image: Fluid filled structures, hemorrhage
- Dark on T2 image: Iron deposits, bone, stones, flow void

Magnetic Resonance Imaging (MRI)



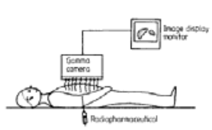
Spine: T1 weighted   Spine: T2 weighted

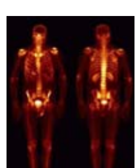
Brain MRI: Multi-modal imaging important for diagnosis

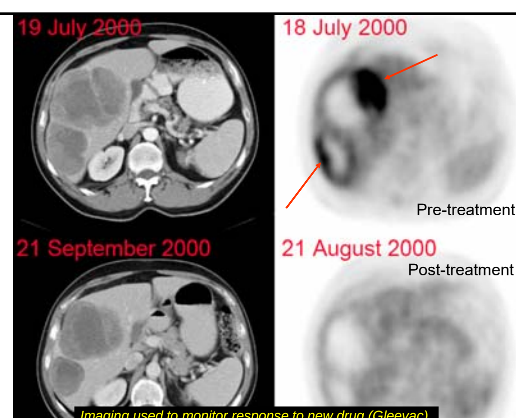


Nuclear medicine (molecular) imaging

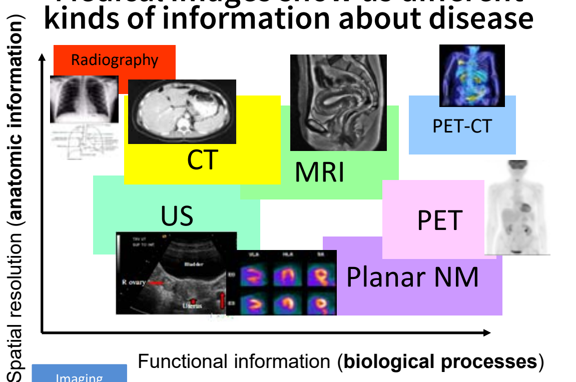
- *Injection of radioactive probes* that are metabolically-active
  - Image wide range of molecular events
  - Varying disease sensitivity/specificity
  - ^{18}FDG glucose: a radio-labeled probe for glucose uptake (increased uptake in tumors and inflammation)
- Imaging with scintillation camera (planar), detector array (SPECT/PET), optical, fluorescence







Medical images show us different kinds of information about disease



Example case

- 55 y.o. runner with hip pain; rule-out fracture
- Plain film: normal; no fracture
- Can we see anything else to explain the pain?

The abnormality is seen better on MRI…

- MRI: more sensitive to marrow changes
  - –T1-weighted image
  - –Dark area=edema

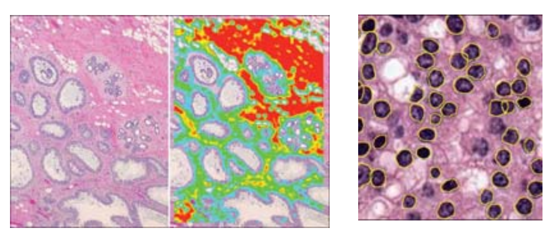
Digital pathology

- Scanner converts whole slide (or tissue microarray) into digital images
- Huge size
- Opportunities for computational analysis, disease classification, etc.

Contrast resolution in digital pathology

- Contrast resolution based on differences staining of tissues
- Different kinds of stains
  - H&E: Most common; nuclei purple, blood red, cellular elements pink
  - Special stains for specific kinds of proteins
- Multi-resolution images (40x – 1x) is unique aspect
- White on the image: non tissue areas
- Color on images: based on tissue stain (nuclei are purple)

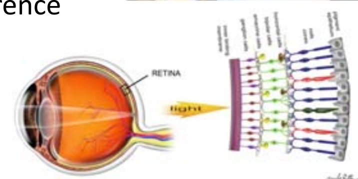
Pathology images: H&E stain



Low power: Tissue regions | Low power: Image segmentation showing different tissue regions | High power: Cell/nuclear regions

