Lecture 3: Medical Imaging

CSCI6410/EPAH6410/CSCI4148

Finlay Maguire (finlay.maguire@dal.ca)

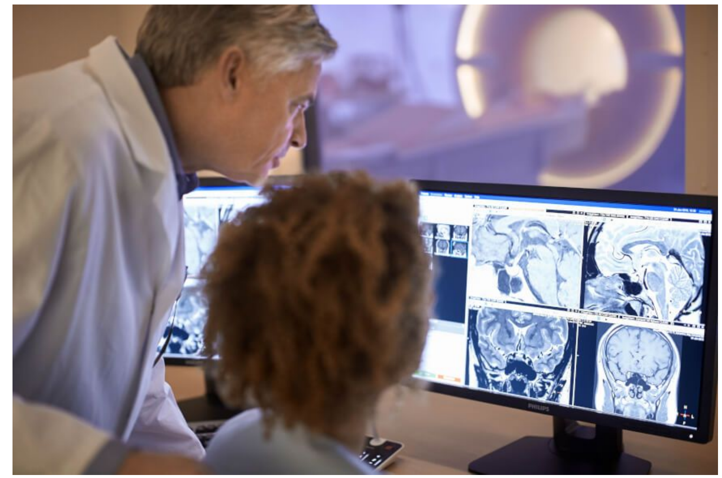
Learning Overview

- Many types of medical imaging data and their respective formats
- DICOM file format is ubiquitous but complex
- Medical imaging data analysis involves many different questions
- Lots of different machine learning paradigms are used to handle challenges of medical image data
- Traditional computer vision approaches
- Deep learning enables learning features/representations
- Convolutions key to capturing spatial relationships
- Augmentation and generative models enable better training with limited data
- Transfer learning and joint-transformer models hugely expand training options
- Machine learning in medical image analysis is very promising but has several major hurdles to broad acceptance

What kind of image data is there in medicine?

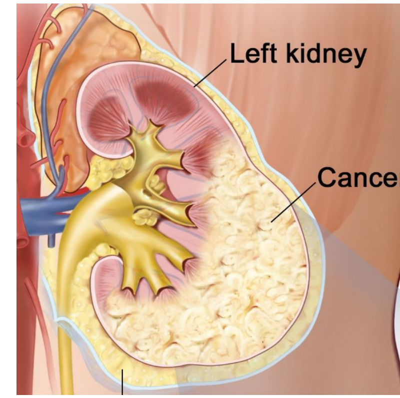
Imaging intensive specialties: Radiology & Pathology

- **Radiologists**: collection and interpretation of medical imaging data (including using imaging to guide minimally-invasive procedures). Often specialise in body parts or types of imaging.
- **Nuclear Medicine:** imaging/intervention involving radioisotopes
- Medical Physicists, Technologists, Sonographers, Technician
- **Pathologists:** study of tissue/samples taken from human body, extensive use of microscopy and staining. Often specialise in body parts and/or analysis methods (e.g., molecular pathology)
- ***Other specialties***: many other specialties use imaging/image data emergency medicine



https://commons.wikimedia.org/wiki/File:Radiologist_interpreting_MRI.jpg

Patient can have many imaging modalities

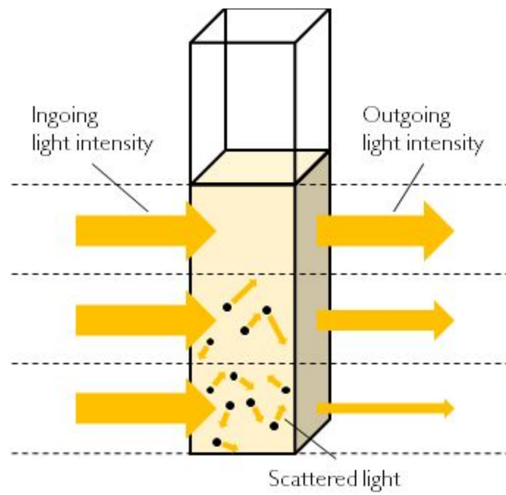


45 year old male presents to *Emergency* with abdominal pain:

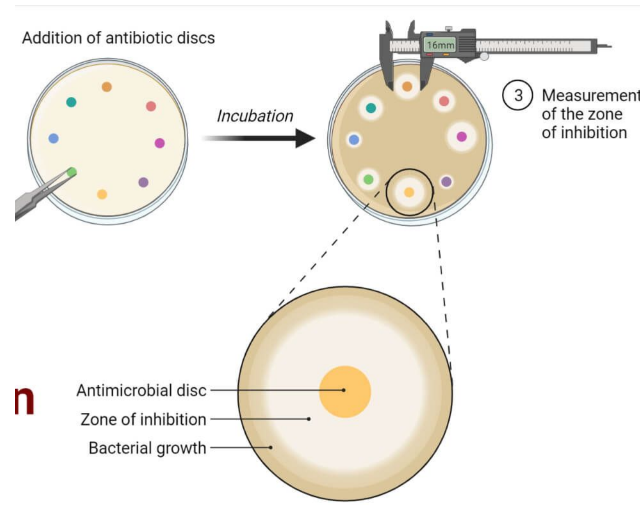
1. *Emergency* performs **ultrasound** finding a kidney lesion
2. *Radiology* performs **CT** on lesion suggestive of renal cancer needing resectioning by *Urology*
3. Given tumour size *Radiology* performs pre-operative **MRI** to guide surgery
4. *Oncology* request **PET-CT** to check for metastasis but no evidence
5. *Urology* resect tumour and send to *Pathology* who perform **histopathological slide imaging** to confirm renal cell carcinoma
6. Due to high-patient risk, regular follow up **PET-CT** performed by *Radiology*
7. Patient returns to *Emergency* with leg weakness 1-2 years later
8. *Radiology* performs **MRI** and identifies a malignant spinal cord compression.
9. *Radiation oncology* perform **CT**-guided emergency radiotherapy

1-dimensional “image data”

