Medical Imaging and Big Data

Big Data Seminar Series
Center for Innovative Design and Analysis

Antonio R. Porras, PhD

Fuyong Xing, PhD

colorado school of public health
UNIVERSITY OF COLORADO
COLORADO STATE UNIVERSITY
UNIVERSITY OF NORTHERN COLORADO

University of Colorado
Anschutz Medical Campus

Outline

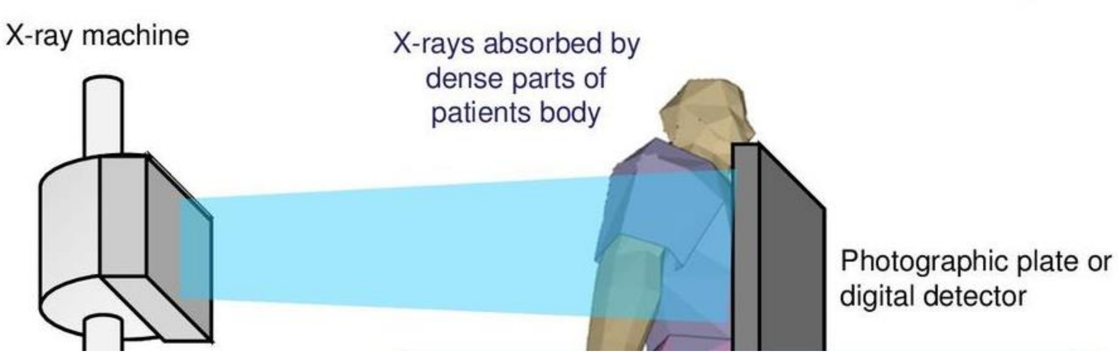
- Overview of medical image modalities

- Medical Image Computing and data availability
    - No datasets
    - Limited datasets
    - Big Data

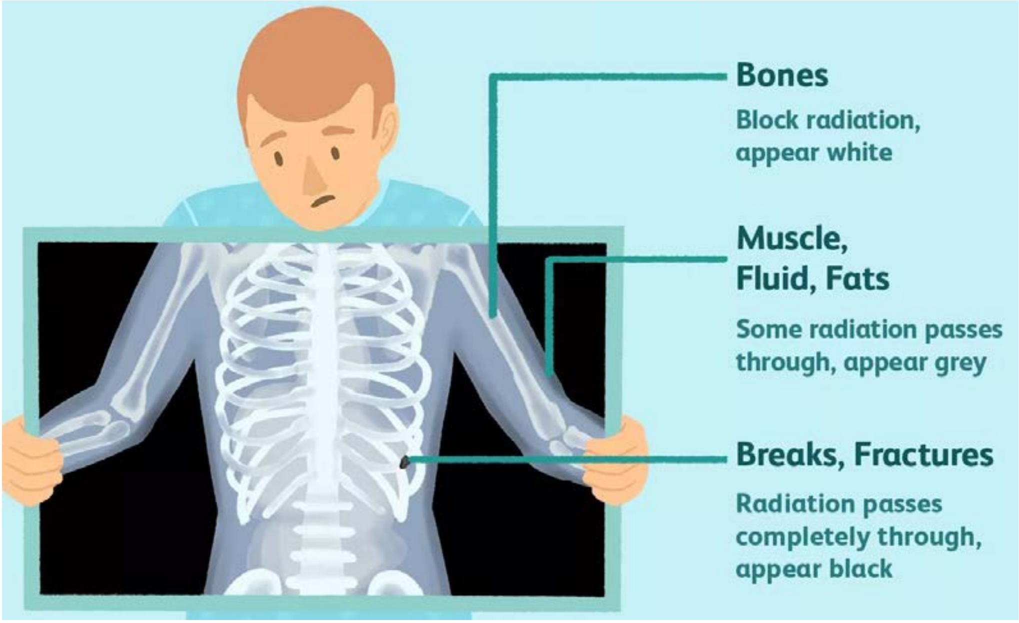
Medical Imaging

- X-ray and CT
- Magnetic Resonance
- Ultrasound
- Nuclear imaging
- Microscopy

X-ray Imaging



www.stem.org.uk



www.verywellhealth.com

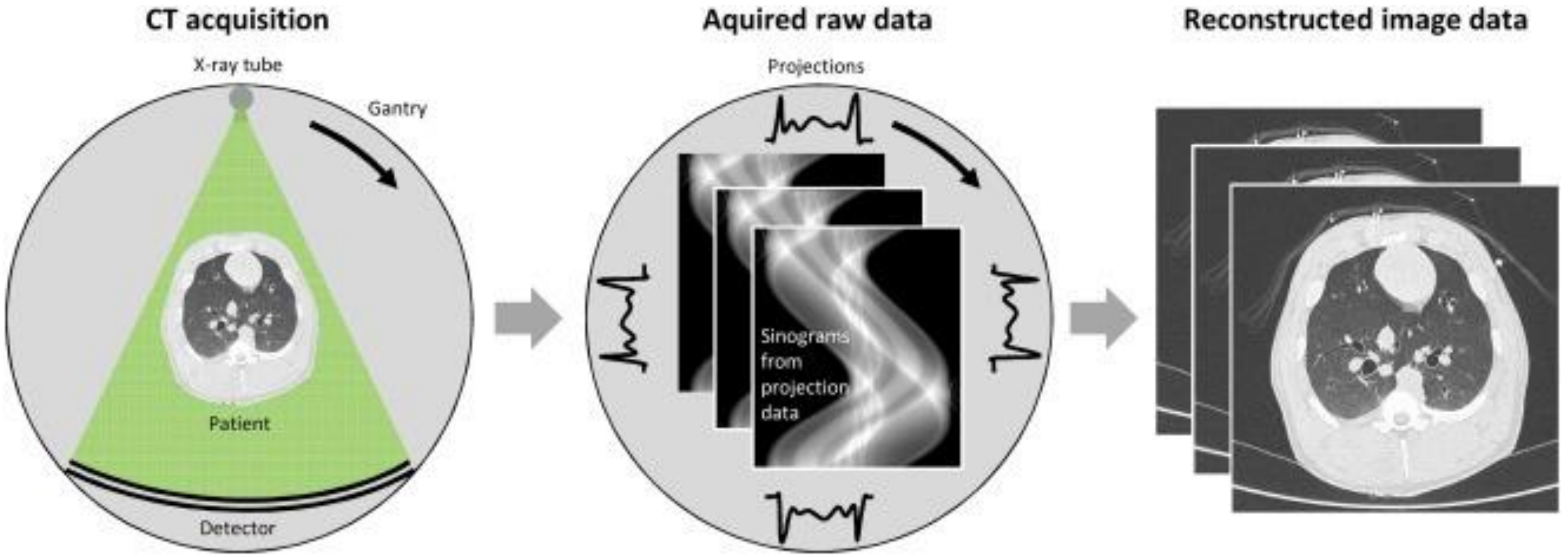
Clear boundaries between tissues with different radiodensity

Fast and cheap

Harmful radiation

Cumulated 2D projection

CT Imaging

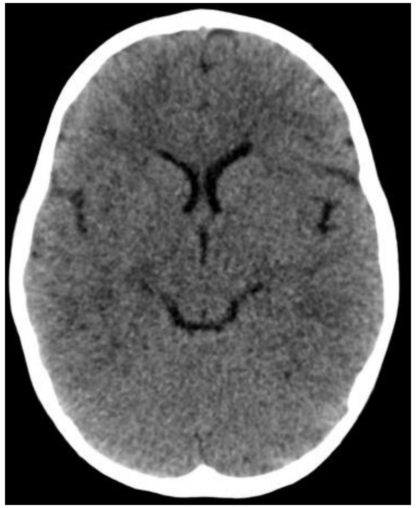


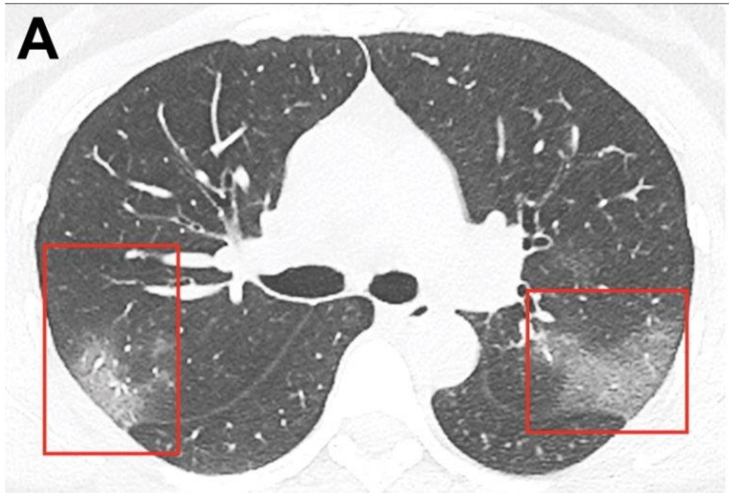
Clear boundaries between tissues with different radiodensity

Relatively fast

Standardized information (Hounsfield scale)

[Stiller, Basics of iterative reconstruction methods in computed tomography: A vendor-independent overview, Eur. J. Radiol., 2008]



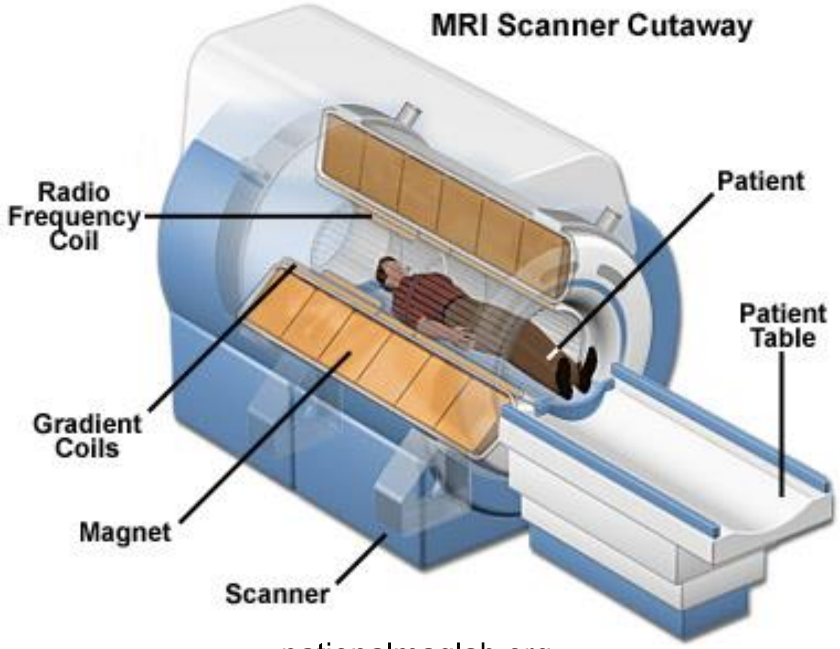


[Lei et al, CT Imaging of the 2019 Novel Coronavirus (2019-nCoV), Pneumonia Radiol, 2020]

