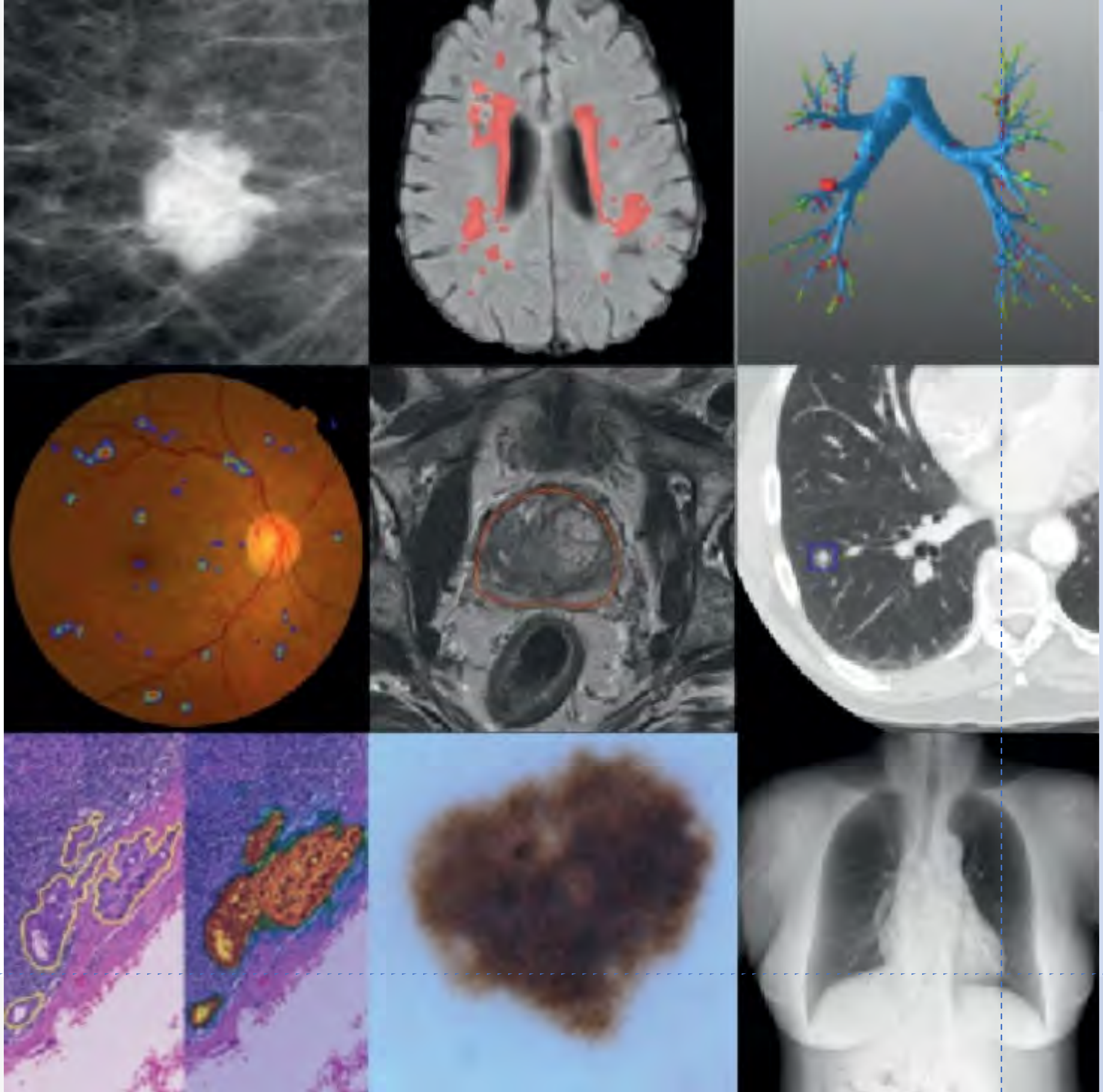
Big Data Seminar on Imaging

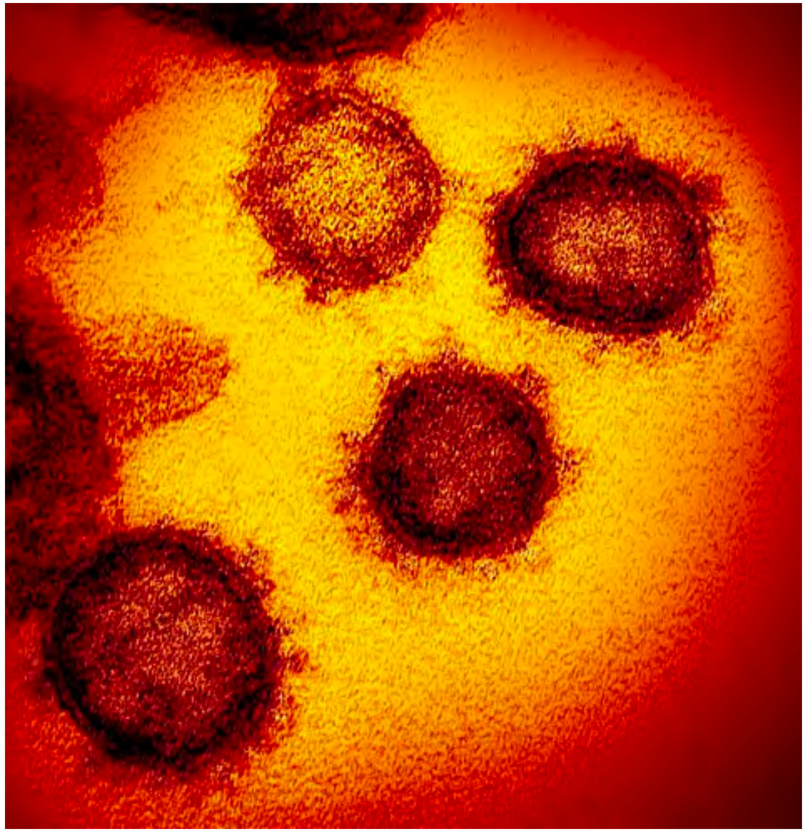
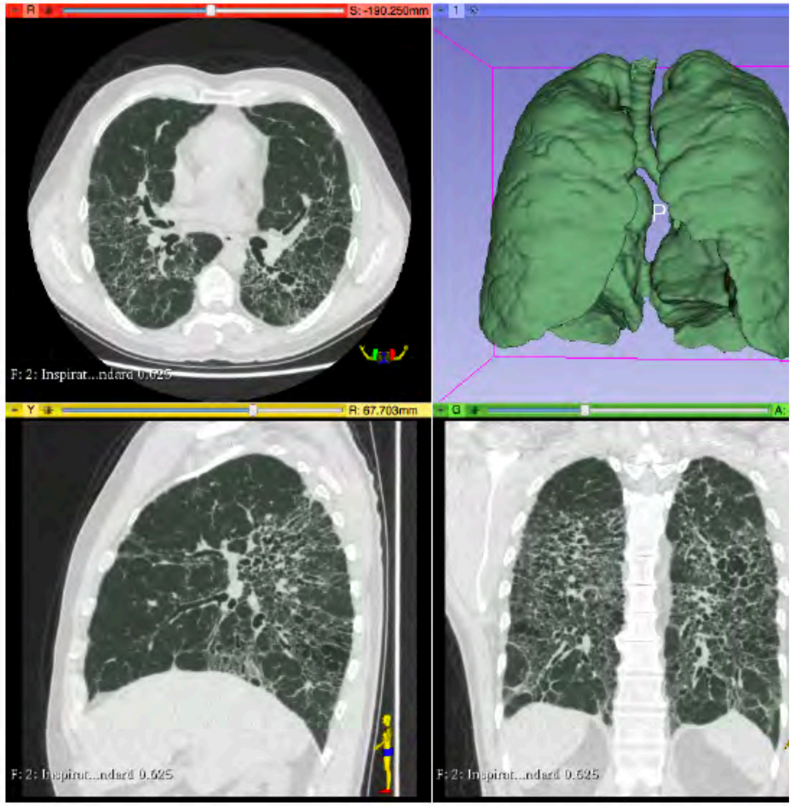
February 18, 2020

Sponsored by the Biostatistics, Epidemiology, and Research Design (BERD) core of the CCTSI



Overview

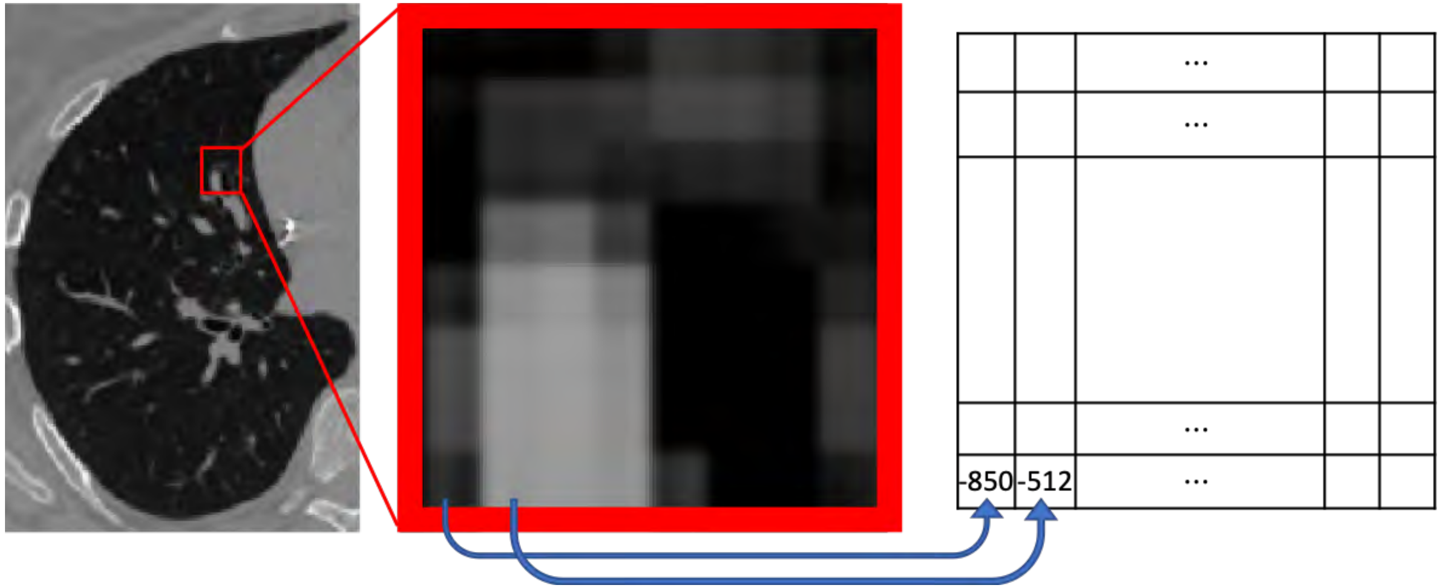
- Introduction to Medical Imaging
  - Sarah M. Ryan, PhD Candidate
- Deep Learning and Its Applications in Medical Image Analysis
  - Fuyong Xing, PhD
- Application of Radiomics to Lung Disease to Develop a Novel Biomarker of Sarcoidosis
  - Lisa A. Maier, MD, MSPH, FCCP

Introduction to Medical Imaging

Imaging Basics

- An image is a visual representation of a mathematical function, say f(x,y), where f may give the intensity or color at point (x,y)
- Each 2D image is comprised of picture elements, called pixels, which form a matrix. In 3D, images are comprised of volumetric elements, called voxels

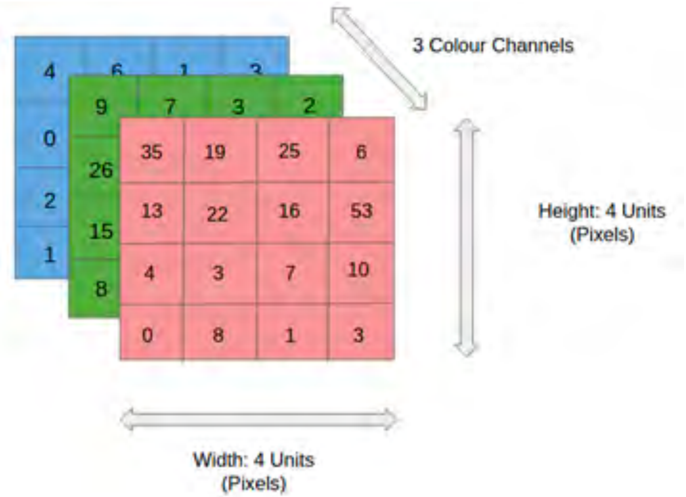


Medical Imaging Modalities

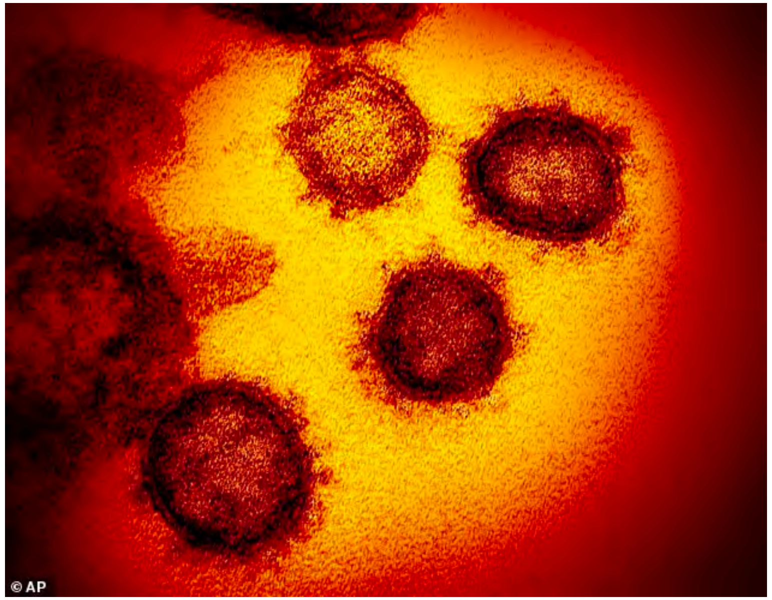
- Technologies that are used to view the anatomical areas in order to diagnose, monitor, or treat medical conditions
- Common imaging modalities include:
  - Microscopy
  - Radiography
  - Computed Tomography (CT)
  - Magnetic Resonance Imaging (MRI)
- Each modality uses different techniques to represent what we see

Microscopy

- A technique using properties of light to illuminate different cellular structures
- Results in a 2D image
  - Grayscale -> one number per pixel
  - RGB -> 3 numbers per pixel



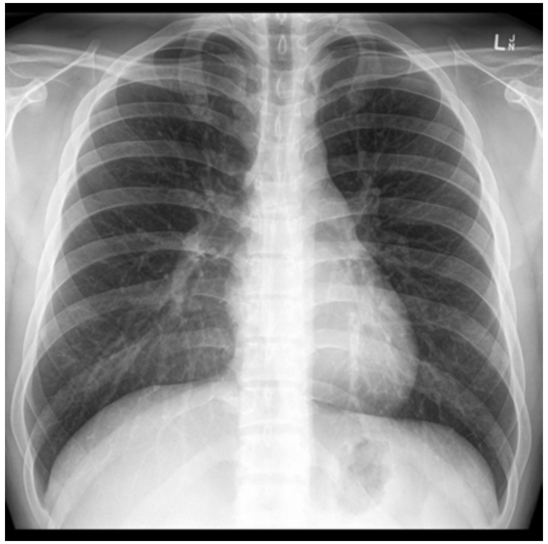
https://medium.com/@raycad.seedotech/convolutional-neural-network-cnn-8d1908c010ab



https://www.niaid.nih.gov/diseases-conditions/coronaviruses

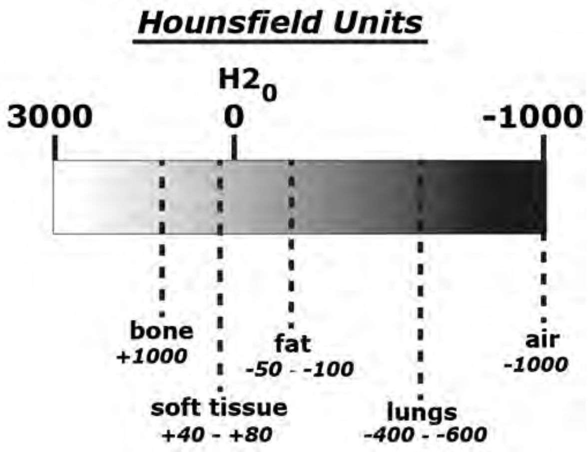
Radiography

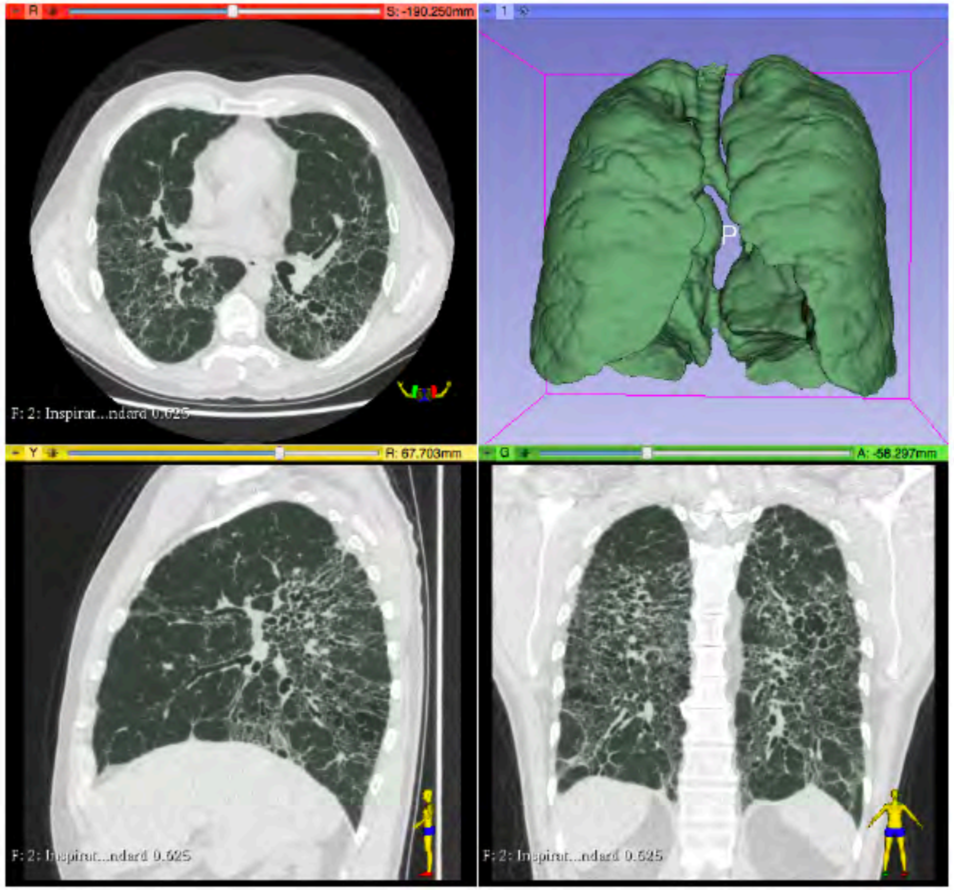
- An imaging technique using x-rays, gamma rays, or similar ionizing (or non-ionizing) radiation to visualize internal anatomical structures
- Commonly used to detect pathology in bones, lungs, and breasts
- Produces a 2D grayscale image



Computed Tomography (CT)

- A computerized x-ray imaging procedure which generates cross-sectional images of the body at various angles
- Results in a 3D image with each voxel corresponding to the radiodensity of tissue, measured in Hounsfield units





Magnetic Resonance Imaging (MRI)

- An imaging procedure which measures the response of atomic nuclei to high-frequency radio waves when placed in a strong magnetic field
- Produces 3D images of internal anatomical structures, commonly brain
  - Structural -> Produces T1 and T2 3D images, corresponding to magnetization states
  - Functional -> Produces multiple 3D images over time

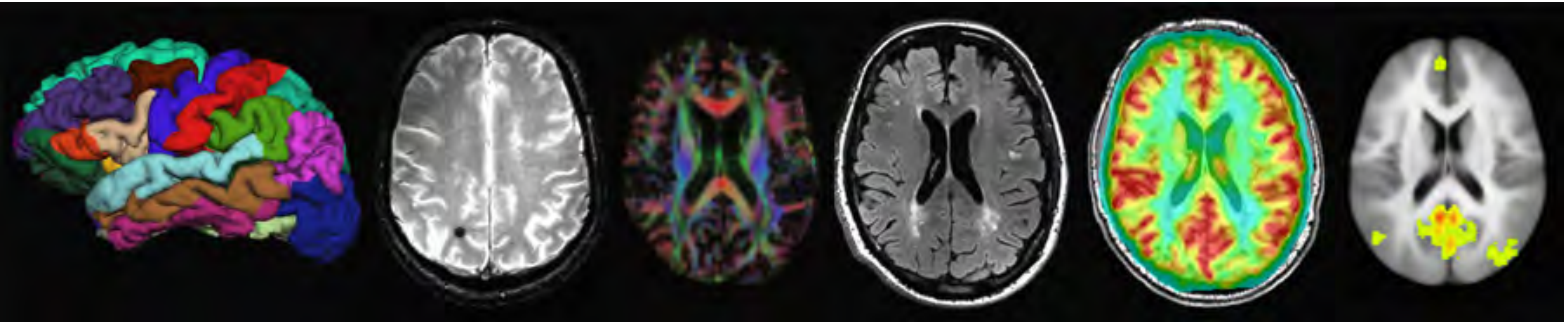
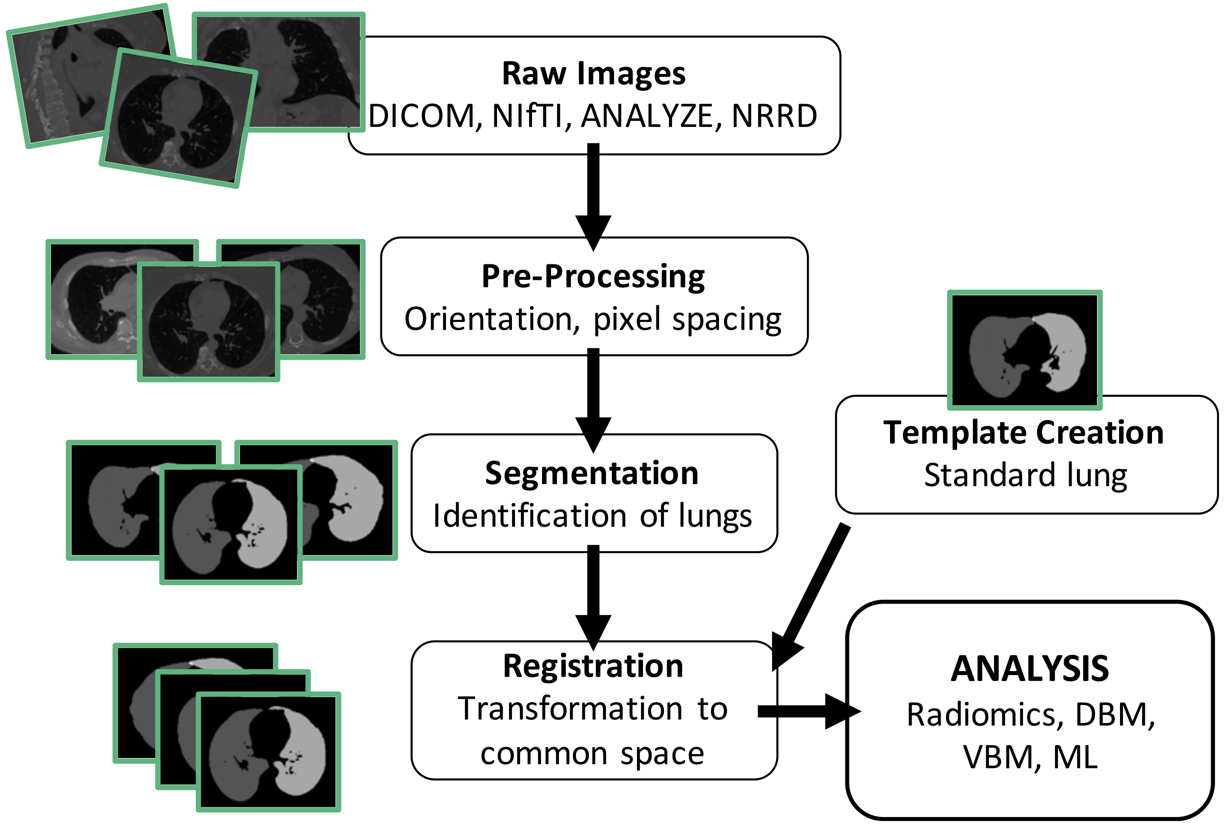