Ophthalmology: Fundus photography

- Photograph of retina
- Full view of retina
- Color images
- Lower resolution than optical coherence tomography



Imaging Modalities

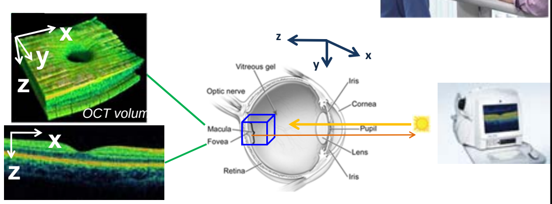
Retinal fundus photography

Healthy Retina | Diabetic Retinopathy

Imaging Modalities

Optical Coherence Tomography (OCT)

- A non-invasive 3D imaging technique
- Standard of care since 1991
- Emit lights into the eye; measure reflectivity of tissues within a target cube
- Visual rendering the retinal structures



Imaging Modalities

Why do we need computer help working with images?

Why do we need computer help?

1. Too many images to look at
2. Variation in image interpretation
3. Image data are *unstructured*

Why need computer help?

Challenge 1: Too many images to look at (data explosion)…

2001 | 2009

Why need computer help?

Challenge 2: Variation in Image Interpretation

- Radiologists are not perfect
  - Missed observations in images
  - Misinterpretation of observations in images
  - Both cause incorrect diagnoses
- Variation in practice produces suboptimal quality of care

Why need computer help?

Radiologist performance is variable

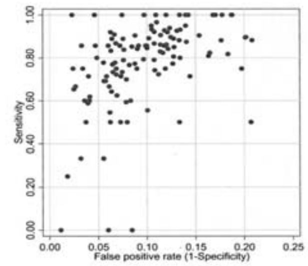
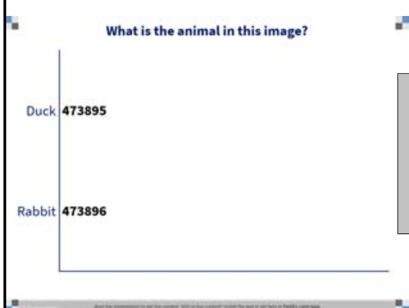


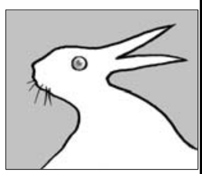
Fig. 1. True positive rate (sensitivity) of the 124 radiologists versus the false positive rate (1 − specificity). Rates not adjusted for patient variables.

Barlow-WE, et al, JNCI 96(24):1840-50, 2004

Why need computer help?

Image interpretation is variable





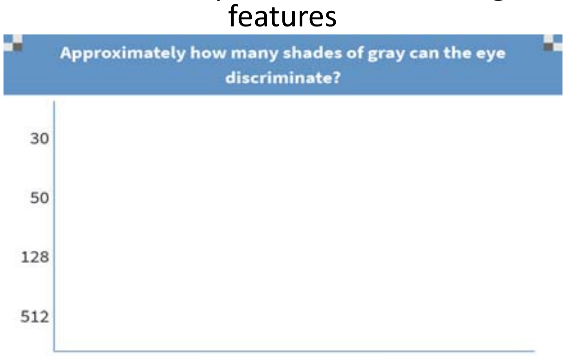
Why need computer help?

Radiology: Errors in human perception…

Click on the middle of abnormality in the image

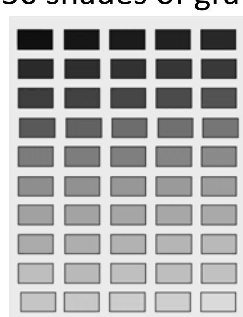
Why need computer help?

Human ability to discriminate image features



Why need computer help?

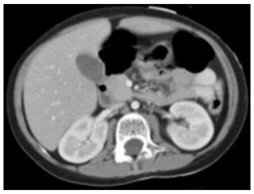
50 shades of gray



Can you distinguish among all pairs?

Why need computer help?