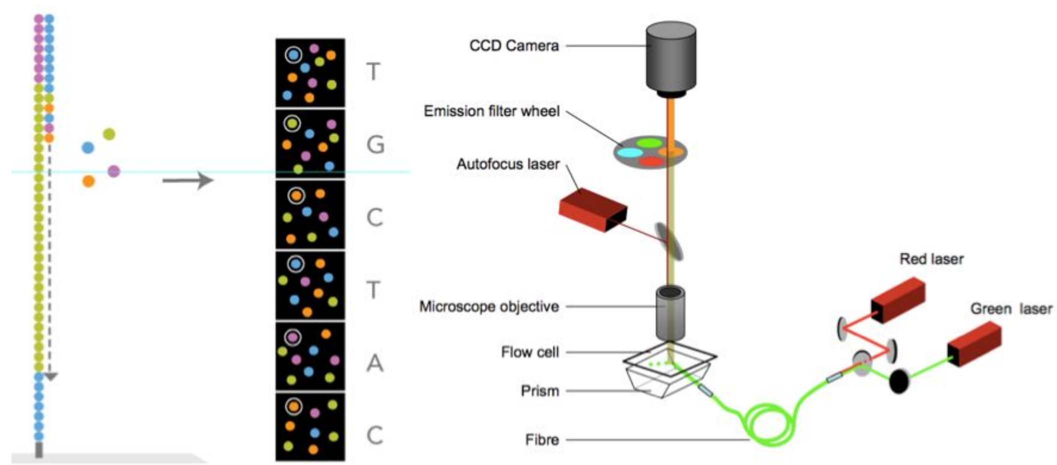
- Many diagnostic tests involve imaging (***culture density***, ***ELISA***, ***antibiotic susceptibility testing***, ***Illumina DNA sequencing***)
- Images can be an *intermediate* format
- Intermediate formats don’t always need stored/further analysed.
- Image analysis trivial and/or hard-coded into machine



aquila biolabs 2020



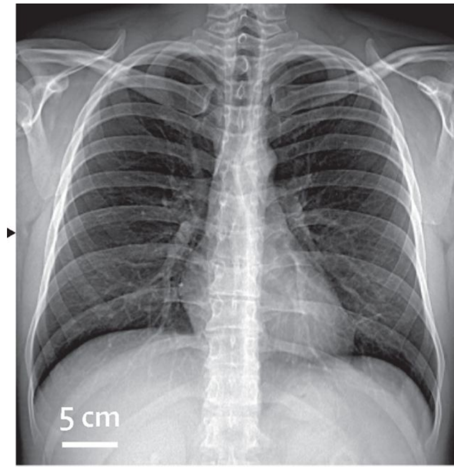
https://microbenotes.com/kirby-bauer-disc-diffusion/

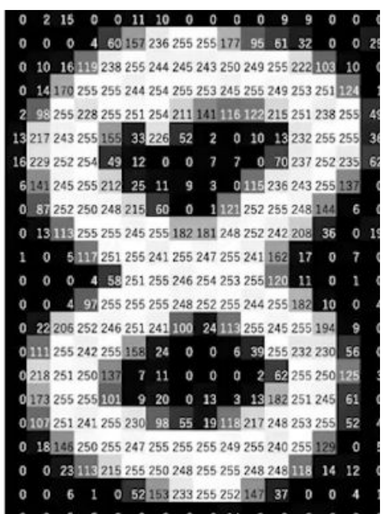


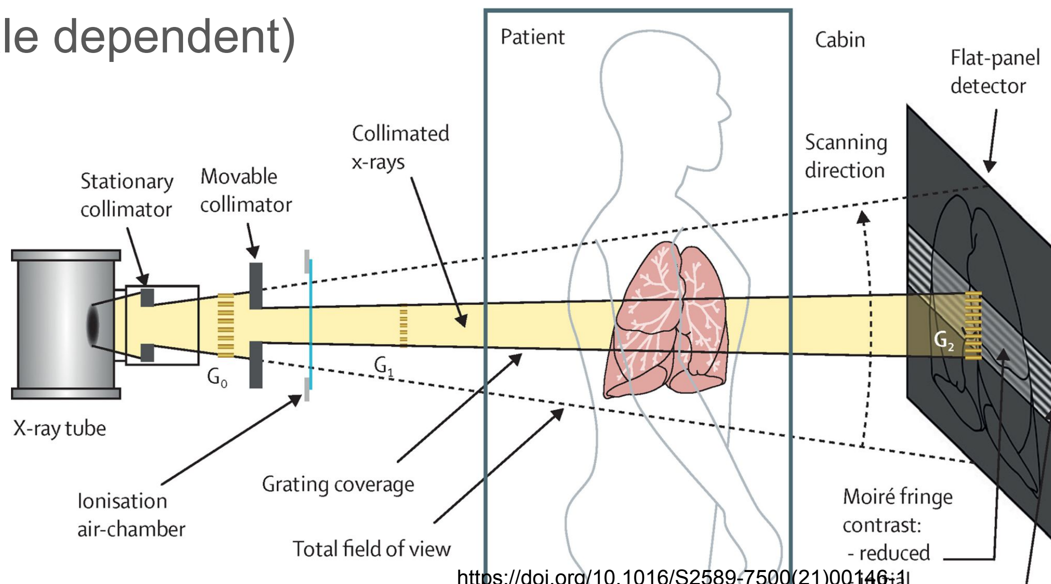
http://enseqlopedia.com/2014/01/nextseq-500s-new-chemistry-described/

2-dimensional grayscale images: X-ray

- Use of **x-rays** to image bone and soft tissue anatomy (special case: mammograms)
- **Pro**: low radiation dose, cheap, common, quick, standardised units
- **Con**: limited tissue density range & metal artefacts
- **Data type**: grayscale image (2D matrix of whiteness intensities), many formats but Digital Imaging and Communications in Medicine (DICOM) standard (size is device/scale dependent)