Image Processing

- Image processing techniques are necessary to enhance image features prior to analysis

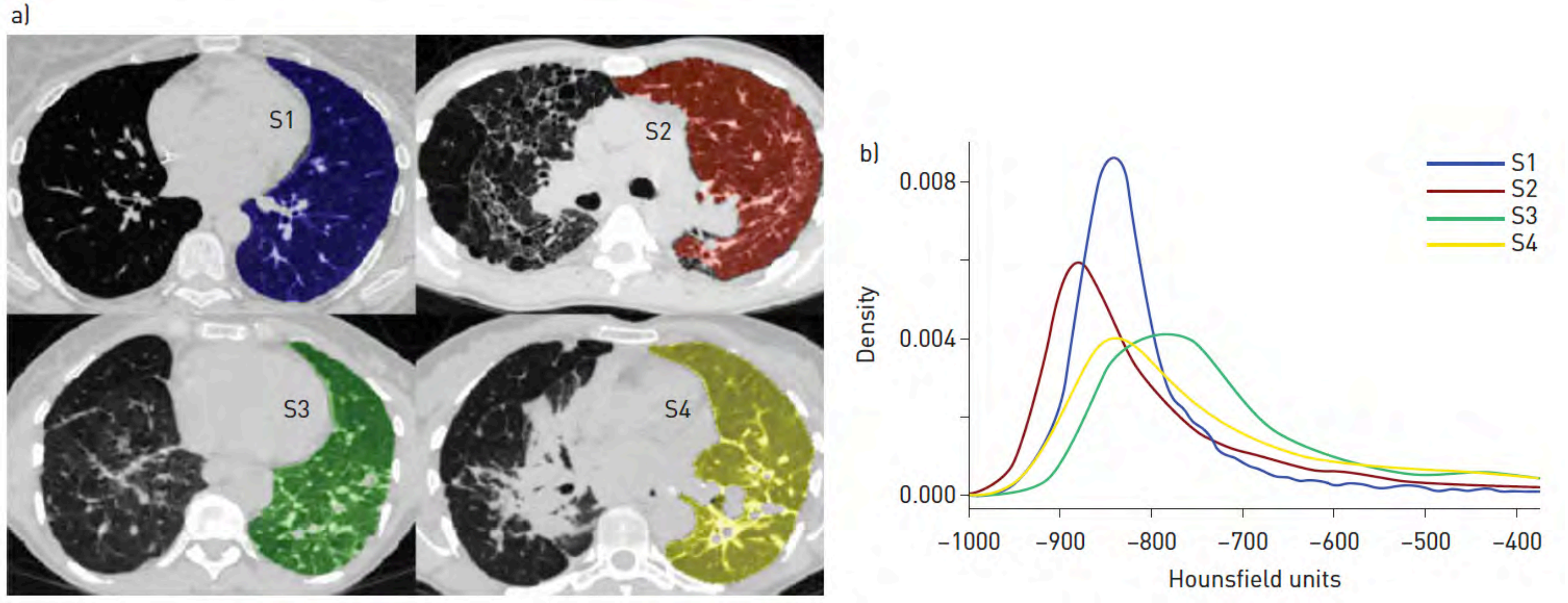
- This may include:
  - Removal of scanner effect
  - Removal of physiological effects (e.g., breathing, heartbeat, movement, etc.)
  - Identification and segmentation of region of interest
  - Alignment of spatial coordinates across images (aka registration)



https://github.com/ryansar/lungct

Image Feature Extraction

- **Radiomics** - An emerging field in which large numbers of quantitative features are computed from medical images, providing a rapid, objective, and sensitive quantification of abnormalities



[Radiomic measures from chest high-resolution computed tomography associated with lung function in sarcoidosis]

Deep Learning and Its Applications in Medical Image Analysis

Fuyong Xing

Department of Biostatistics and Informatics
Colorado School of Public Health
University of Colorado Anschutz Medical Campus

Big Image Data

https://medium.com/abraia/first-steps-with-transfer-learning-for-custom-image-classification-with-keras-b941601fcad5

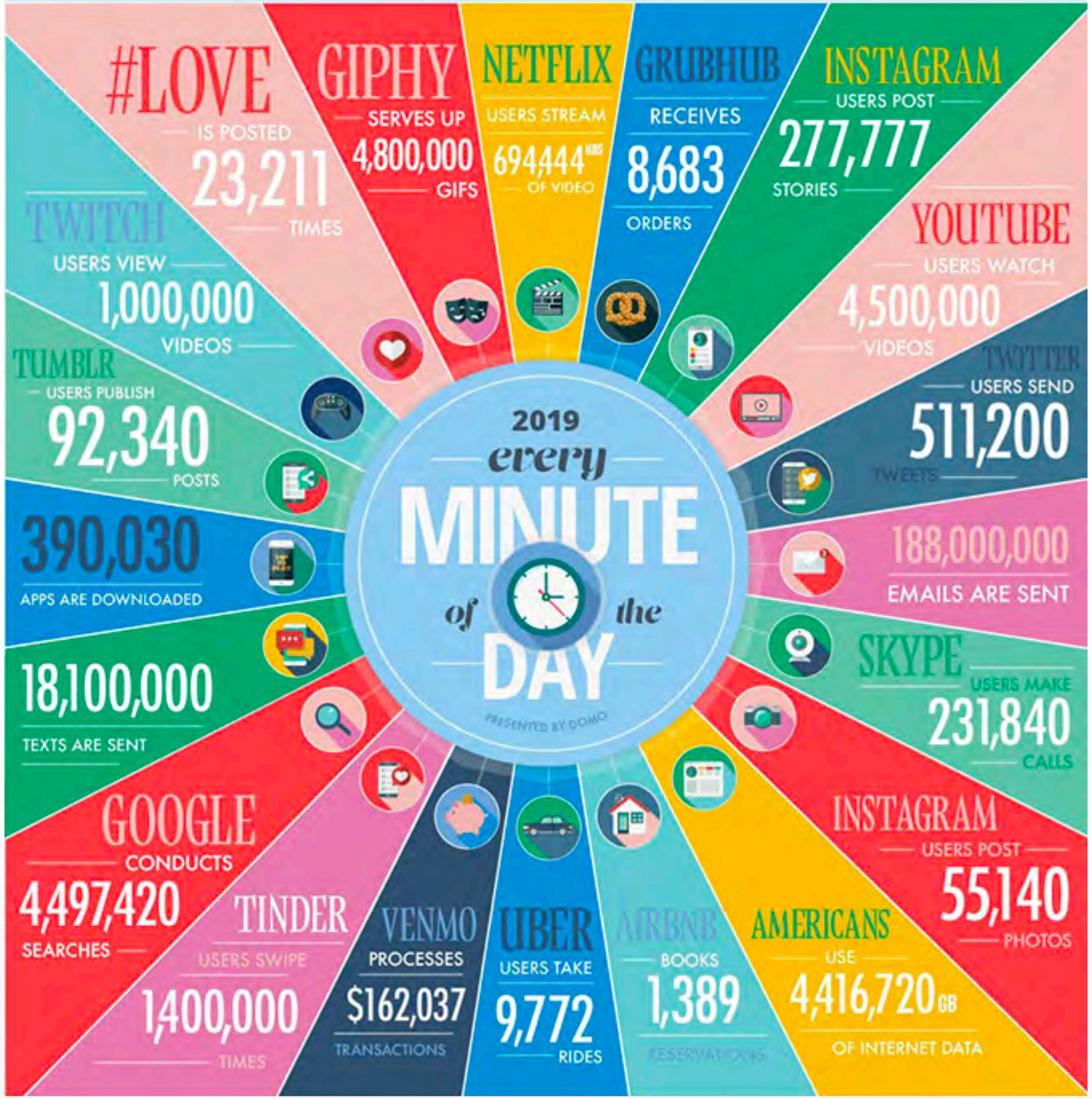
Big Image Data

- ImageNet: a large-scale image dataset for computer vision
- 10,000,000+ labeled images, 20,000+ object categories
- Annual ILSVRC Challenges (up to 2017):
  - 1000 object categories
  - 1.2M training, 50k validation, and 100k testing images

O. Russakovsky *et al.* “ImageNet Large Scale Visual Recognition Challenge”, *IJCV*, 2015

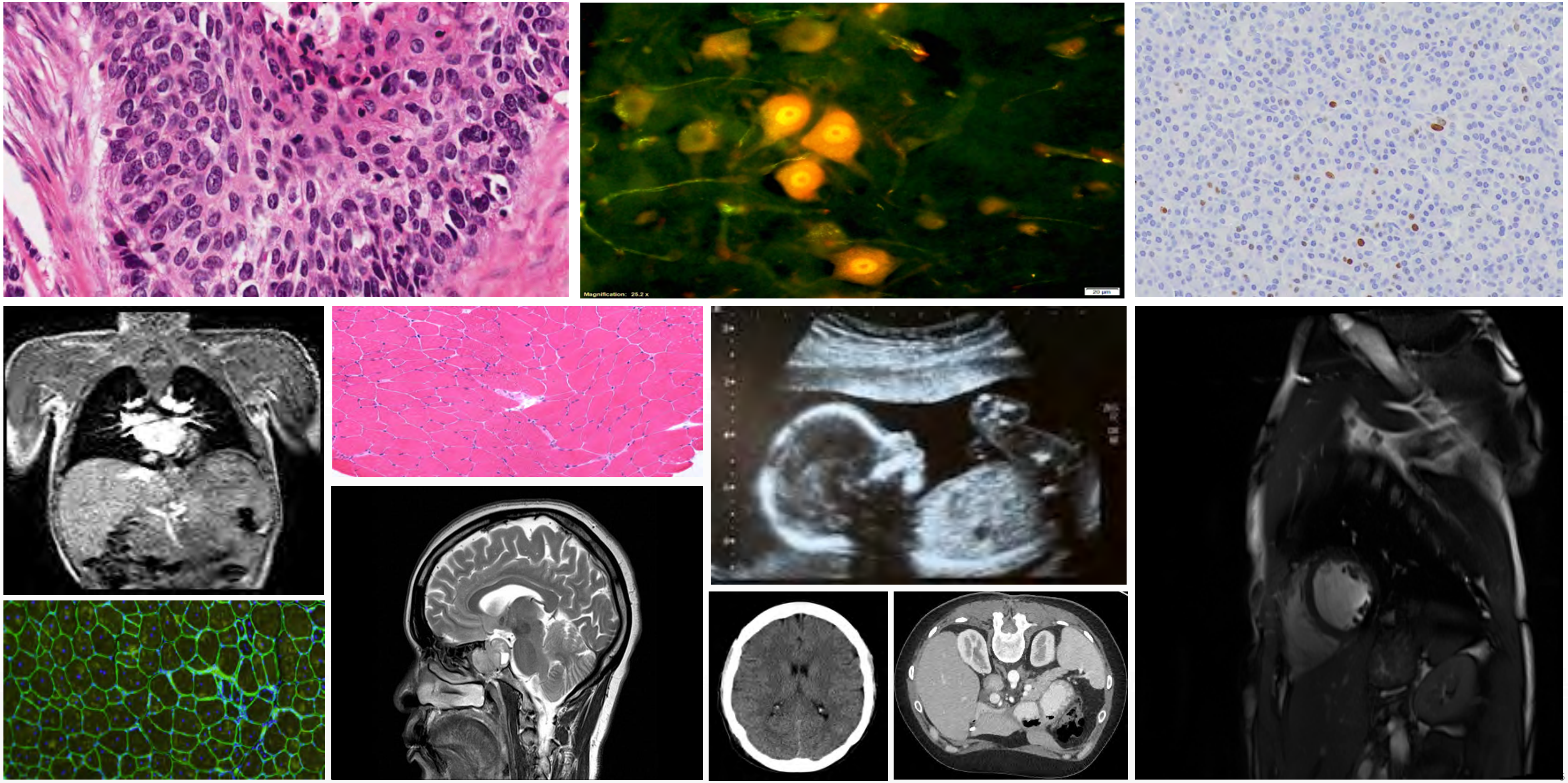
Data Never Sleeps

- How much data is generated every minute?
  - Instagram: over 55000 photos shared



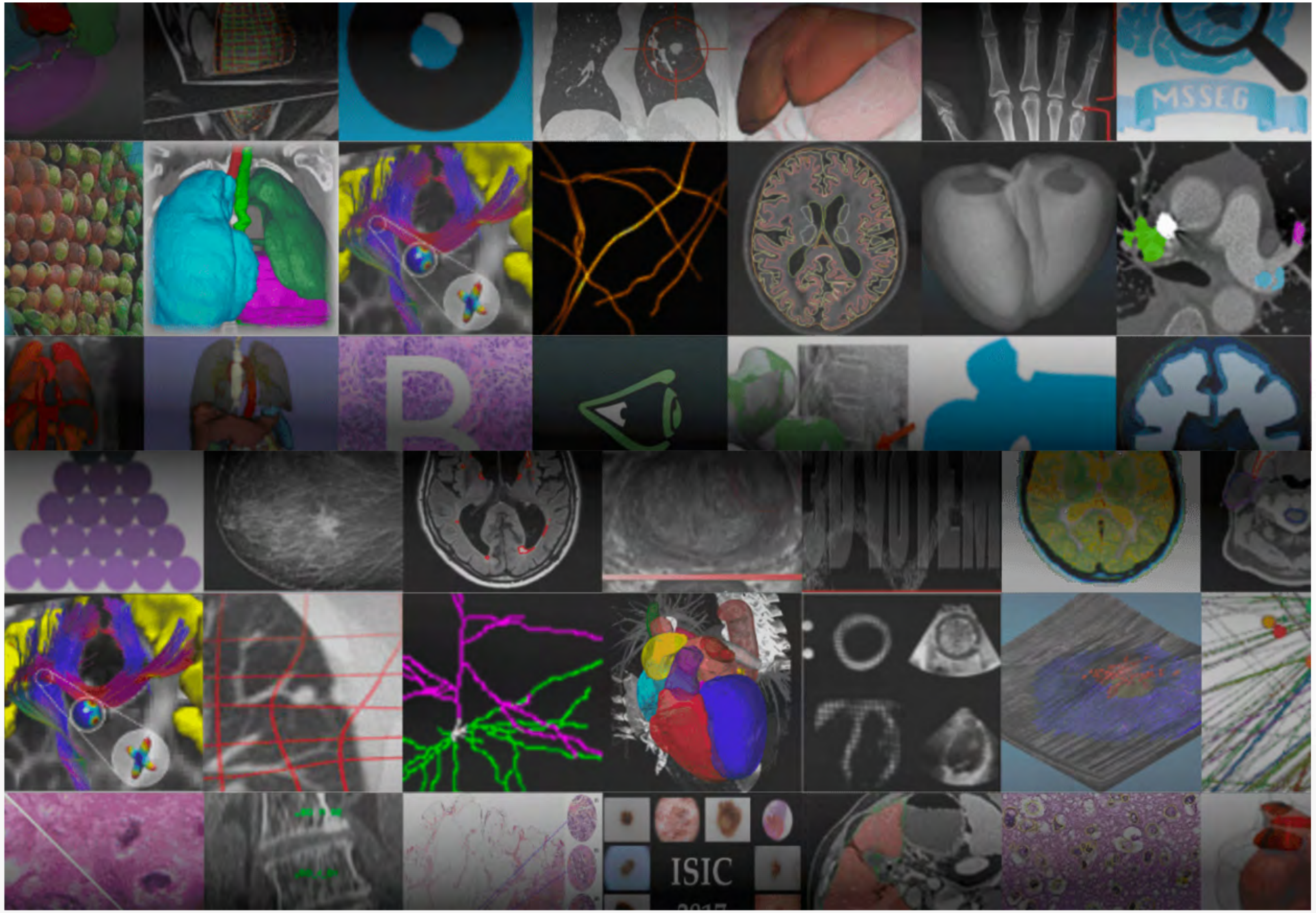
https://www.domo.com/learn/data-never-sleeps-7

Biomedical Images Are Everywhere



Grand Challenges in Biomedical Image Analysis

https://grand-challenge.org/



What Methods Used to Analyze Large-Scale Image Data?

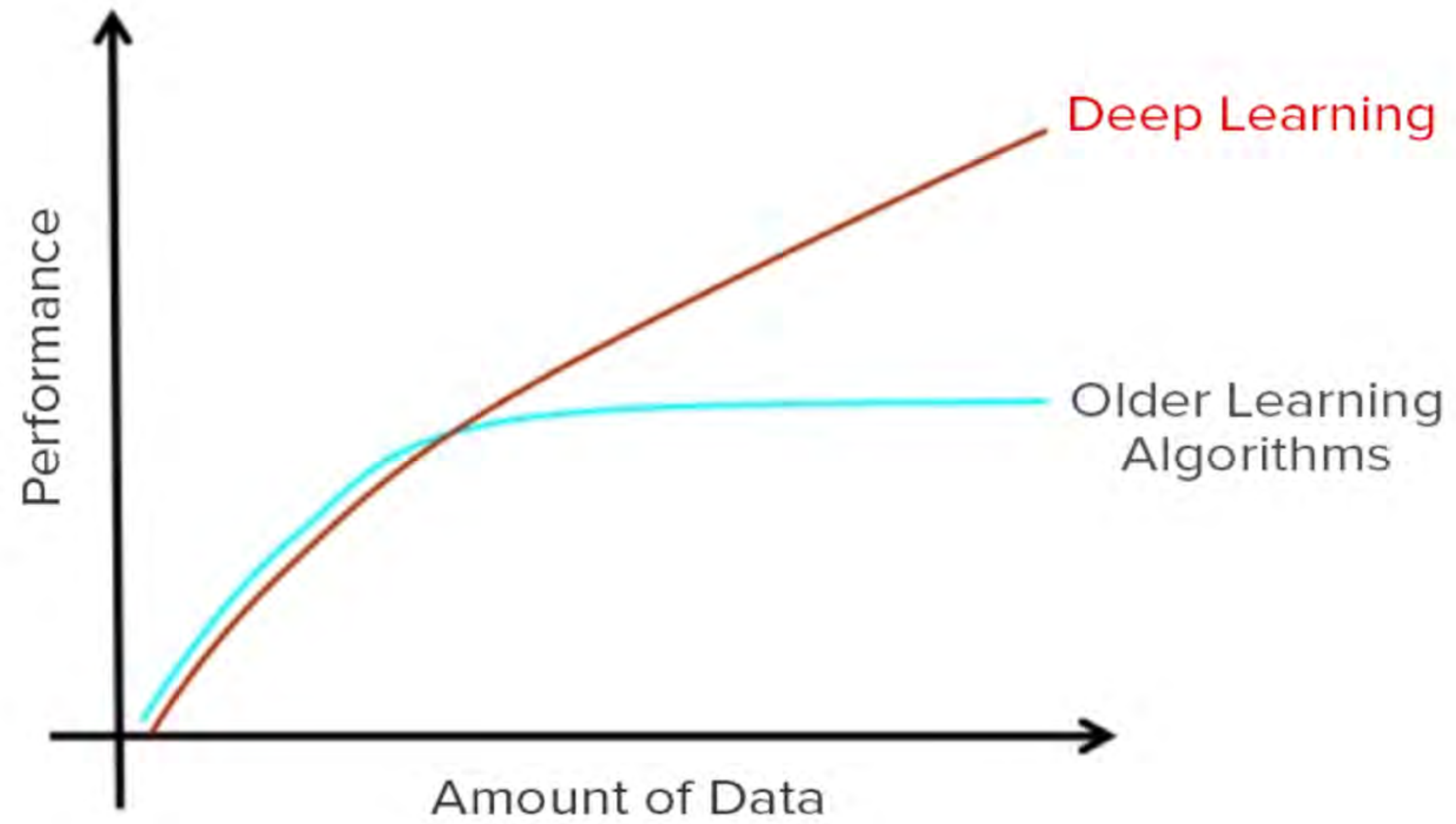
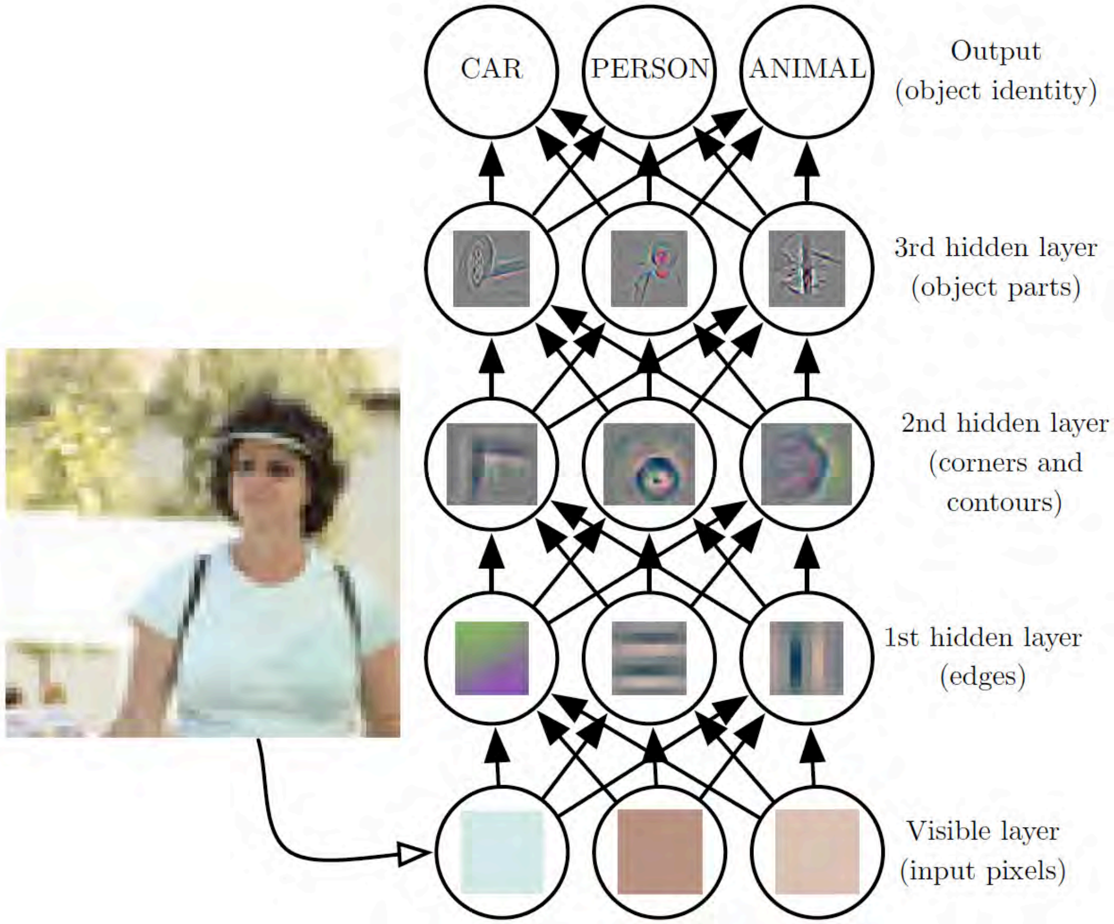