Harmful radiation

Sometimes poor contrast between soft tissues

Magnetic Resonance Imaging



nationalmaglab.org

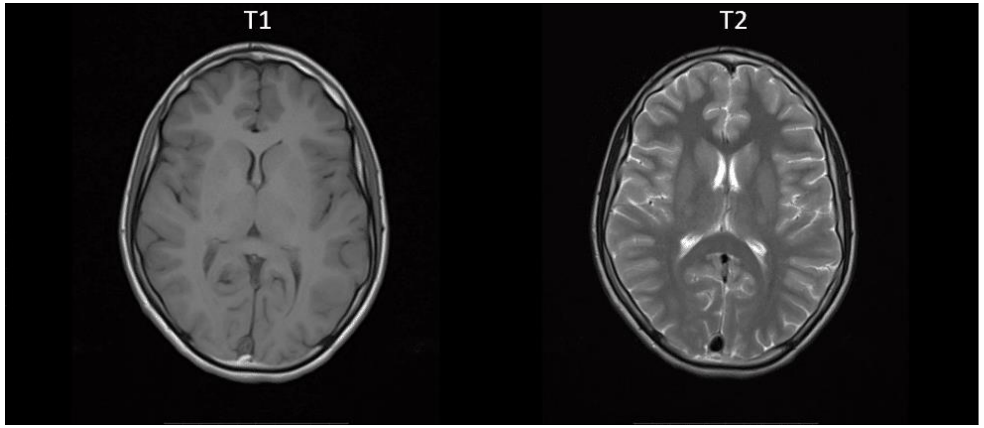
Great for soft tissue

Different protocols can highlight different properties

Non-invasive

Time-consuming

Electro-magnetic field interacts with metals

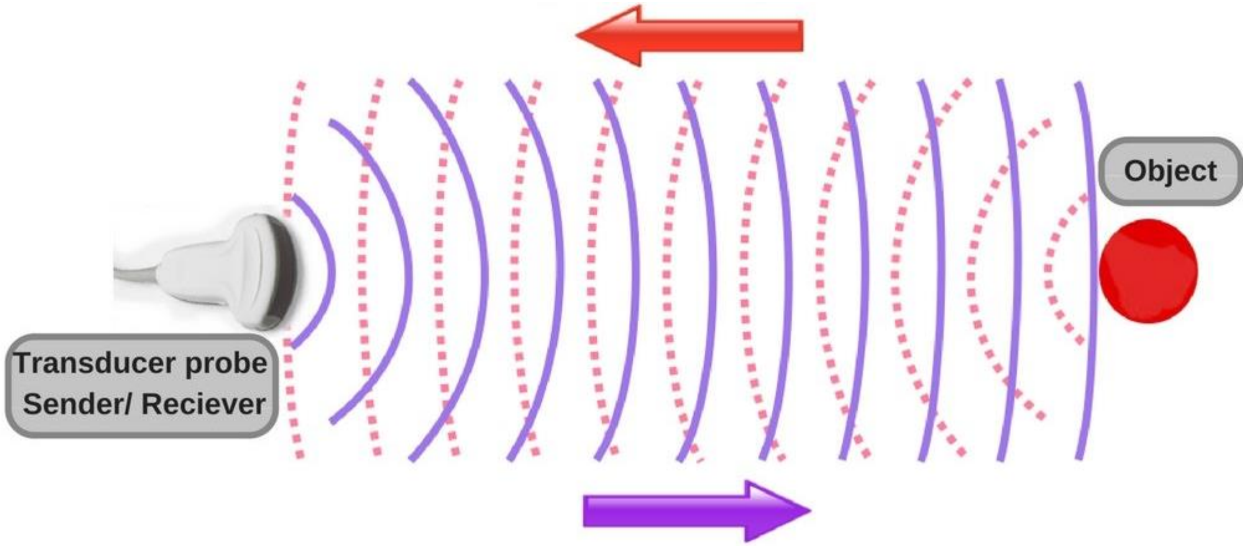


T1-weighted: time of realignment of protons with magnetic field

T2-weighted: time of transversal signal to decay

Functional MRI (BOLD): magnetic properties of deoxygenated hemoglobin

Ultrasound Imaging

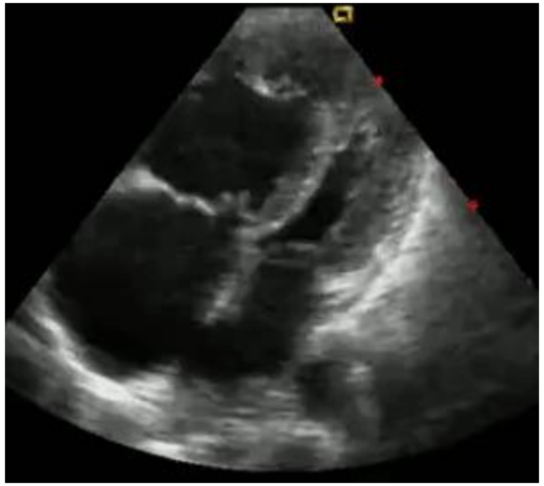


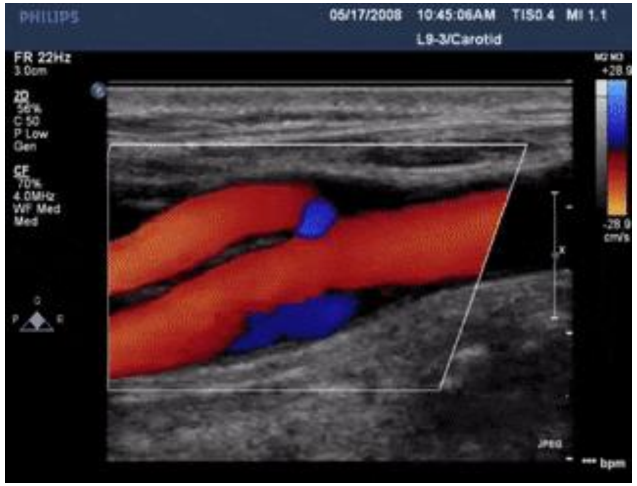
Real time

Objective velocity measurements

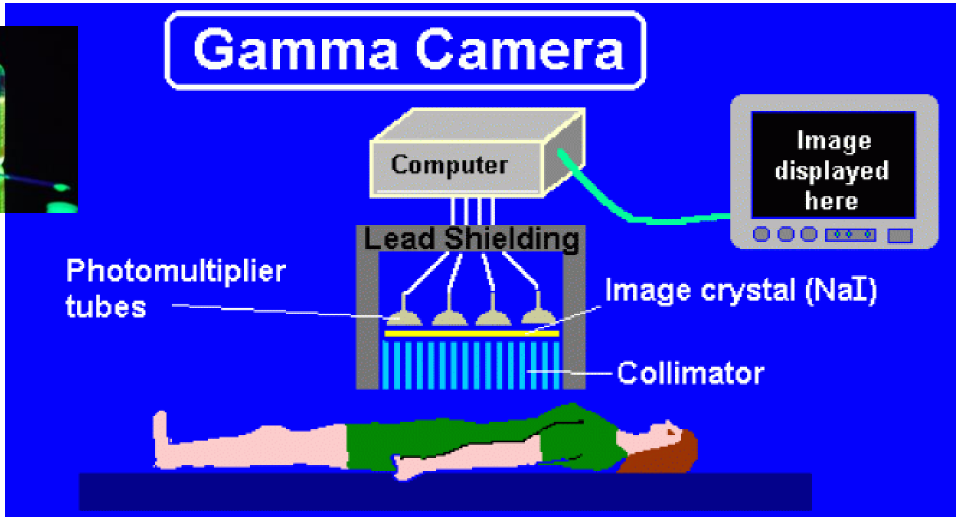
Poor image quality

Limited field of view

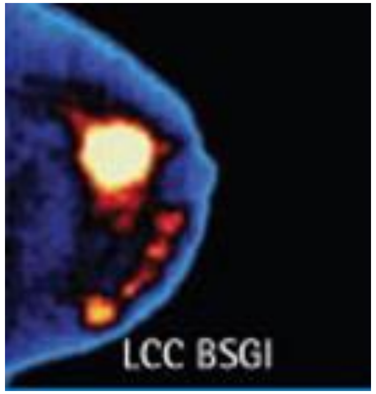


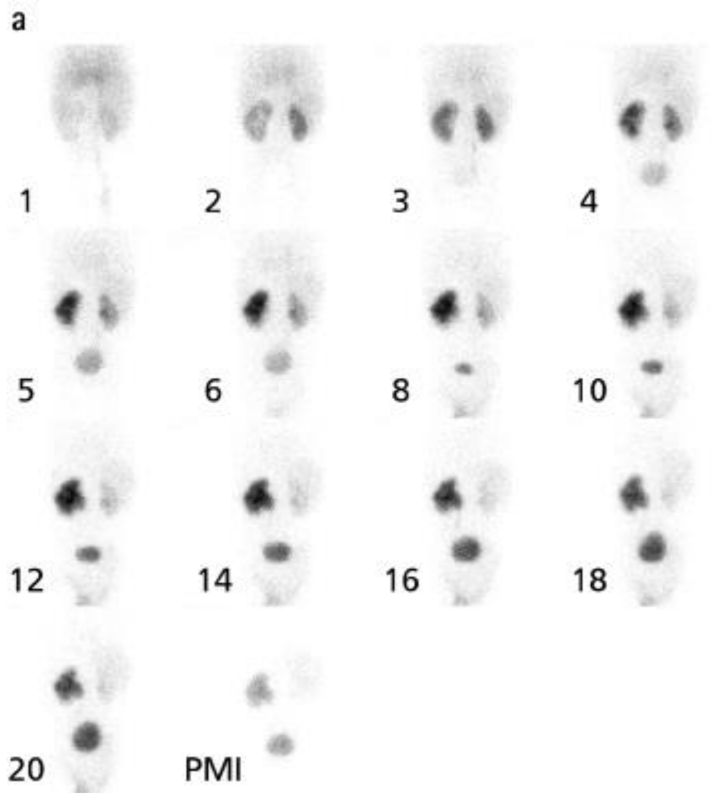


Nuclear Imaging



www.theevolutionofimagingtechnology.net





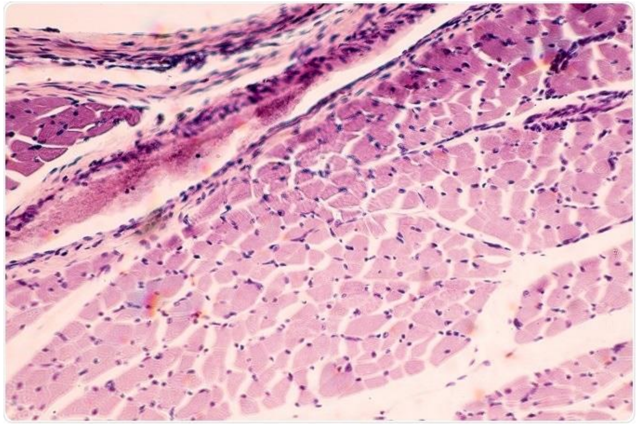
Metabolic information

Functional information

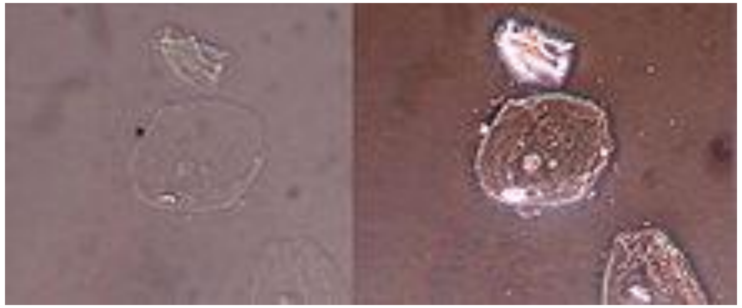
Toxicity

Microscopy

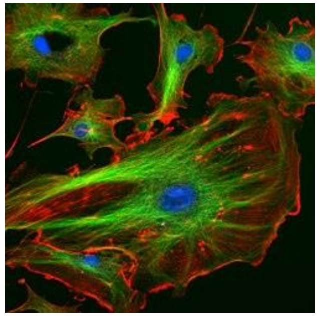
Optical microscopy



Phase contrast microscopy

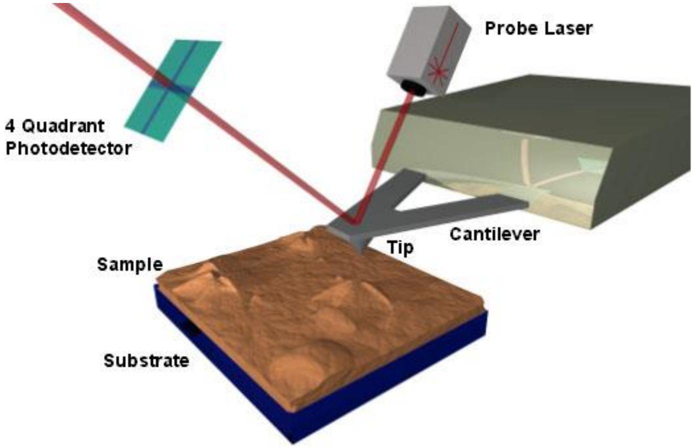


Fluorescent microscopy



Scanning electron microscopy

Atomic forces microscopy

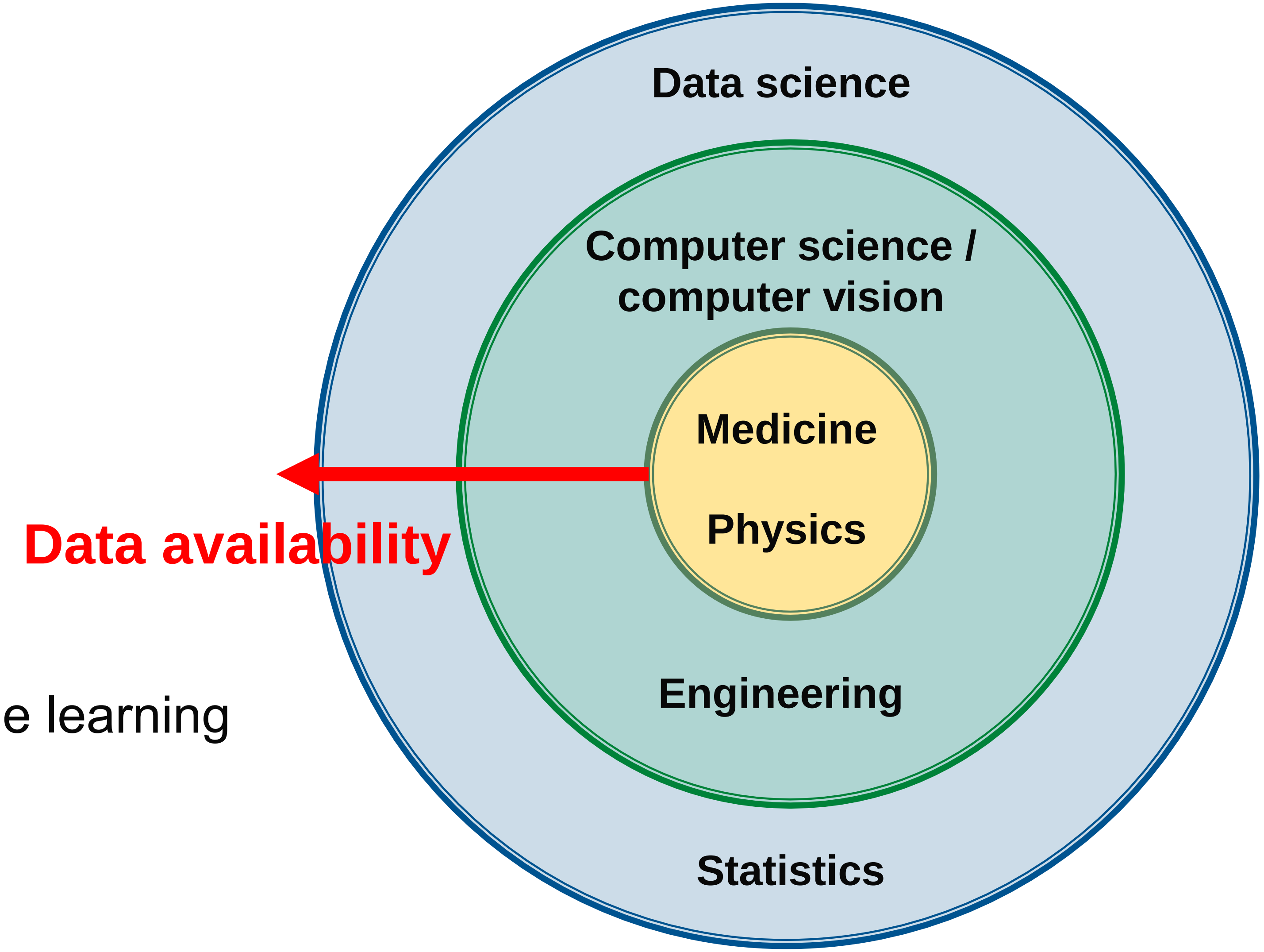


Challenges in Medical Imaging

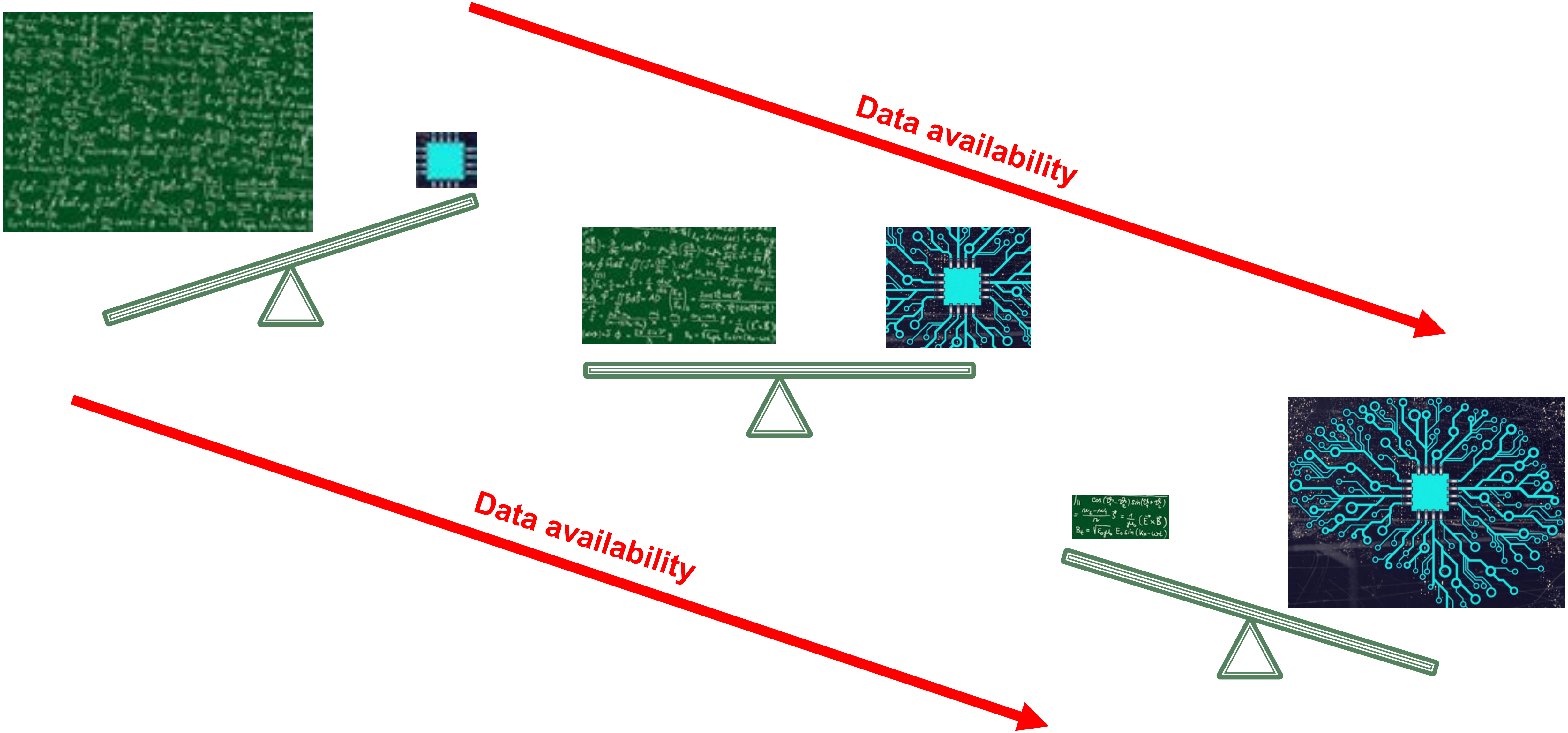
- Every medical image modality only provides partial information
- Images provide much more than the information we are interested in
- Every person has a unique anatomy: variable observations are the norm
- Radiographic image assessment is highly subjective

Medical Image Computing

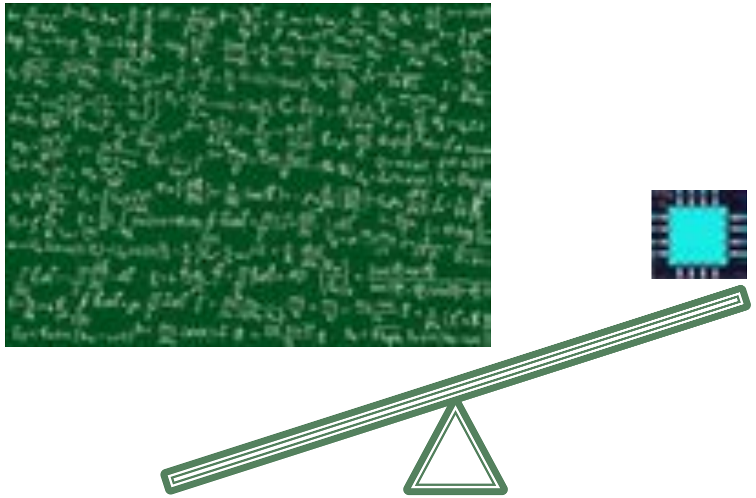
- Segmentation
- Registration
- Modeling
- Radiomics
- Handcrafted feature-based machine learning
- Neural networks



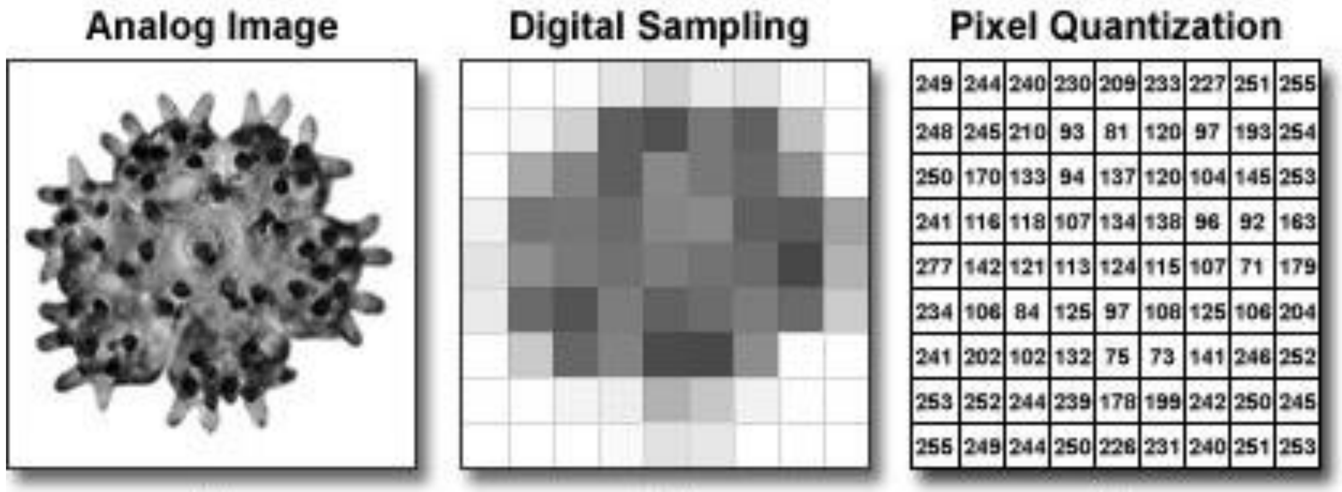
Human vs. data-driven knowledge

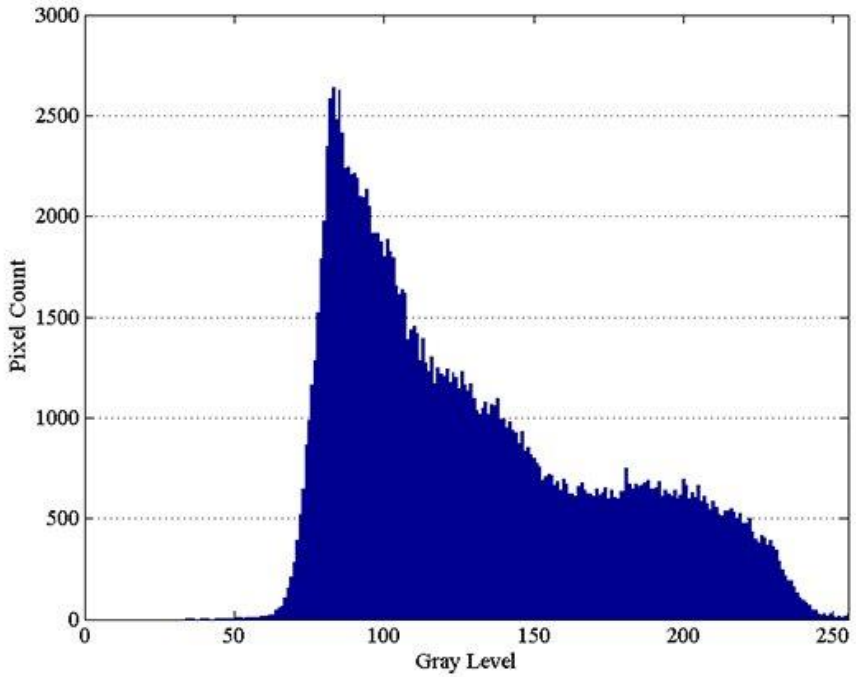


No data



Enhancement



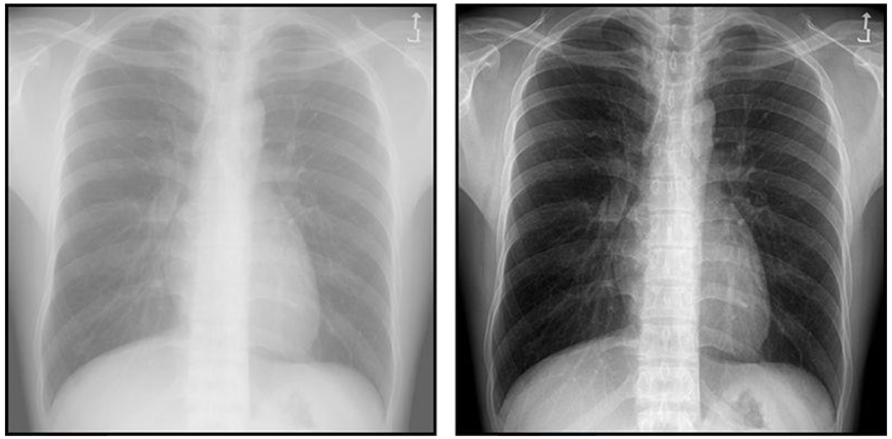


Histogram operations: equalization

Spatial operations: smoothing, sharpening

Frequency-domain operations: high-pass, low-pass, band-pass

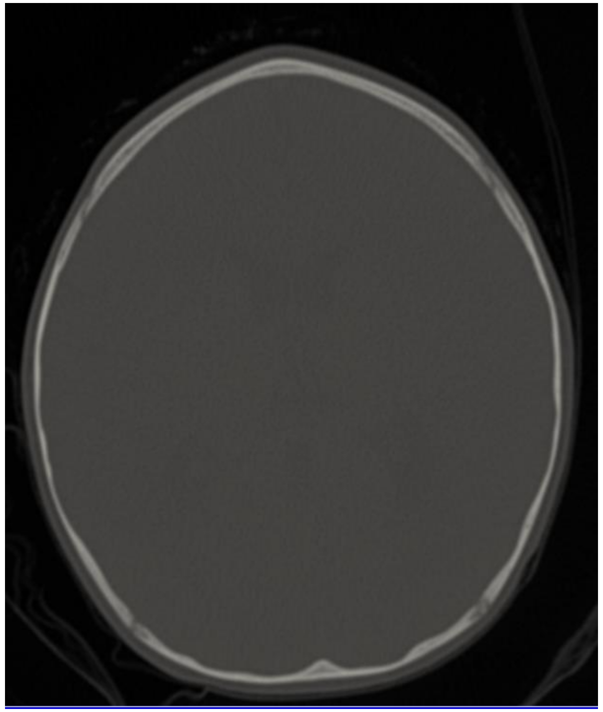
Multi-resolution methods: pyramid-based, wavelets…



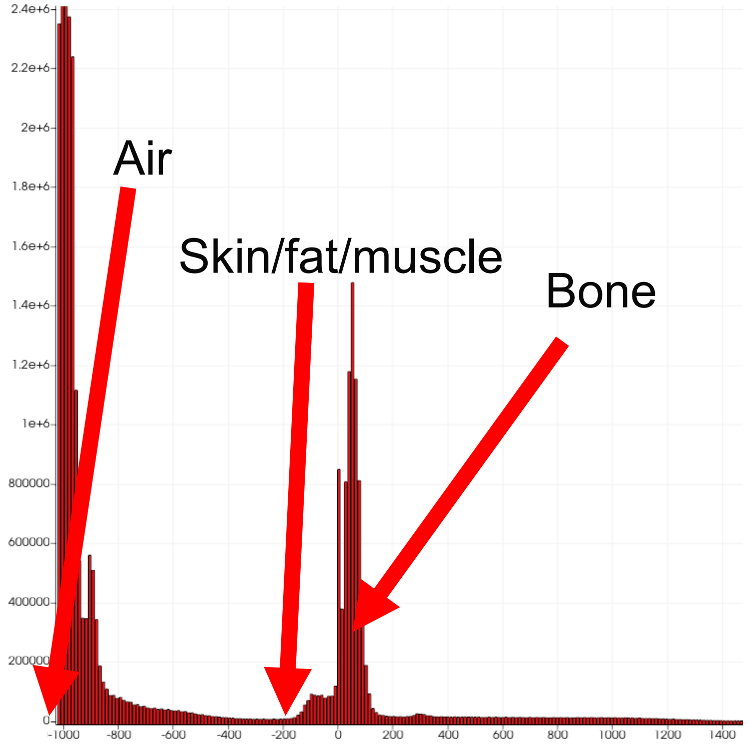
Segmentation

Domain-specific knowledge: intensity range

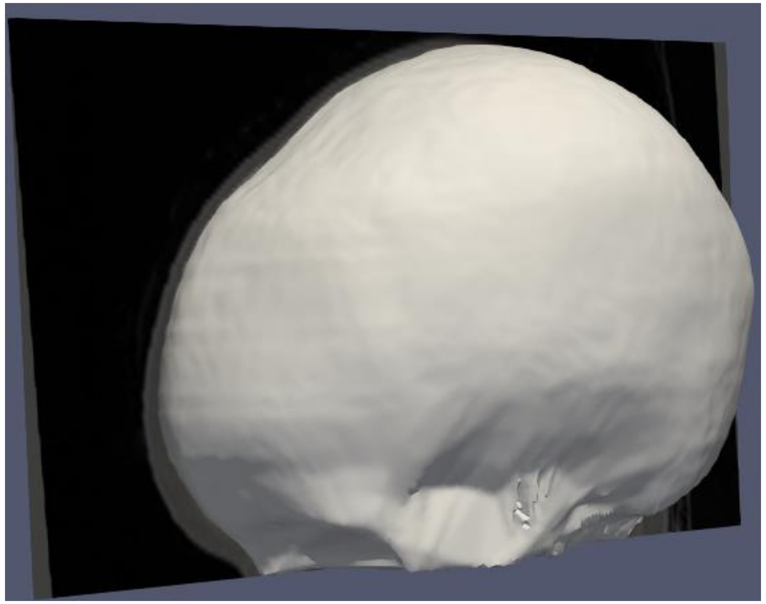
Thresholding



CT image



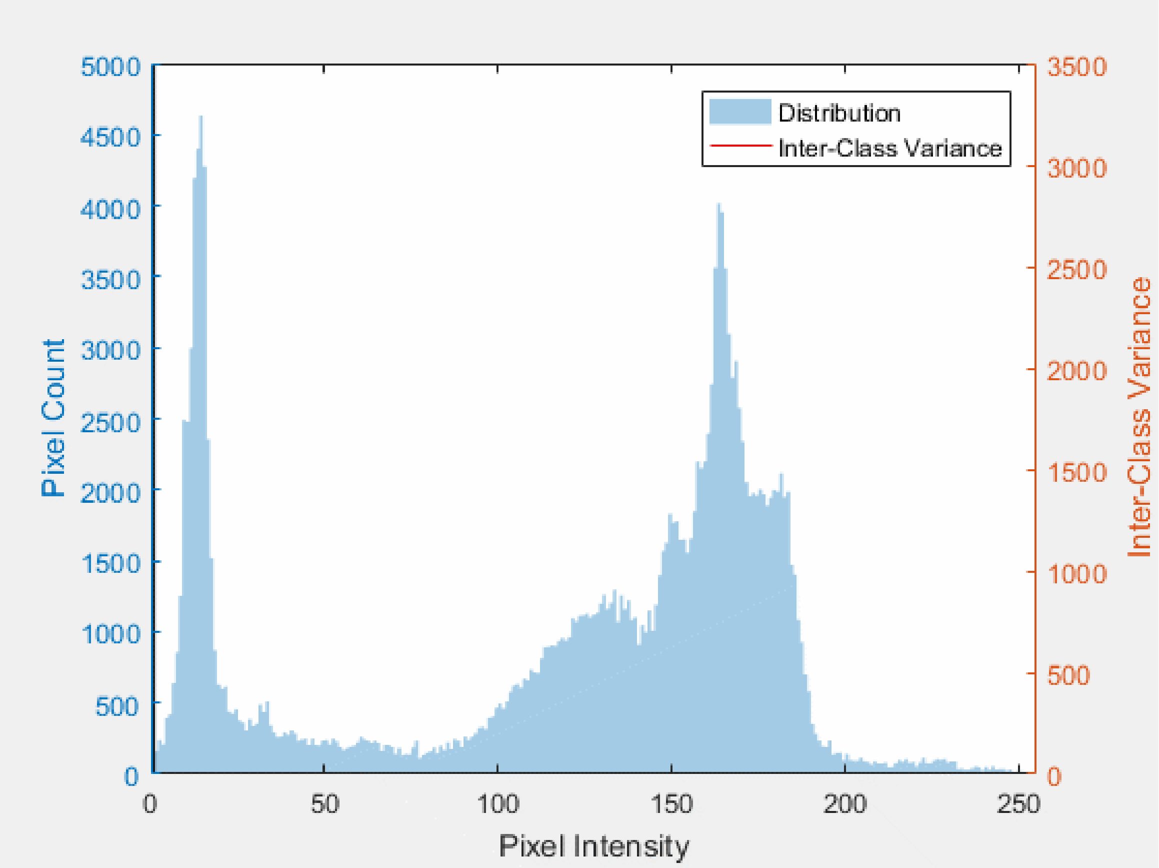
Histogram (Hounsfield units)