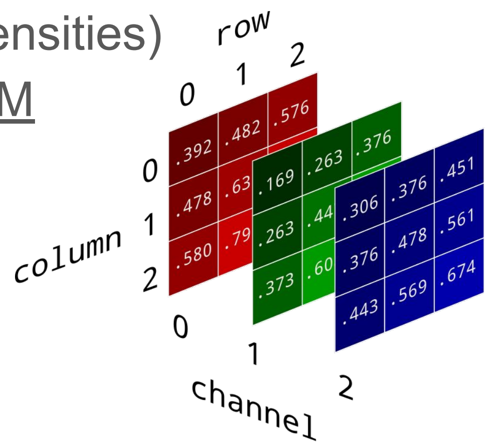


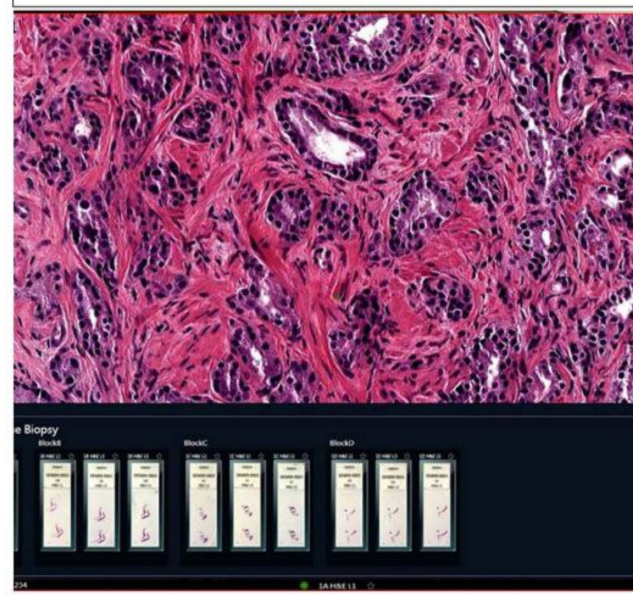


https://doi.org/10.1016/S2589-7500(21)00146-1

2-dimensional colour images: Whole Slide Imaging

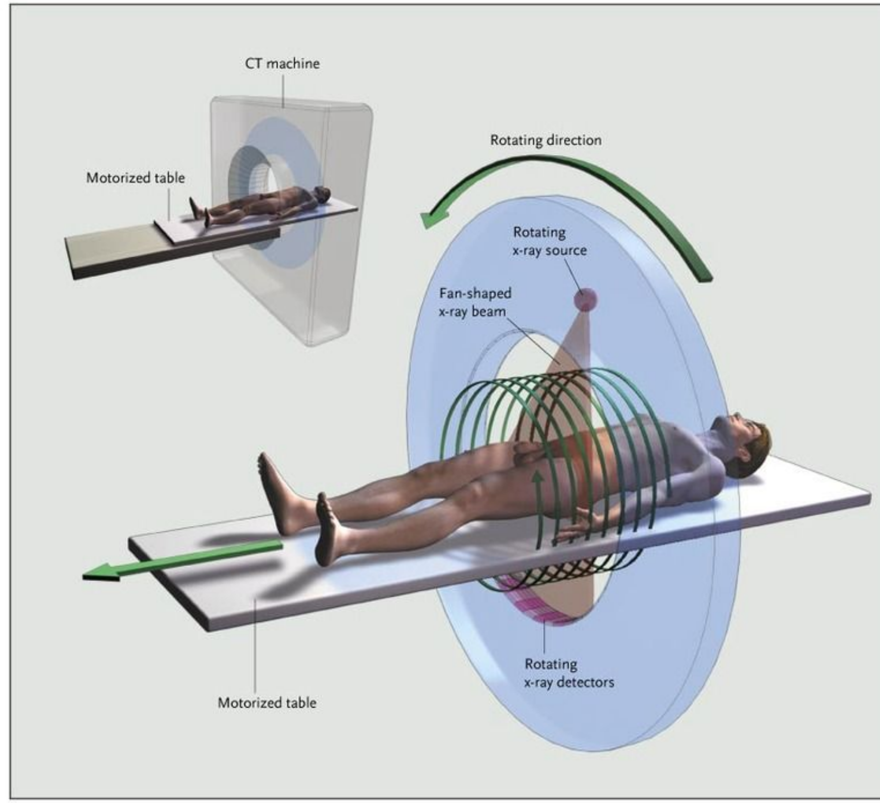
- Digital **scanning/photography** of microscopy slides (including staining/fluorescence)
- **Pro**: sharing/embedded reports, automated analyses, reproducibility
- **Con**: large images (>10GB), sensitive to scanner/preparation
- **Data type:** 3-channel colour image (3D tensor with R,G,B intensities) TIF/JPEG2000/DICOM





Tomographic/2D slices: Computed Tomography (CT)

- **X-rays** in 1000-5000 slices/helical pattern computationally processed into pseudo-3D for any organ system
- **Pro**: great capture of anatomical detail, fast, broader than X-ray (contrast capture of organ/vessels), standardised units
- **Con**: higher radiation dose, relatively expensive & metal artefacts
- **Data Type:** reconstructed grayscale image (3D matrix of whiteness intensities); multiple individual 2D grayscale slides; DICOM