Errors in interpretation: What's the abnormality?



Normal CT

Abnormal CT

Challenge 3: Image data are *unstructured*

- **Structured data**
  - Data whose meaning is explicit
  - Machine-accessible and interpretable
  - Format: attribute-value pairs, controlled terms
  - Clinical data, biological data
- **Unstructured data**
  - Lack explicit meaning; no inherent structure
  - Limited machine-accessibility
  - Format: binary objects, narrative text, numbers
  - Images and tests (the majority of biomedical data!)

Why need computer help?

Examples of structured data

- **Biological data**
  - A, C, T, G, U (DNA, mRNA)
  - Ala, Asx, Lys, Pro, Gly,… (proteins)
  - (gene expression)

CompositeSequence Identifier	dChip Normalized(TSK15503)(H_TSK15503)	dChip Normal
AFFX-BioB-5_at	319.57	
AFFX-BioD-M_at	365.678	
AFFX-BioB-3_at	360.075	

  - ATOM PRO 21.2 21.5 17.4 (protein structure)
- **Medical data**
  - K=2.4, Cl=112, Hgb=13.2,… (lab values)
  - ICD 112.3, CPT 11707 (diagnostic codes)
  - BP 119/80, HR 77, RR 12 (vital signs)

Why need computer help?

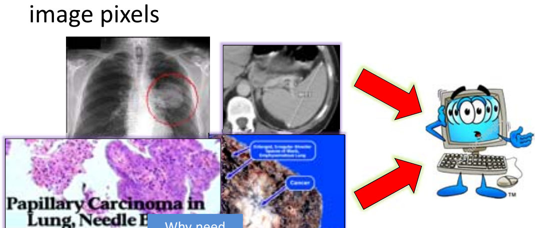
Two key types of unstructured data related to images

- Images themselves (sets of pixels)
- Free text (e.g., radiology and pathology reports that describe *qualitative features* and *diagnoses*)

Why need computer help?

Image contents are unstructured

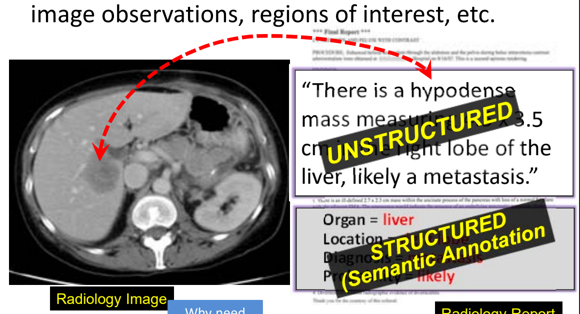
- Images are just pixels; lack *knowledge* about their contents
- Computer vision methods access/process image pixels



Why need computer help?

Texts related to images are unstructured

Things radiologists say about images: Anatomy, image observations, regions of interest, etc.



Why need computer help?

In Summary

Exploding data…

Challenging features…

Variable interpretation…

Unstructured data…

What to do?

Why need computer help?

What are techniques for computer understanding of image data?

Imaging informatics:

"The application of computer science methods to the challenges of medical imaging"

(This includes "AI")

Tackling the challenges

We will learn methods for deriving structure from unstructured data

- **Image (and text) annotation by humans ("semantic annotation") to structure human-observed image features**
- **Image processing (and natural language processing) techniques to extract features from images/texts**

Tackling the challenges

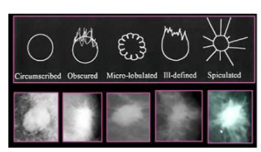
Extracting quantitative image features

- Two approaches for deriving structure from images:
  - **Extract pre-defined (hand-crafted) features**
  - **Unsupervised feature learning** (data-driven)
- N.B.: Image classification using deep learning is an example of deriving structure from images

Tackling the challenges

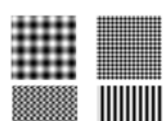
Pre-defined image features

Shape:



Edge:

Texture features: (characterize lesion interior)