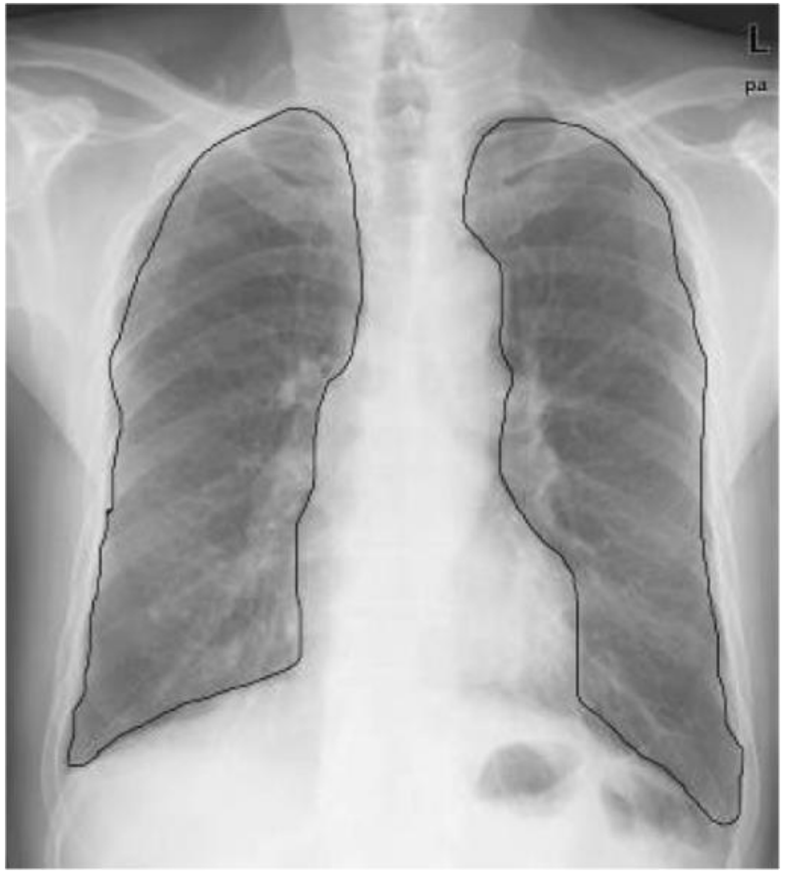


[Porras et al., Locally Affine Diffeomorphic Surface Registration and Its Application to Surgical Planning of Fronto-Orbital Advancement, IEEE Trans. Med. Imaging, 2018]

Segmentation

Domain-specific knowledge: relative intensity range (2 classes)





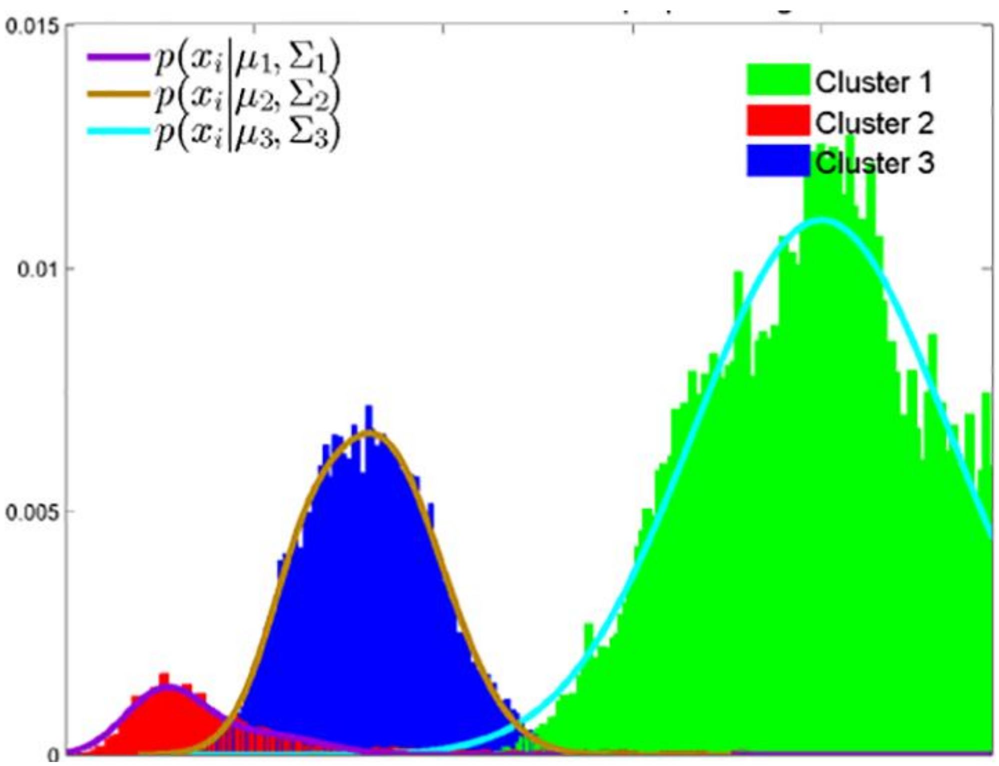
Otsu's algorithm

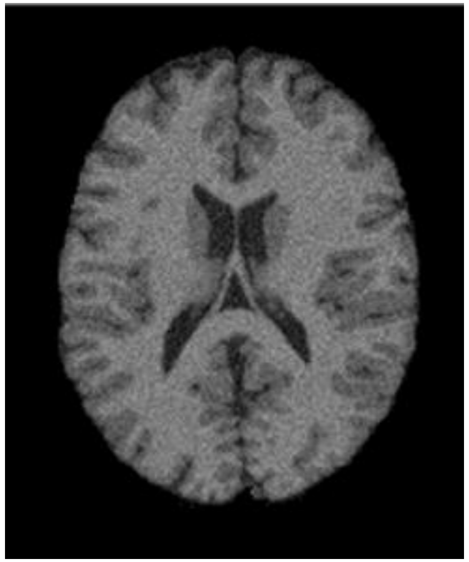
[Zhu et al, The development and evaluation of a computerized diagnosis scheme for pneumoconiosis on digital chest radiographs, Biomed. Eng. Online, 2014]

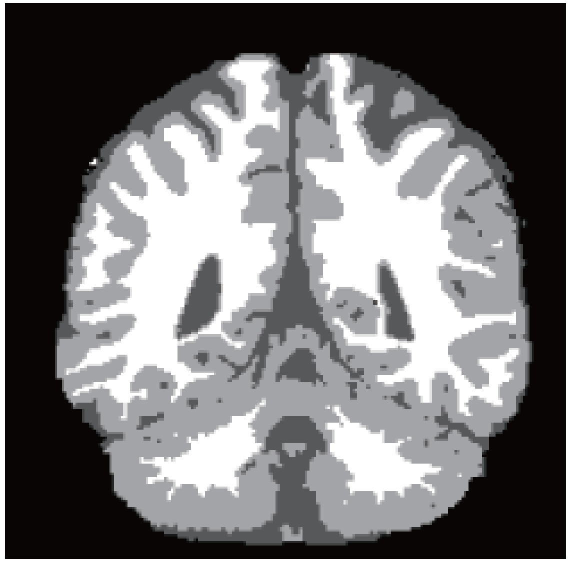
Segmentation

Domain-specific knowledge: intensity distribution (k classes)

Gaussian mixture models and clustering



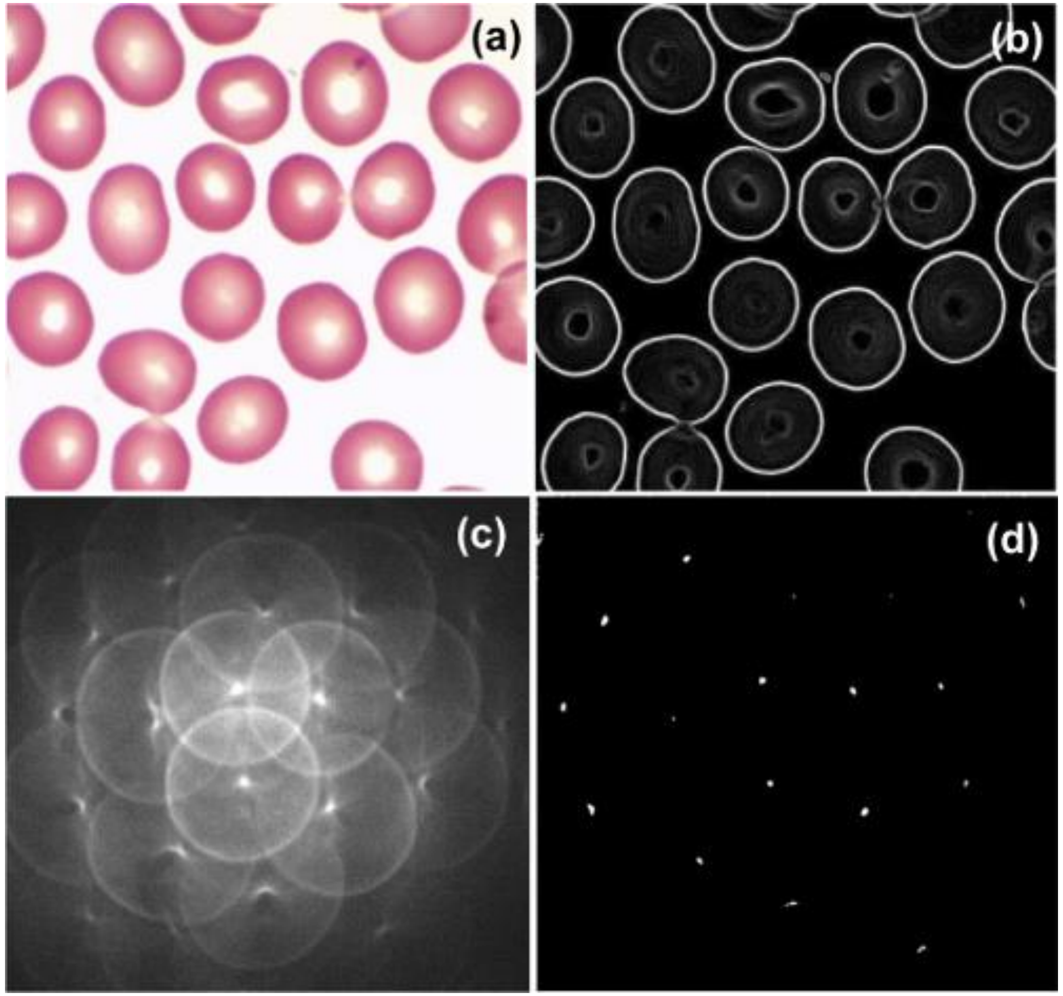




[Ji et al, A Rough Set Bounded Spatially Constrained Asymmetric Gaussian Mixture Model for Image Segmentation, PLOS One, 2017]

Cell identification

Domain-specific knowledge: expected basic shape



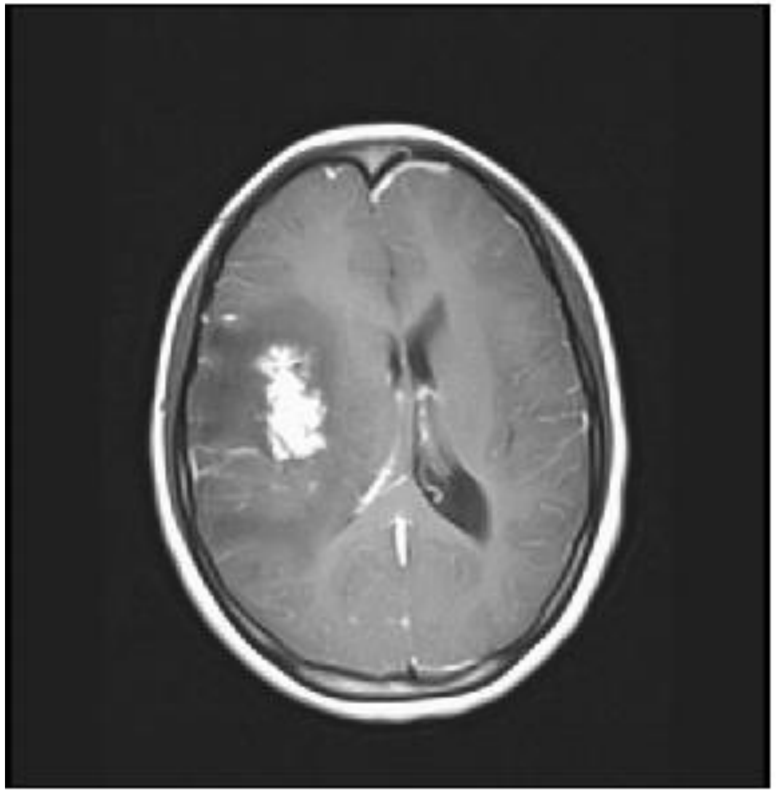
Edge detector

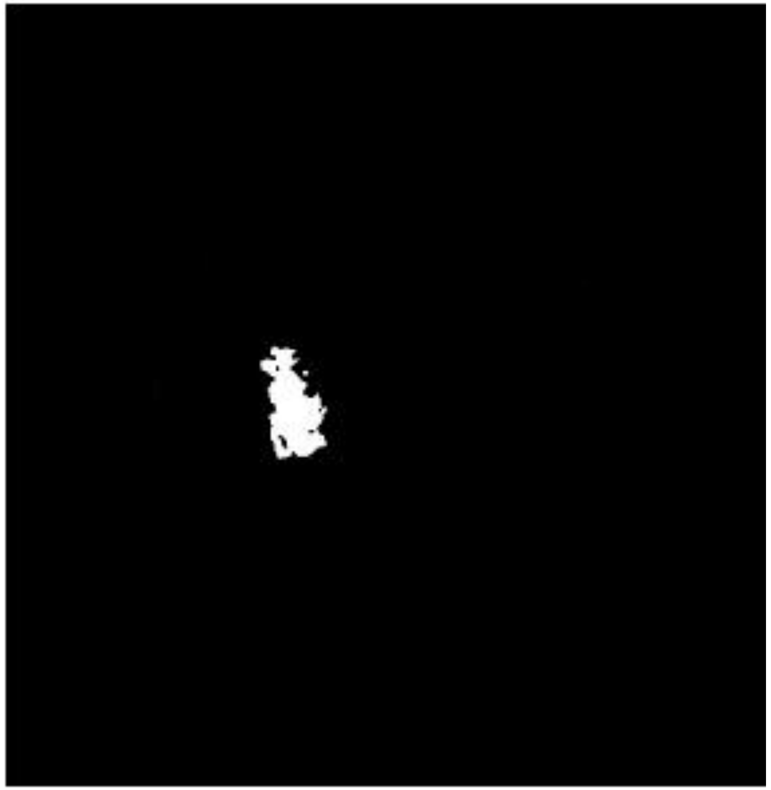
Hough transform

Segmentation

Domain-specific knowledge: location

Region growing methods

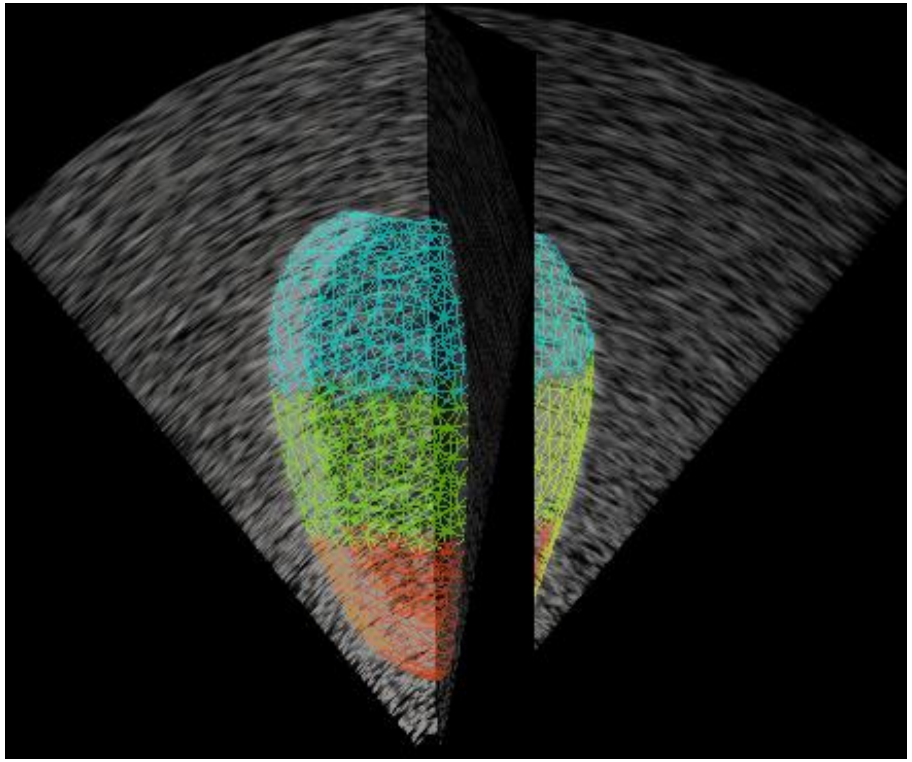


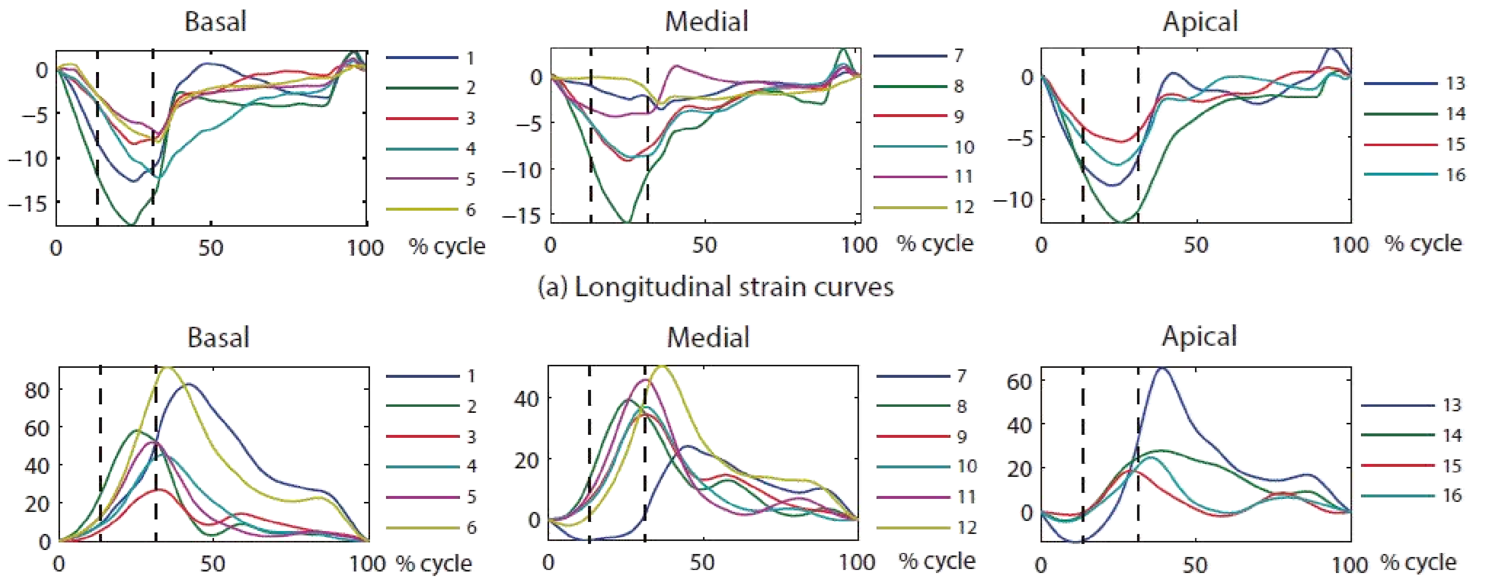


[Deng et al, MRI brain tumor segmentation with region growing method based on the gradients and variances along and inside of the boundary curve, Biomed. Eng. Inform., 2010]

Temporal registration

Domain-specific knowledge: temporal consistency





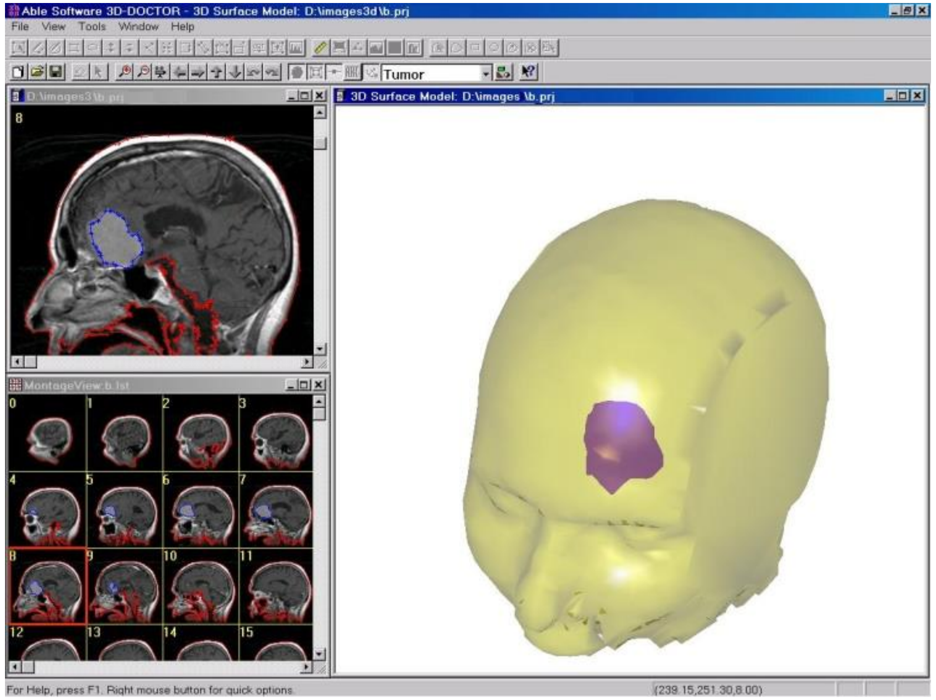
(a) Longitudinal strain curves

(b) Radial strain curves

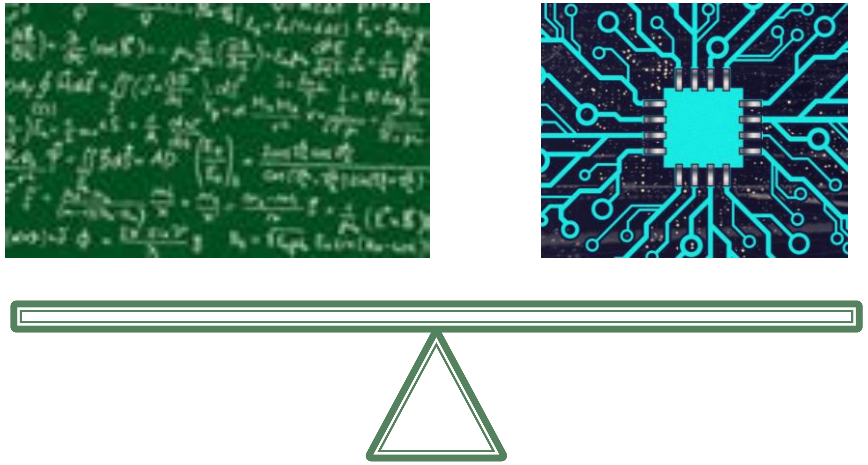
[Porras et al, Integration of multi-plane tissue Doppler and B-mode echocardiographic images for left ventricular motion estimation. IEEE Trans. Med. Imag. 2016]

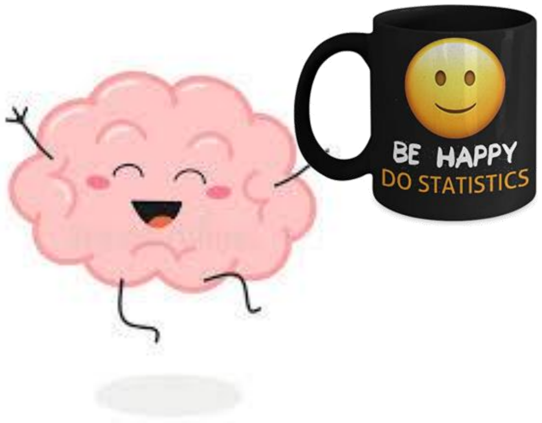
[Porras et al., Improved Myocardial Motion Estimation Combining Tissue Doppler and B-Mode Echocardiographic Images, IEEE Trans. Med. Imaging, 2014]

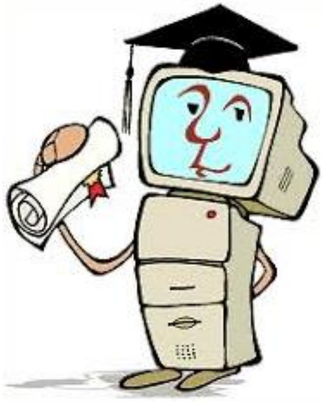
What can we do?