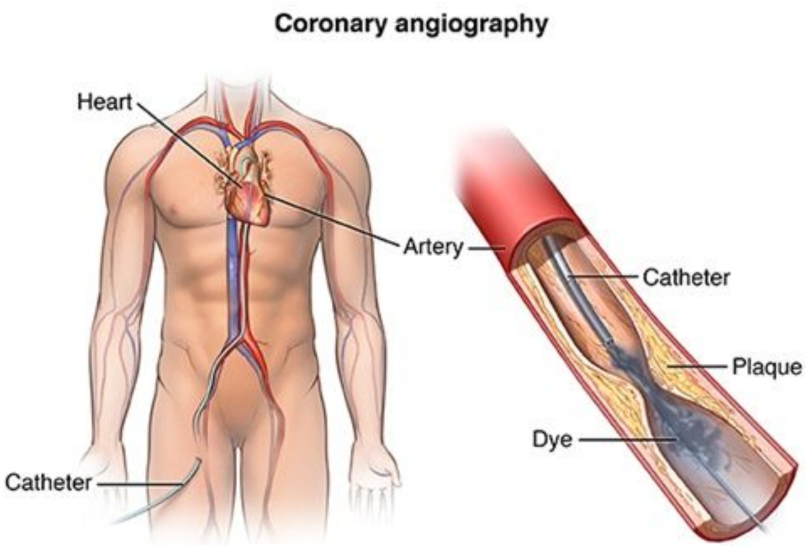


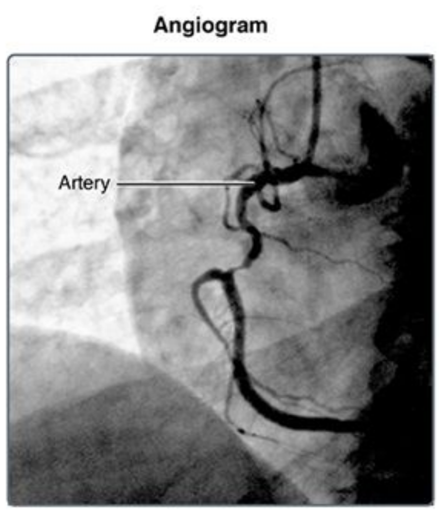
10.1056/NEJMra072149

Pseudo-coloured Dual Energy Contrast CT

Clínica Universidad de Navarra, Pamplona, Spain / Siemens Healthineers

2D video: Fluoroscopy



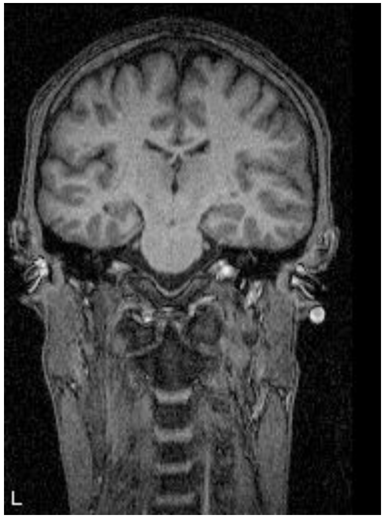


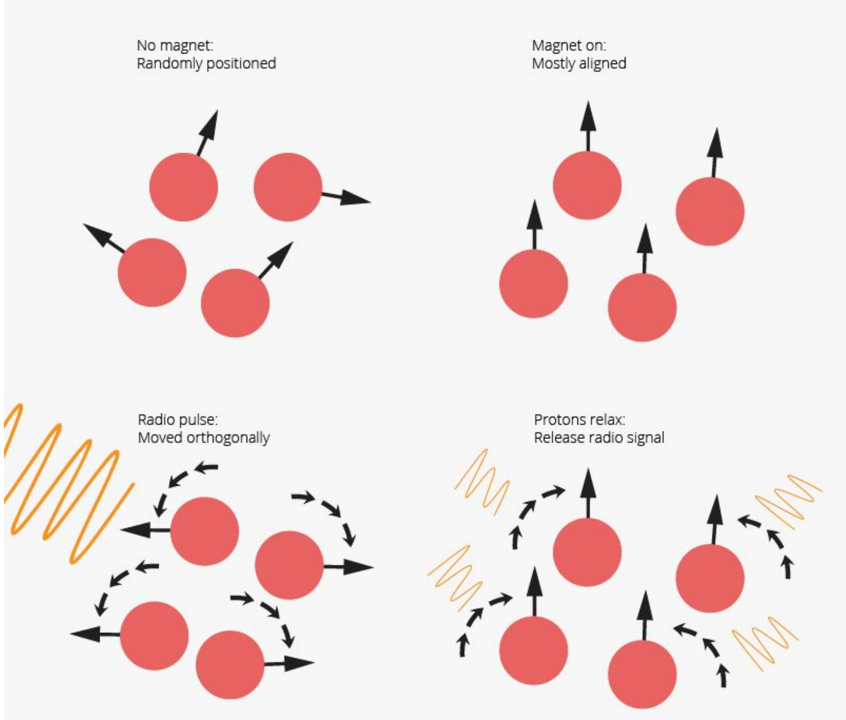
https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/cardiac-catheterization

- **X-ray** collected in a **time-series** with contrast medium used to guide procedures or evaluate change over time (angio.
- **Pro**: real-time imaging, widely available, relatively cheap
- **Con**: more expensive and higher radiation dose than X-ray alone
- **Data**: Series of 2D grayscale (3D tensor): DICOM

3D imaging: Magnetic Resonance Imaging (MRI)

- 3D images constructed from **radiofrequency pulse** perturbation of body atoms (with a magnetic moment protons>neutrons) aligned by strong **magnetic field** (with or without contrast). Can be optimised for blood flow (fMRI)
- **Pro:** detailed multiplanar/3D imaging without contrast, better depiction of soft-tissue than CT, no radiation, painless
- **Con:** Expensive, tissue-specific noise, non-standardised units, lots of required space, susceptible to patient movement => can require sedation.
- **Data:** 3D grayscale tensor; 2D grayscale slices; 4D timeseries; DICOM

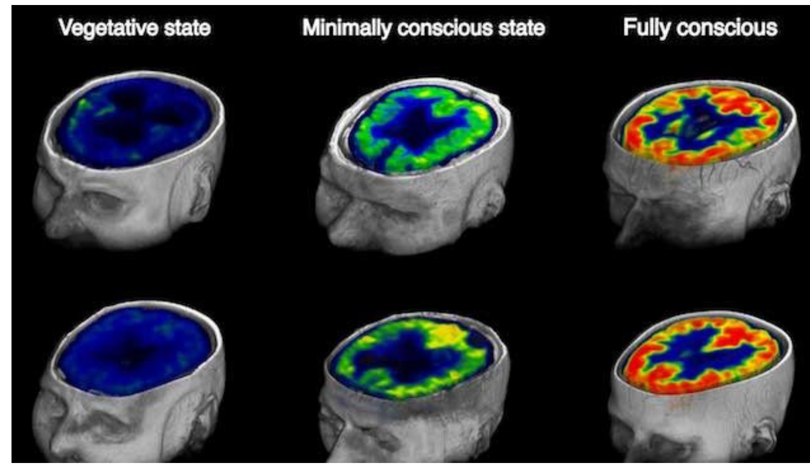




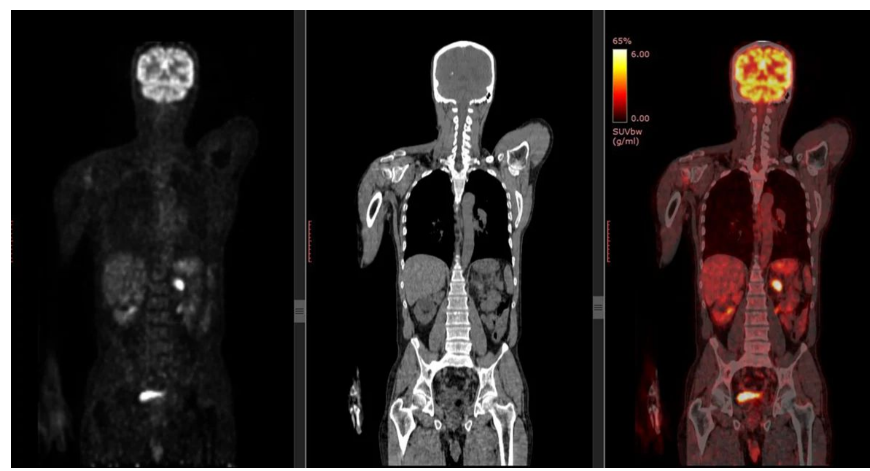
https://imotions.com/blog/eeg-vs-mri-vs-fmri-differences/

4D imaging: Positron Emission Tomography (PET)

- Time-series of detected **positrons** from radiolabelled fludeoxyglucose or O-15 to detect tissue/lesions with high metabolic activity
- **Pro:** measures function, painless
- **Con:** poor anatomy resolution (combine with CT/MRI), very expensive, radioactive tracer
- **Data:** 4D tensor DICOM (5D colour PET-CT/MRI or 3D fixed time-point).
- SPECT and fMRI also offer dynamic functional imaging options