Image credit: Andrew Ng

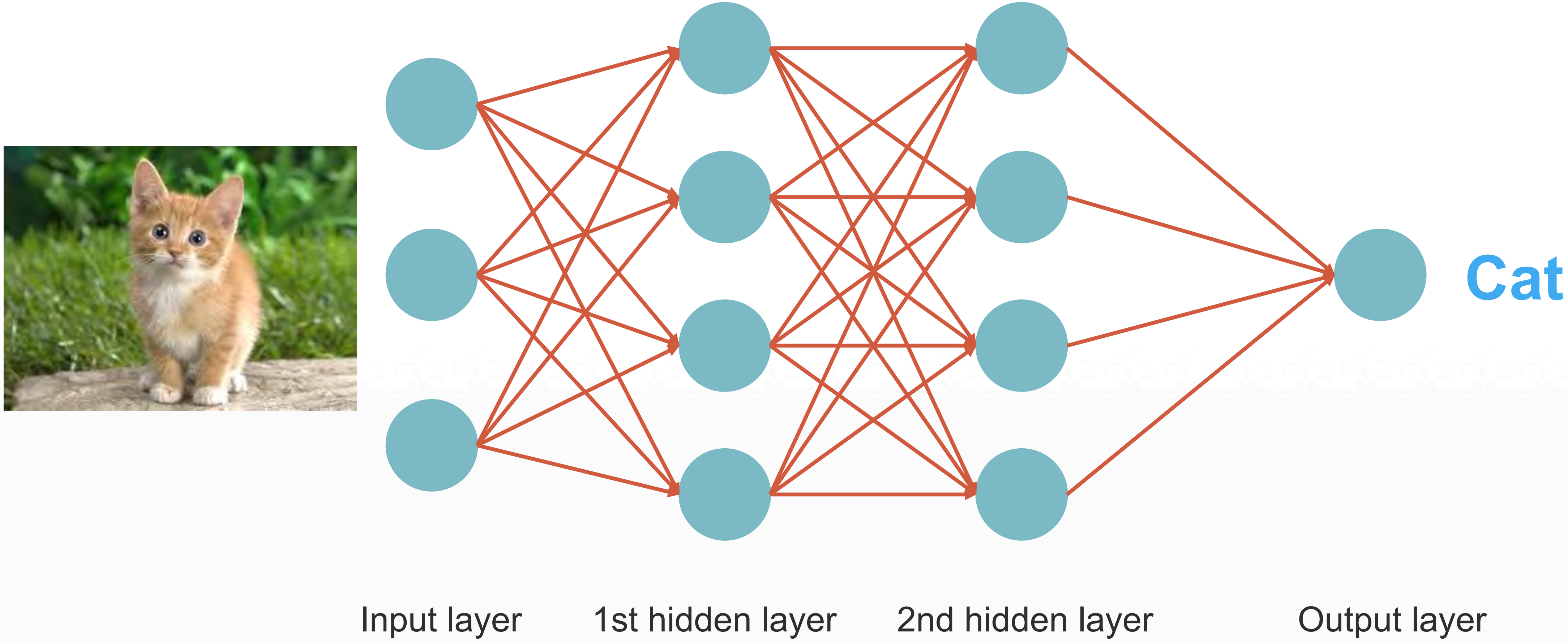
Deep Learning

- Deep learning is a class of machine learning techniques that exploit multiple layers of non-linear information processing for supervised or unsupervised feature extraction and transformation, and for pattern analysis and classification.



I. L. Deng and D. Yu. "Deep Learning: Methods and Applications". *Foundations and Trends in Signal Processing*, 2014
II. Ian Goodfellow, Yoshua Bengio, Aaron Courville. "Deep Learning", *MIT Press*, 2016

Deep Fully-connected Neural Network



Convolutional Neural network (CNN)

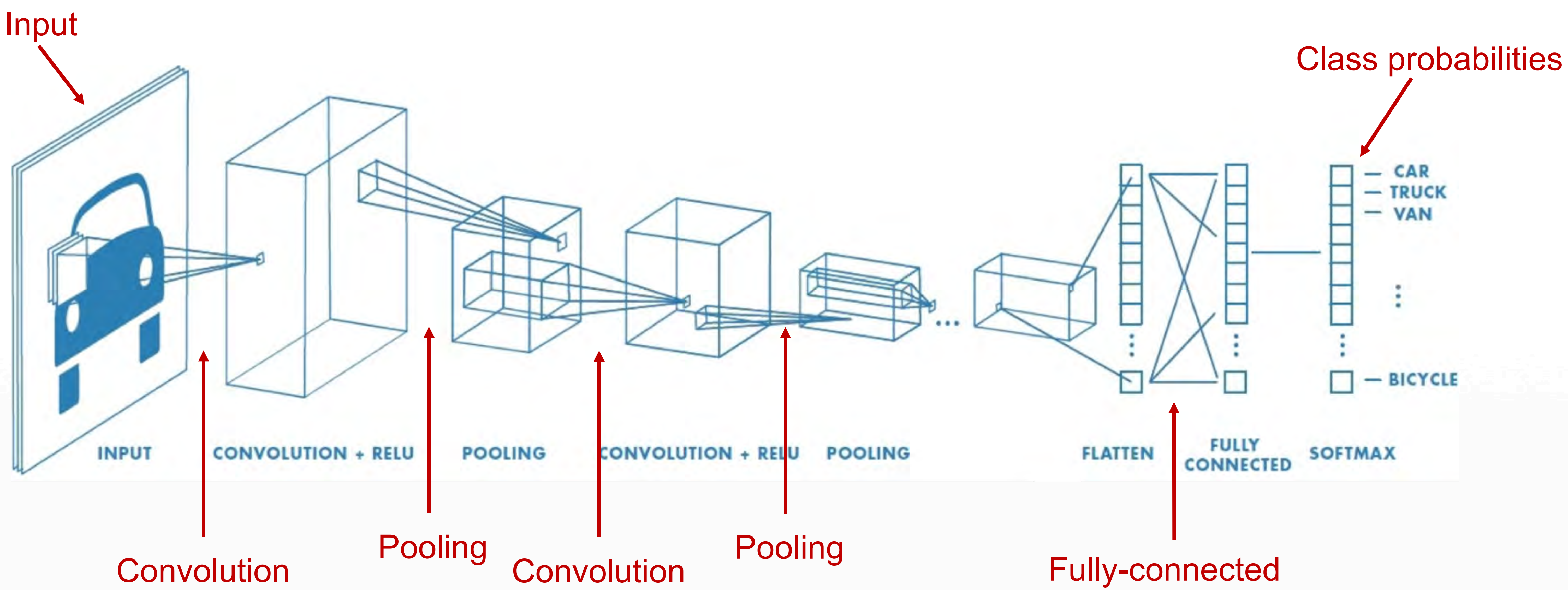
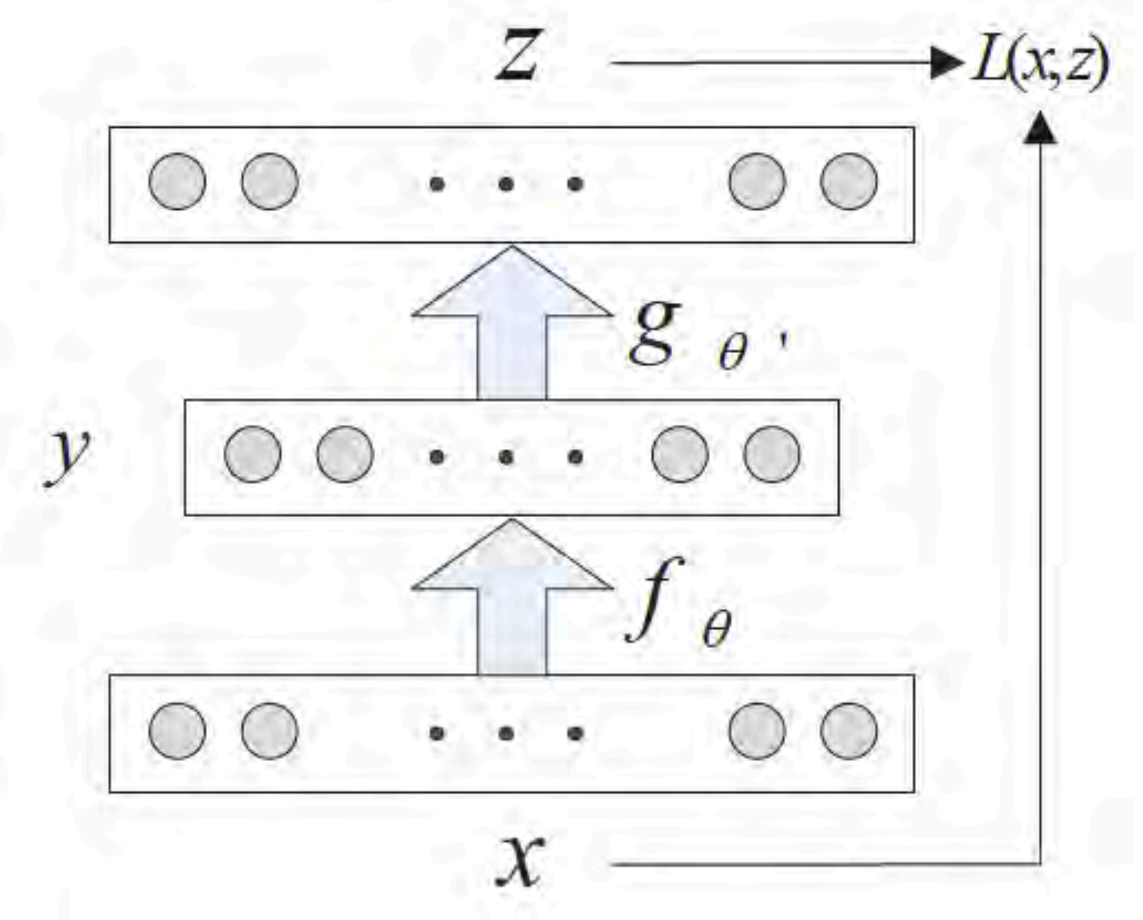


Image credit: Mathworks

Stacked Autoencoders

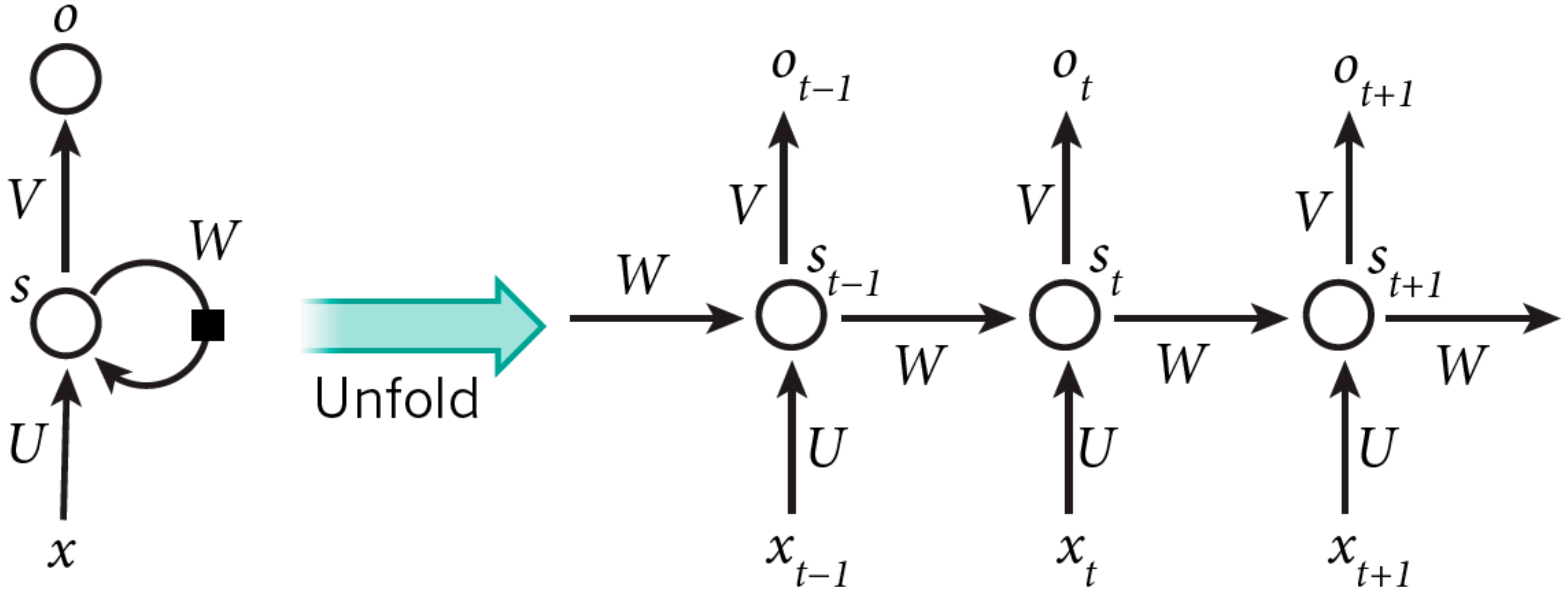
- Stacked autoencoders: stack multiple autoencoders to form a multi-layer neural network
- Typically trained in a layer-wise fashion: train one layer at a time
- Can be further fine-tuned in a supervised learning manner



An autoencoder

Figure credit: F. Xing *et al.*. "Deep Learning in Microscopy Image Analysis: A Survey", *IEEE TNNLS*, 2018

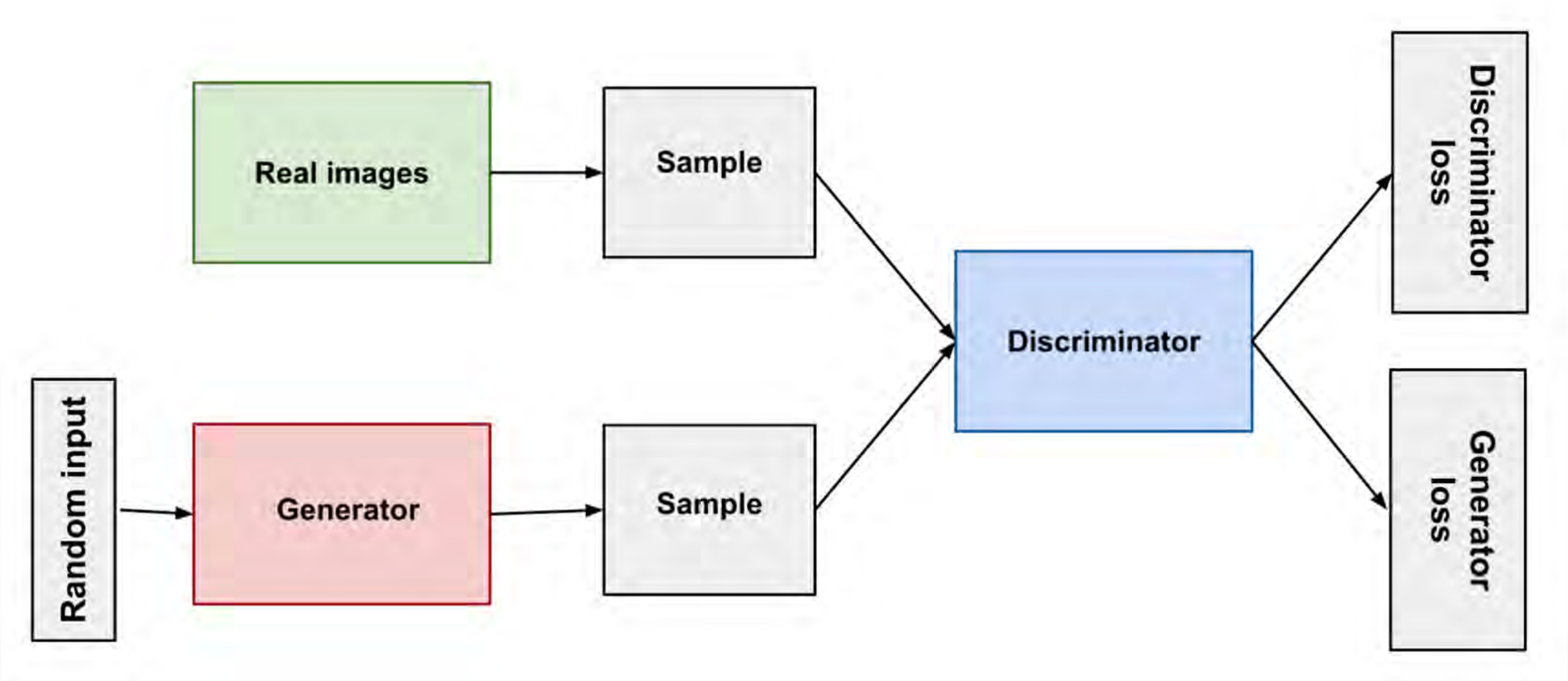
Recurrent Neural Network (RNN)



$$s_t = f(Ux_t + Ws_{t-1})$$
$$\sigma_t = g(Vs_t)$$

Y. LeCun *et al.*, “Deep Learning”, *Nature*, 2015

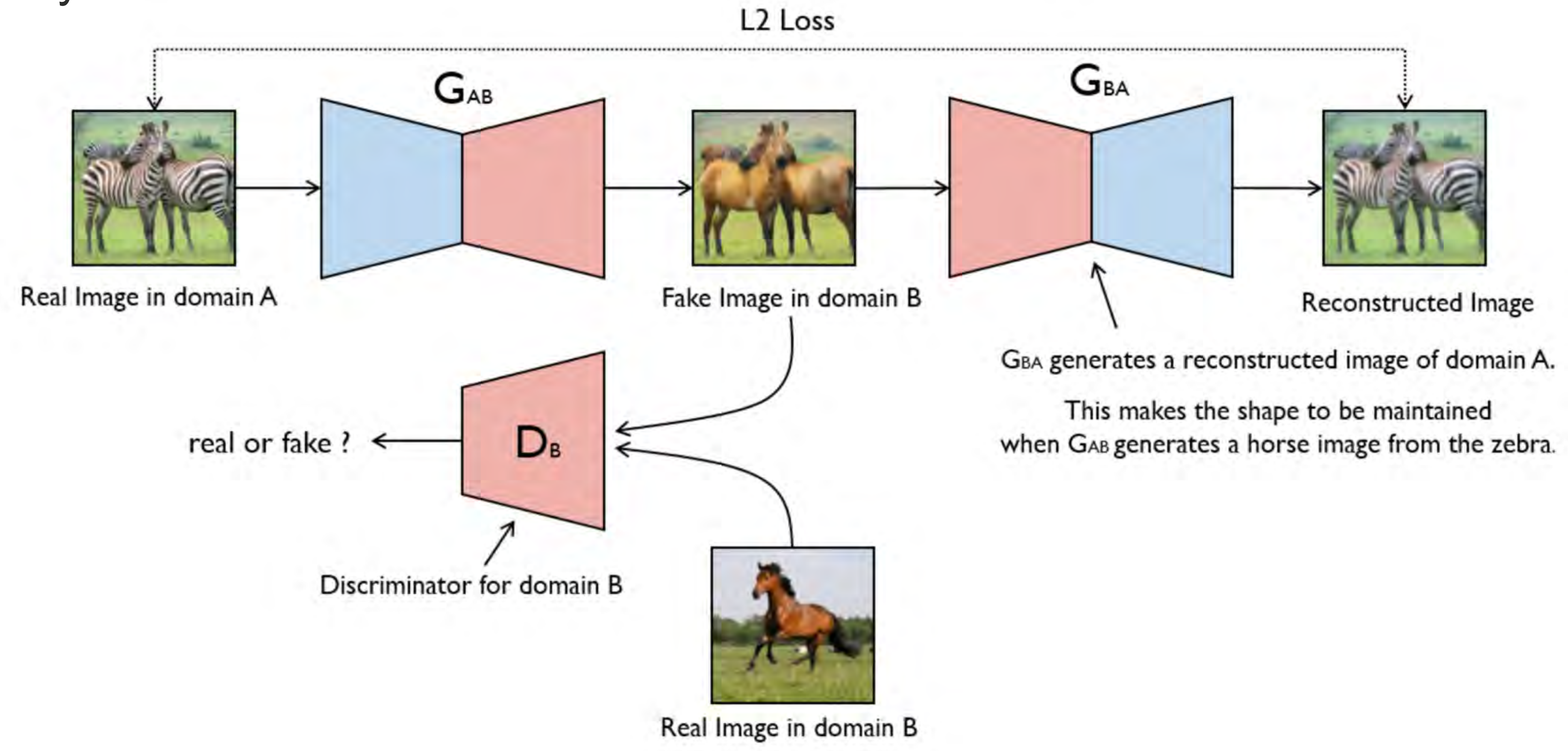
Generative Adversarial Networks (GANs)



I. Goodfellow *et al.*, “Generative Adversarial Nets”, *NeurIPS*, 2014

Generative Adversarial Networks (GANs)

- CycleGAN



I. J. Zhu *et al.*, "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks", *ICCV*, 2017
II. https://towardsdatascience.com/image-to-image-translation-using-cyclegan-model-d58cfff04755