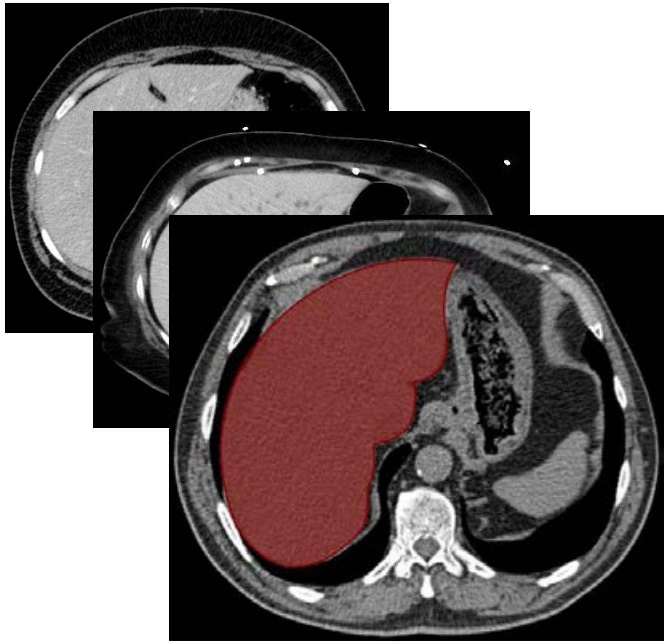
Datasets become available

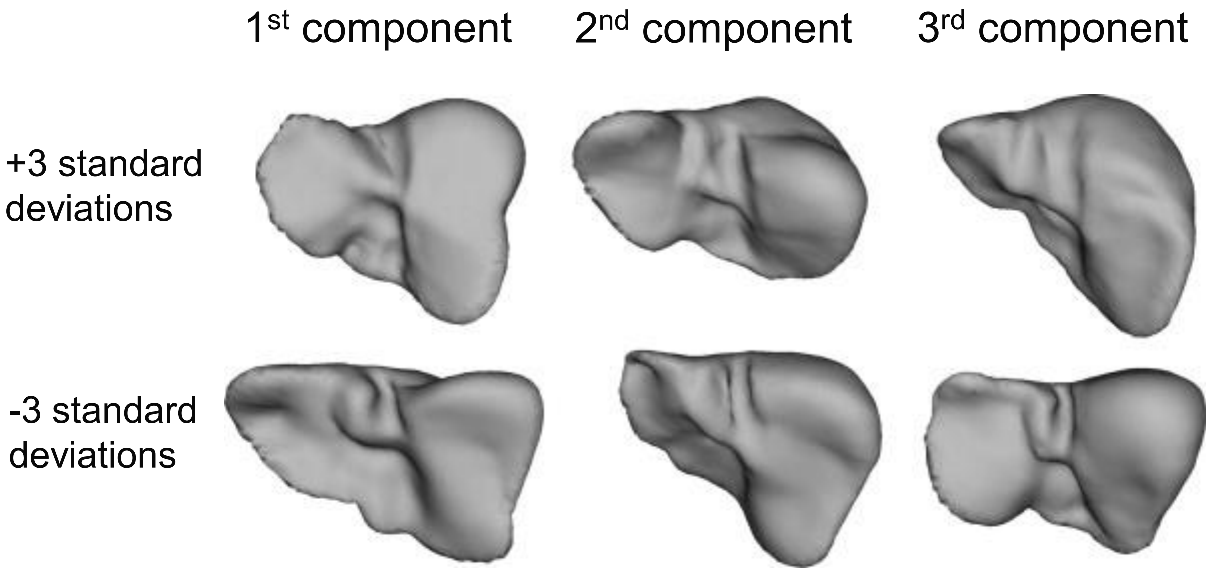






Statistical shape models



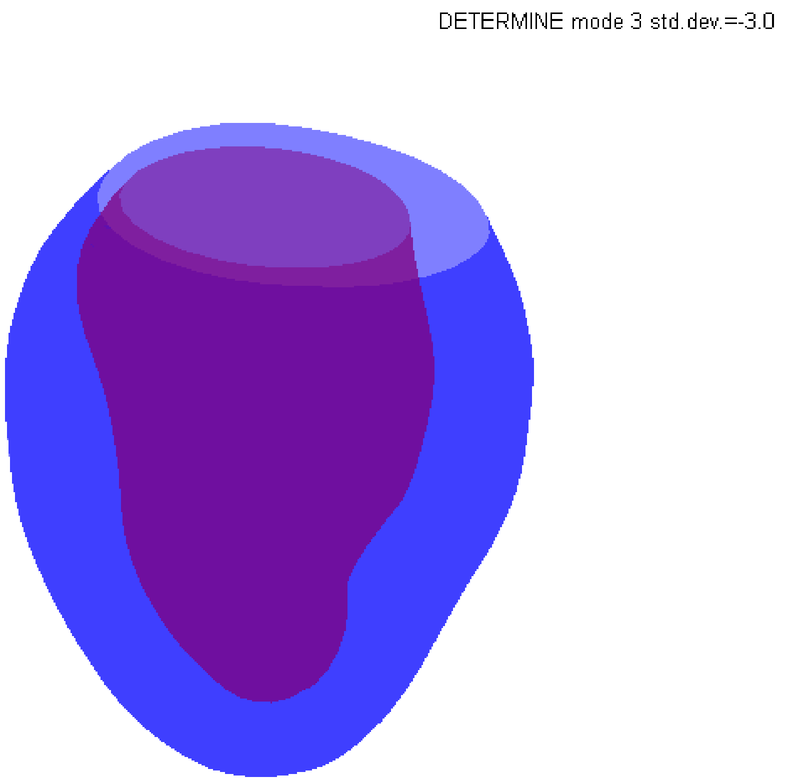


Registration / alignment

Representation function (PDM)

[Heimann and Meinzer, Statistical shape models for 3D medical image segmentation: A review, Med. Imag. Anal., 2009]

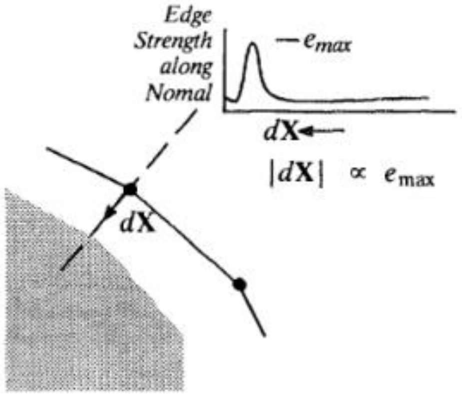
Statistical shape models

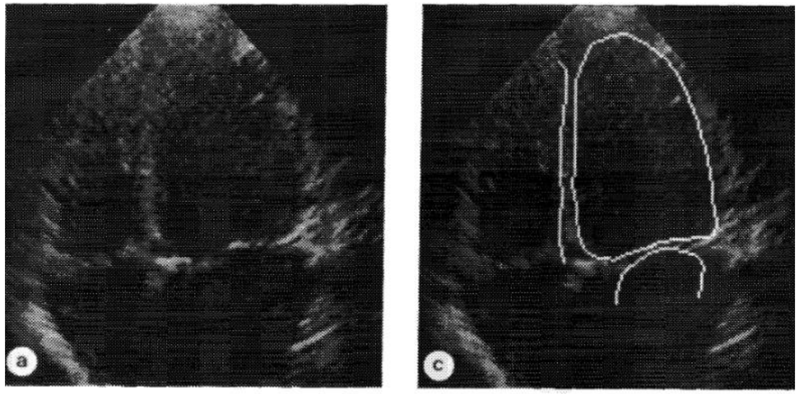


https://www.cardiacatlas.org/

Active shape / appearance models

Domain-specific statistical knowledge: shape





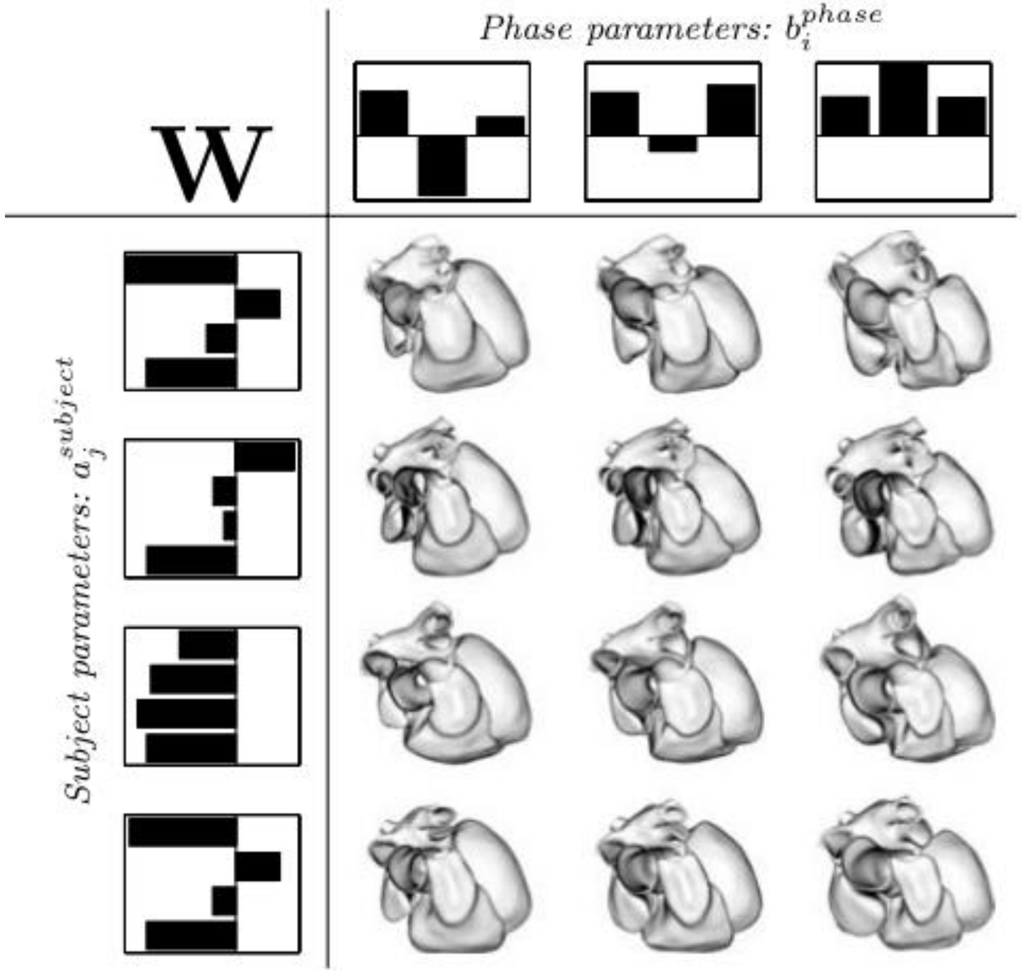
Domain-specific statistical knowledge: appearance

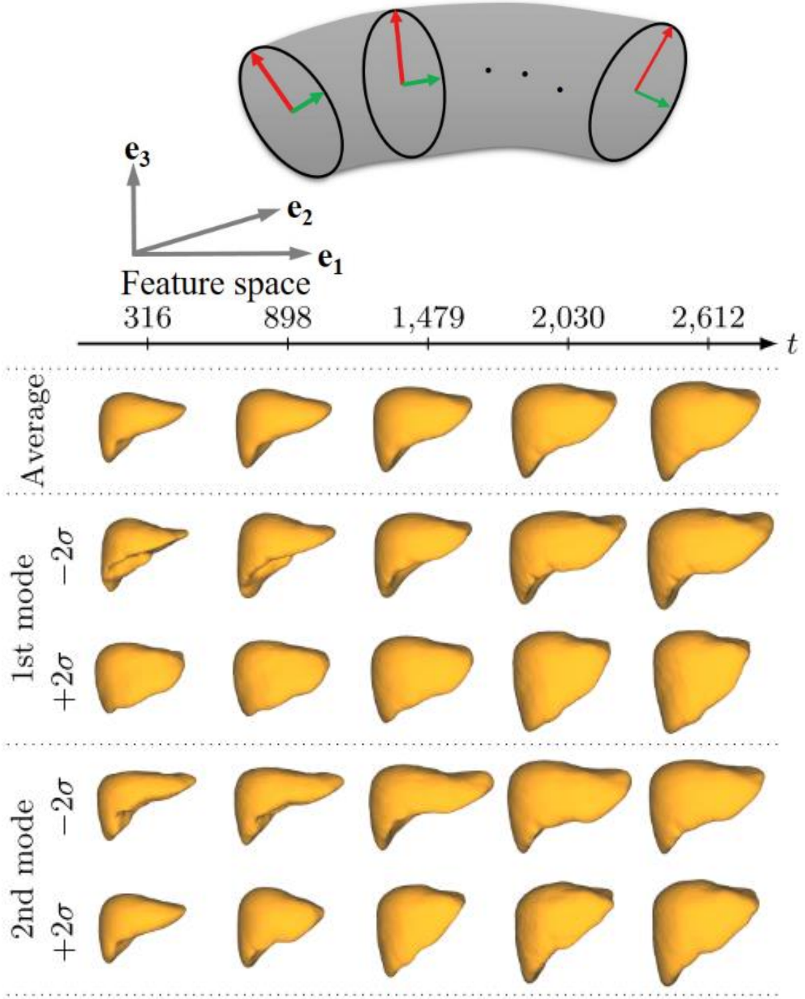
Iterations

[Cootes, et al, Active shape models - their training and application, Computer Vision and Image Understanding, 1995]

[Cootes et al, Active appearance models, IEEE Trans on Patt Anal. Mach. Intel., 2001]

Spatiotemporal models



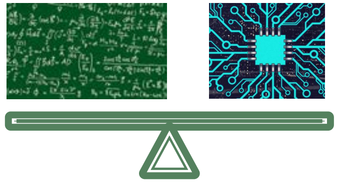


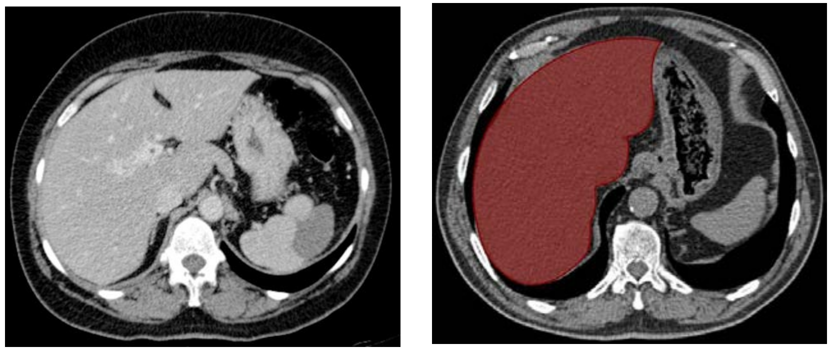
[Hoogendoorn et al, Bilinear models for spatio-temporal point distribution analysis :Application to extrapolation of left ventricular, biventricular and whole heart cardiac dynamics, Int. J. Comput. Vis., 2009]

[Porras et al, Interventional Endocardial Motion Estimation from Electroanatomical Mapping Data: Application to Scar Characterization, IEEE Trans. Biomed. Eng., 2013]

[Saito et al, Construction of a Spatiotemporal Statistical Shape Model of Pediatric Liver from Cross-Sectional Data, Med Image Comput Comput Assist Interv, 2018]

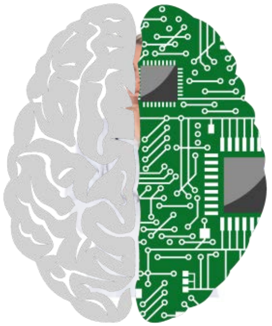
Classification and diagnosis





Context

Quantitative phenotyping:
- Shape / anatomy (volumes, distances…)
- Appearance (avg. intensity, texture...)

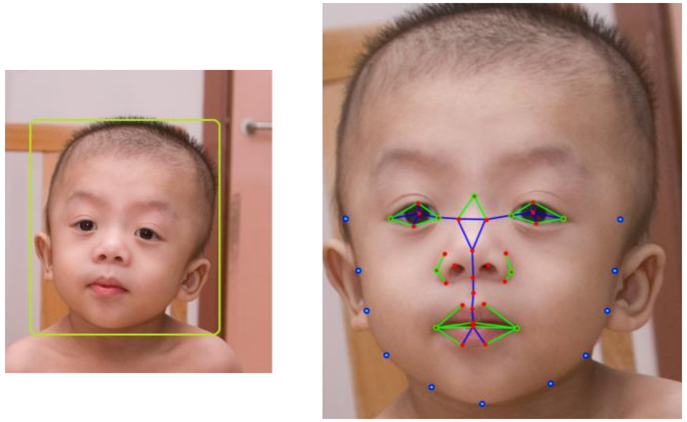


Decision

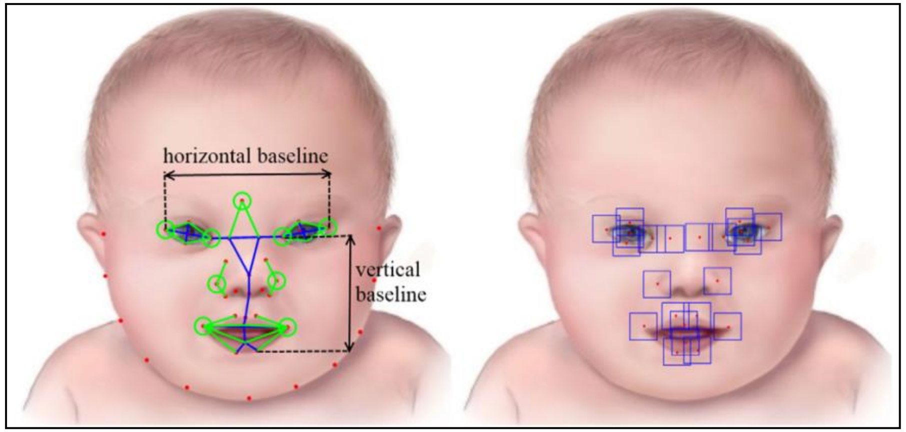
Images → Segmentation → Feature Quantification → Feature Normalization → Feature selection → Model/predictor training

Classification and diagnosis

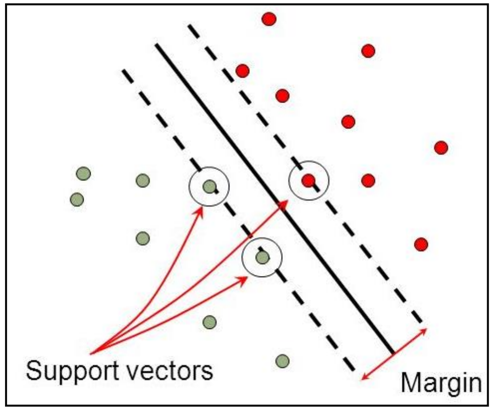
Facial landmark detection



Feature quantification



Classifier



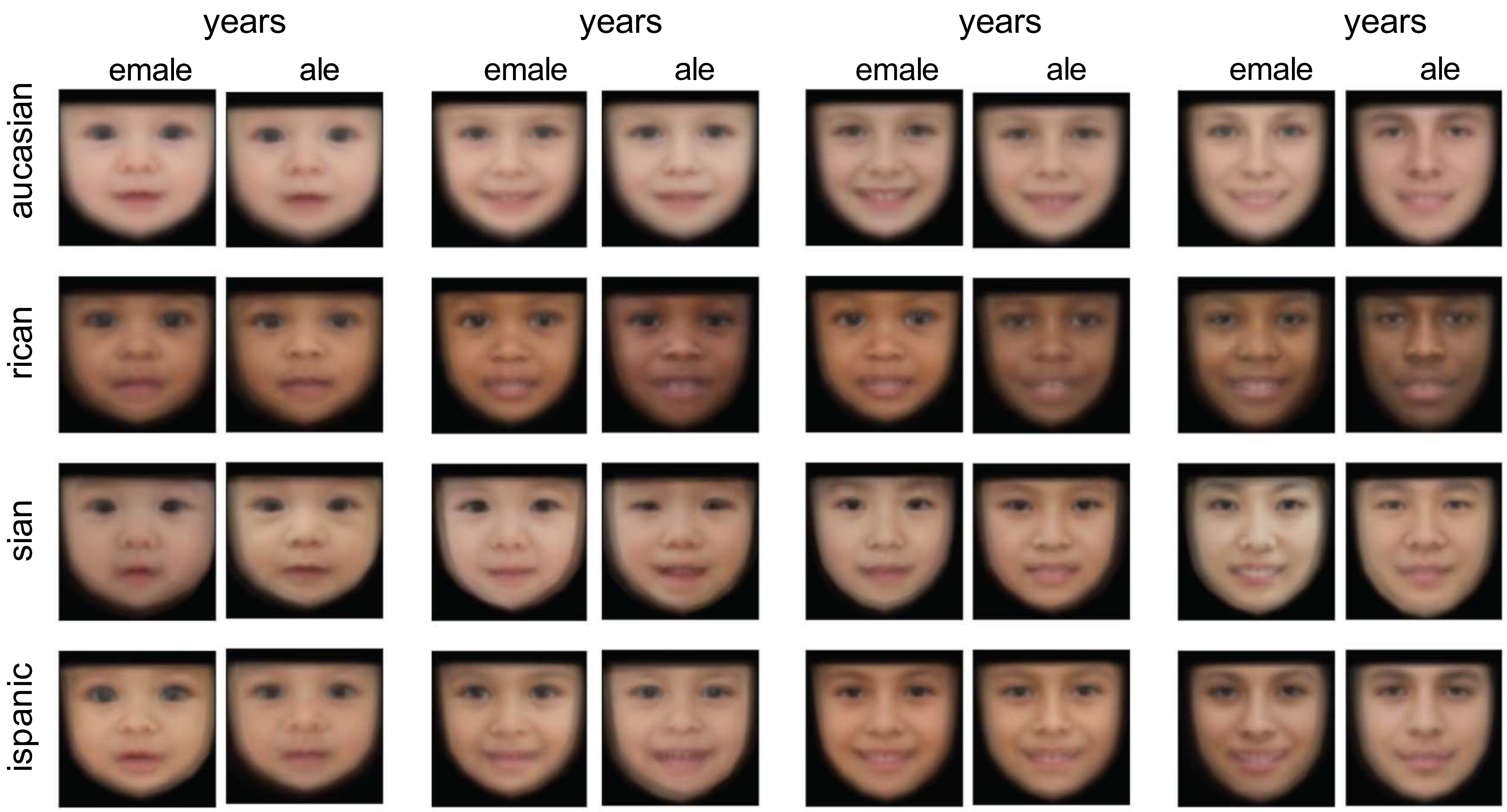
Shape and appearance model

Clinical knowledge

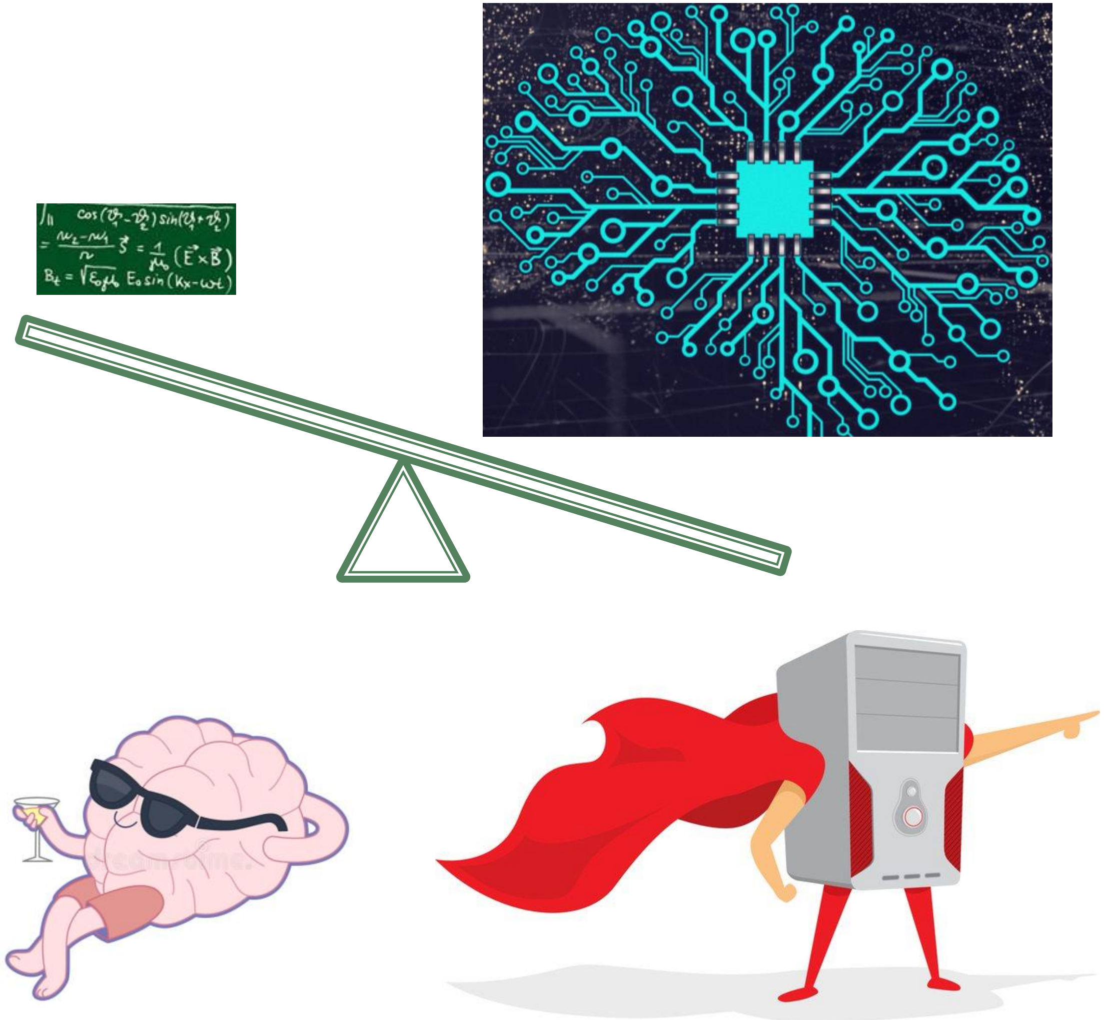
[Kruszka et al, Down syndrome in diverse populations, Am. J. Med. Genet. *2017*]