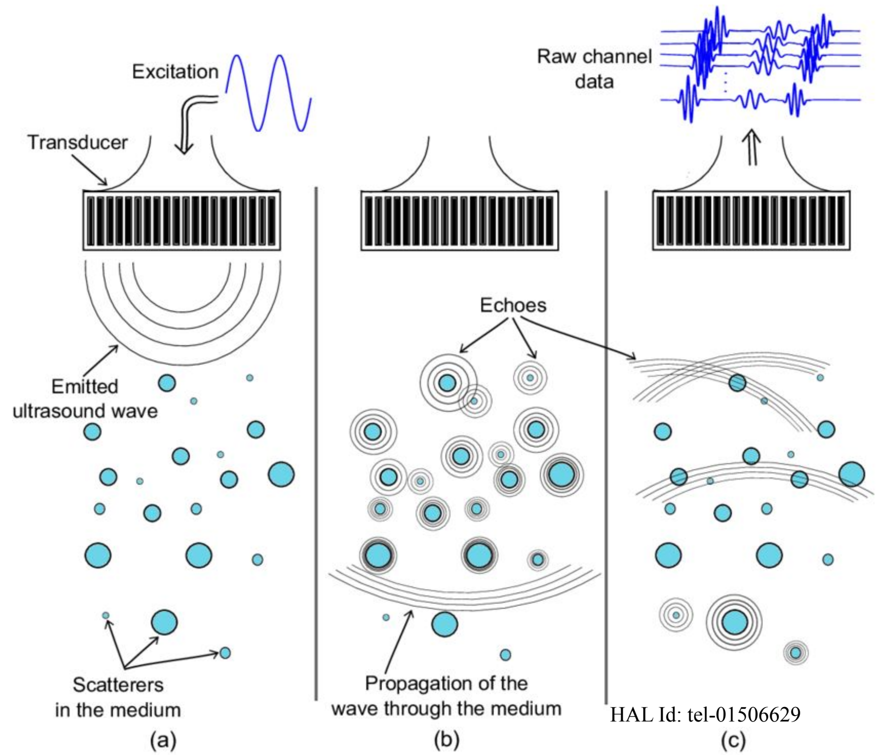


https://sciencebasedmedicine.org/pet-scans-predict-coma-outcome/

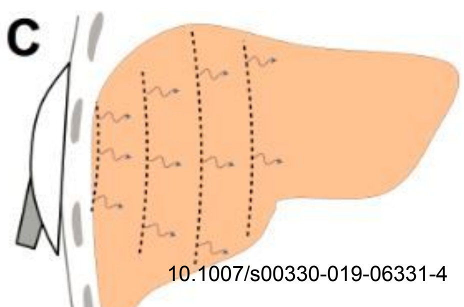


Many formats: Ultrasound

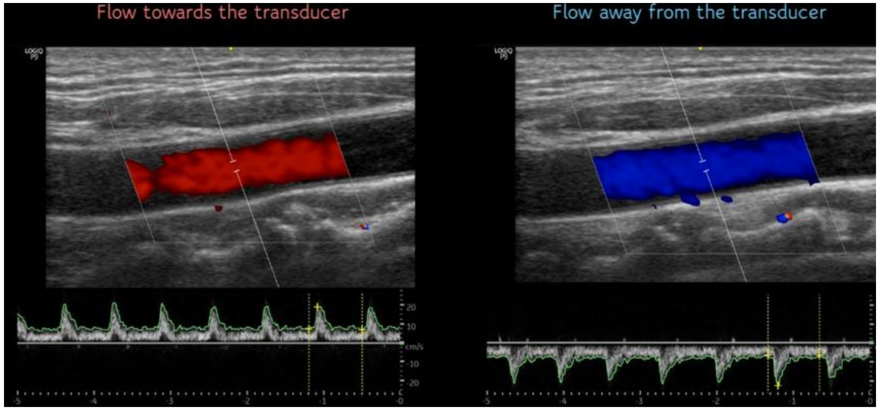
- Uses **sound waves** and echo pattern to image internal structures (soft tissue/organ/vessels)
- **Pro**: real-time imaging, no radiation, portable
- **Con**: operator skill dependent, huge image variation depending on operator, patient anatomy, more expensive than X-ray
- **Many types**: *elastography*, *doppler*, *triplex*, *transvaginal*, *endoscopic*
- **Data type:** be rendered as 2D, 2D-slices, 3D or 4D! DICOM



HAL Id: tel-01506629



10.1007/s00330-019-06331-4



https://www.renalfellow.org/2020/12/07/basics-of-doppler-ultrasound-for-the-nephrologist-part-2/

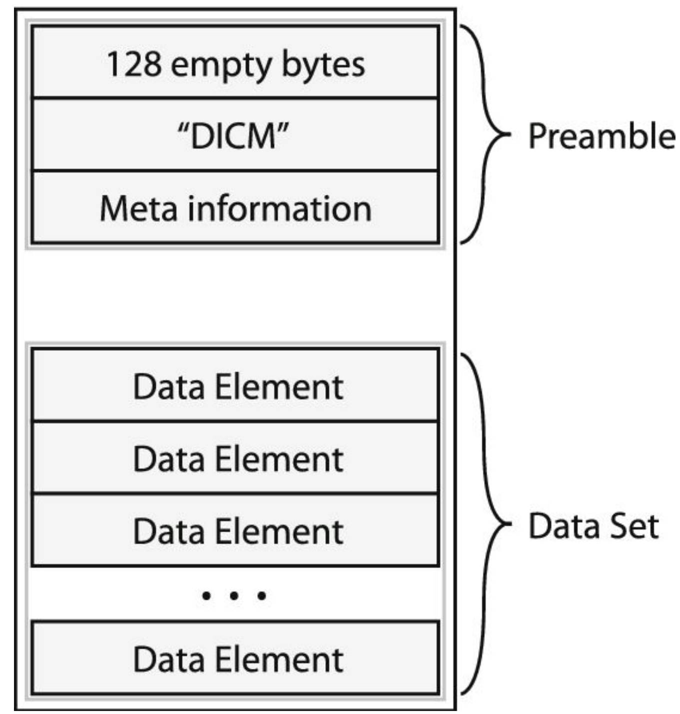
Overview of medical image data

- Point measure (1D - single value): optical density/turbidity/fluorescence
- Project/Planar imaging (2D matrix of pixels): X-ray
- Tomographic/Multislice Imaging (3D tensor of pixels): a series of images representing slices through a volume: CT, Optical Coherence Tomography Angiography
- 3-dimensional/Volume (3D tensor of voxels): MRI
- Dynamic Series (4D tensor of voxels): fMRI/PET-CT/MRI

- 45 year old male presents to *Emergency* with abdominal pain: 100-1000s of GB of imaging data in a variety of different imaging formats/modalities (mostly as DICOM files)