Deep Learning Frameworks

PYTORCH

mxnet

DEEPLEARNING4J

Microsoft Cognitive Toolkit
aka.ms/CognitiveToolkit
Microsoft

theano

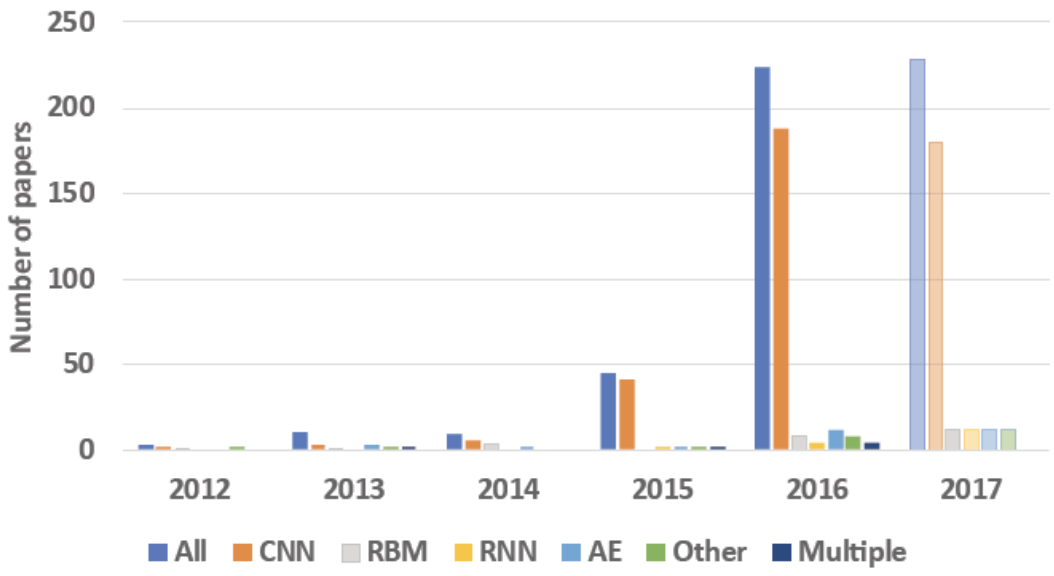
Caffe

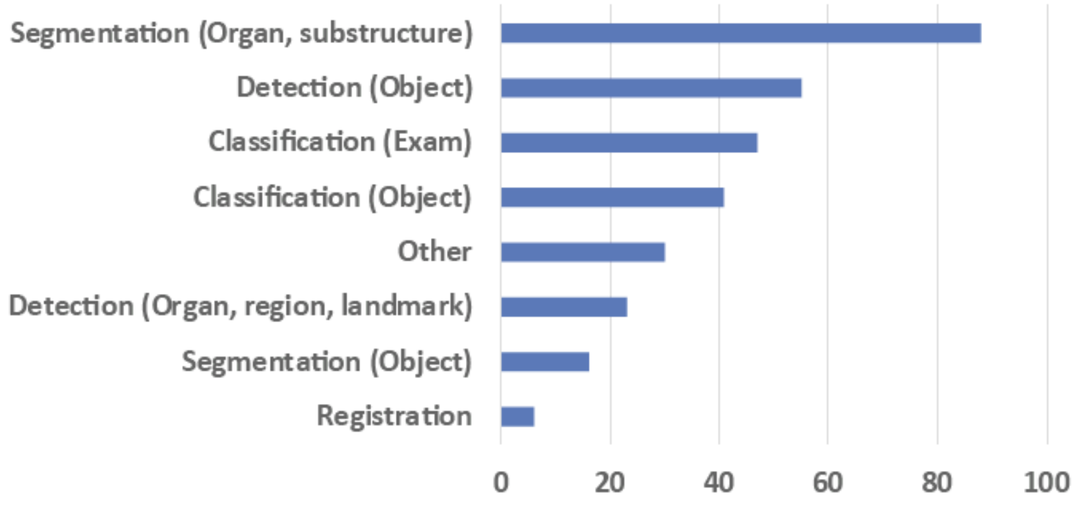
TensorFlow

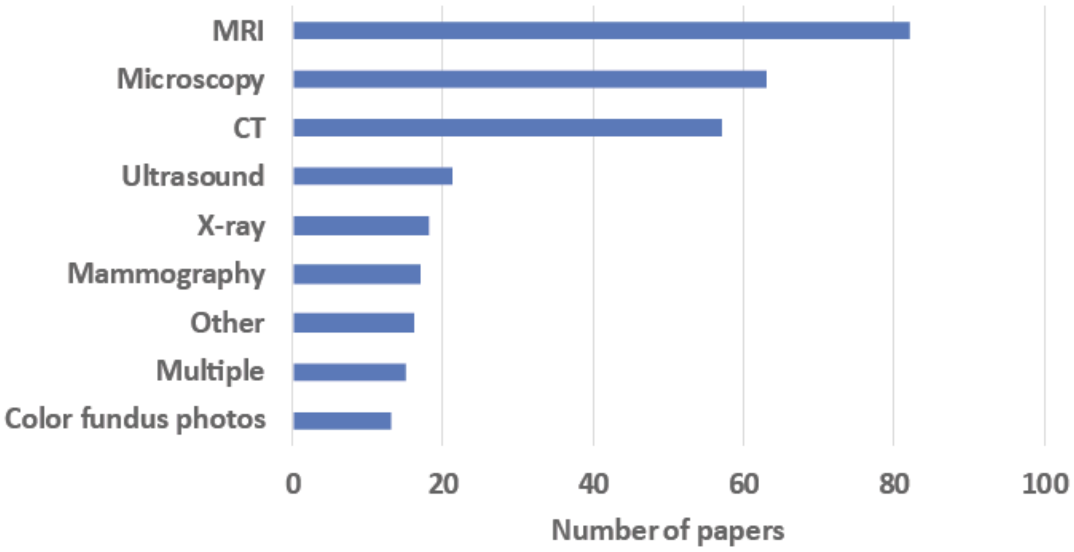
Deep Learning in Medical Image Analysis

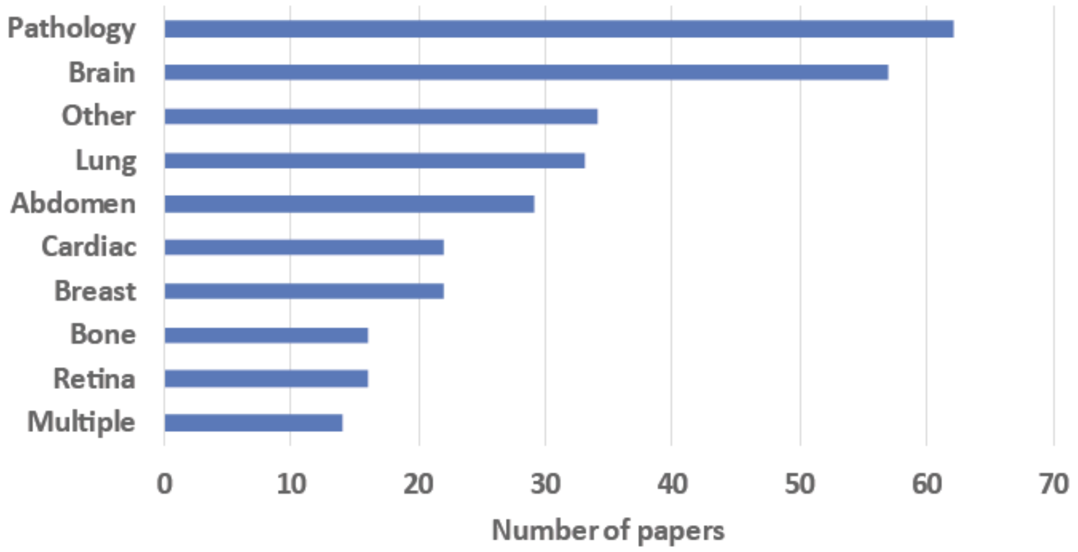
- Automated detection of objects, regions, landmarks, etc.
- Automated segmentation of organs, substructures, etc.
- Automated classification of objects, diseases, etc.
- Image registration
- Image retrieval
- Biomarker discovery
- Medical Imaging report generation
- More …

Deep Learning in Medical Image Analysis





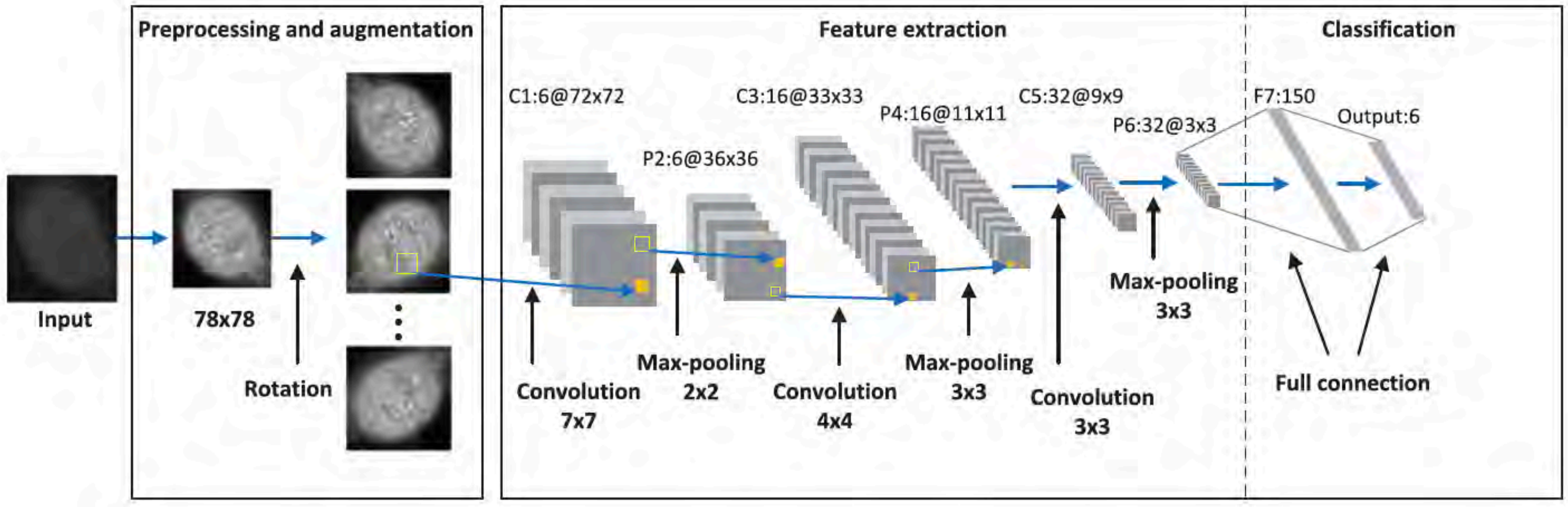




G. Litjens *et al.*, “A survey on deep learning in medical image analysis”, *Medical Image Analysis*, 2017

Image Classification

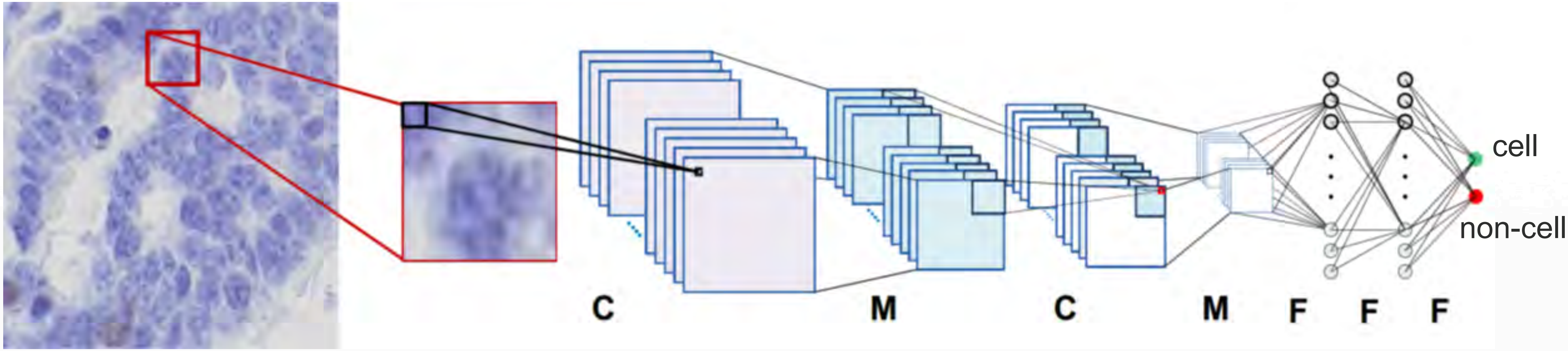
- CNN-based classification of HEp-2 cell images



Z. Gao *et al.*, "HEp-2 cell image classification with deep convolutional neural networks," *IEEE JBHI*, 2017

Object Detection

- CNN-based individual nucleus/cell localization in pancreatic neuroendocrine tumor images

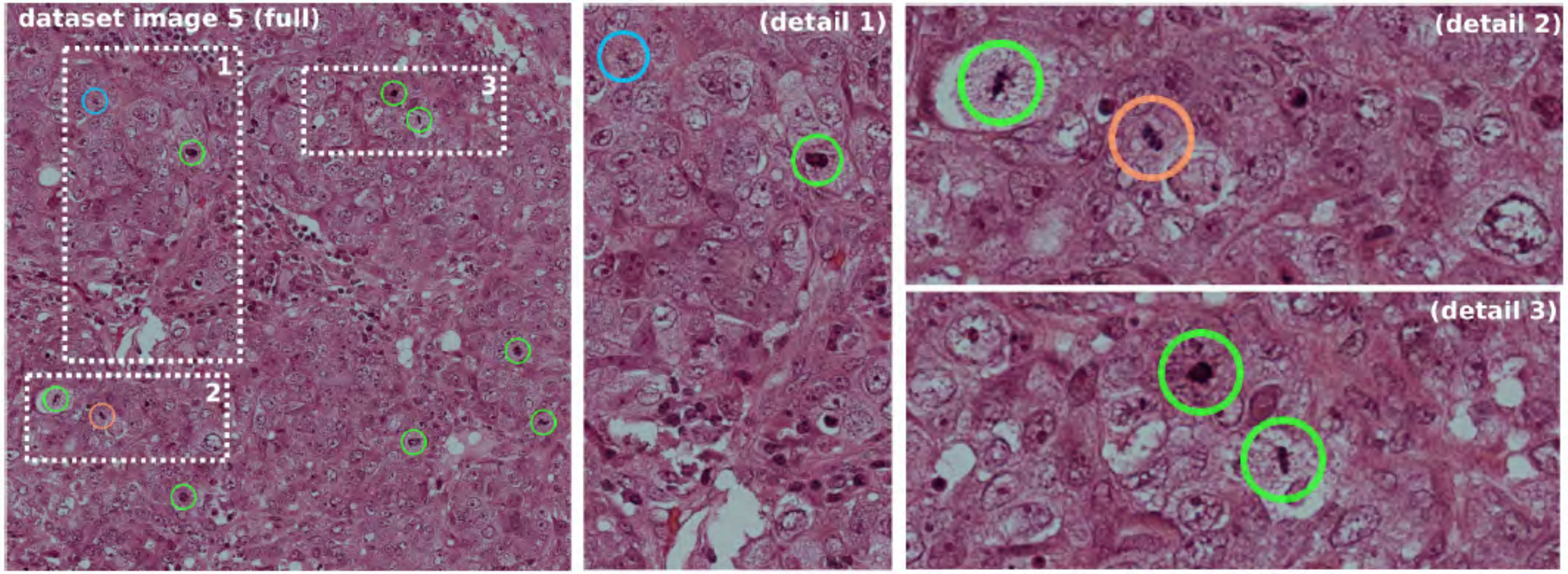


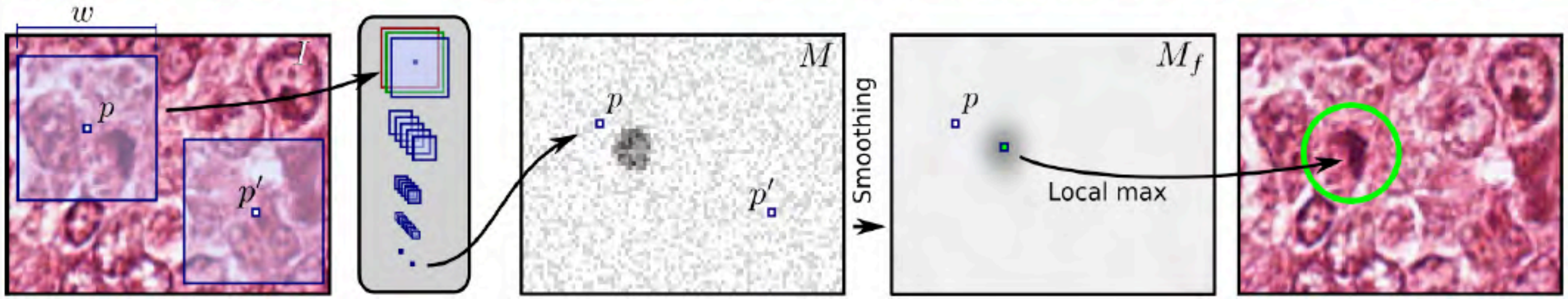
C: convolutional layers; M: max-pooling layers; F: fully-connected layers. The last layer has two units for binary classification.

F. Xing *et al.* “An Automatic Learning-Based Framework for Robust Nucleus Segmentation”, *IEEE TMI*, 2016.

Object Detection

- CNN-based mitosis detection in breast cancer histology images





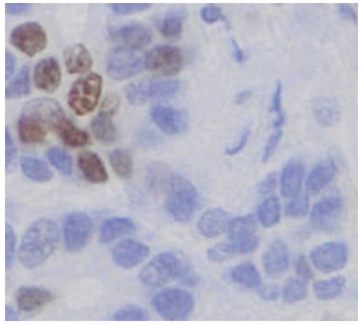
D.C. Ciresan *et al.*, "Mitosis Detection in Breast Cancer Histology Images with Deep Neural Networks", *MICCAI,* 2013

Object Detection

- Nucleus/cell detection with fully convolutional networks

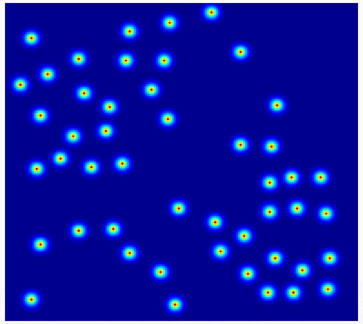
W x H

Input



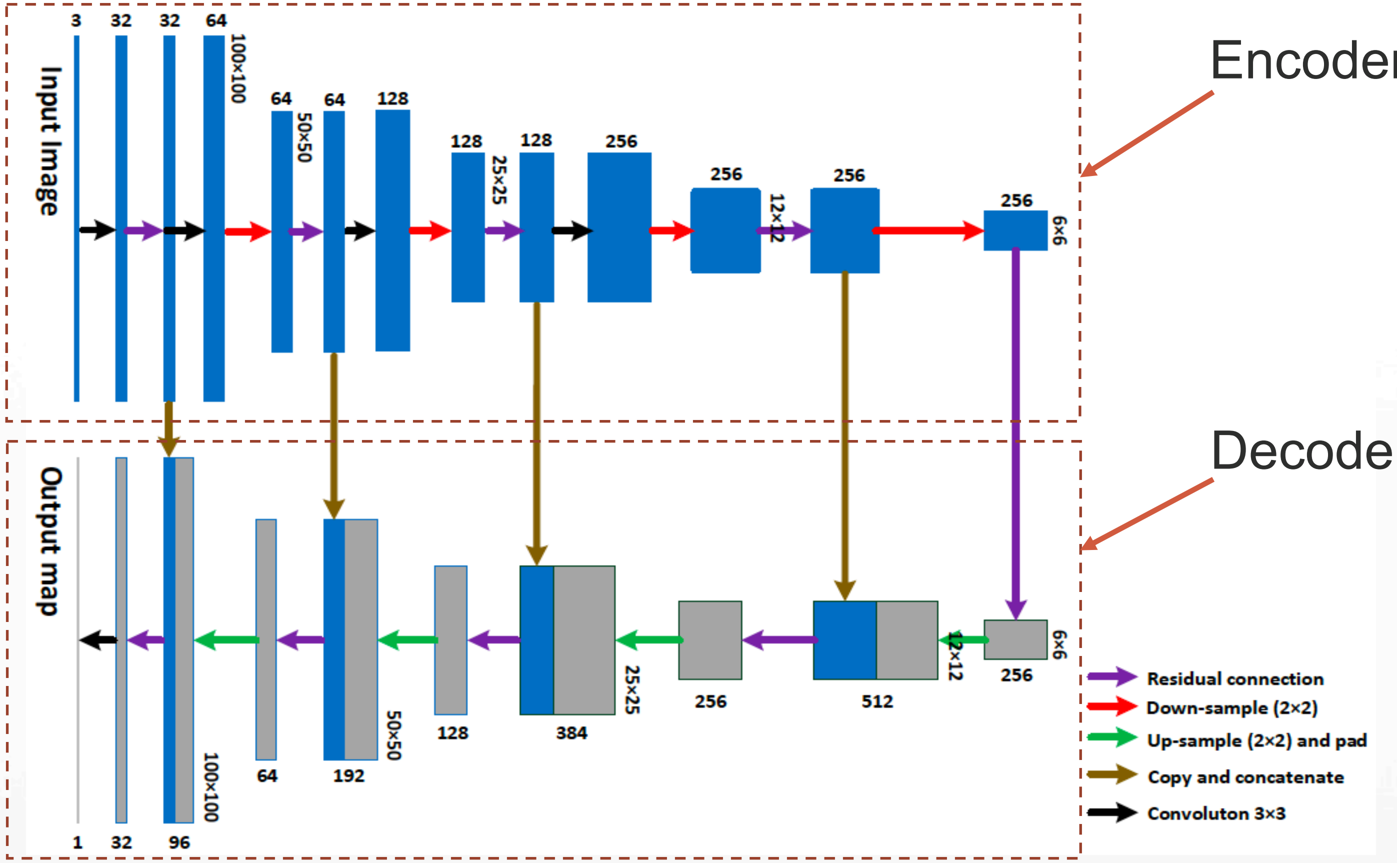
W x H

Output



Encoder

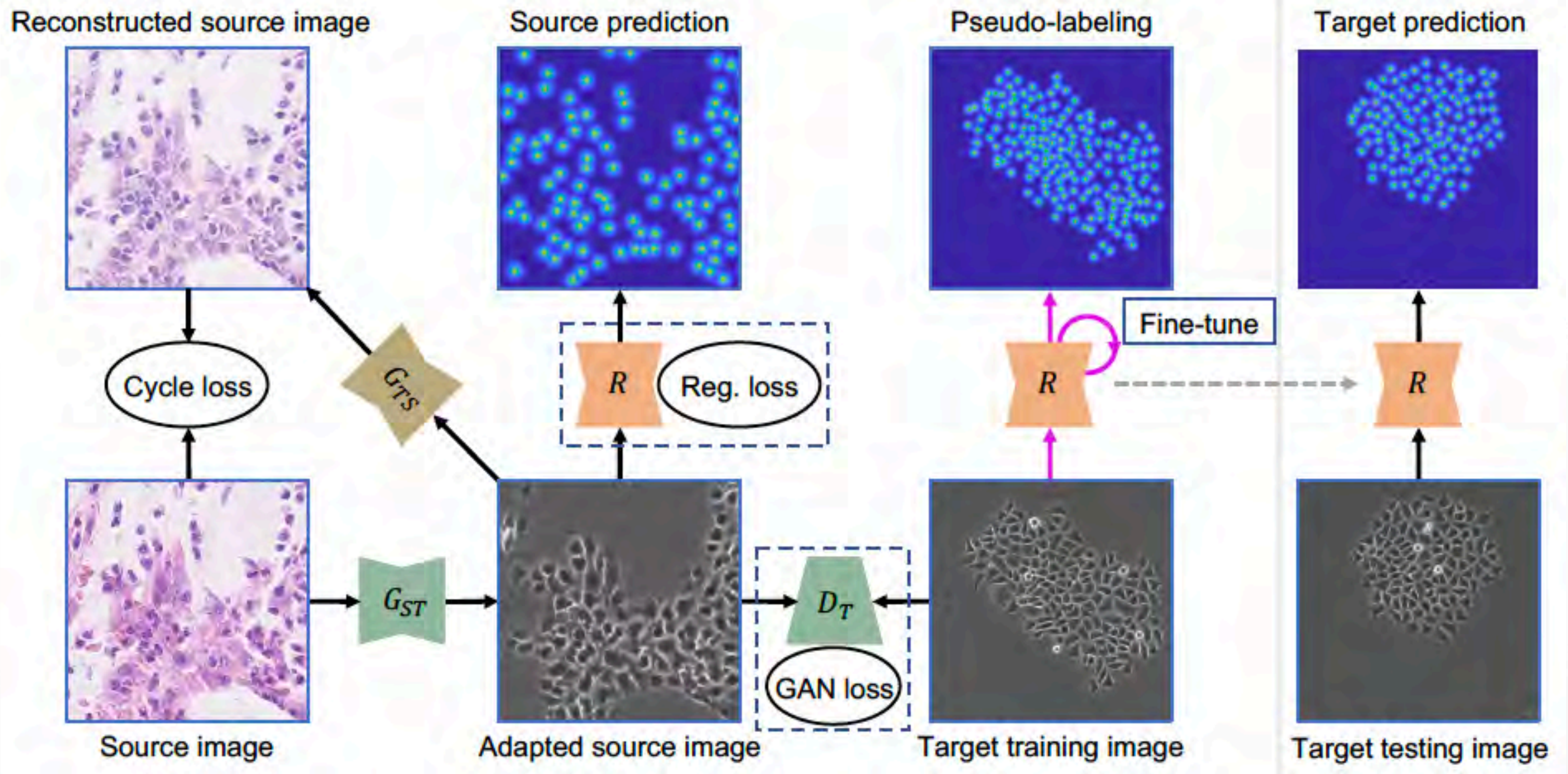
Decoder



Y.. Xie *et al.*, "Efficient and robust cell detection: A structured regression approach", *Medical Image Analysis,* 2018