[Porras et al, Objective differential diagnosis of Noonan and Williams-Beuren syndromes in diverse populations using quantitative facial phenotyping, Mol. Gen. Genom. Med., *In press*]

Classification and diagnosis



Big data



antonio.porras@cuanschutz.edu

https://sites.google.com/view/medicalimagephenotyping

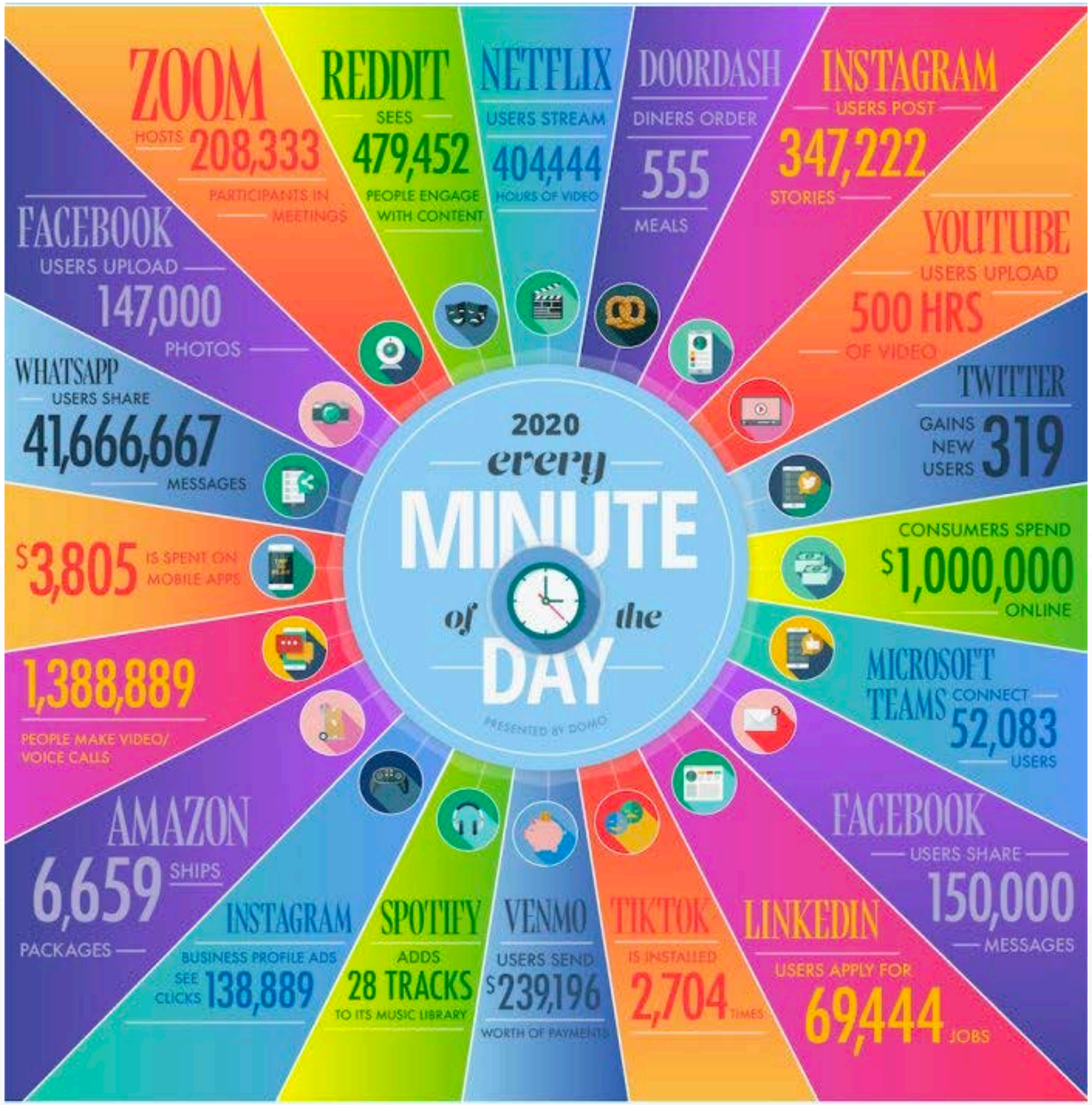
Medical Imaging and Big Data

Antonio R. Porras, Fuyong Xing

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

Data Never Sleeps

- How much data is generated every minute?
  - Facebook: users upload about 147,000 photos



https://www.domo.com/learn/data-never-sleeps-8

Big Image Data

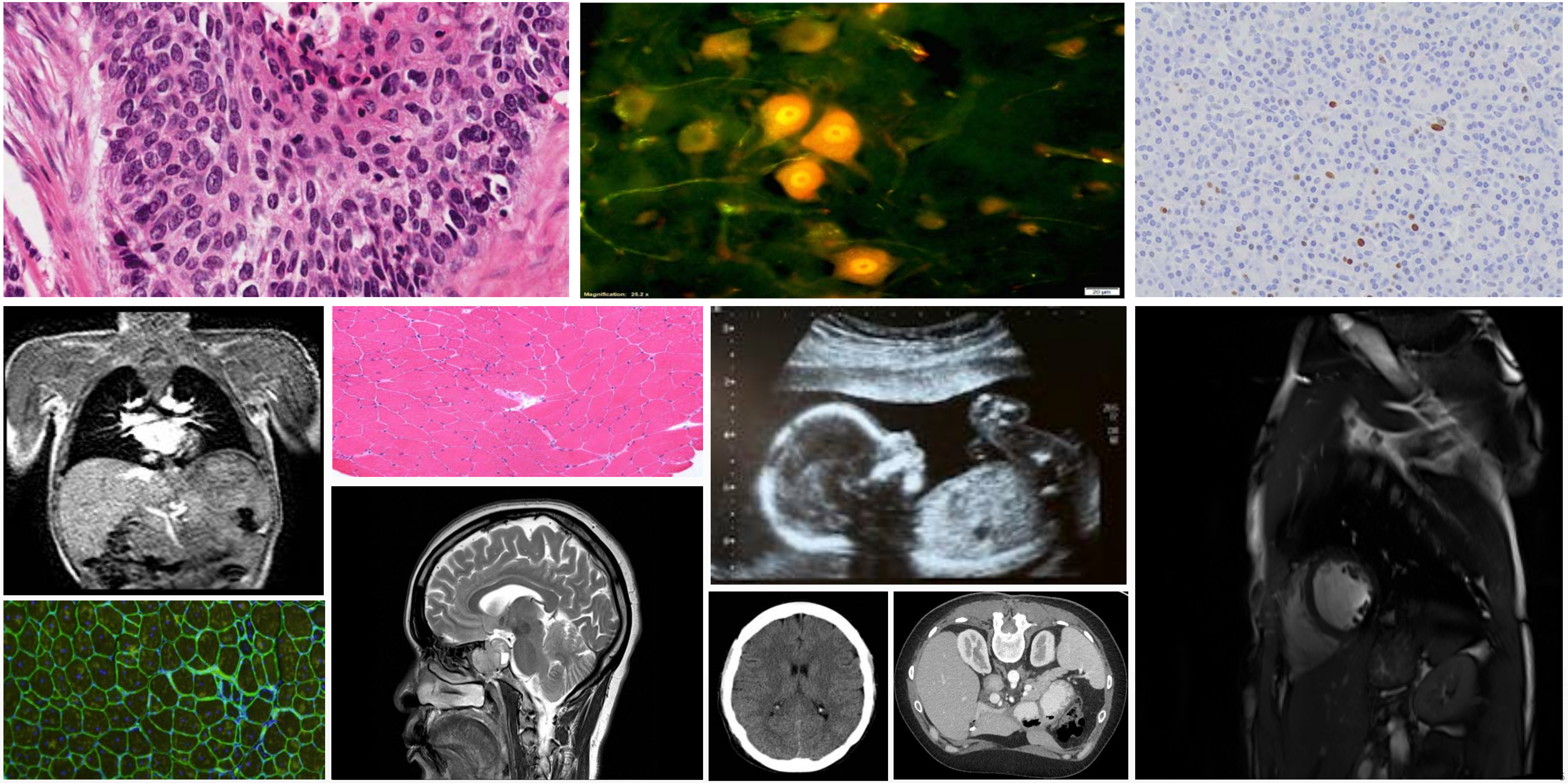
https://medium.com/abraia/first-steps-with-transfer-learning-for-custom-image-classification-with-keras-b941601fcad5

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

Biomedical Images Are Everywhere

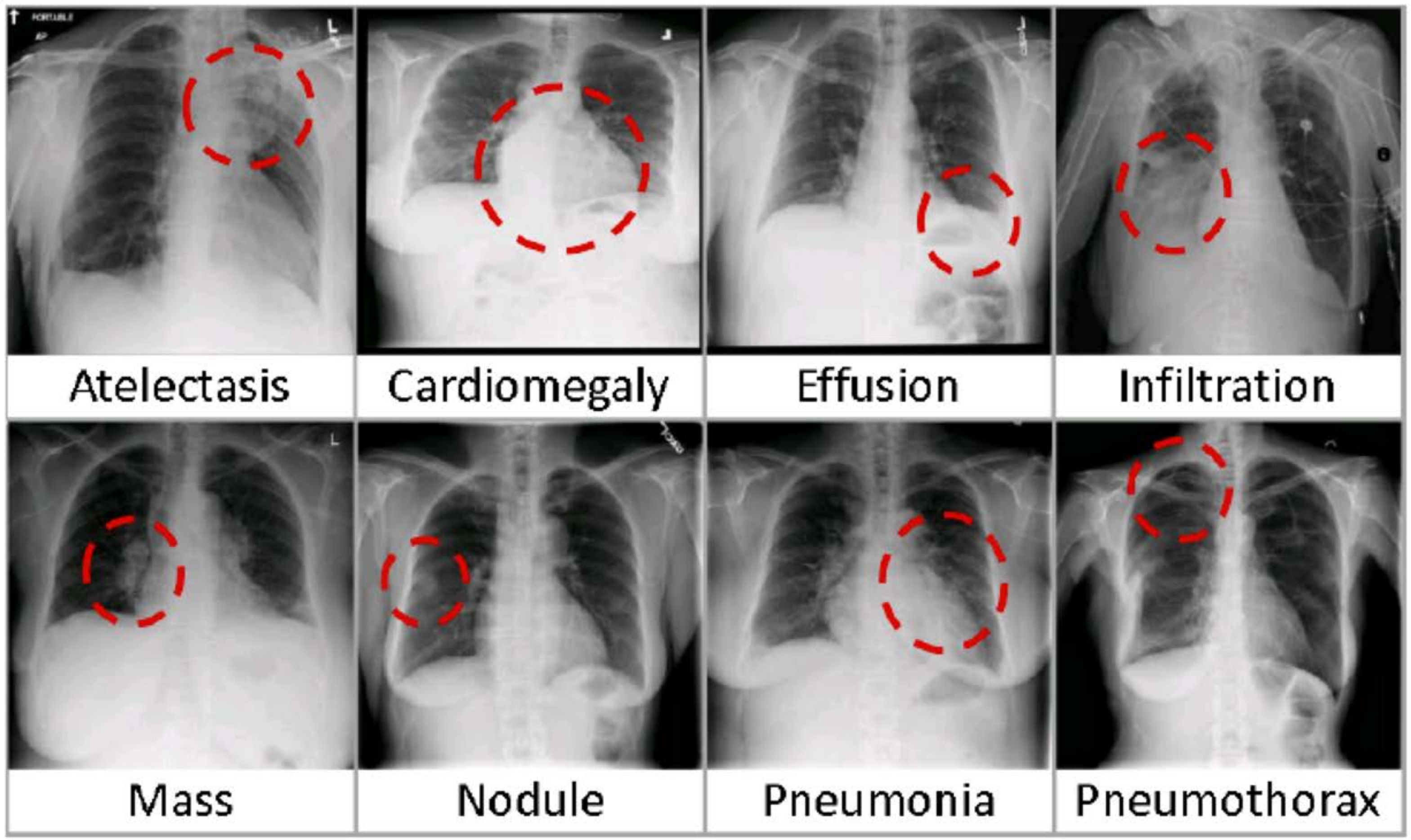


Example Medical Image Datasets

- NIH Chest X-Ray-14 dataset for classification and localization of thorax diseases

- 112,120 frontal images from over 30,000 unique patients

https://nihcc.app.box.com/v/ChestXray-NIHCC



Wang et al., ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases, CVPR 2017

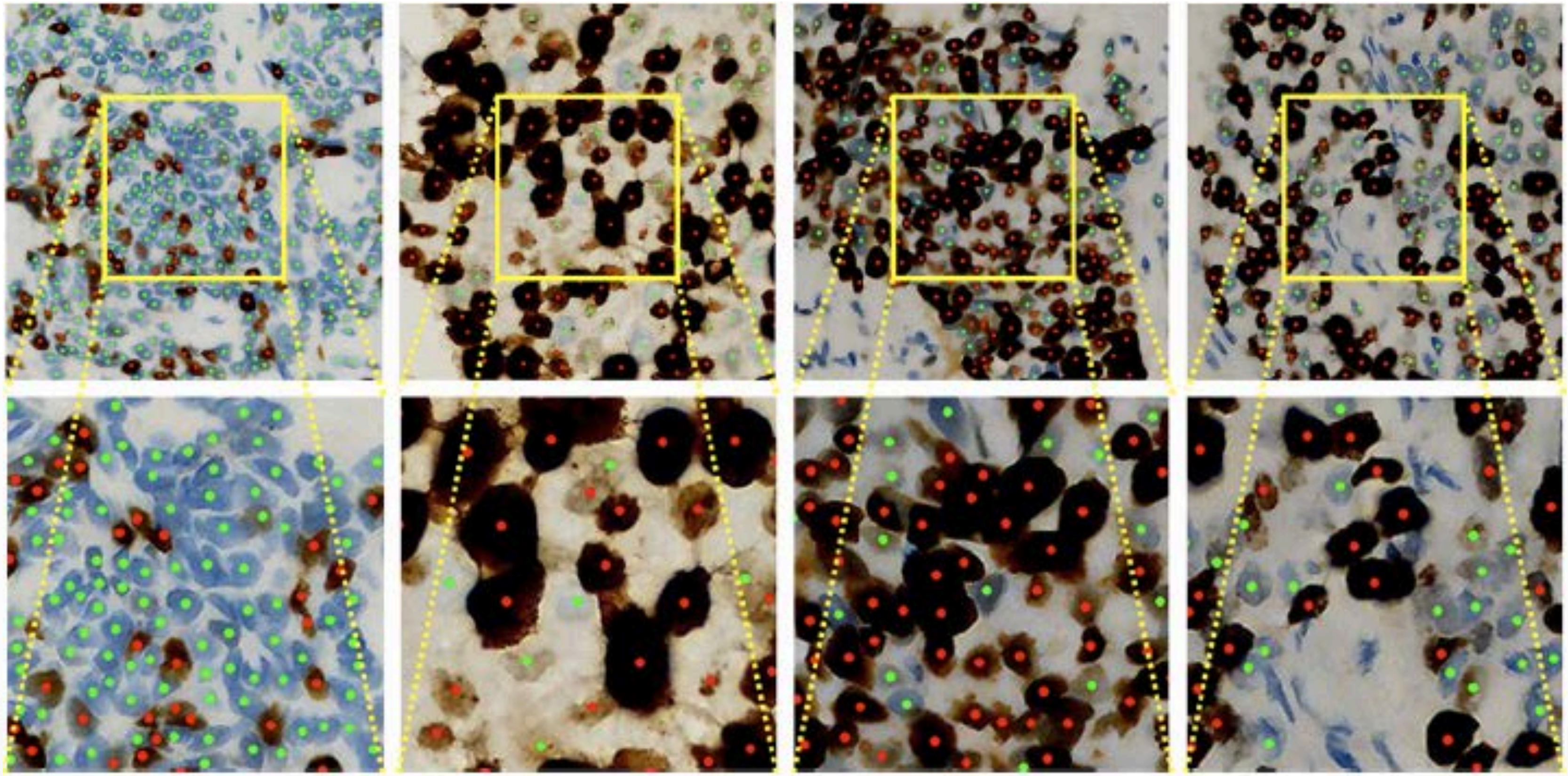
Example Medical Image Datasets

- A public Ki67 immunohistochemistry-stained breast cancer image dataset for nuclei detection and classification