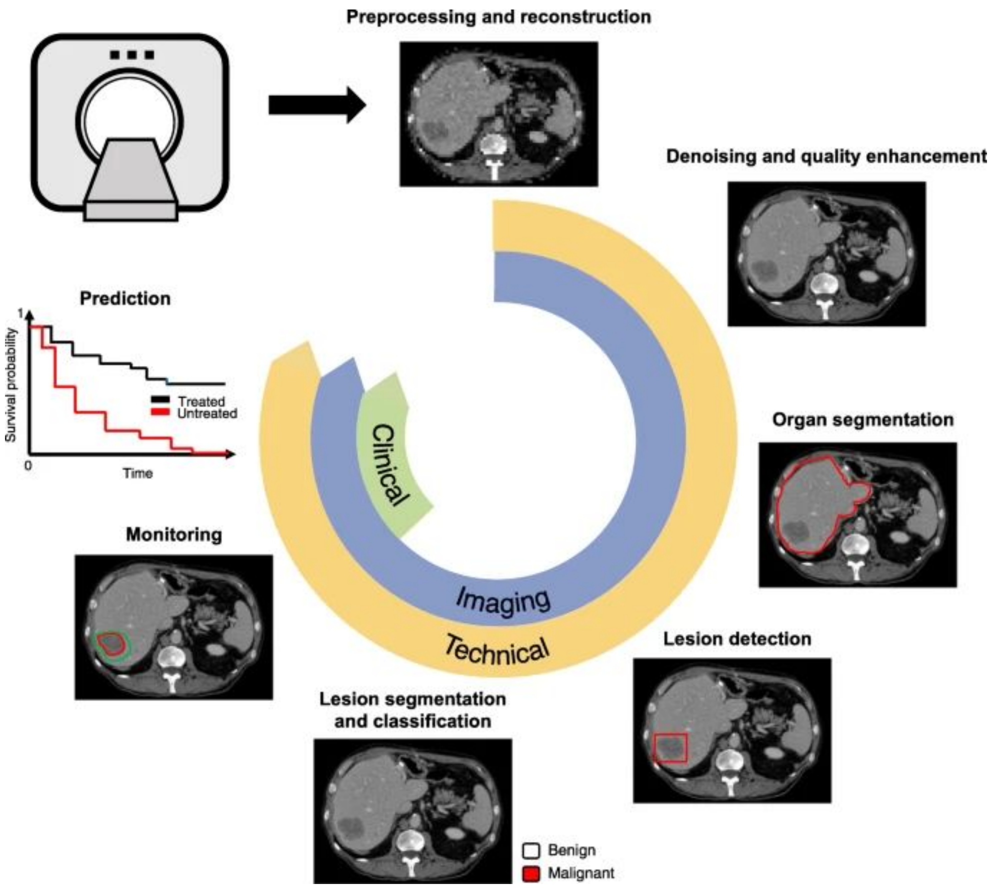
DICOM Data Format

- Standardised file format split into preamble and image information
- Preamble contains key metadata:
    - **Pixel depth**: number of bits encoding each pixel/voxel (e.g., 8/32/64-bits)
    - **Samples per Pixel/Number of Channels**: number of values encoding each pixel/voxel e.g., monochrome = 1 channel, colour = 3 channels (R,G,B)
    - **Spatial resolution**: size of smallest discernible feature
    - **Other Embedded Metadata**: capture frequency/contrast/capture model/patient size (important for normalising across your data)
- Image data: image matrix/tensors (nominally integer only although scaling is possible)
- Many other proprietary formats exist but DICOM is mostly successful as a standard



What kind of analysis task would we want to do using these images?

Lots of things we may want to do using medical images



- **Image preprocessing**: super resolution, densification
- **Registration:** align spatial coordinates of images into 1 common system (PET + MRI)
- **Detection:** highlighting specific elements (anomaly/lesion/artefacts)
- **Segmentation:** delineation or volume extraction of target object (organ/lesion)
- **Classification:** distinguish classes of objects (benign vs malignant lesion)
- **Monitoring:** longitudinal measurement of lesion (% of organ impacted by lesion)
- **Prediction:** predicting outcome based on image (success of chemotherapy, age of a brain)

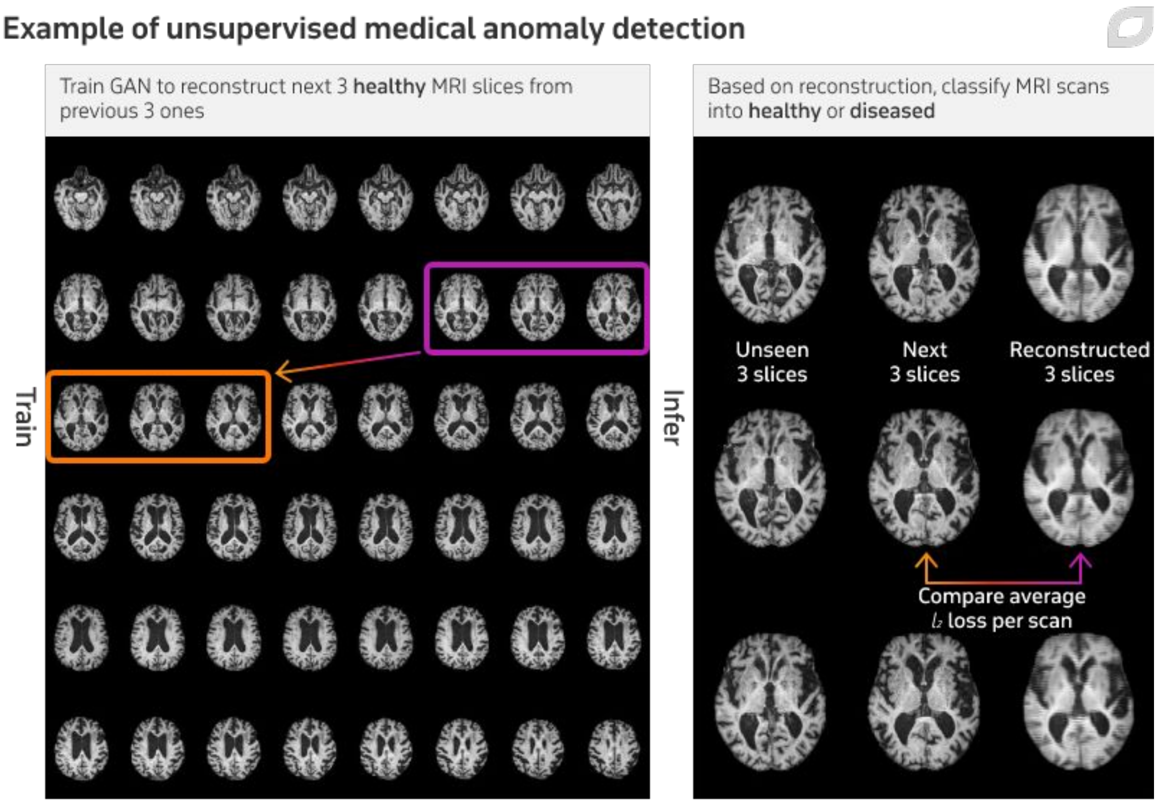
https://doi.org/10.1186/s13244-019-0832-5

Medical image data is intrinsically challenging

- Lots of modalities with very large image size (but small datasets)
- Non-standardised acquisition (varied devices, set-ups etc)
- Disease patterns in images are very long-tailed
- Labels are sparse and noisy
- Samples are heterogeneous and imbalanced
- Subjectivity in ground-truth
- Can be impossible to de-identify e.g., facial scans