Object Detection

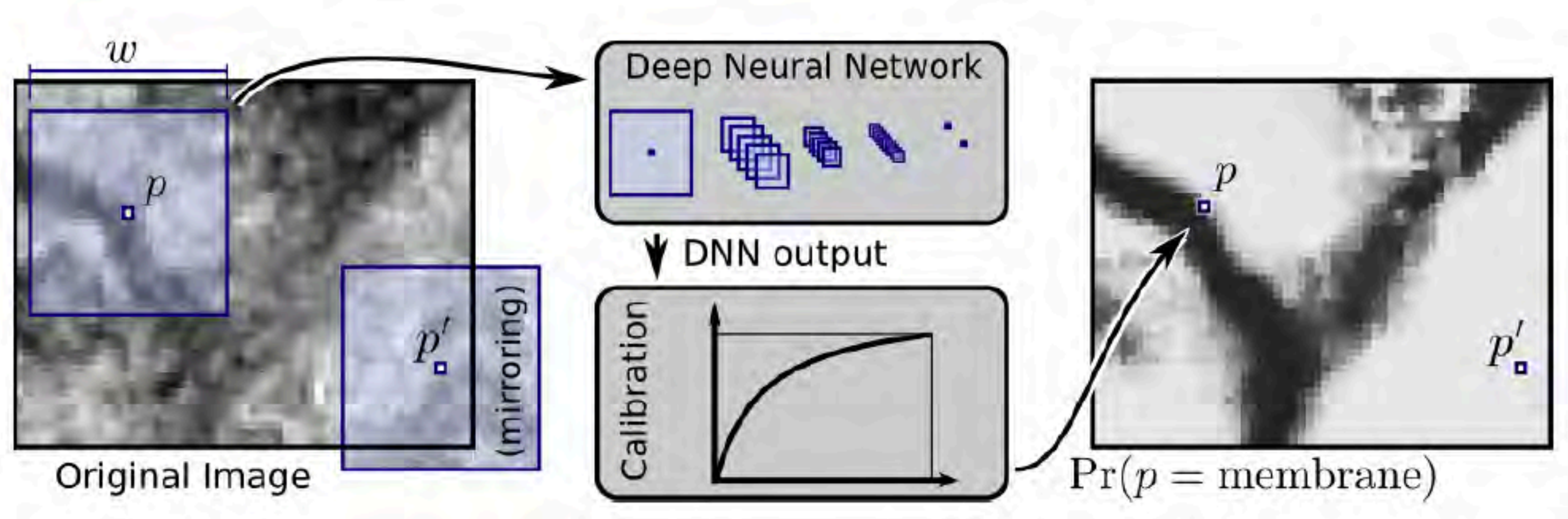
- Nucleus/cell detection generative adversarial networks (GANs)



F. Xing *et al.*, "Adversarial Domain Adaptation and Pseudo-Labeling for Cross-Modality Microscopy Image Quantification", *MICCAI,* 2019

Image Segmentation

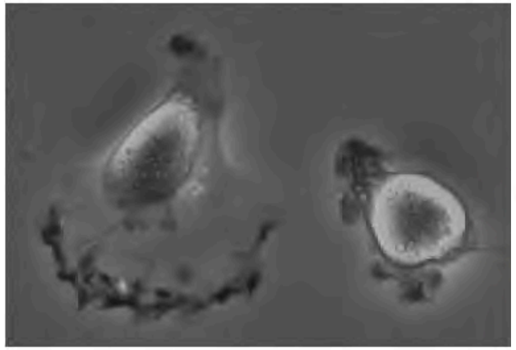
- CNN-based neuronal membrane segmentation electron microscopy images

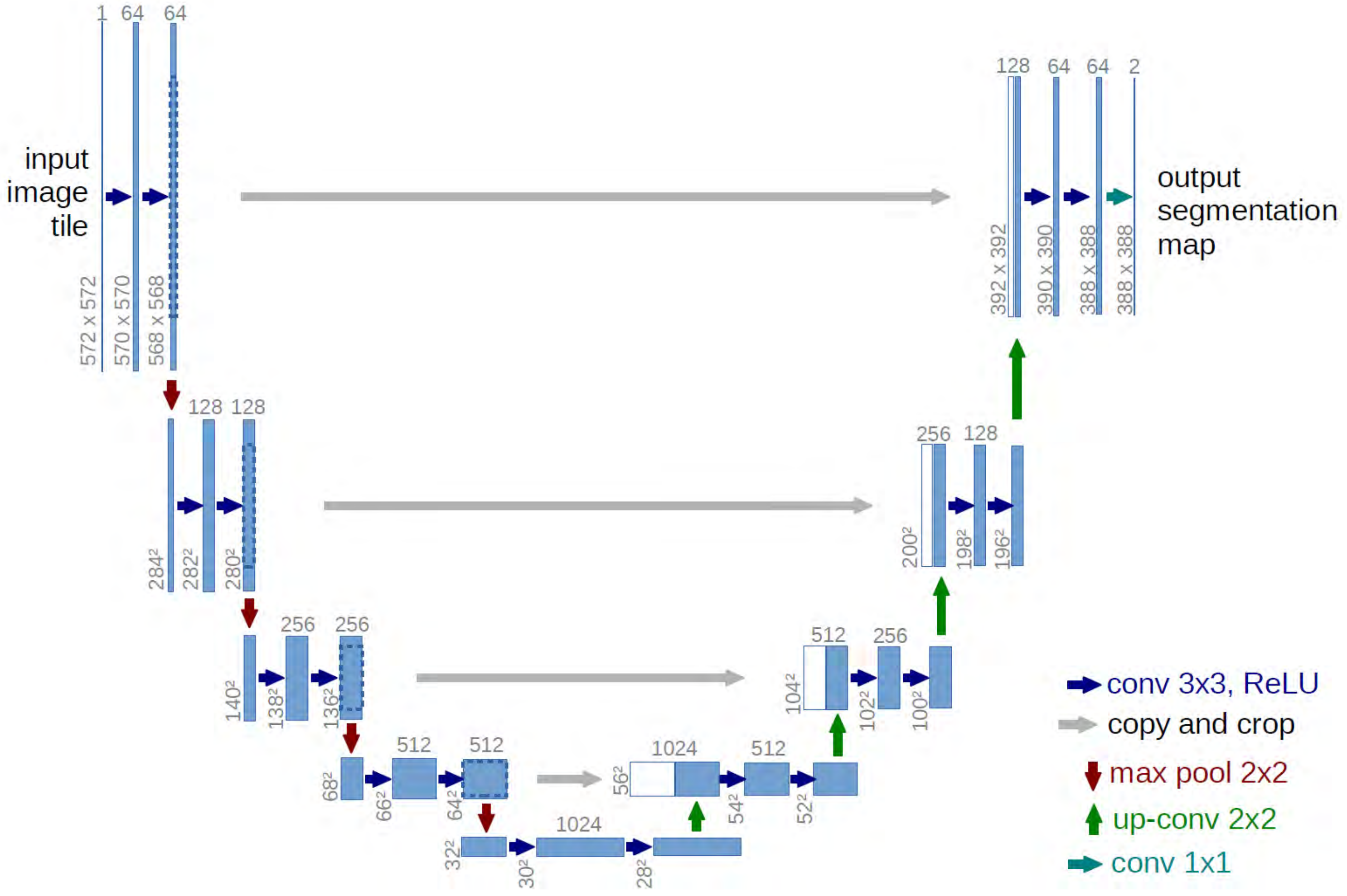


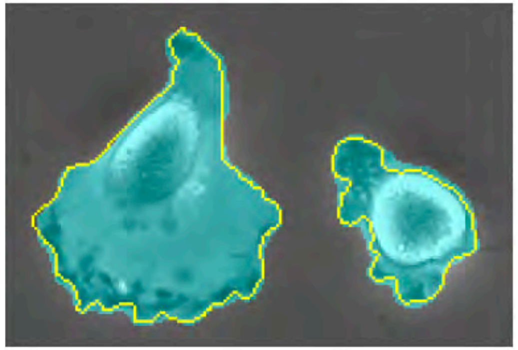
D.C. Ciresan *et al.*, "Deep Neural Networks Segment Neuronal Membranes in Electron Microscopy Images", *NeurIPS,* 2012

Image Segmentation

- Cell segmentation with U-Net (an encoder-decoder network)



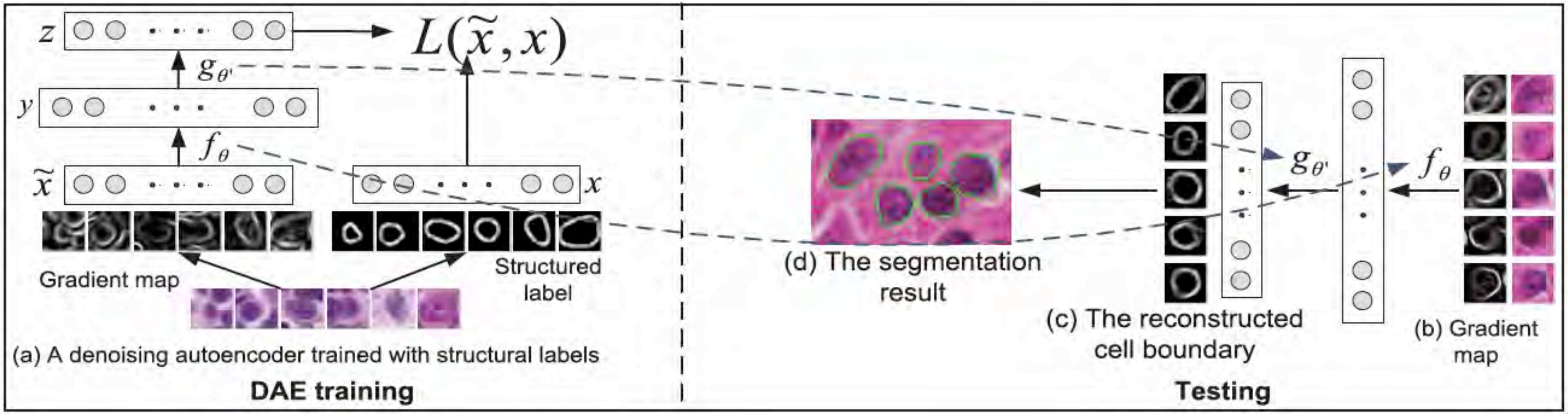




O. Ronneberger *et al.*, “U-Net: Convolutional Networks for Biomedical Image Segmentation”, *MICCAI,* 2015

Image Segmentation

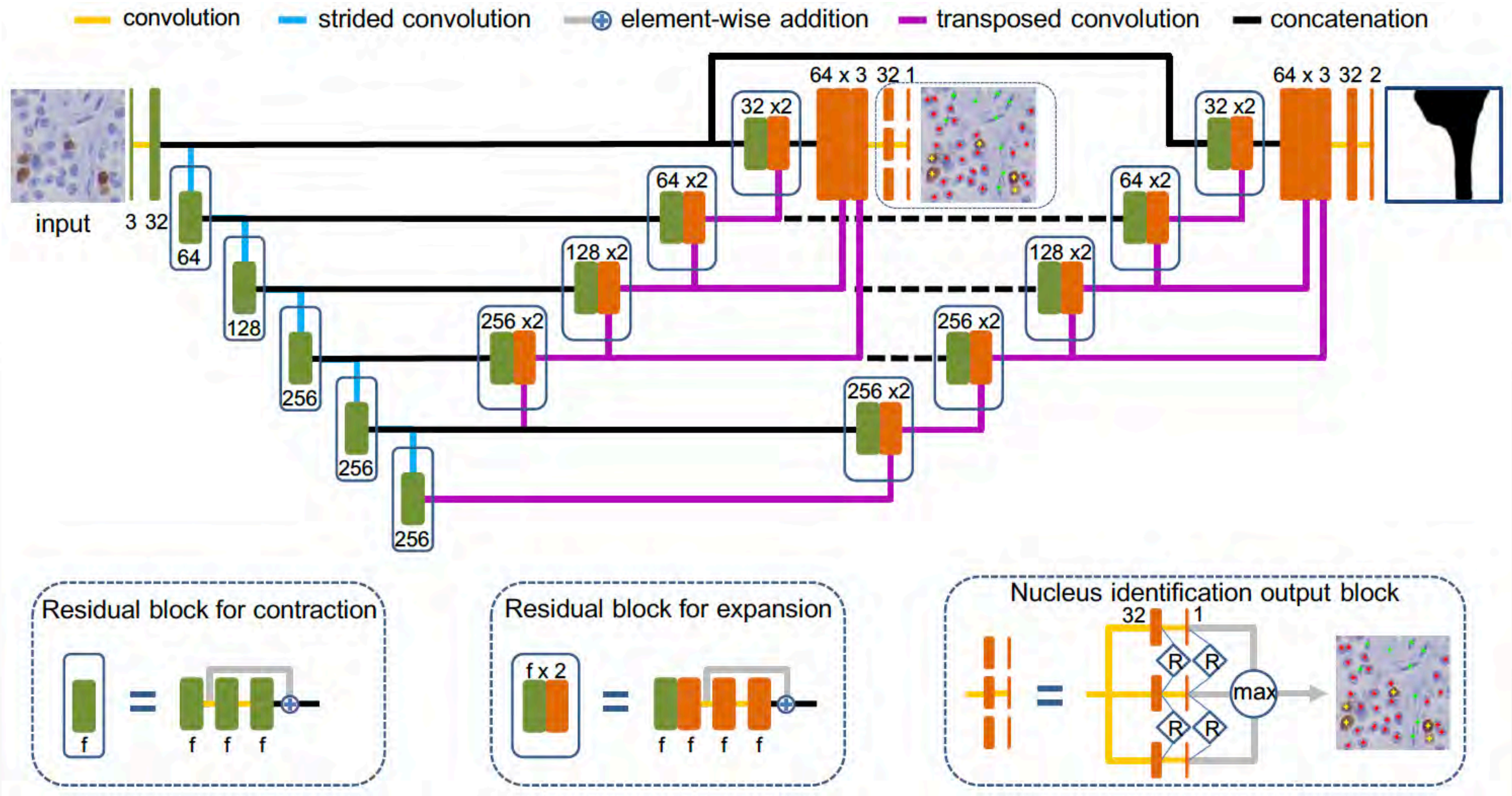
- Nucleus segmentation in histopathology images using stacked denoising autoencoders



H. Su *et al.*, "Robust cell detection and segmentation in histopathological images using sparse reconstruction and stacked denoising autoencoders", *MICCAI,* 2015

Object Recognition

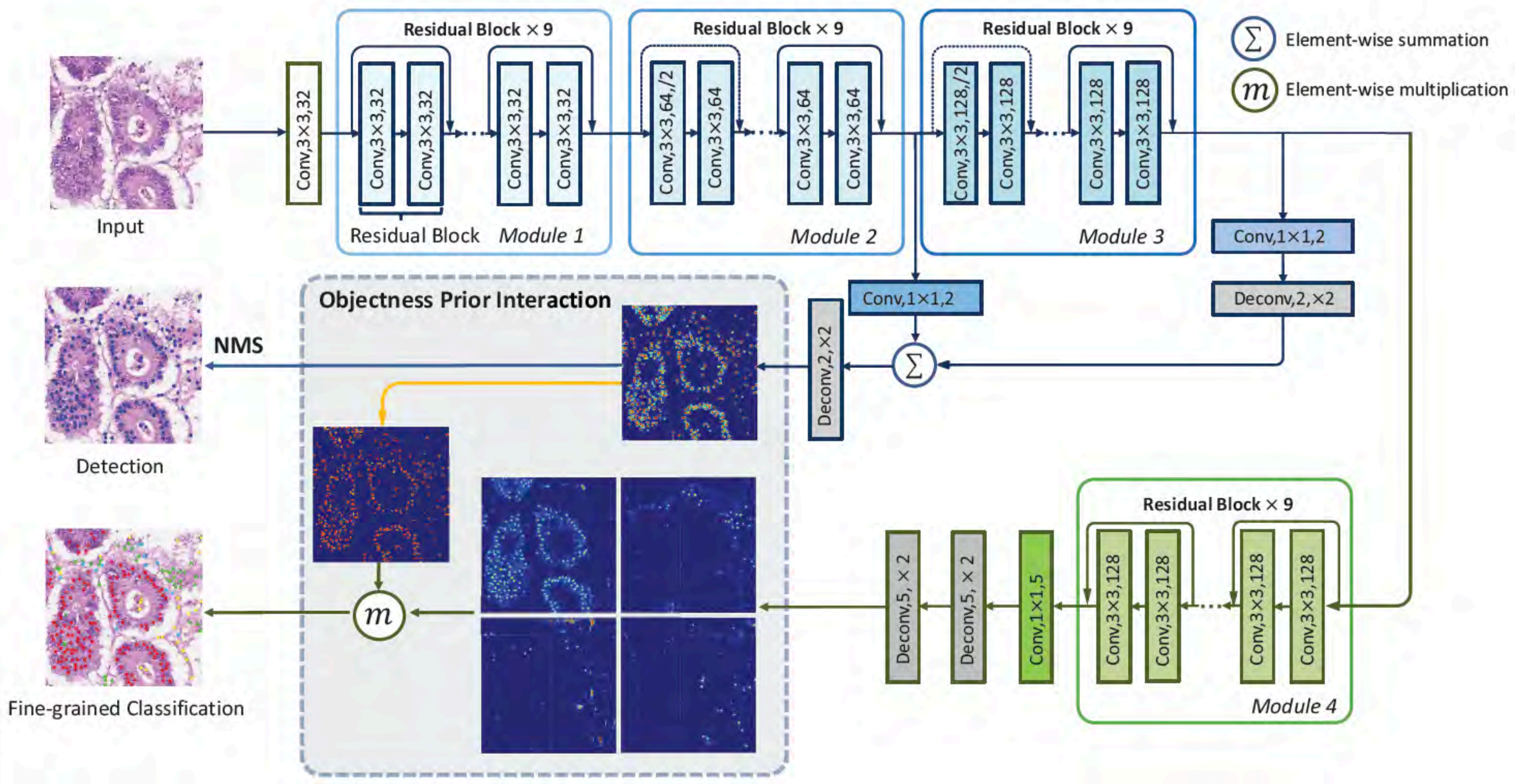
- Nucleus recognition with fully convolutional networks



F. Xing *et al*., "Pixel-to-Pixel Learning With Weak Supervision for Single-Stage Nucleus Recognition in Ki67 Images", *IEEE TBME,* 2019

Object Recognition

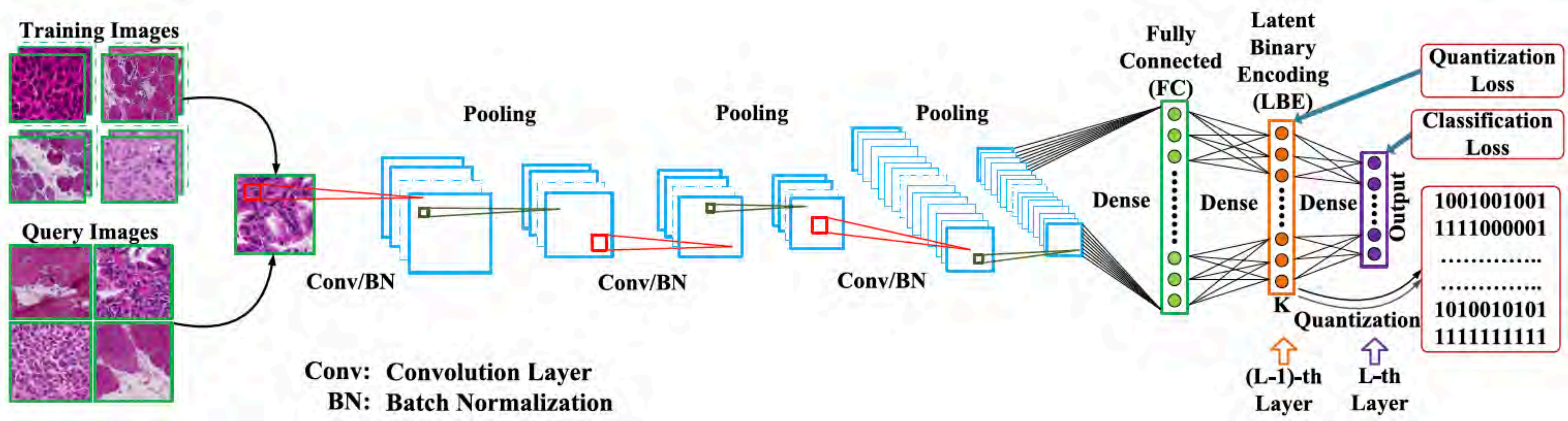
- Nucleus classification with fully convolutional networks