- 1,338 images from 394 subjects/cases

- 181,074 annotated nuclei

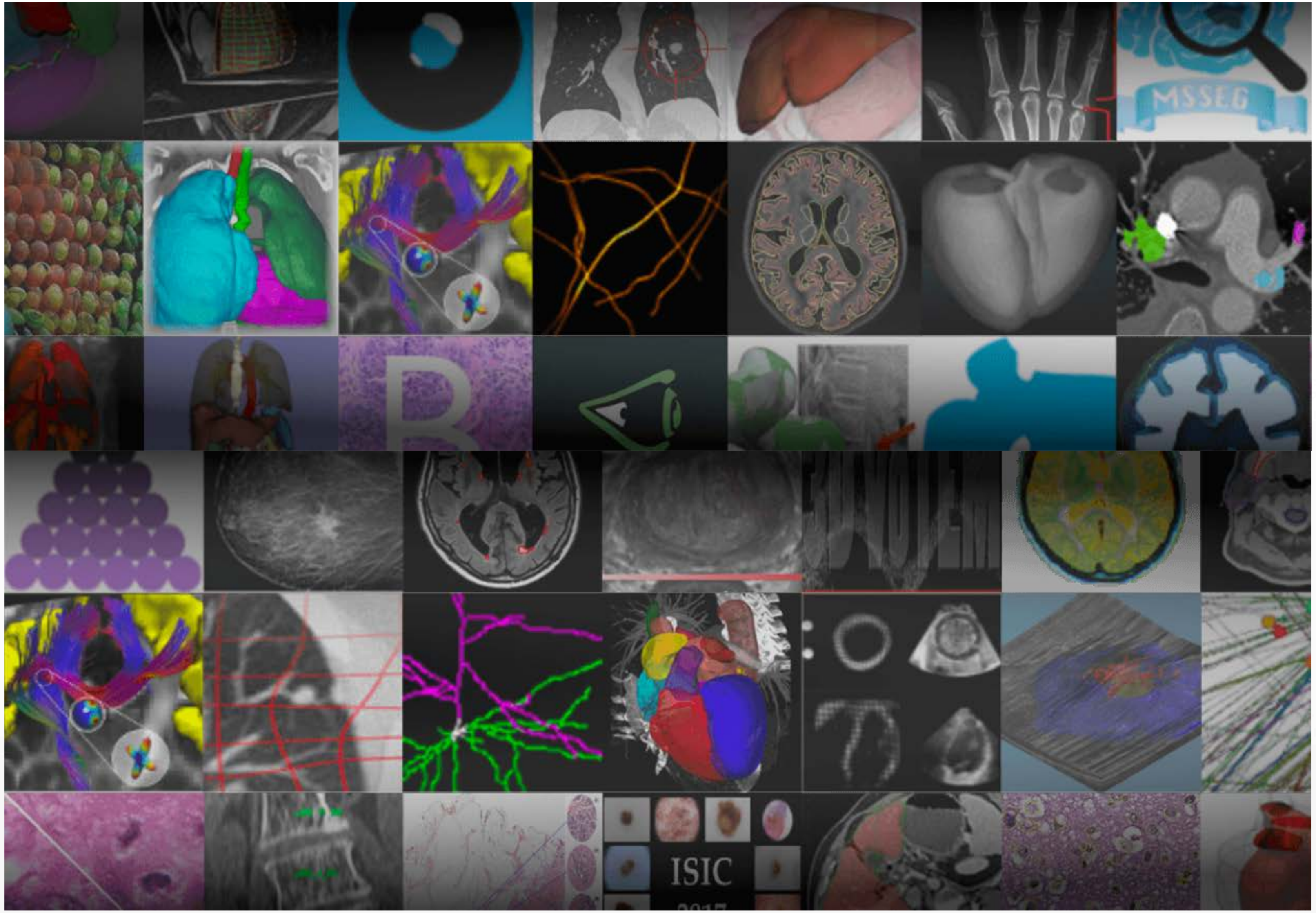
https://sites.google.com/view/bcdataset



Z. Huang, et al., “BCData: A Large-Scale Dataset and Benchmark for Cell Detection and Counting”, *MICCAI, 2020*.

More Biomedical Image Datasets

https://grand-challenge.org/challenges/



Methods for Large-Scale Image Data

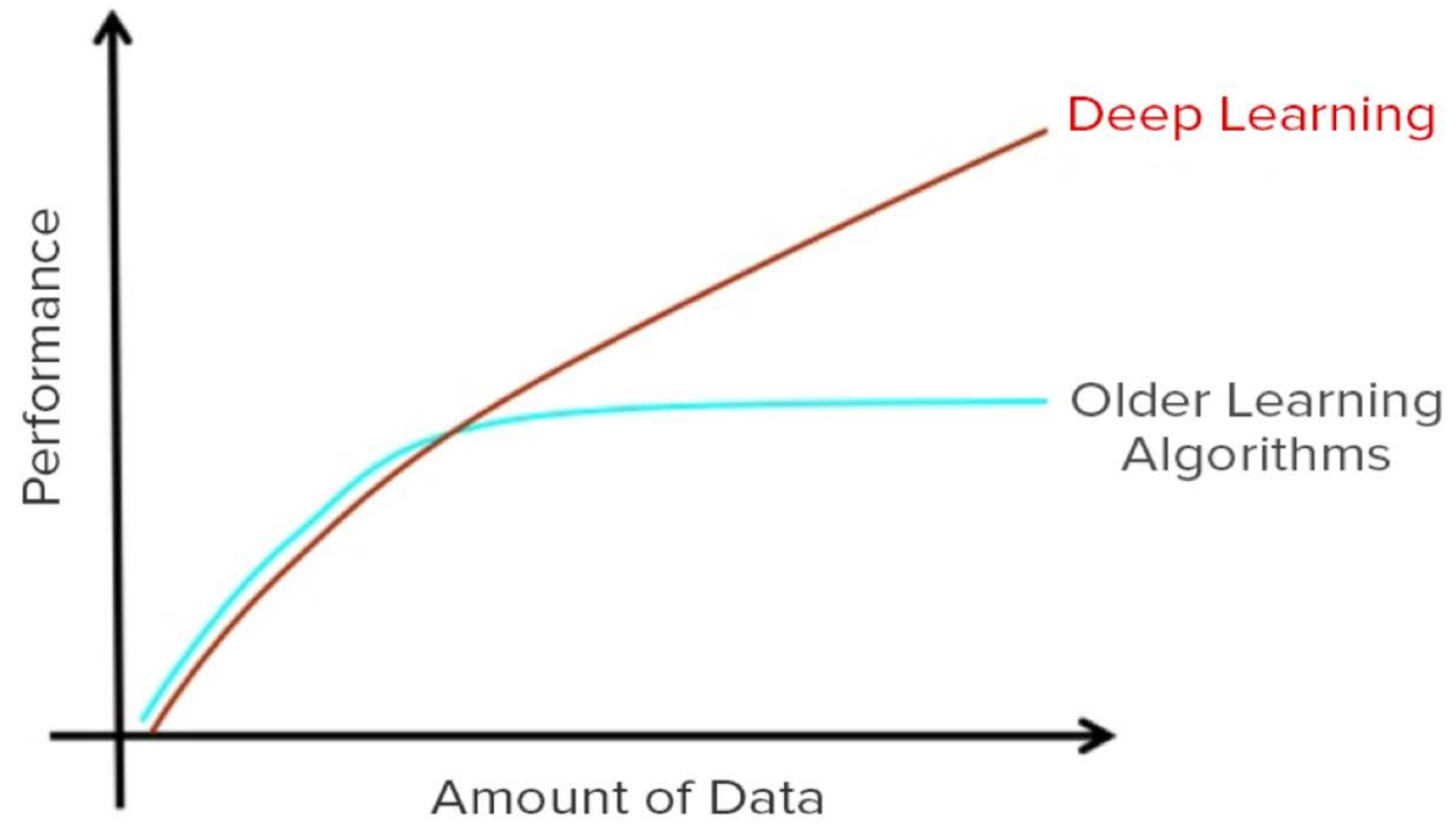
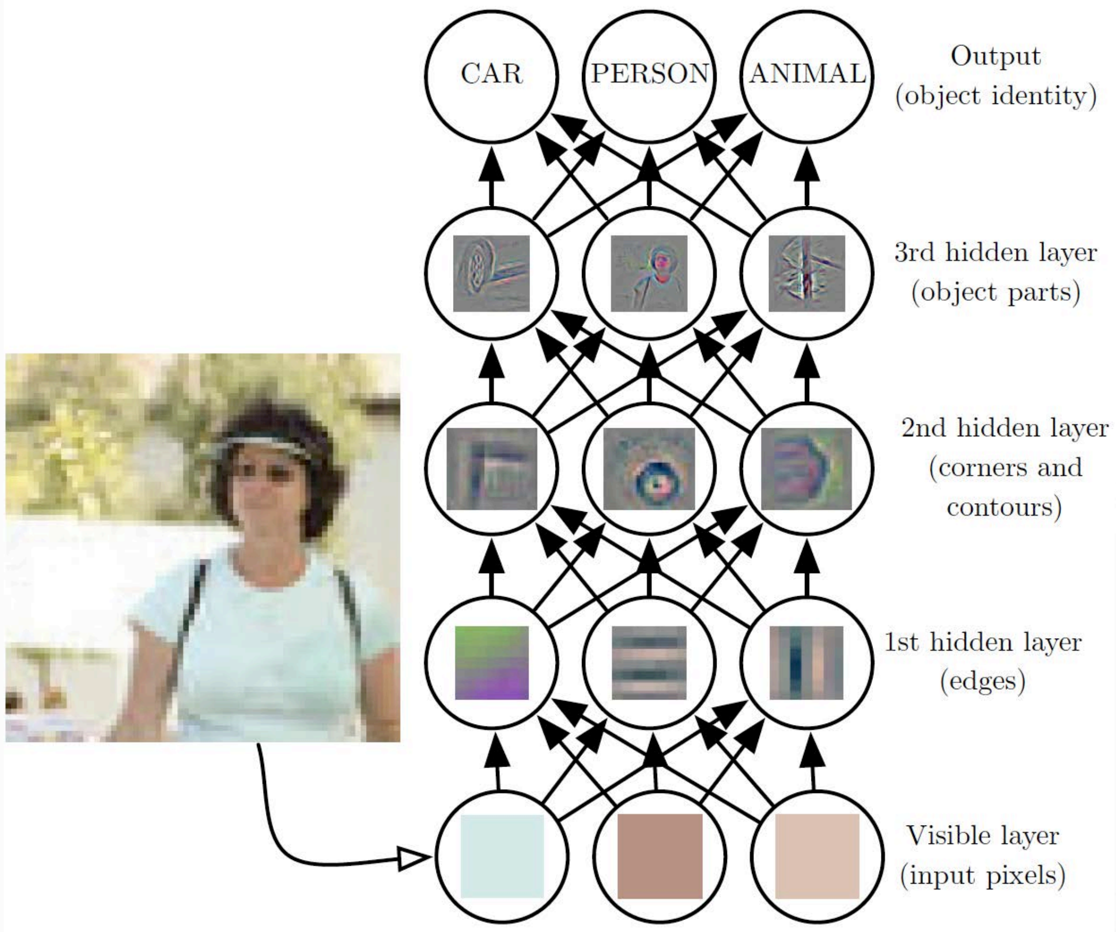


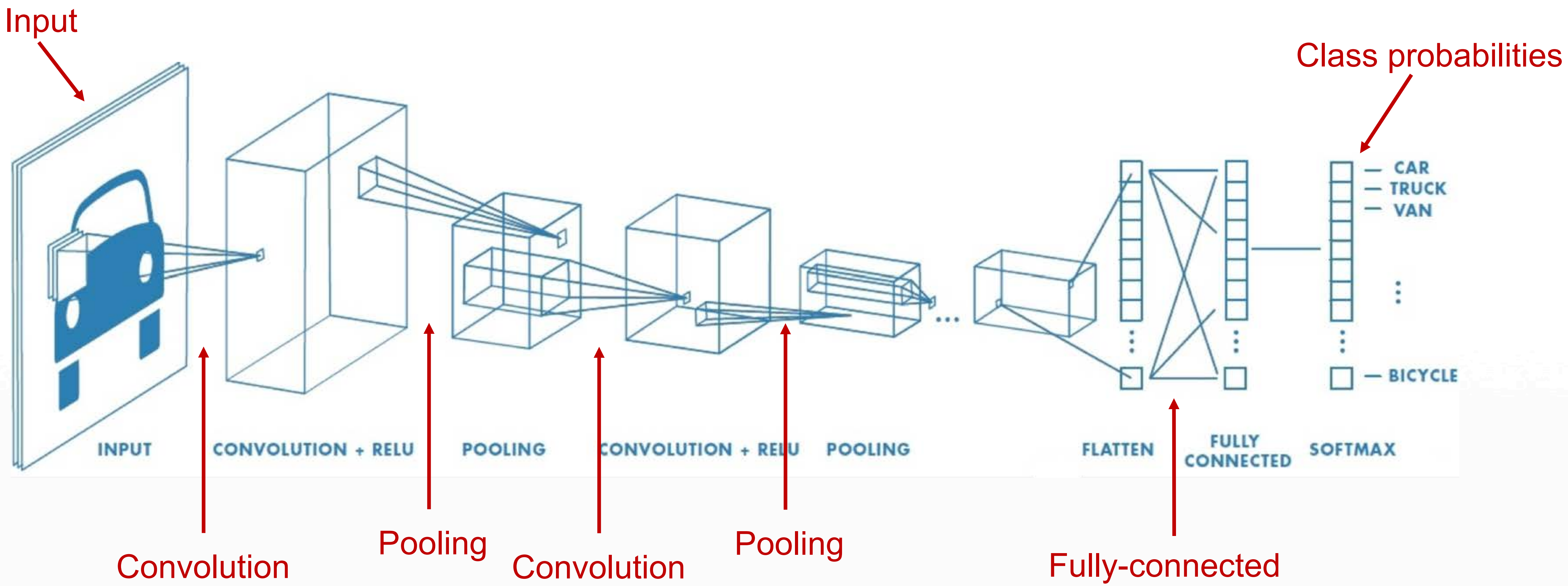
Image credit: Andrew Ng

Deep Learning



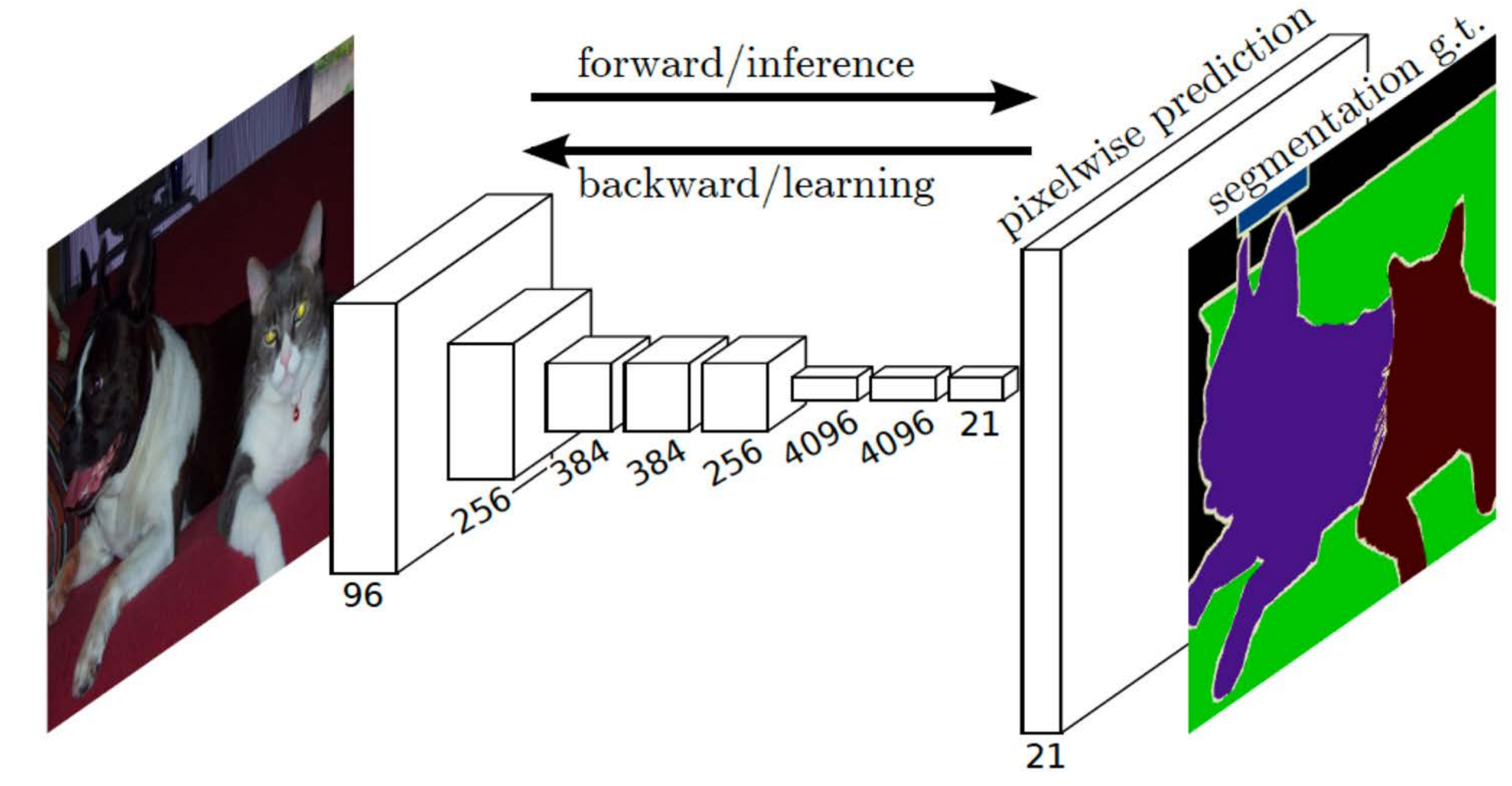
Goodfellow *et al*. "Deep Learning", *MIT Press*, 2016

Convolutional Neural Networks (CNNs)



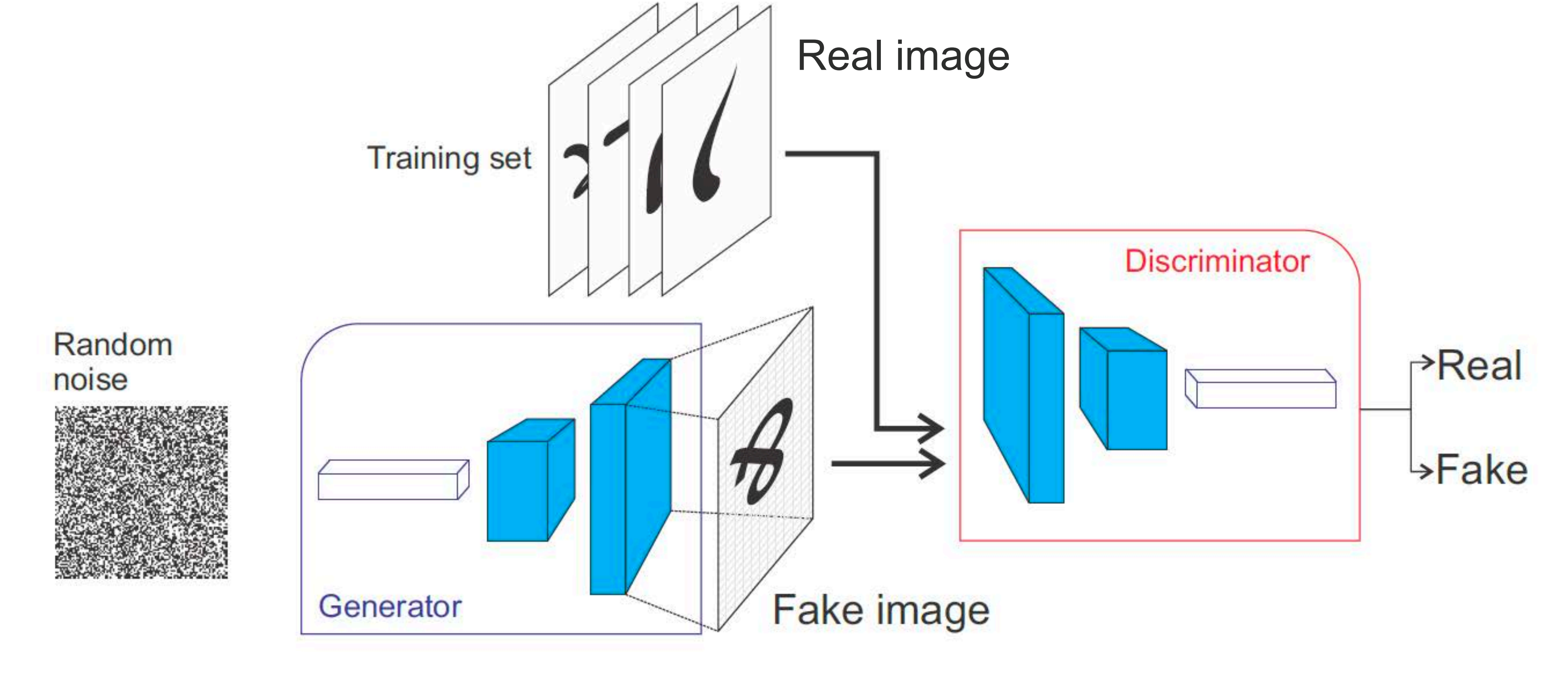
https://www.mathworks.com/discovery/convolutional-neural-network-matlab.html

Fully Convolutional Networks (FCNs)



Long et al., Fully Convolutional Networks for Semantic Segmentation, CVPR 2015

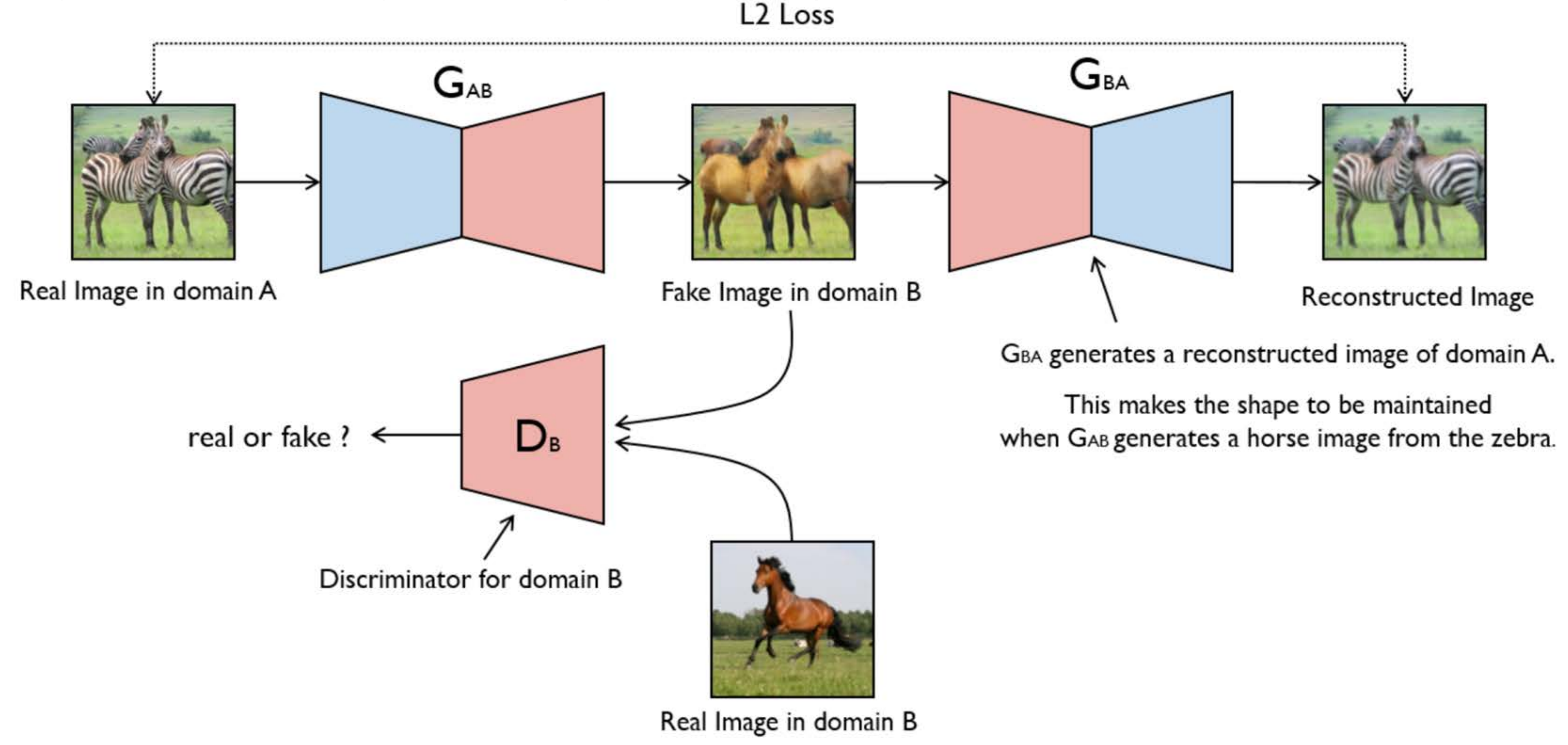
Generative Adversarial Networks (GANs)



I. Goodfellow *et al.*, “Generative Adversarial Nets”, *NeurIPS*, 2014
II. https://www.oreilly.com/library/view/java-deep-learning/9781788997454/60579068-af4b-4bbf-83f1-e988fbe3b226.xhtml

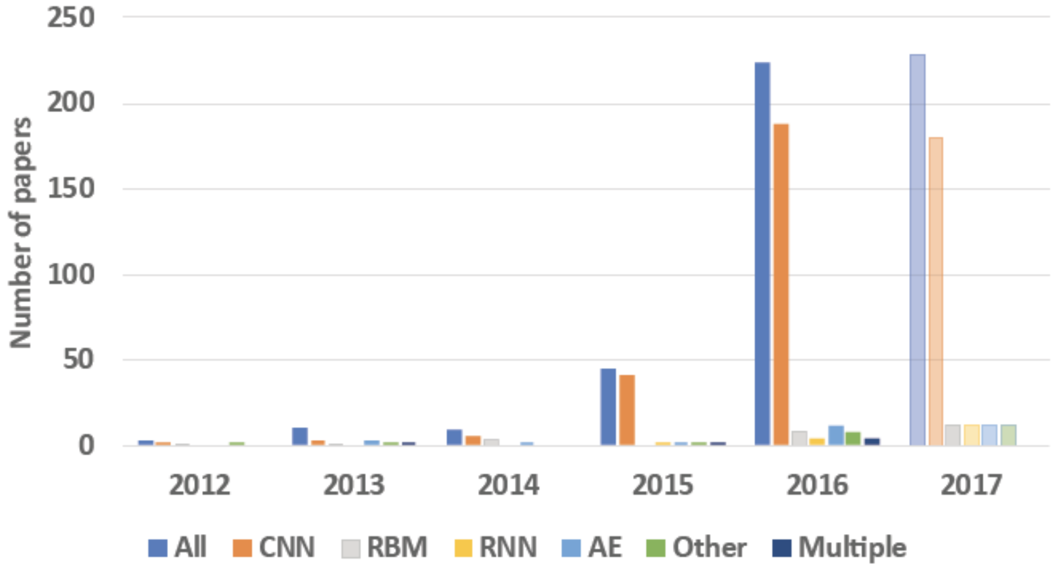
Generative Adversarial Networks (GANs)