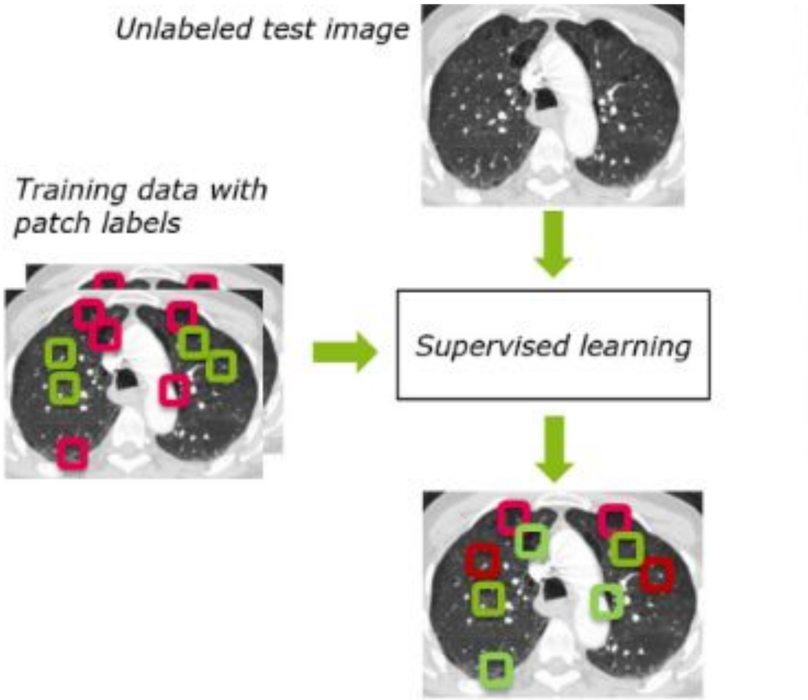
- BUT: **generally informative priors** support computational analysis
  - Body to millimeter scale,
  - Structural hierarchy
  - Bilateral symmetry
  - Morphological similarities
  - Known tissue type density/texture/composition signatures
  - Predictable growth
  - Anatomical landmarks

Unsupervised learning



Data source: bmcbioinformatics.biomedcentral.com—MADGAN: unsupervised medical anomaly detection GAN using multiple adjacent brain MRI slice reconstruction, 2021

Label challenges require alternative training paradigms

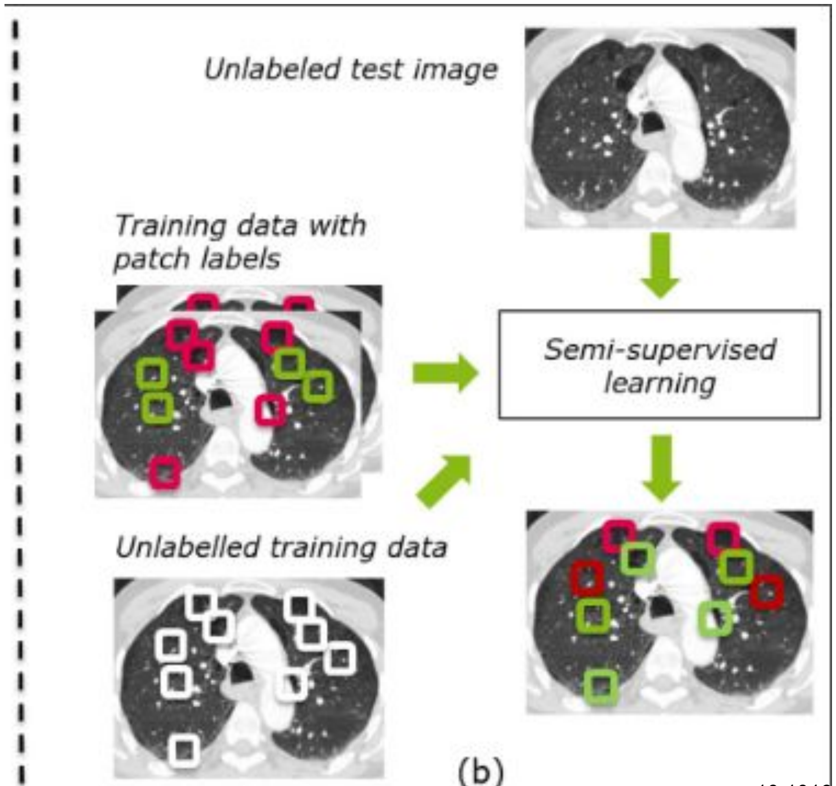


(a)