Y. Zhou *et al.*, "SFCN-OPI: Detection and fine-grained classification of nuclei using sibling FCN with objectness prior interaction," *AAAI*, 2018

Image Retrieval

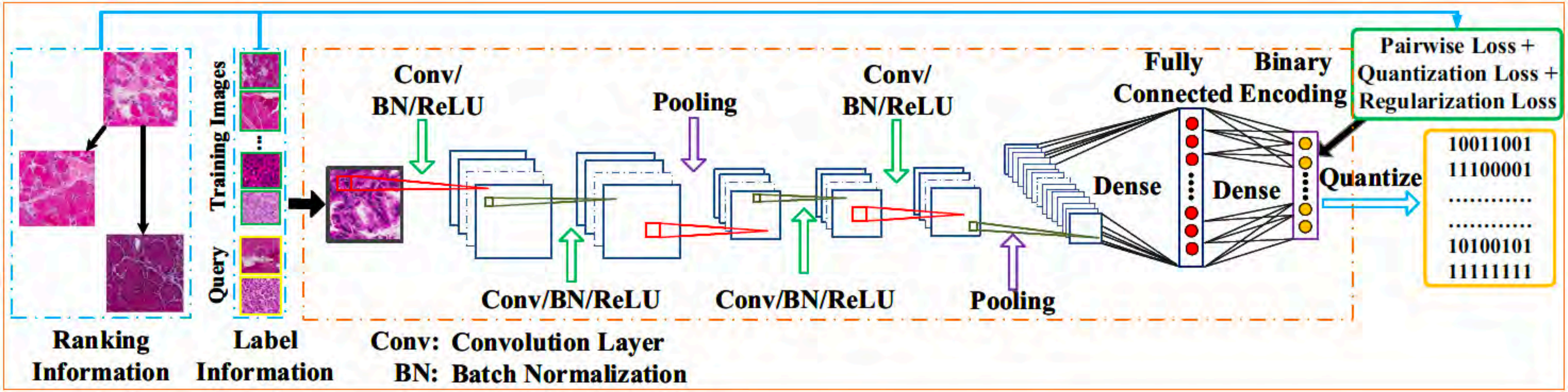
- Skeletal muscle image retrieval with CNNs



M. Sapkota *et al.*, "Deep Convolutional Hashing for Low-Dimensional Binary Embedding of Histopathological Images," *IEEE JBHI*, 2019

Image Retrieval

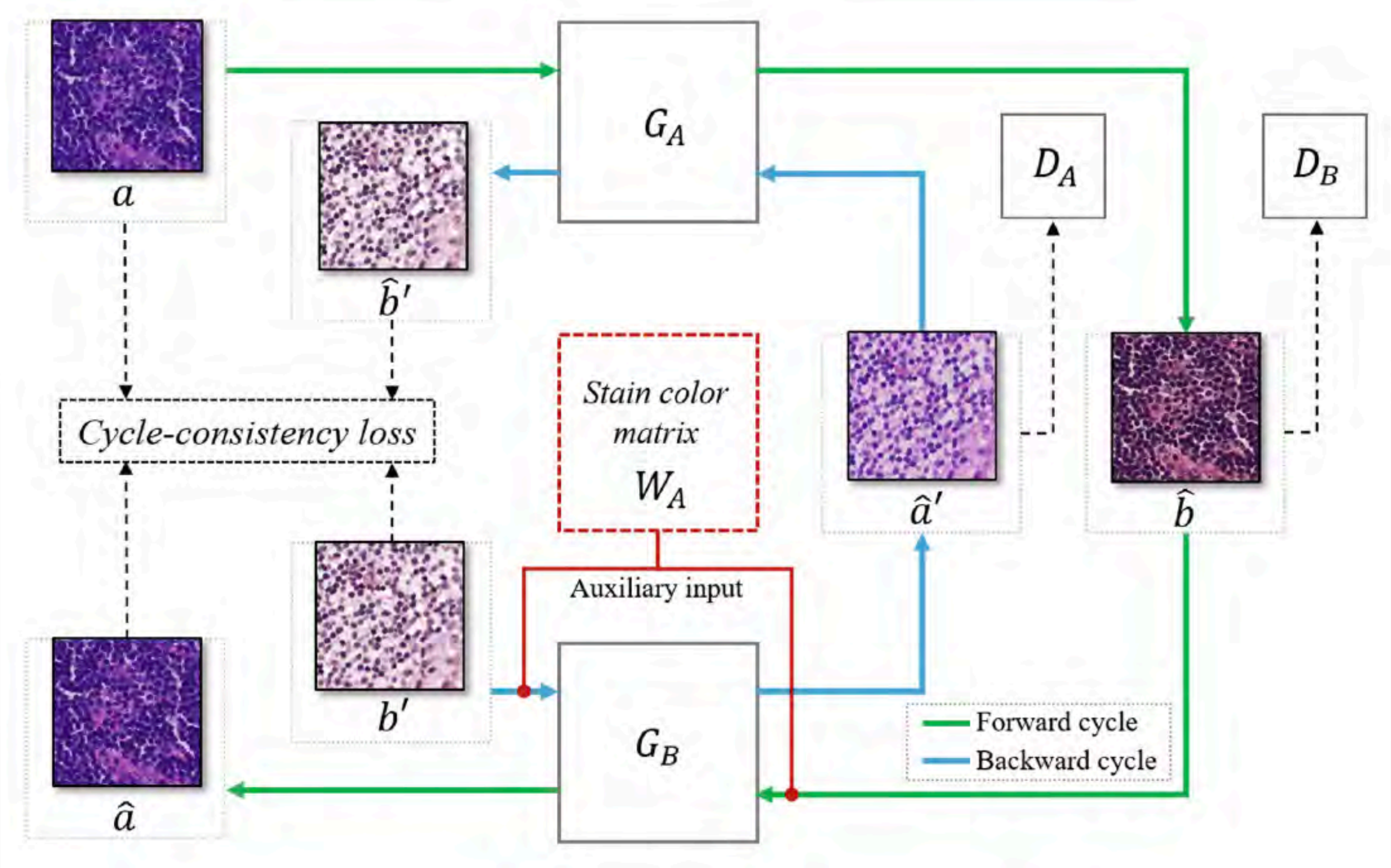
- Skeletal muscle and lung cancer image retrieval with CNNs



X. Shi *et al.*, "Pairwise based Deep Ranking Hashing For Histopathology Image Classification and Retrieval," *Pattern Recognition*, 2018

Stain/Color Normalization

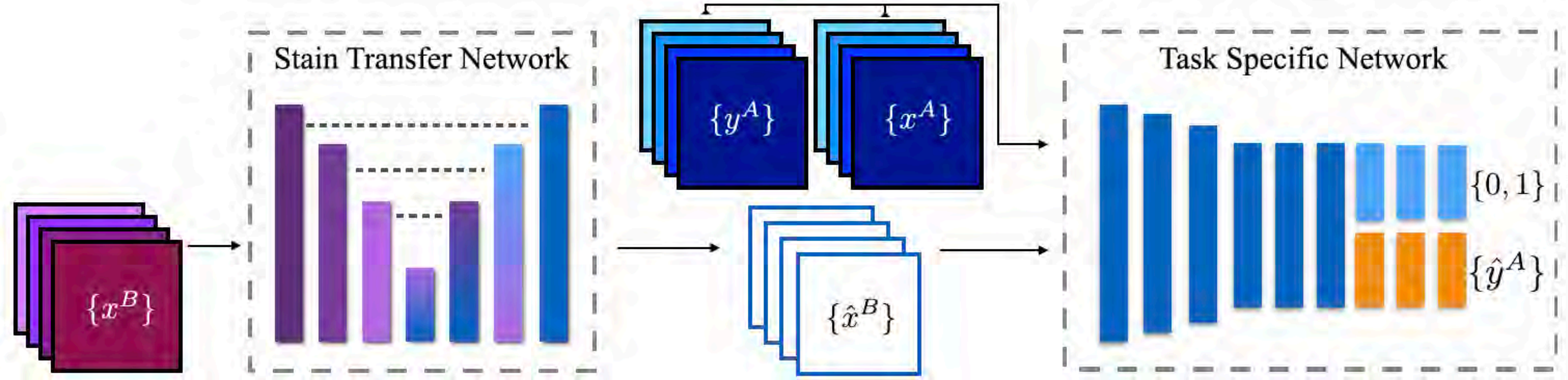
- Stain/color normalization with generative adversarial networks (GANs)



N. Zhou *et al.*, "Enhanced Cycle-Consistent Generative Adversarial Network for Color Normalization of H&E Stained Images," *MICCAI*, 2019

Stain/Color Normalization

- Stain/color normalization with generative adversarial networks (GANs)

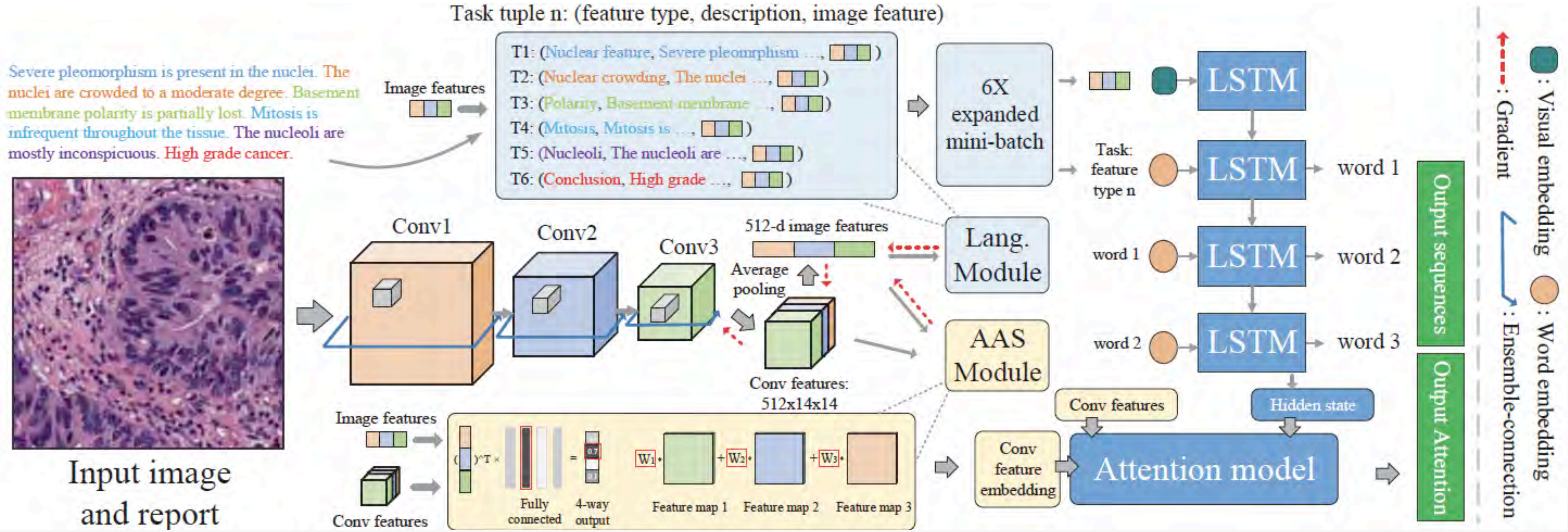


A. BenTaieb and G. Hamarneh, "Adversarial Stain Transfer for Histopathology Image Analysis," *IEEE TMI*, 2018

Text Generation

- Text generation from pathology images

  Link: https://www.youtube.com/watch?v=yy7NUrc3KI0



Z. Zhao *et al.*, “MDNet: A Semantically and Visually Interpretable Medical Image Diagnosis Network”, *CVPR,* 2017.

Thank You

- **Web:** https://fuyongxing.github.io
- **Email:** fuyong.xing@cuanschutz.edu
- **Address:**

Department of Biostatistics and Informatics
Colorado School of Public Health
University of Colorado Anschutz Medical Campus

National Jewish Health®

University of Colorado Anschutz Medical Campus

Application of Radiomics to Lung Disease to Develop a Novel Biomarker of Sarcoidosis

Lisa Maier, MD MSPH, Maierl@njhealth.org

Professor and Chief, Division of Environmental and Occupational Health Sciences

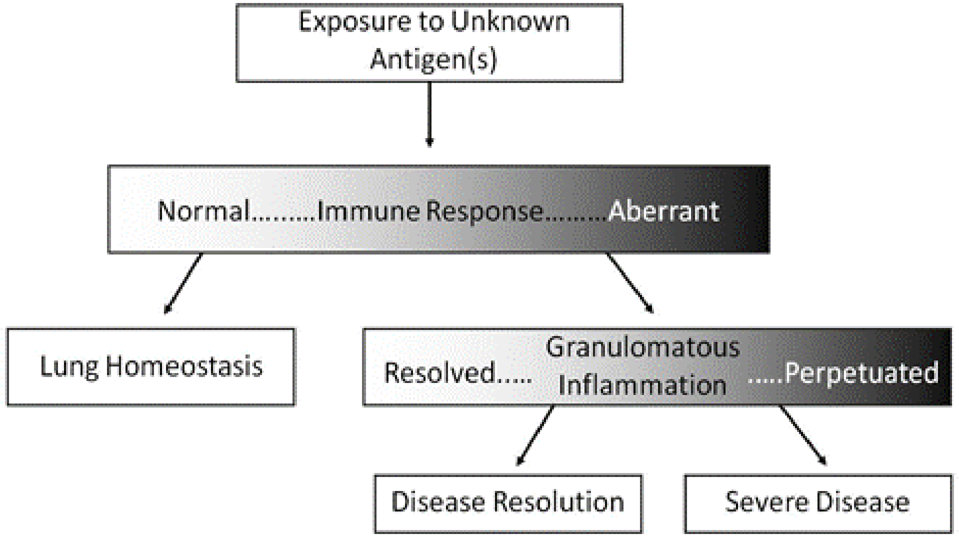
National Jewish Health, University of Colorado

GRADS

Genomic Research in Alpha-1 Antitrypsin Deficiency and Sarcoidosis

Sarcoidosis

- Burden of disease unknown, but may be increasing
  - Increasing hospitalizations
  - Higher costs of hospitalizations vs non-sarc (Ungprasert Lung 2019)
- NHLBI Symposium “Levaraging Disparities..” (Maier AATS 2017)
  - Conduct “Omics” and systems biology research
  - Develop biomarkers of severe disease
  - Pulmonary, cardiac, neurologic disease



Genetic variants may influence progression from each stage to next

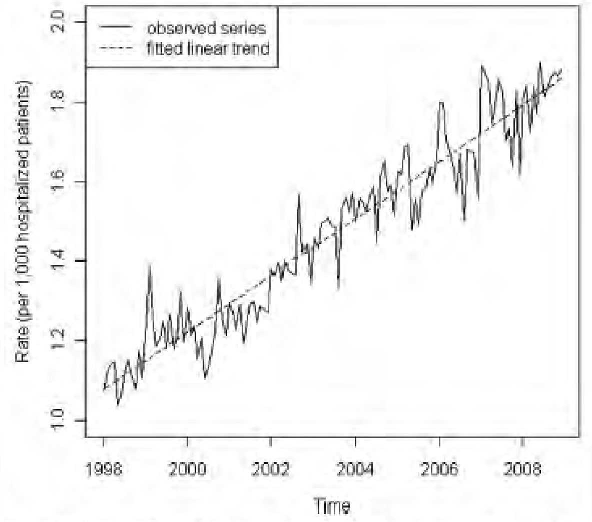
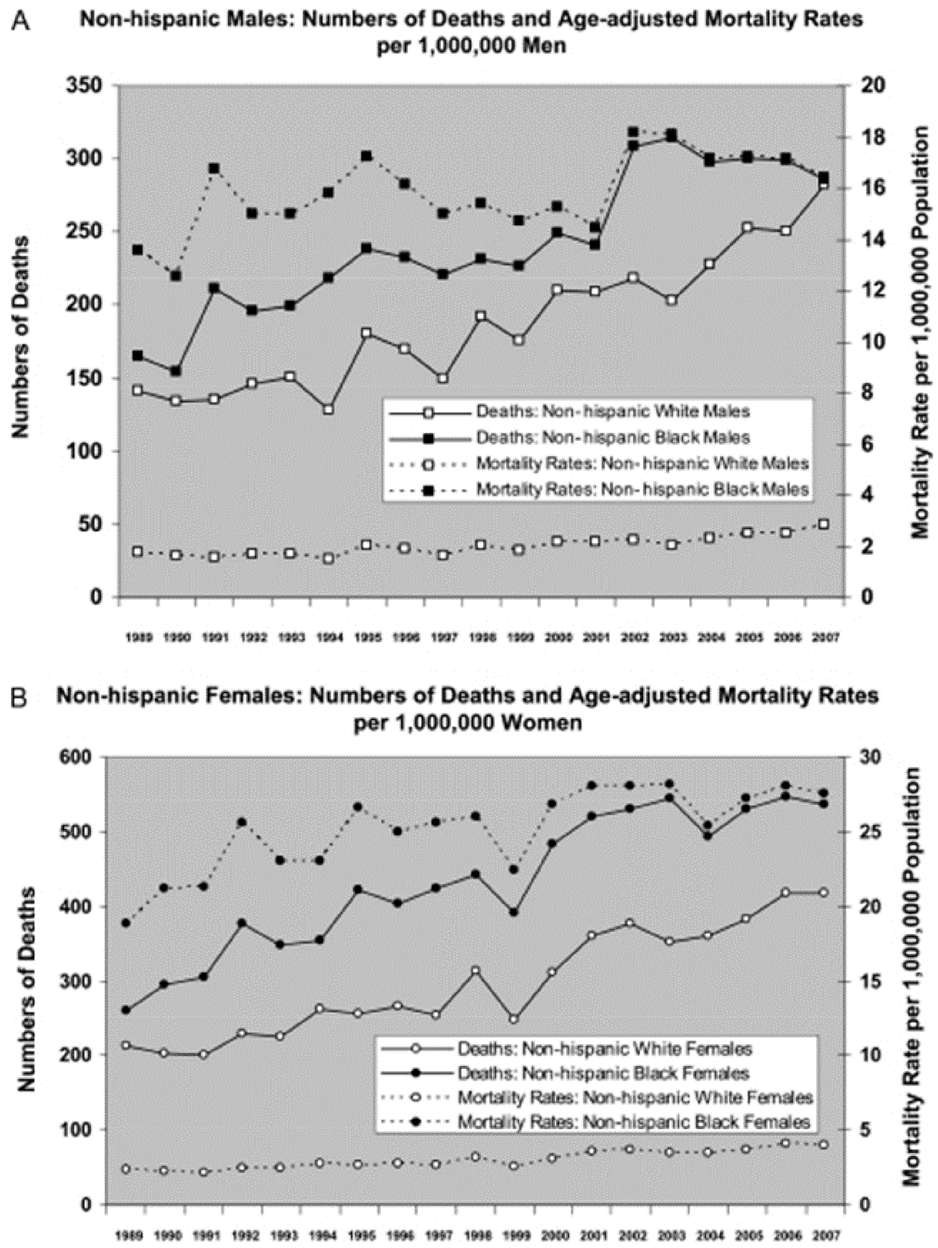


Figure 1 Monthly incidence rates of hospitalizations among patients with sarcoidosis (Nationwide Inpatient Sample, 1998–2008). The solid line represents the observed rate over time. The dashed line is the fitted linear trend ($p < 0.0001$).