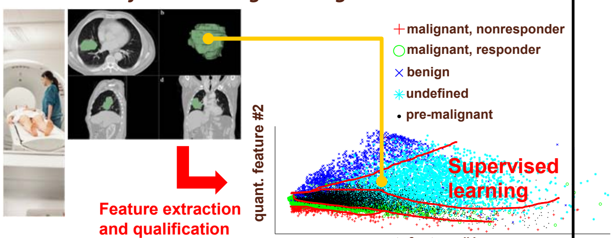
scale

directions

Tackling the challenges

Machine learning with pre-defined image features

***Radiomics*: 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.

Tackling the challenges

Unsupervised feature learning

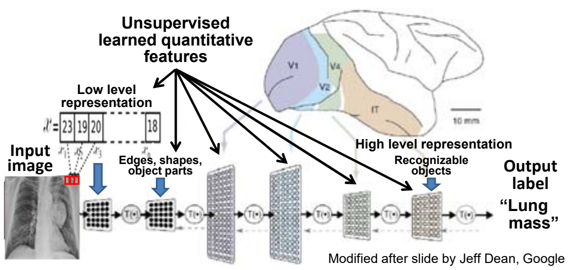
- *Raw image pixel data* input into a model
  - Image patch analysis
  - Deep learning
  - Word embeddings

} **Image data** (Image patch analysis, Deep learning)
} **Text data** (Word embeddings)

Tackling the challenges

Unsupervised feature learning: Deep learning

- High-level abstractions of image features (hierarchical, non-linear transformations)
- Inspired by hierarchical visual processing by the brain
- Higher-level features (layers) are defined from lower-level ones, and represent higher levels of abstraction



Modified after slide by Jeff Dean, Google

Tackling the challenges

Word embeddings

f(CT HEAD WITHOUT CONTRAST: X/X/XXXX XX:XX PM CLINICAL HISTORY: 114 years of age, Unknown, Stroke Code. COMPARISON: None. PROCEDURE COMMENTS: CT of the head was performed without IV contrast. Dose information: Based on a 16 cm phantom, the estimated radiation dose (CTDIvol [mGy]) for each series in this exam is 0.29, 0.29, 58.14 The estimated cumulative dose (DLP [mGy-cm]) is 1066. FINDINGS: Parenchyma: Diffuse subarachnoid hemorrhage. Intraparenchymal hemorrhage in the inferior right frontal lobe. Ventricles and extra-axial spaces: Mild hydrocephalus. Mastoid air cells: Clear. Bones: No focal abnormality. Additional comment: Right forehead scalp soft tissue swelling. IMPRESSION: 1. Diffuse subarachnoid hemorrhage. 2. Intraparenchymal hemorrhage in the right inferior frontal lobe. 3. Mild hydrocephalus. Discussed with Dr. XXXX by Dr. XXXX on X/X/XXXX at XX:XX PM. I have personally reviewed the images for this examination and agreed with the report transcribed above.) = y

Word embedding provides vector-based representation of text (learned using unsupervised methods);

e.g., to permit learning a classifier for document *x* being classified to label y

Tackling the challenges

What are example applications?

You will build such applications too!

Key clinical uses of unsupervised feature learning

- Disease detection
- Lesion segmentation
- Diagnosis

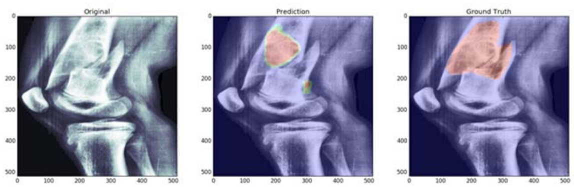
} ***Current applications***

- Treatment selection
- Response assessment
- Clinical prediction (of treatment response or future disease)