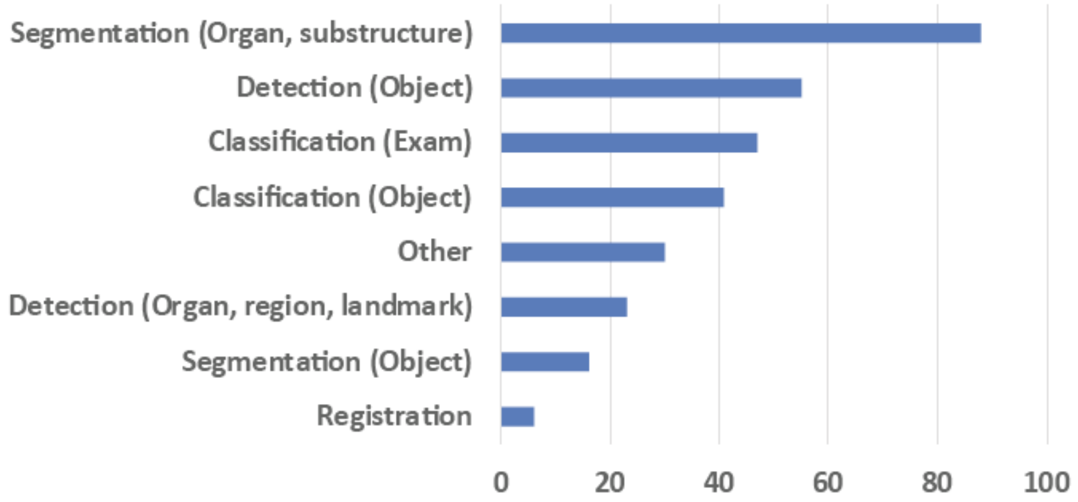
- Cycle-consistency GANs (CycleGANs)

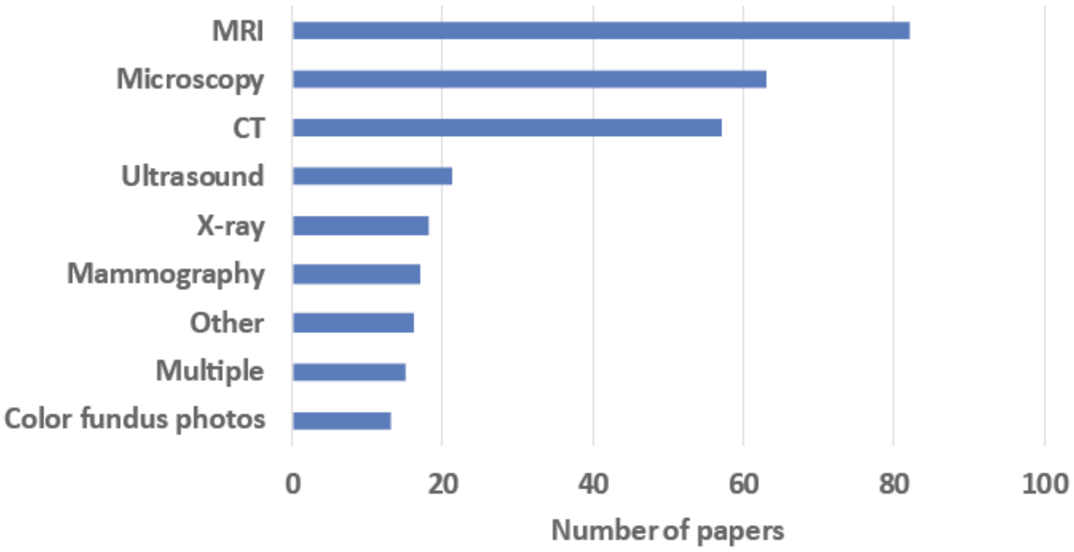


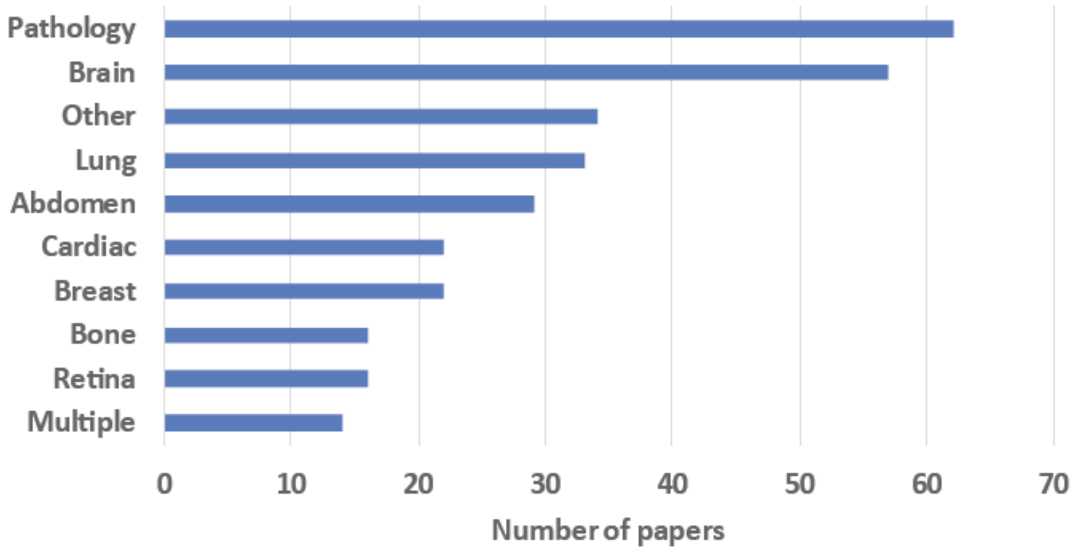
1. Zhu *et al.*, "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks", *ICCV*, 2017
2. https://towardsdatascience.com/image-to-image-translation-using-cyclegan-model-d58cfff04755

Deep Learning in Medical Image Analysis





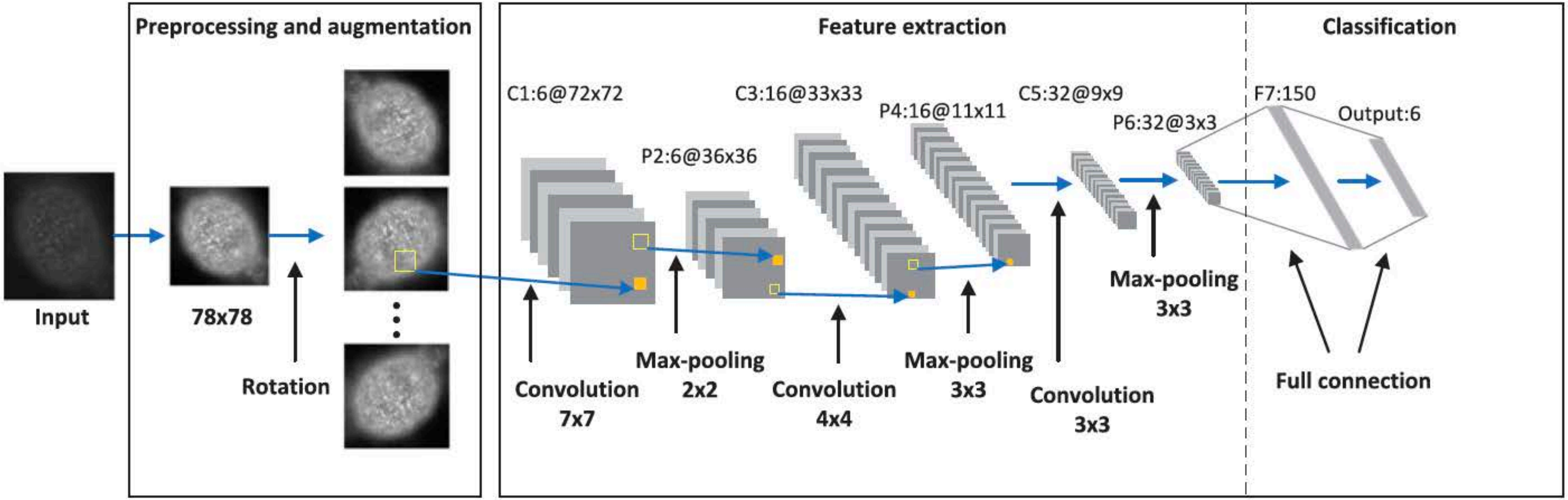




Litjens *et al.*, “A survey on deep learning in medical image analysis”, *Medical Image Analysis*, 2017

Image Classification

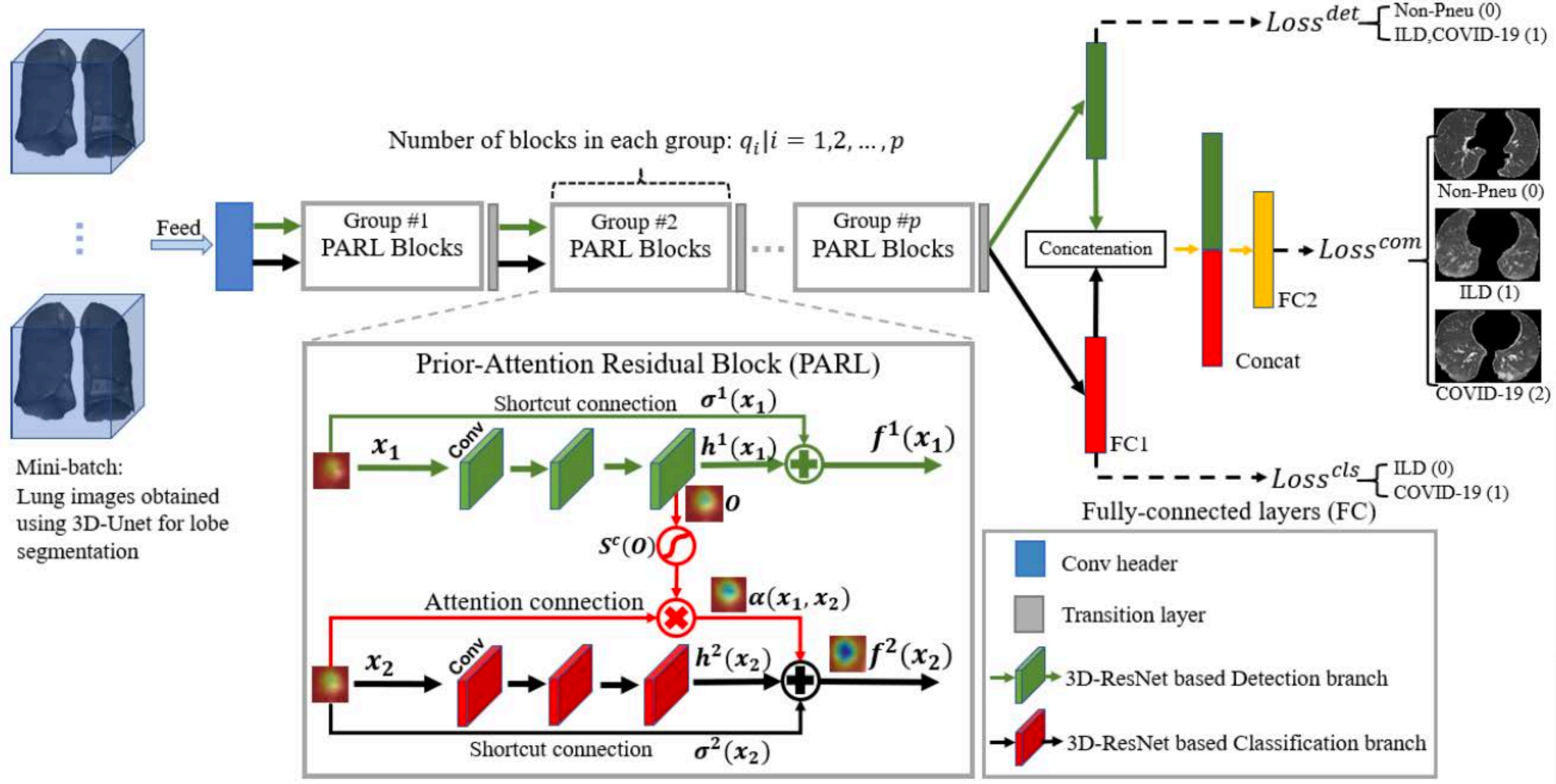
- CNN-based classification of HEp-2 cell images



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

Image Classification

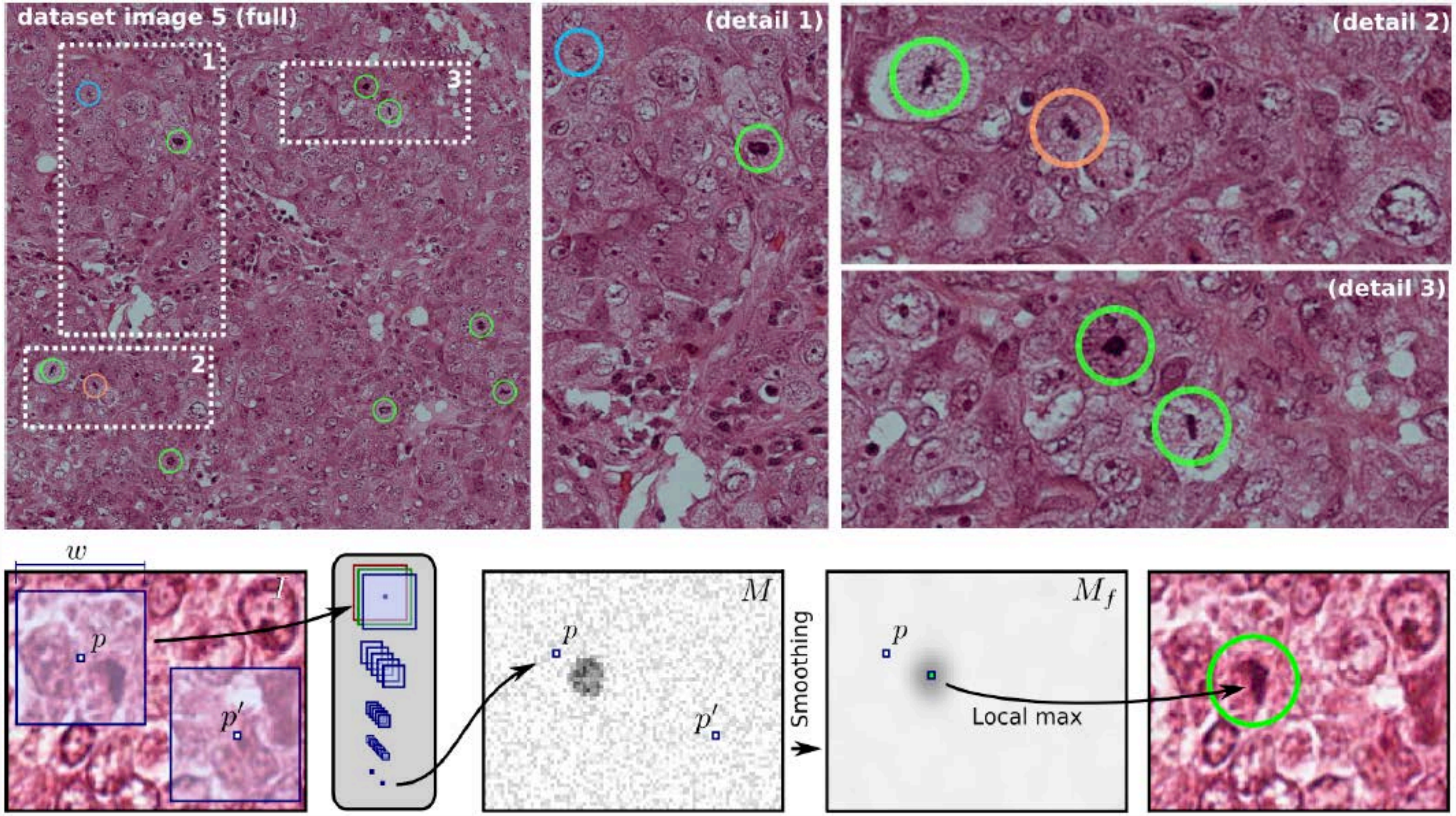
- COVID-19 screening



Wang et al., “Prior-Attention Residual Learning for More Discriminative COVID-19 Screening in CT Images”, *IEEE TMI, 2020*.

Object Detection

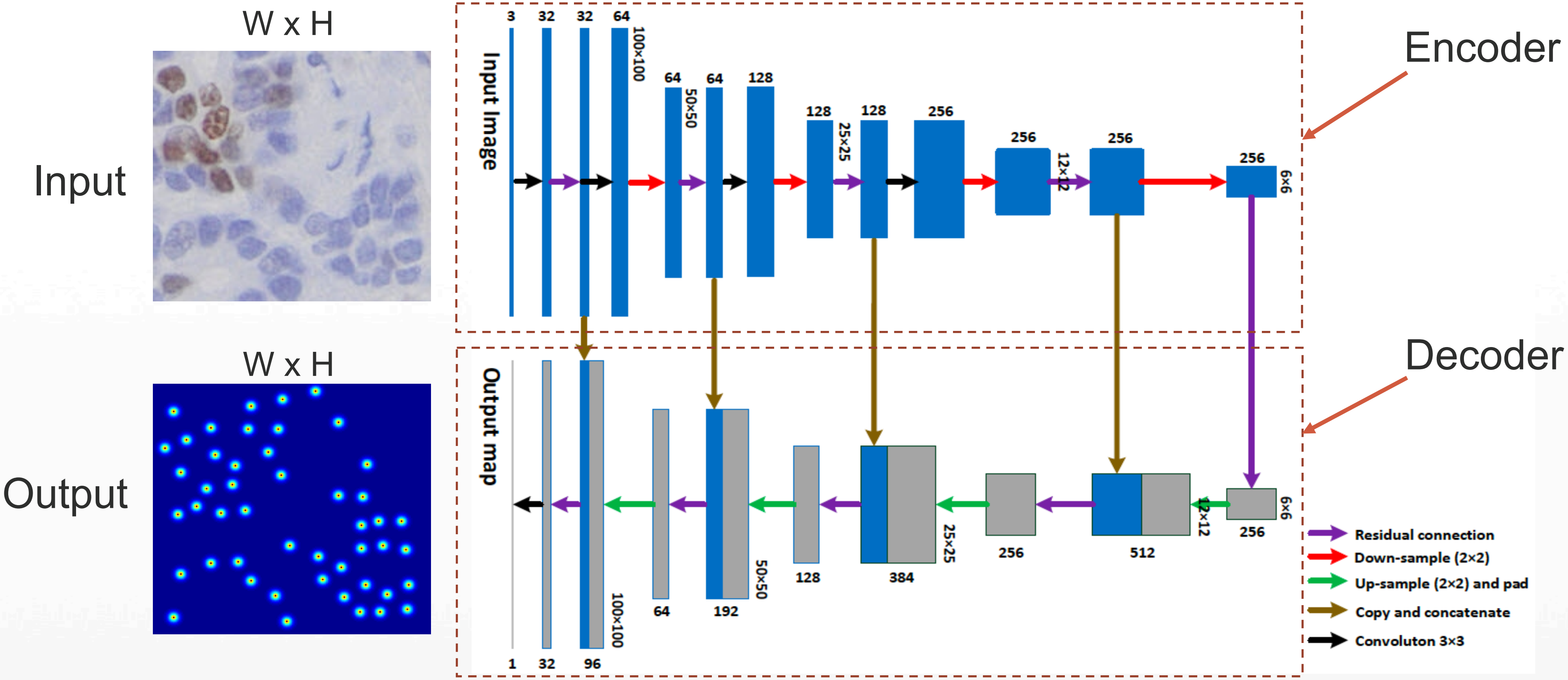
- CNN-based mitosis detection in breast cancer histology images



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

Object Detection

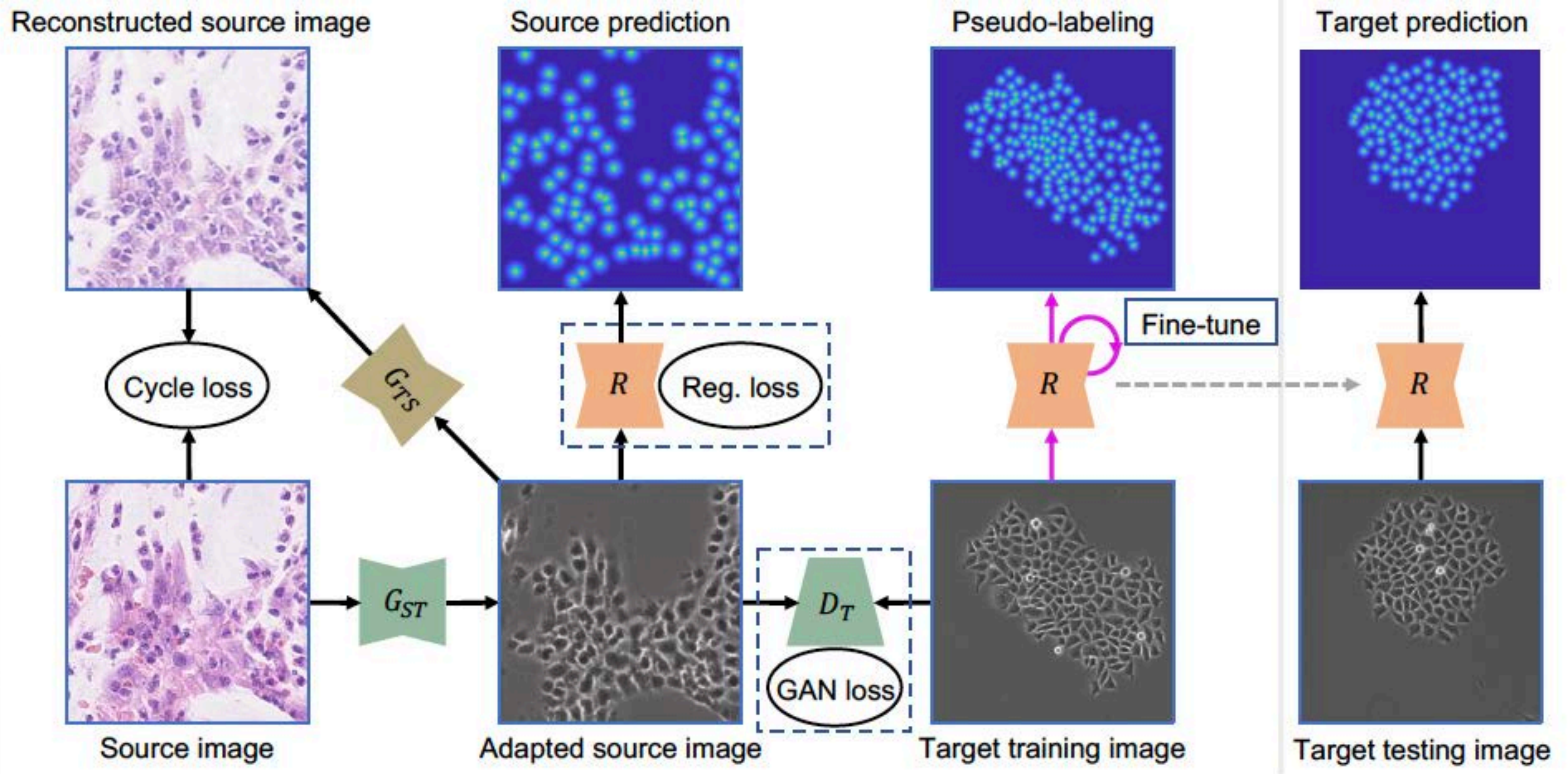
- Nucleus/cell detection with fully convolutional networks (FCNs)



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

Object Detection

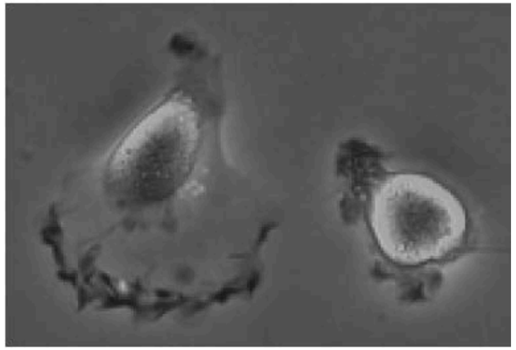
- Nucleus/cell detection with generative adversarial networks (GANs)