Gerke et al BMC Pulm

Pulmonary Sarcoidosis

- Primary cause of death
- Impacts individuals in the prime of their lives (mean age at death= 55 years)
- Response to therapy variable
- No definitive prognosticators- genetic, genomics proposed
- Designing treatment and longitudinal studies is problematic



Swigris AJRCCM 2011

National Jewish Health®

Risk factors for advanced pulmonary sarcoidosis unclear

Risk Factors for Disease Persistence

- CXR at presentation
  - Absence of lymphadenopathy
  - Ascending Scadding Stage
- Architectural distortion of the airways or cystic changes
- Multiple organ involvement
- Splenomegaly
- Need for systemic therapy
- Older age
- Female gender
- Black race

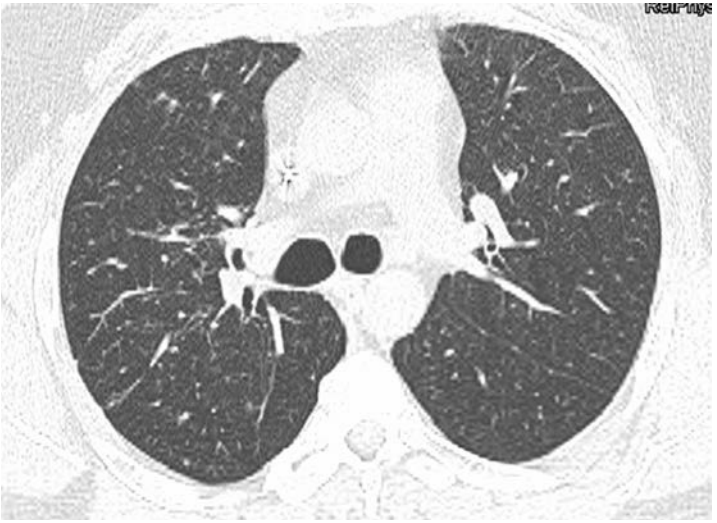
Risk Factors for Clinically Bothersome Disease

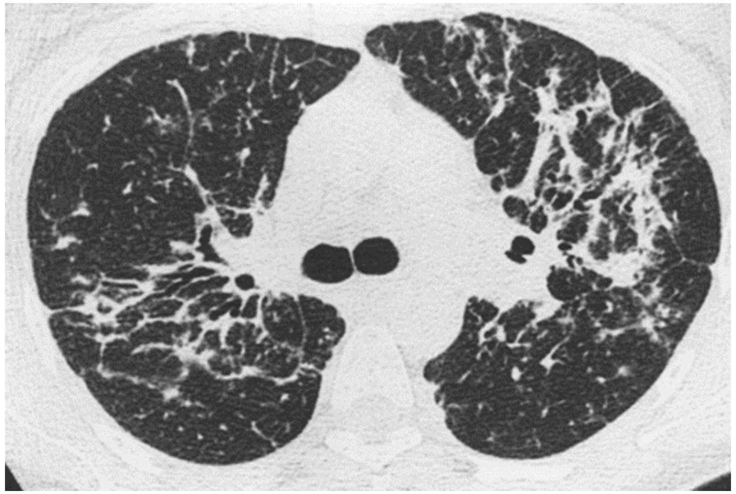
- Worse CXR
- Dyspnea at time of diagnosis
- Need for treatment in 1st 6 months
- Multiple organ involvement
- Lower socioeconomic status
- Black race

National Jewish Health

Biomarkers Diagnosis and Prognosis

- No biomarkers for diagnosis
  - Based on biopsy
  - Diagnosis currently one of exclusion
- No biomarkers to prognosticate patients
  - No predictors severe phenotype
    - Remitting versus progressive types
  - Who will or will not respond to Rx



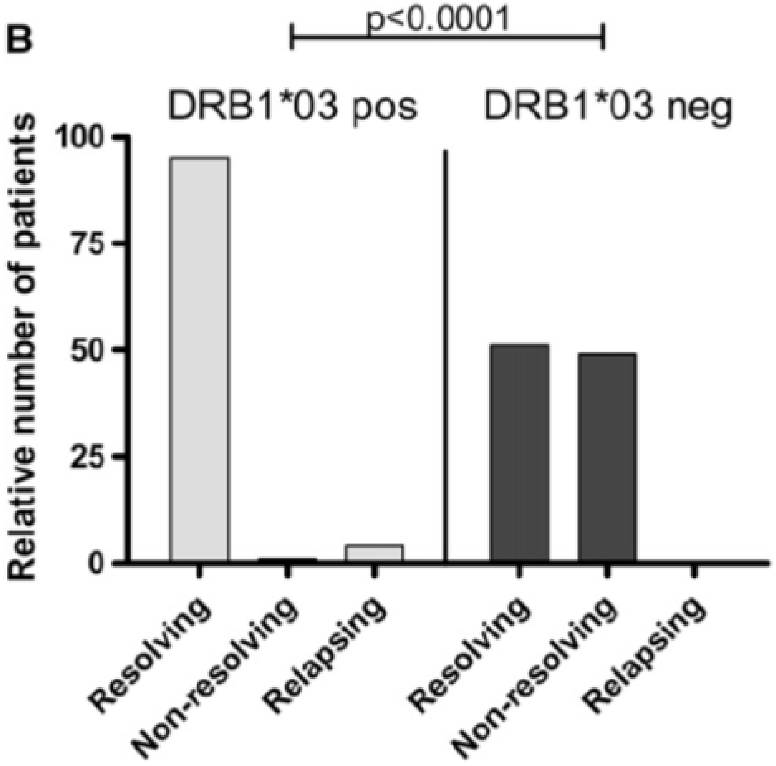


HLA Class II: Ethnicity, Phenotypes and Prognosis

HLA DRB1	Risk	Groups
*0301	Lofgren's	Swedish, Finnish, Dutch
	Protective	AA
	Resolving	Swedish
*1101	Overall Disease	US White, AA, Chinese
	Persistent Disease	AA
*1201	Overall Disease	European, US, AA
	Pulmonary	European, Japanese
*1401	Disease Risk	European, US, AA
	Progressive pulm	European, Scandinavian
*1501	Disease Risk	European, US, AA
	Progressive pulm	Scandinavian

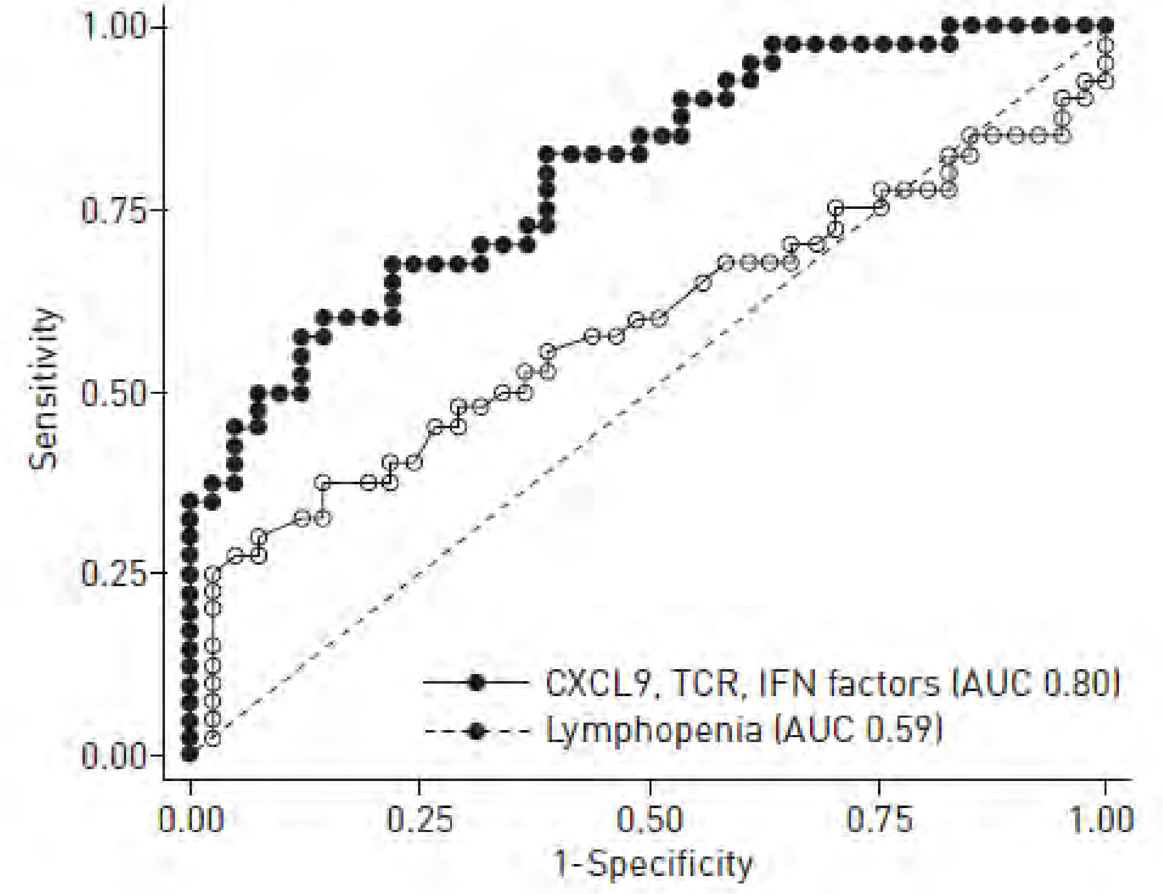
Fingerlin, Hamzeh, Maier Clin Chest Med 2015

Lofgren's Syndrome



Grunewald AJRCCM 2009

Blood Gene Transcripts Predict Remitting vs Progressive Sarcoidosis

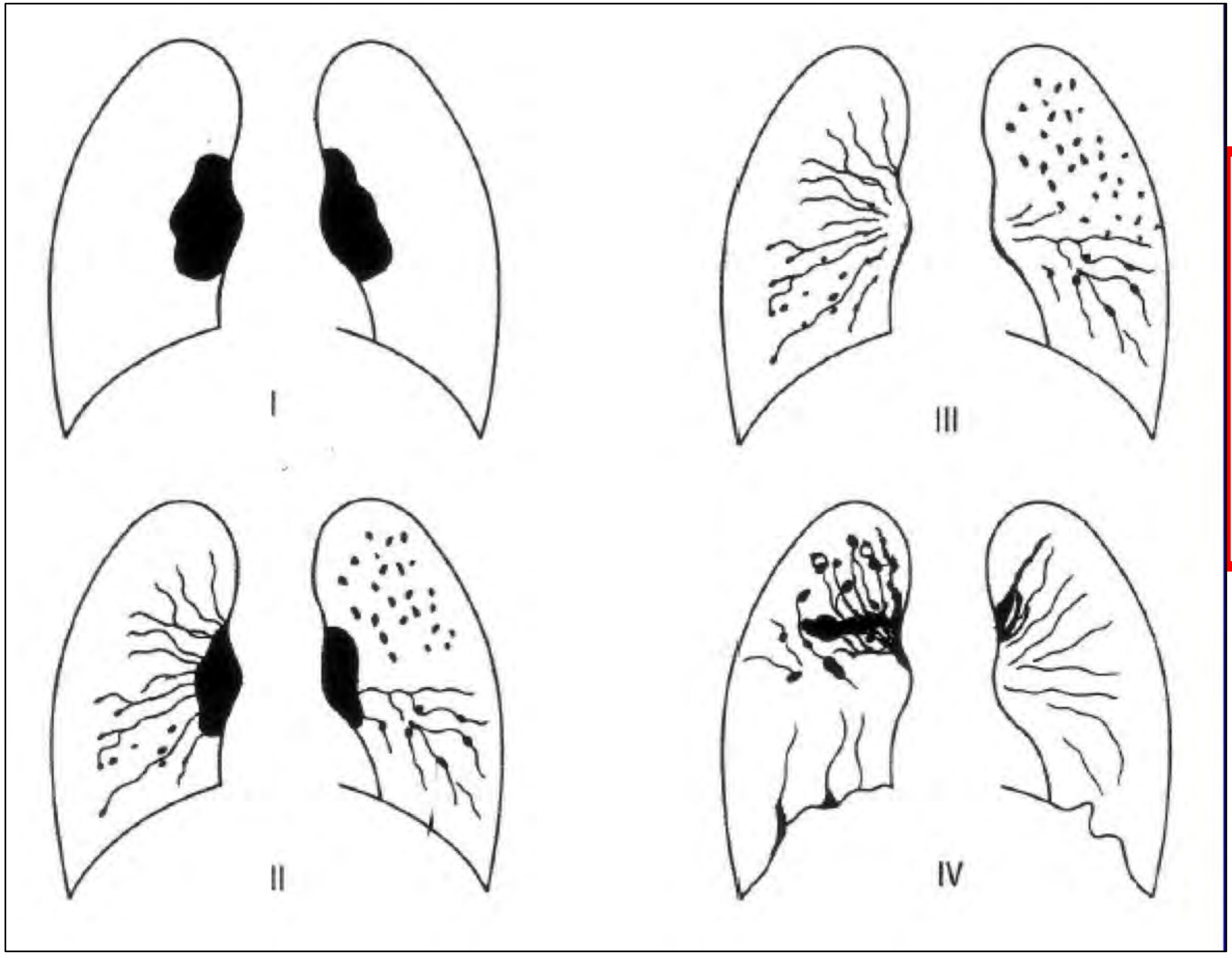


- Su et al identified pathways associated with DE genes IFN, TCR, and CXCL9 in Sarc blood vs controls
- These gene factors act as predictors of disease course in a longitudinal Sarc cohort
- Initial high CXCL9 expression associated w/ more severe disease longitudinally

Su et al. Eur Resp J 2014

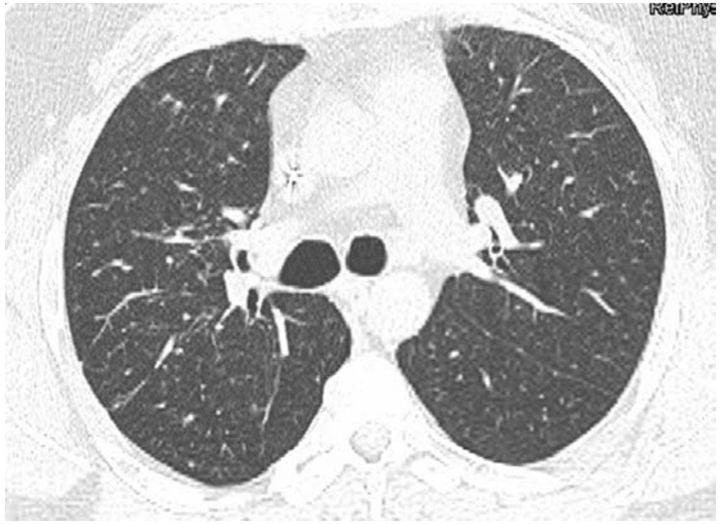
Scadding Staging

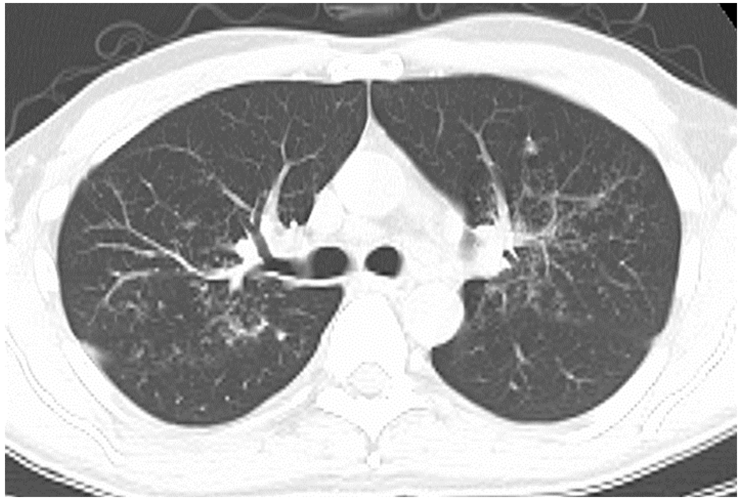
- Prognostic value for remission
- Easy to use

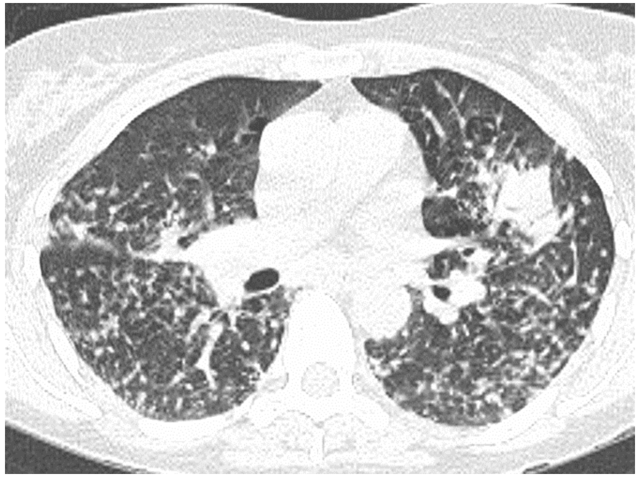


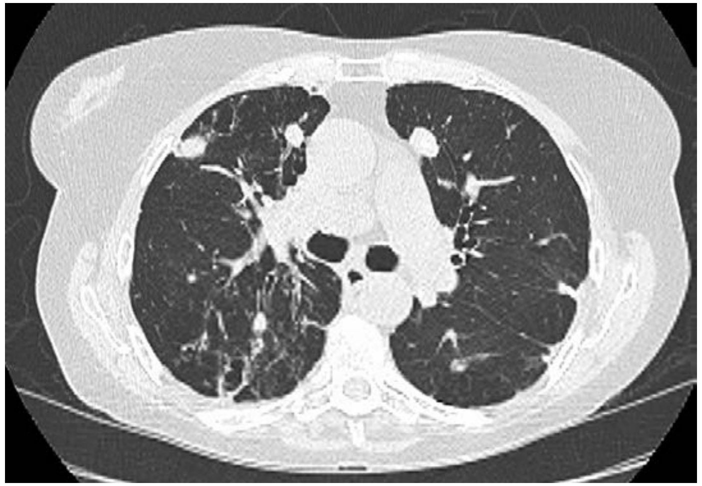
- Not linear
- Interobserver variability
- Low sensitivity/specificity for fibrosis
- Correlation with lung function variable

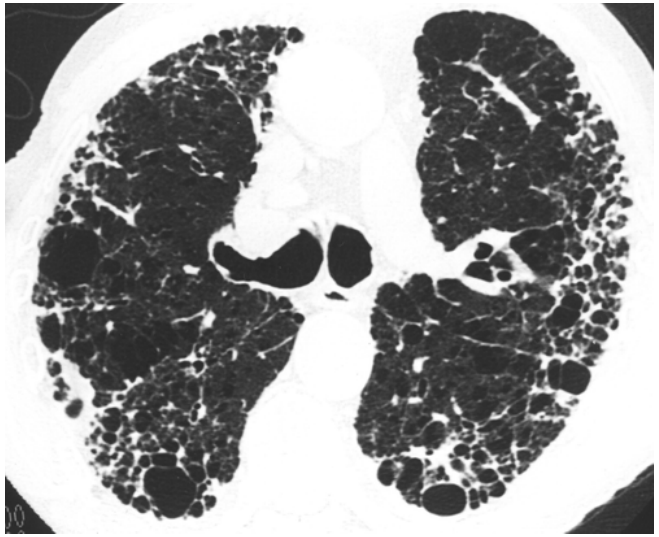
Variable Manifestations of Sarcoidosis on HRCT





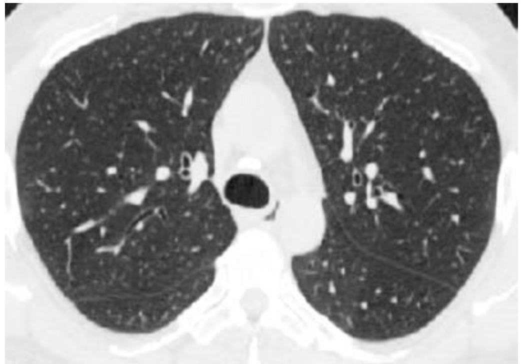






National Jewish Health®

Can information from HRCT better inform disease classification and prognosis?

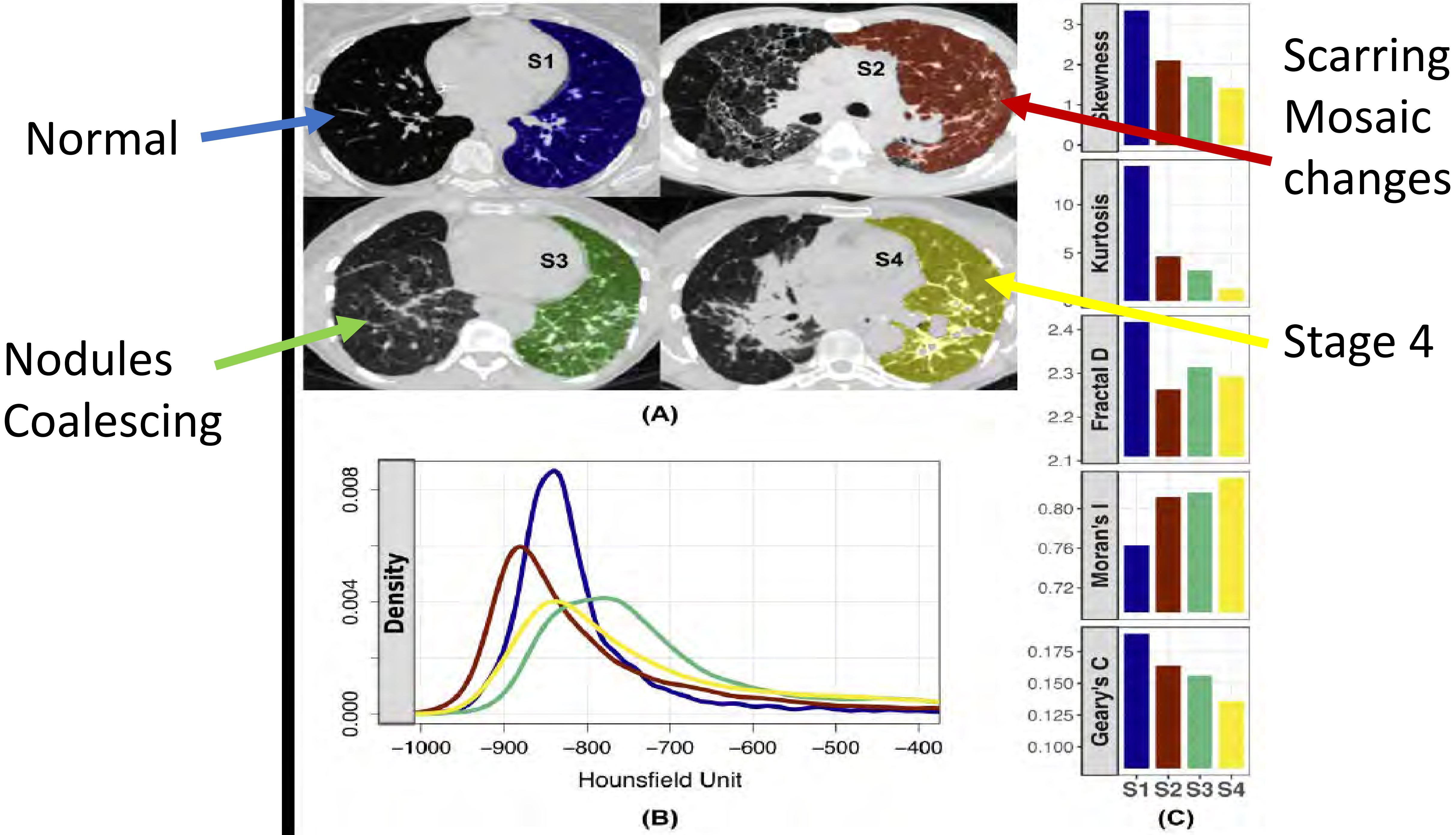


- **Radiomics** –extracts large quantities of texture and related measures from medical images
- Good for detecting differences in CT patterns; lung cancer, COPD

Hypothesis: Detailed radiomic analysis of lung CT images in sarcoidosis will identify new more refined subtypes of pulmonary disease

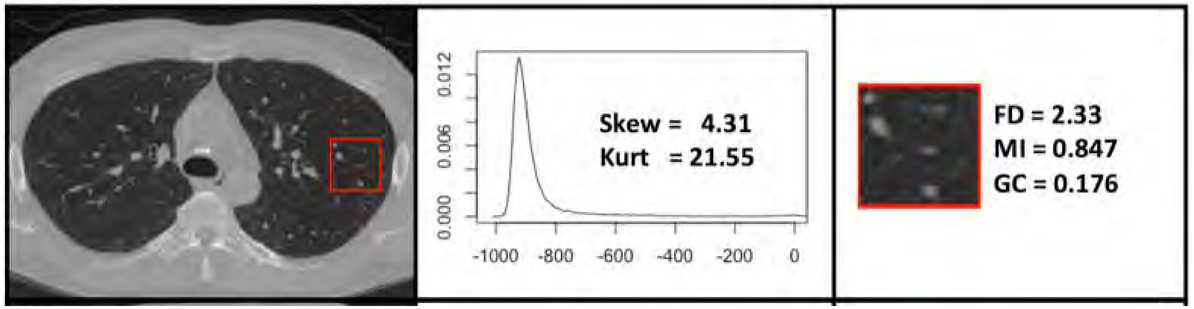
Study Design:

- Exploratory case-control study using NJH GRADS sarcoidosis cases (N=73) and COPD Non-smoking healthy controls (N=78)
- Classification study using only GRADS sarcoidosis cases (N=321)