} ***Active research area***

Applications

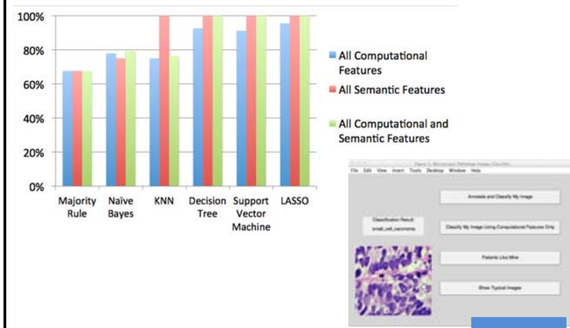
Automated detection of bone tumors



Yi, 2016

Applications

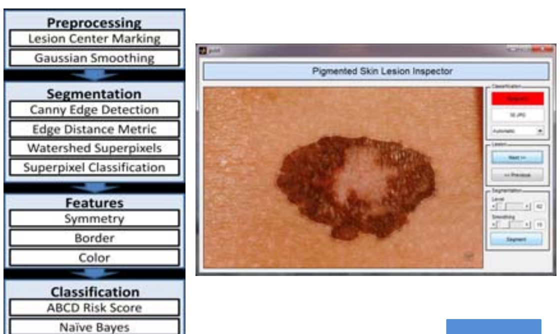
Classification of Malignant vs. Benign Tumors with Semantic and Computational Features



Yu, Tuo, 2013

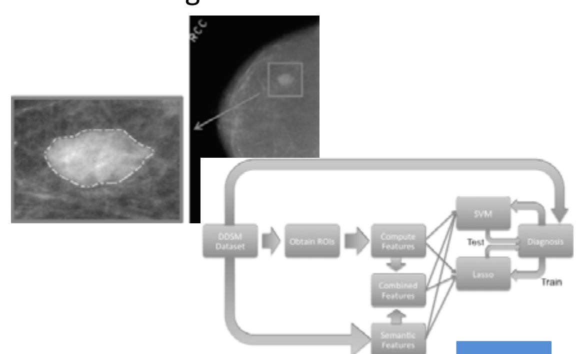
Applications

A skin lesion diagnostic tool



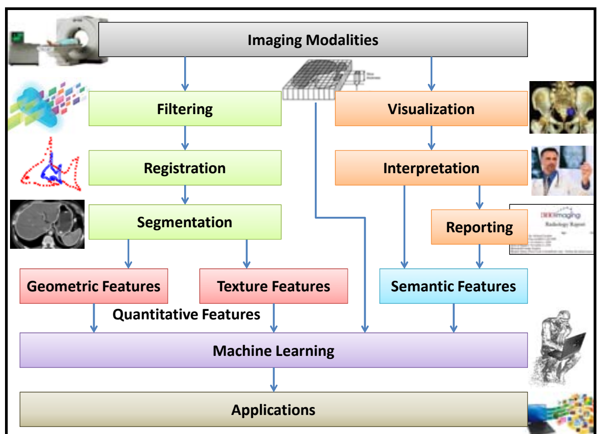
Applications

Automated segmentation and classification



Applications

Putting this in the larger context…



To Summarize…

- Medical imaging is broad
- Many different imaging modalities
- The basis for image contrast varies with modality

Medical Imaging?

- We need computer help because:
  - Too many images to look at
  - Variation in image interpretation
  - Image data are *unstructured*

Why need computer help?

For computers to understand images:

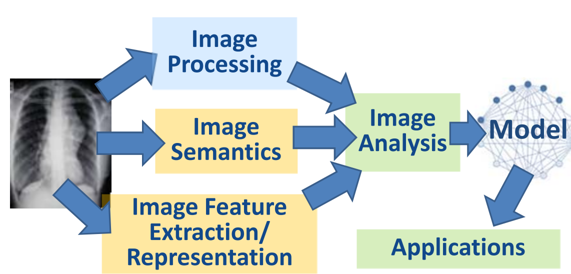
1. They must recognize and extract **image features**
2. These image features are both **quantitative** and **qualitative**
3. Extracting these features is challenging and requires specialized techniques (which you will learn in this class…)

Tackling the challenges

Many cool imaging applications can be created to meet important clinical needs

Applications

From Images to Understanding



Thank you!

Next time:
Image visualization