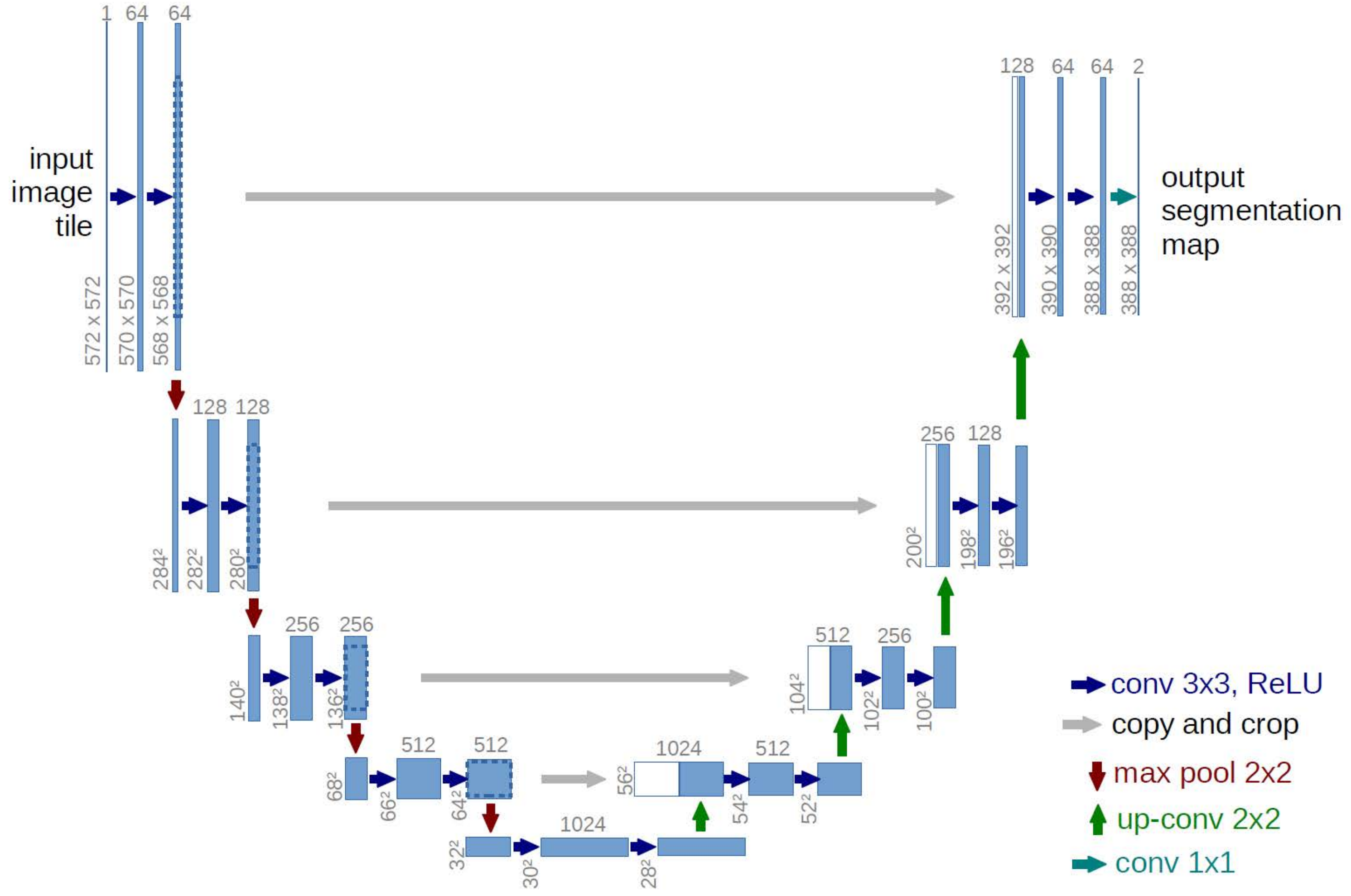


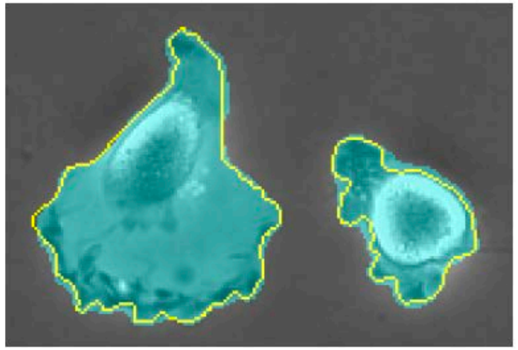
Xing *et al.*, "Adversarial Domain Adaptation and Pseudo-Labeling for Cross-Modality Microscopy Image Quantification", *MICCAI,* 2019

Image Segmentation

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



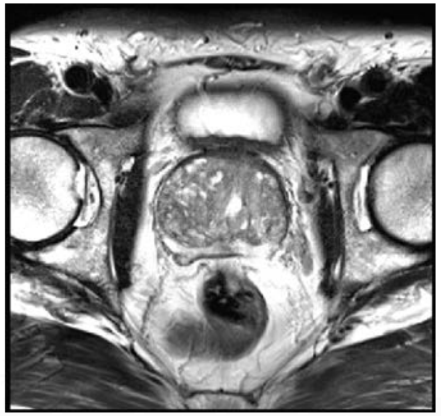


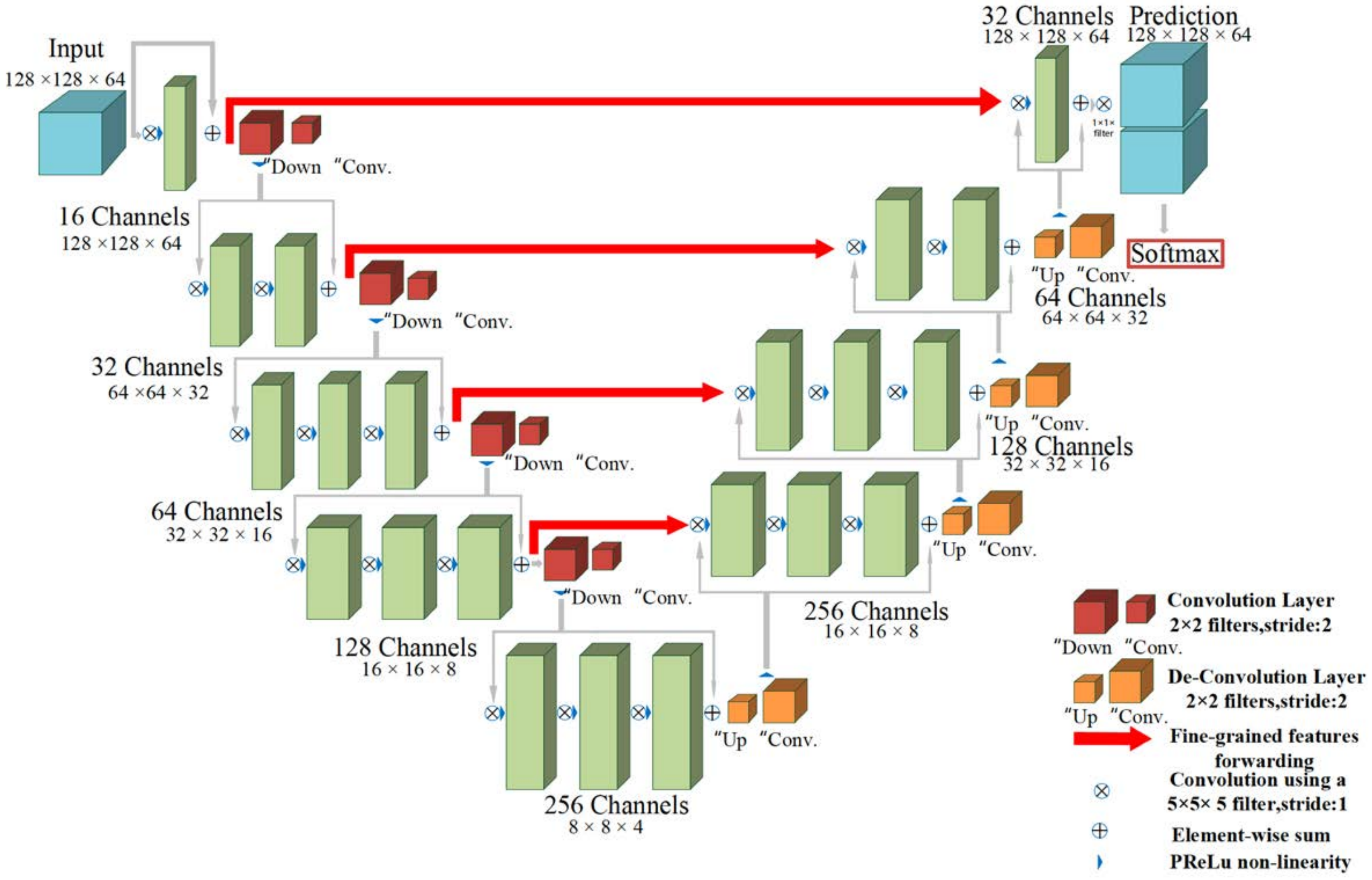


Ronneberger *et al.*, “U-Net: Convolutional Networks for Biomedical Image Segmentation”, *MICCAI,* 2015

Image Segmentation

- 3D MRI prostate image segmentation

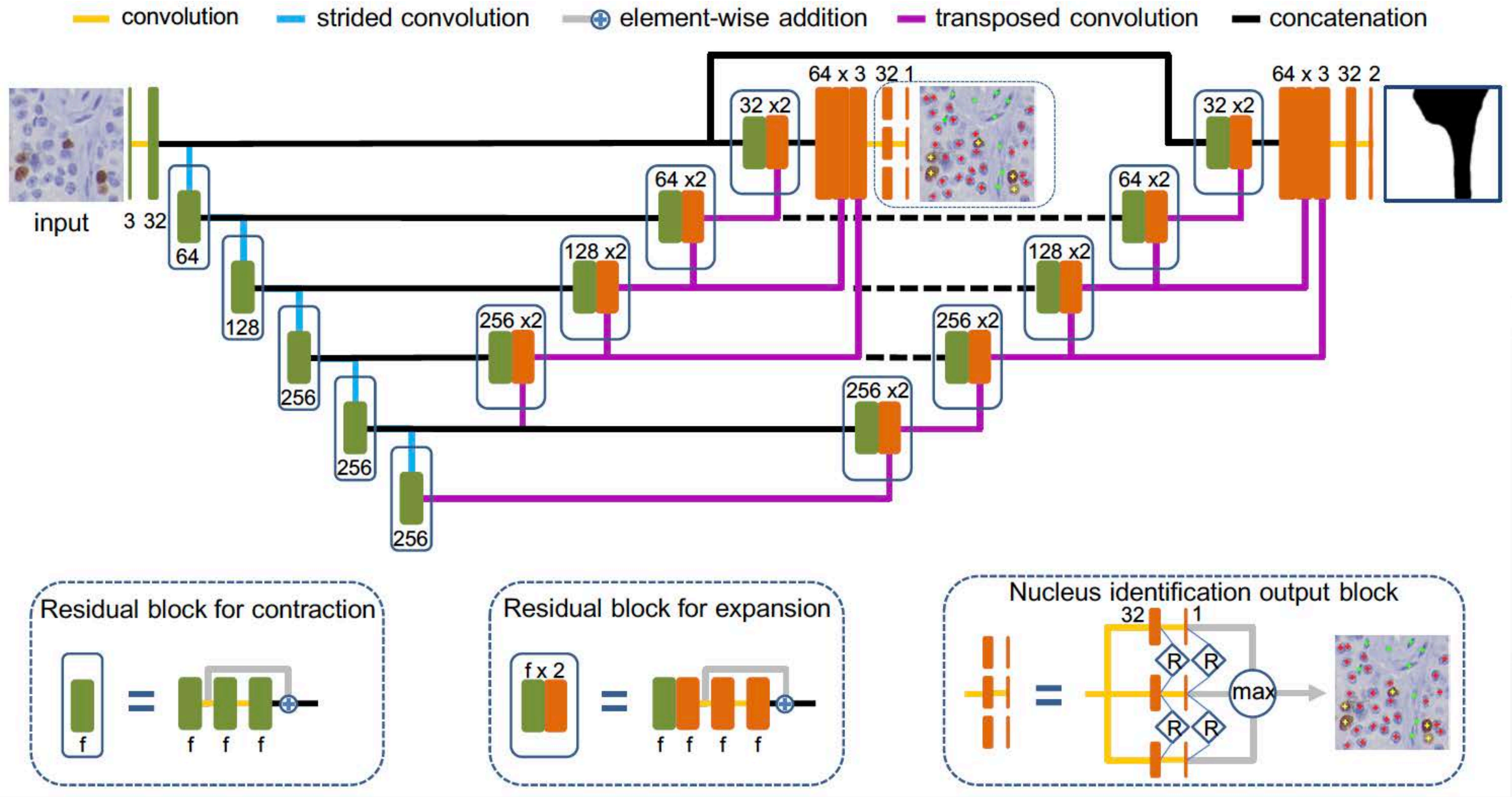




1. Lei *et al.*, "Medical Image Segmentation Using Deep Learning: A Survey", *arXiv,* 2020
2. Milletari et al., "V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation", 3DV, 2016

Object Recognition

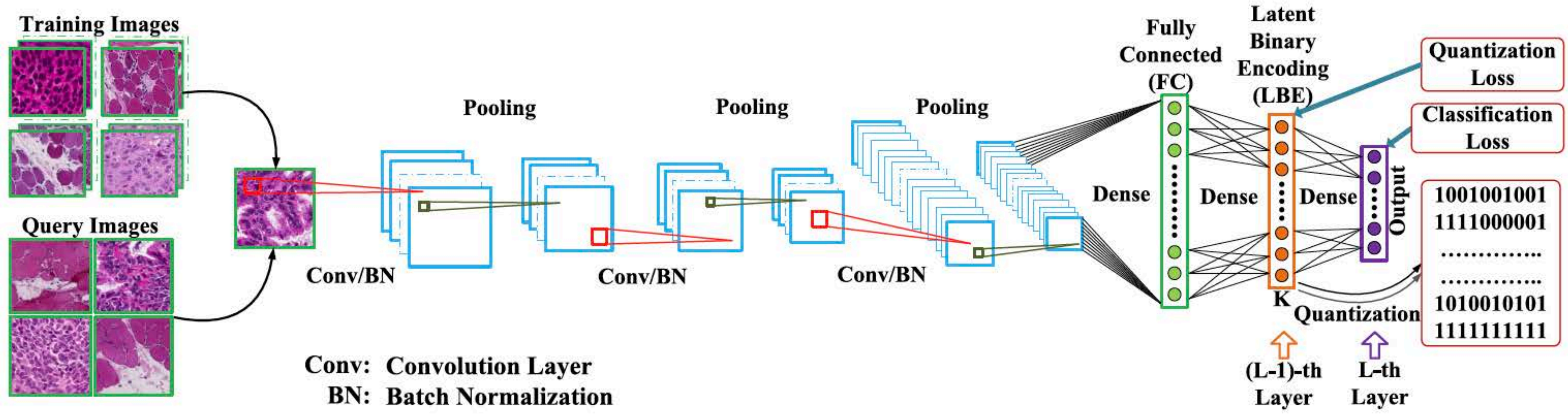
- Nucleus recognition in Ki67 IHC-stained pancreatic neuroendocrine tumors with fully convolutional networks (FCNs)



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

Image Retrieval

- Skeletal muscle image retrieval with CNNs

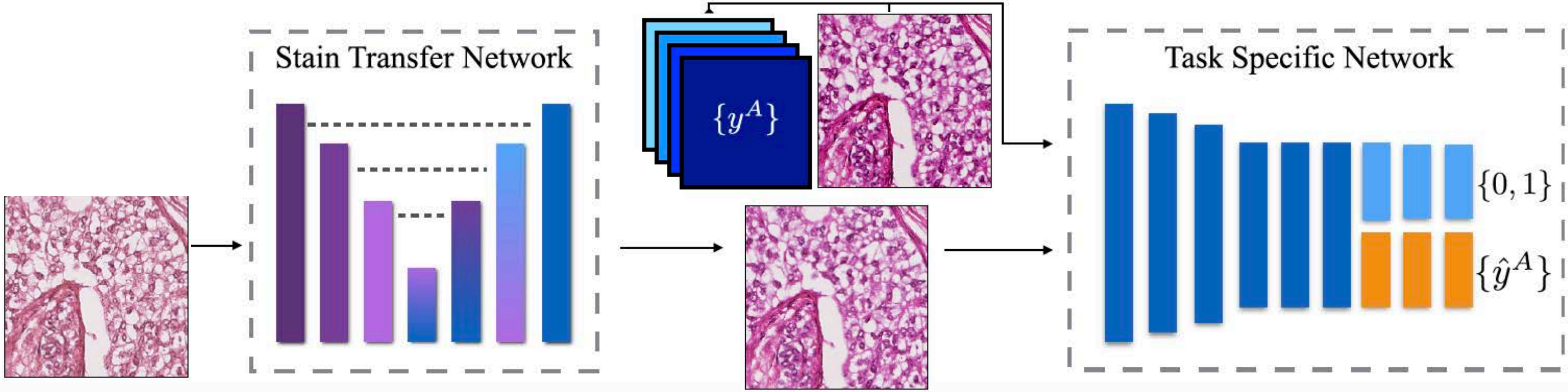


Sapkota *et al.*, “Deep Convolutional Hashing for Low-Dimensional Binary Embedding of Histopathological Images,” *IEEE JBHI*, 2019

Stain/Color Normalization

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

Real images
in Dataset A



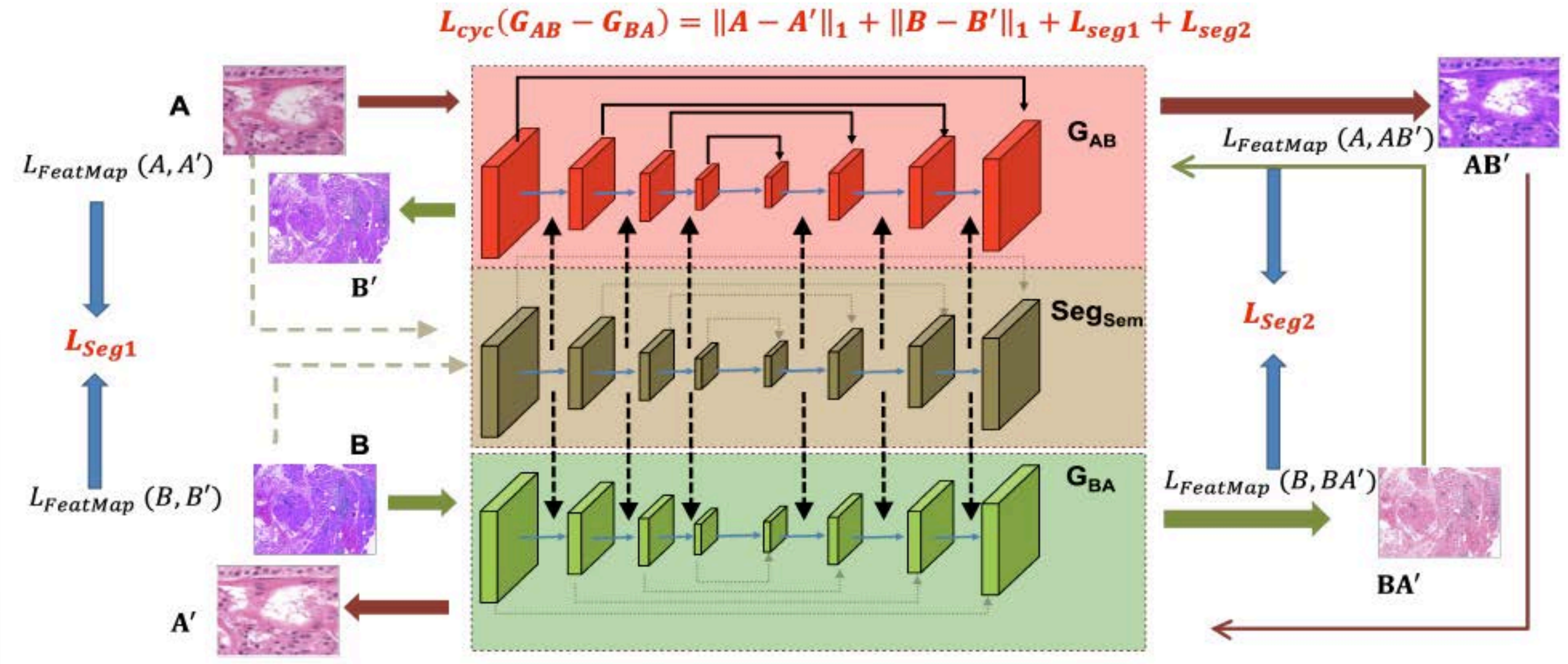
Original images
in Dataset B

Fake images
from Dataset B

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

Stain/Color Normalization

- Stain normalization in digital pathology with cycle-consistency GANs

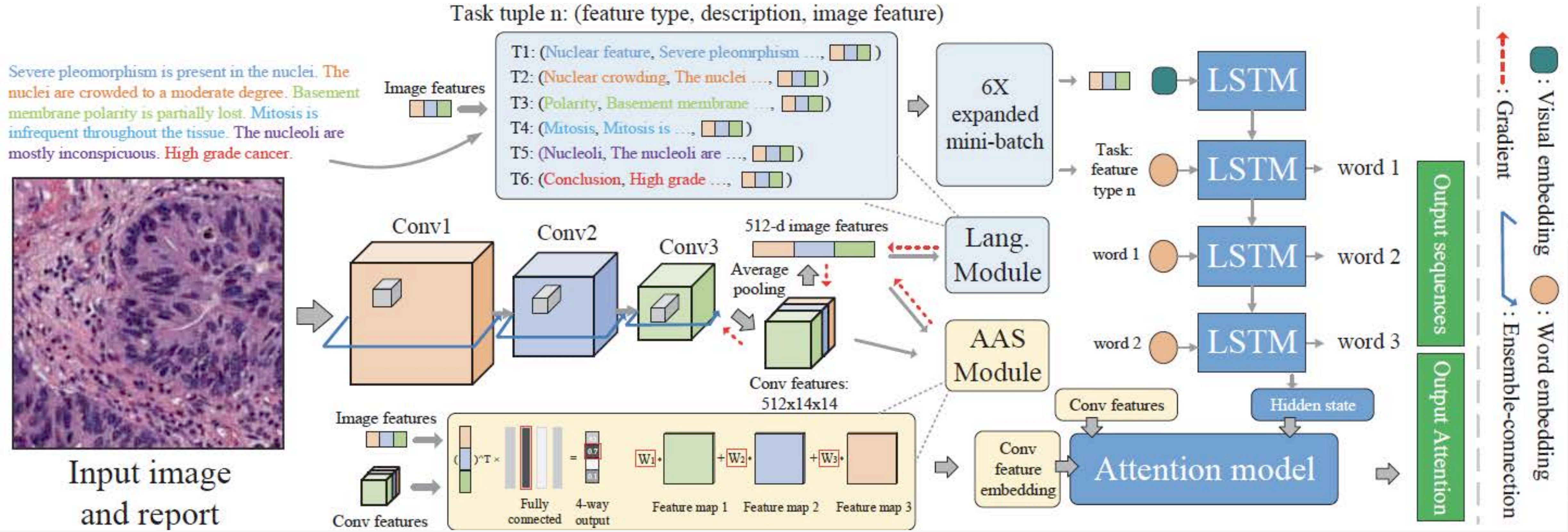


Mahapatra et al., “Structure Preserving Stain Normalization of Histopathology Images Using Self-Supervised Semantic Guidance”, *MICCAI, 2020*.

Text Generation

- Text generation from bladder cancer pathological images

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



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

Thank You