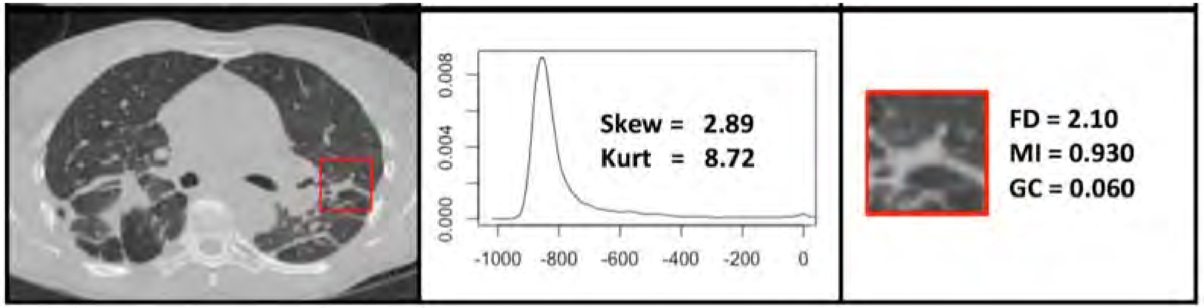

Normal
Nodules
Coalescing
Scarring
Mosaic
changes
Stage 4
S1
S2
S3
S4
(A)
Density
0.000
0.004
0.008
−1000
−900
−800
−700
−600
−500
−400
Hounsfield Unit
(B)
Skewness
0
1
2
3
Kurtosis
5
10
Fractal D
2.1
2.2
2.3
2.4
Moran's I
0.72
0.76
0.80
Geary's C
0.100
0.125
0.150
0.175
S1 S2 S3 S4
(C)

National Jewish Health®

Differences in Radiomic Measures between Sarcoidosis & Controls and Scadding Stage



Healthy



Sarcoidosis

Radiomic Feature	Control (N = 78)	Sarcoidosis (N = 73)	P-value
Skewness	3.615 (0.037)	3.252 (0.064)	<0.001
Kurtosis	16.12 (0.299)	13.0 (0.451)	<0.001
Fractal D	2.269 (0.005)	2.236 (0.005)	<0.001
Moran's I	0.815 (0.002)	0.838 (0.003)	<0.001
Geary's C	0.135 (0.002)	0.120(0.002)	<0.001

*adj for age, gender, BMI

Ryan ERJ 2019

National Jewish Health®

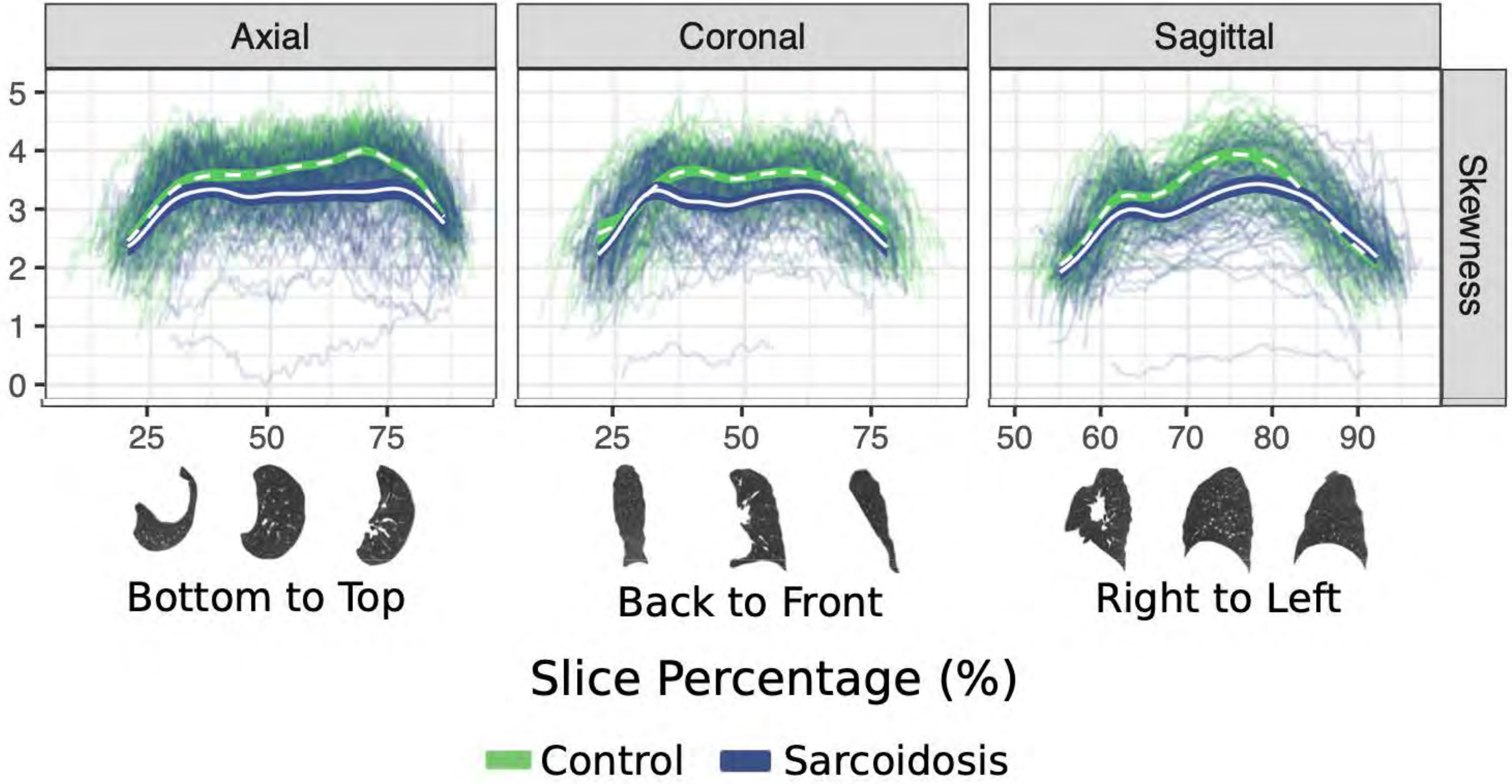
Differences in Radiomic Measures by Scadding Stage

* vs Stage IV; † vs Stage III

Radiomic Feature	Stage 0 (N = 9)	Stage I (N = 8)	Stage II (N = 28)	Stage III (N = 11)	Stage IV (N = 17)	P-value
Skewness	3.356 (0.099)*	3.631 (0.072)*	3.301 (0.103)*	3.479 (0.095)*	2.791 (0.150)	<0.001
Kurtosis	13.71 (0.785)*	16.06 (0.521)*	13.33 (0.769)*	14.65 (0.850)*	9.57 (0.793)	<0.001
Fractal D	2.259 (0.011)*	2.246 (0.008)*	2.234 (0.006)*	2.254 (0.015)*	2.210 (0.010)	<0.001
Moran's I	0.827 (0.008)*	0.816 (0.003)*	0.840 (0.005)*	0.819 (0.007)*	0.863 (0.007)	<0.001
Geary's C	0.131 (0.005)*	0.126 (0.004)*	0.117 (0.003)*†	0.136 (0.007)*	0.105 (0.005)	<0.001

National Jewish Health®

Differences in Spatial Radiomic Measures Between Sarcoidosis and Controls



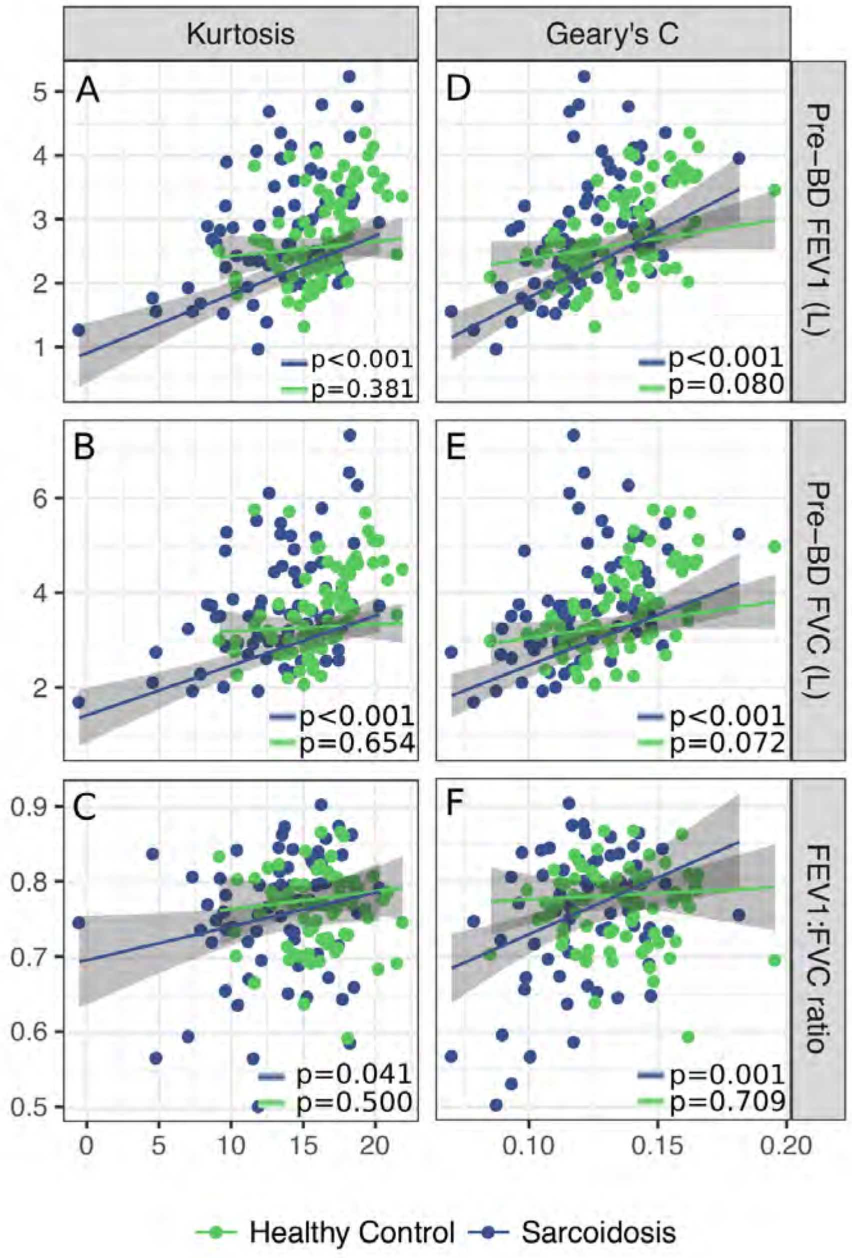
National Jewish Health®

Spirometry is Associated with Radiomic Measures

- Associations stronger and statistically associated with sarcoidosis compared to controls

- Radiomics explained 70.3% of the variability in FVC vs 51.4% for Scadding staging alone

Ryan ERJ 2019



Developing New Radiomic Phenotypes

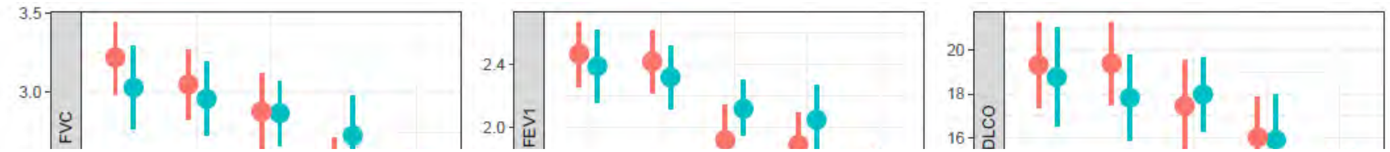
- Entire GRADS Cohort
- Used first and second order radiomic features from each lung
- Corrected for Scanner Effect
- K-means variable selection clustering- for high dimensional, dependent data
- Gap statistic to determine the optimal number of clusters n=5
- Describe clinical aspects of clusters: Lung Function, Qs, Biomarkers
- Compare to Scadding Stage

Cluster Analysis of Radiomic Features: 5 clusters

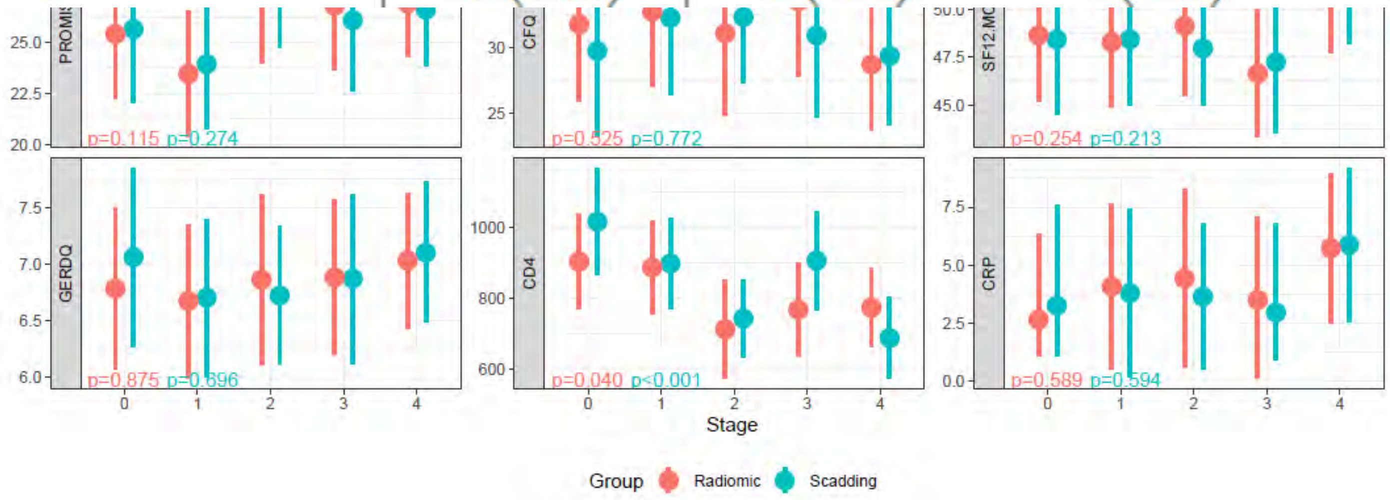
	Overall	0	1	2	3	4	p
Sample Size	321	83	54	58	45	81	
Age	52.91 (9.91)	53.28 (9.96)	49.76 (10.97)	52.49 (10.36)	52.56 (8.86)	55.12 (8.92)	0.043
Sex	147 (45.8)	44 (53.0)	20 (37.0)	26 (44.8)	17 (37.8)	40 (49.4)	0.286
Race							0.005
White	233 (73.0)	72 (87.8)	42 (77.8)	38 (66.7)	34 (75.6)	47 (58.0)	
Black	77 (24.1)	8 (9.8)	12 (22.2)	16 (28.1)	10 (22.2)	31 (38.3)	
Other	9 (2.8)	2 (2.4)	0 (0.0)	3 (5.3)	1 (2.2)	3 (3.7)	
Height	66.98 (4.16)	67.83 (4.55)	66.72 (4.04)	66.43 (3.93)	66.76 (4.09)	66.81 (3.98)	0.293
BMI	30.62 (6.49)	32.89 (6.04)	29.36 (6.84)	32.51 (6.33)	26.88 (5.98)	29.84 (5.89)	<0.001
Scadding							<0.001
0	43 (13.5)	21 (25.6)	7 (13.0)	5 (8.8)	5 (11.4)	5 (6.2)	
1	63 (19.8)	27 (32.9)	17 (31.5)	8 (14.0)	4 (9.1)	7 (8.6)	
2	92 (28.9)	19 (23.2)	18 (33.3)	23 (40.4)	12 (27.3)	20 (24.7)	
3	44 (13.8)	13 (15.9)	10 (18.5)	9 (15.8)	6 (13.6)	6 (7.4)	
4	76 (23.9)	2 (2.4)	2 (3.7)	12 (21.1)	17 (38.6)	43 (53.1)	

Demographics ordered from least severe (0) to most severe (4) based on FVC percent predicted.

Clinical Aspects of “Radiomic Biomarker” Class



	Overall	0	1	2	3	4	p
Sample Size	321	83	54	58	45	81	
Scadding							<0.001
0	43 (13.5)	21 (25.6)	7 (13.0)	5 (8.8)	5 (11.4)	5 (6.2)	
1	63 (19.8)	27 (32.9)	17 (31.5)	8 (14.0)	4 (9.1)	7 (8.6)	
2	92 (28.9)	19 (23.2)	18 (33.3)	23 (40.4)	12 (27.3)	20 (24.7)	
3	44 (13.8)	13 (15.9)	10 (18.5)	9 (15.8)	6 (13.6)	6 (7.4)	
4	76 (23.9)	2 (2.4)	2 (3.7)	12 (21.1)	17 (38.6)	43 (53.1)	



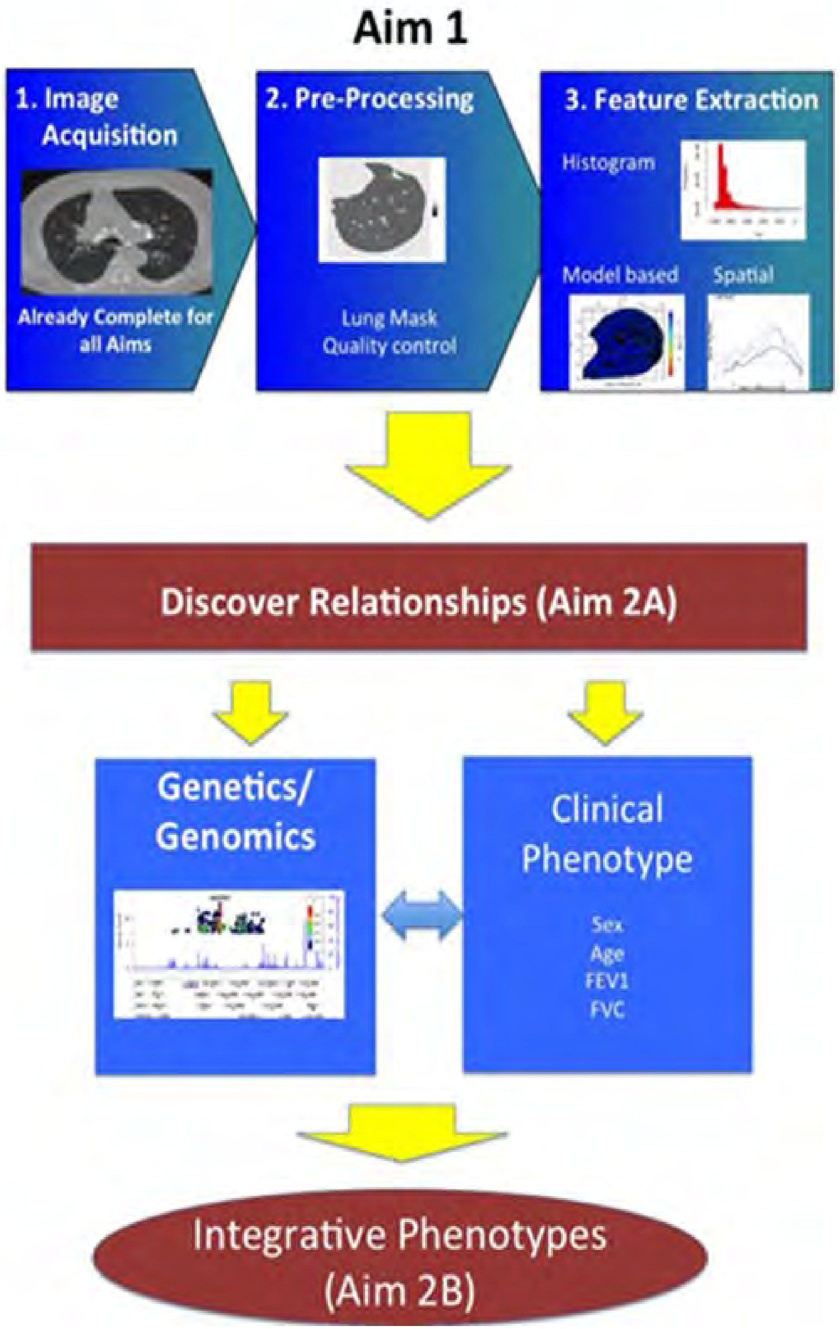
- Scadding Stage similar

visual scoring of CT

National Jewish Health®

Future Directions: Integrated Biomarker of Sarcoidosis

- GRADS Cohort
  - Visual Scores- 2 readers
  - BAL transcriptome
  - GWAS data
- Validation cohort NJH and CCF
- Determine function over time



National Jewish Health®

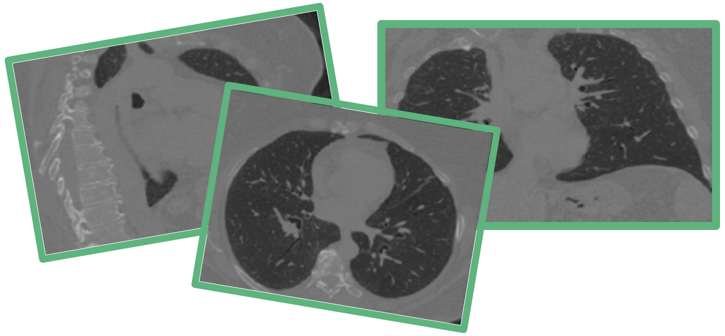
Conclusions:

- Radiomic analyses of Sarcoidosis Chest CT :
  - Differentiates cases and controls
  - Associate with lung function and PROs
  - Classify sarcoidosis subtypes and may be useful as a biomarker
- Potential as a biomarker to augment and or standardize use of CT in clinic
- Integrative Omics approaches show promise
- There is now publicly-available, open-source software to perform many of these analyses for your own data

National Jewish Health®

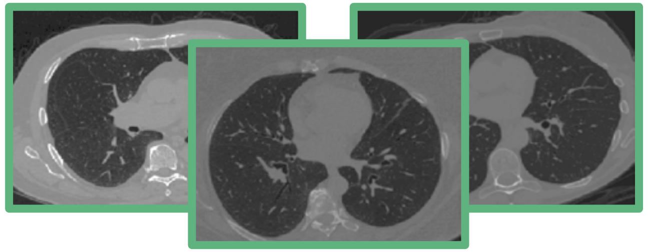
Processing Chest CTs

Software: `lungct` R package
https://github.com/ryansar/lungct



Functionality

1. Convert DICOM to NIfTI
2. Resample to 1x1x1 mm^3
3. Segment the left and right lungs
4. Calculate radiomic features on 2D slice or 3D lung
5. Register lungs to the standard lung template
6. Create study-specific templates
7. Calculate the amount of lung shrinkage
8. And more!



Acknowledgements

National Jewish Health

- Tasha Fingerlin, PhD
- Clara Restrepo, MD
- Li Li, MD PhD
- Nabeel Hamzeh, MD
- Shu-Yi Liao MD, ScD
- Andy Levy, MD
- Peggy Mroz, MSPH
- Bri Barkes, MPH
- Christina Riley
- May Gillespie
- Kristyn MacPhail
- Jill Elliott

UCD

- Nichole Carlson, PhD
- Sarah Ryan, MS, PhD Student

CCF

- Dan Culver, MD

NIH:

1R01 HL114587
1R01 HL142049
UL 1TR002535
U01 HL112695, U01 HL112707

Our patients!