10.1016/j.media.2019.03.009

Supervised learning

Label challenges require alternative training paradigms

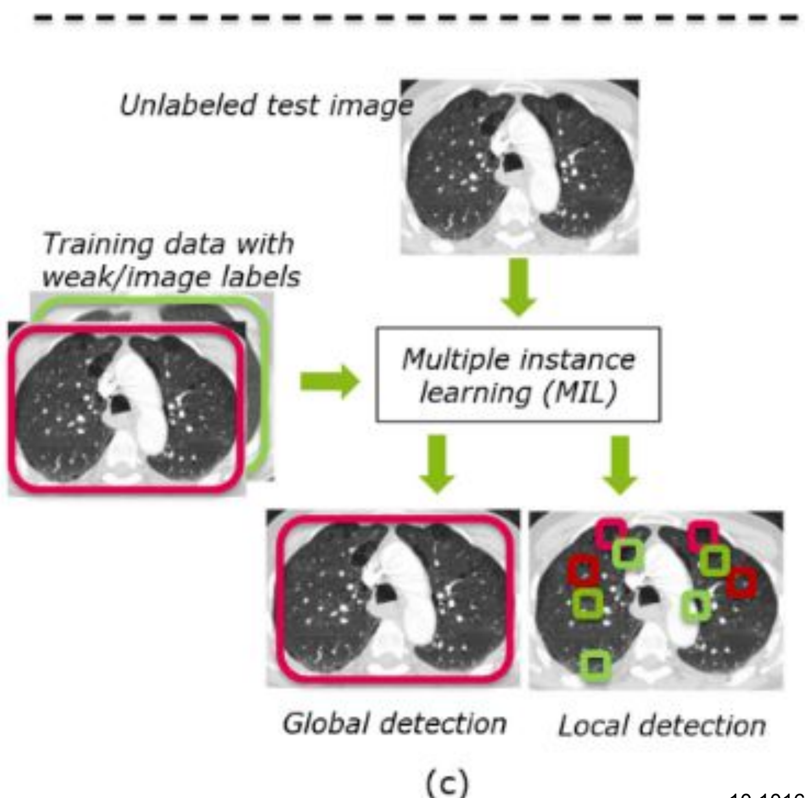


10.1016/j.media.2019.03.009

Supervised learning

Semi-supervised learning

Label challenges require alternative training paradigms



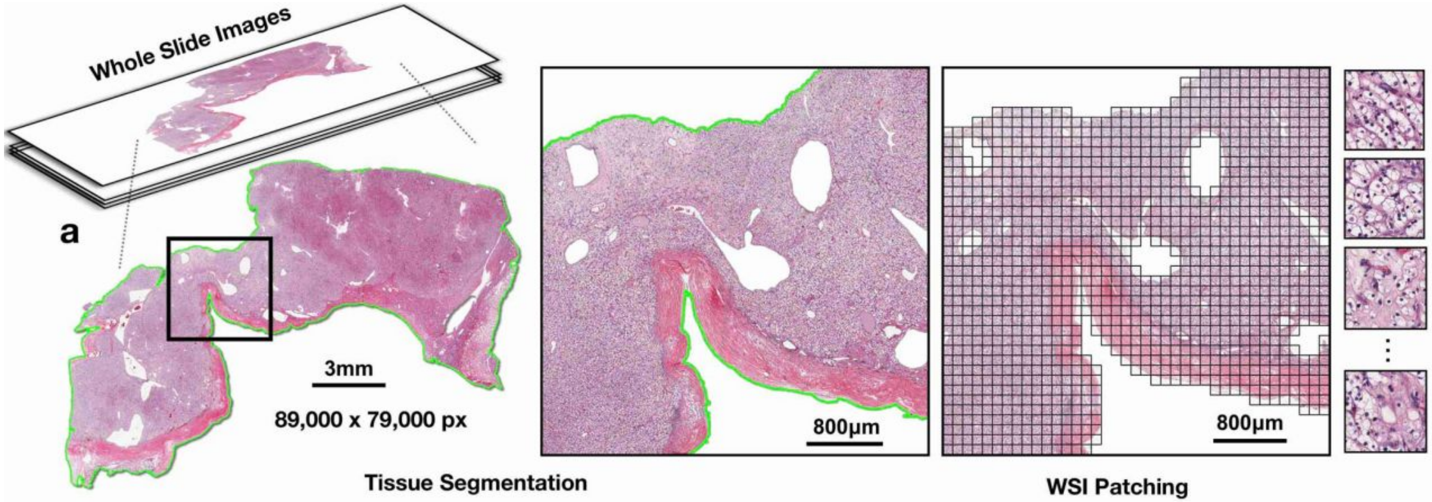
10.1016/j.media.2019.03.009

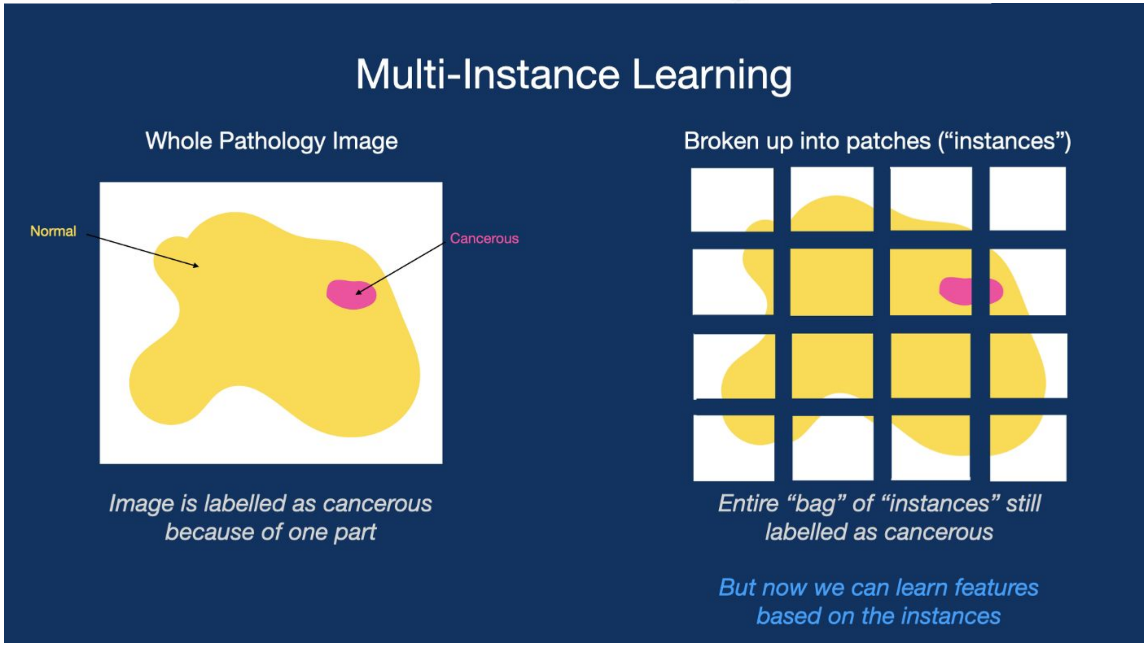
Supervised learning

Semi-supervised learning

Multiple Instance Learning

Label challenges require alternative training paradigms



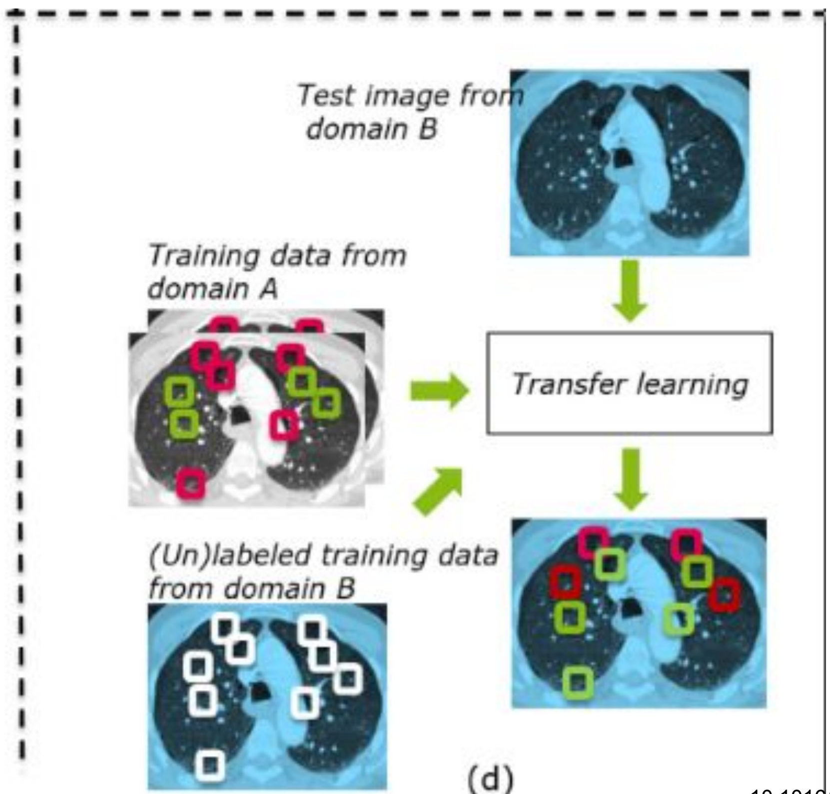


Supervised learning

Semi-supervised learning

Multiple Instance Learning

Label challenges require alternative training paradigms



10.1016/j.media.2019.03.009

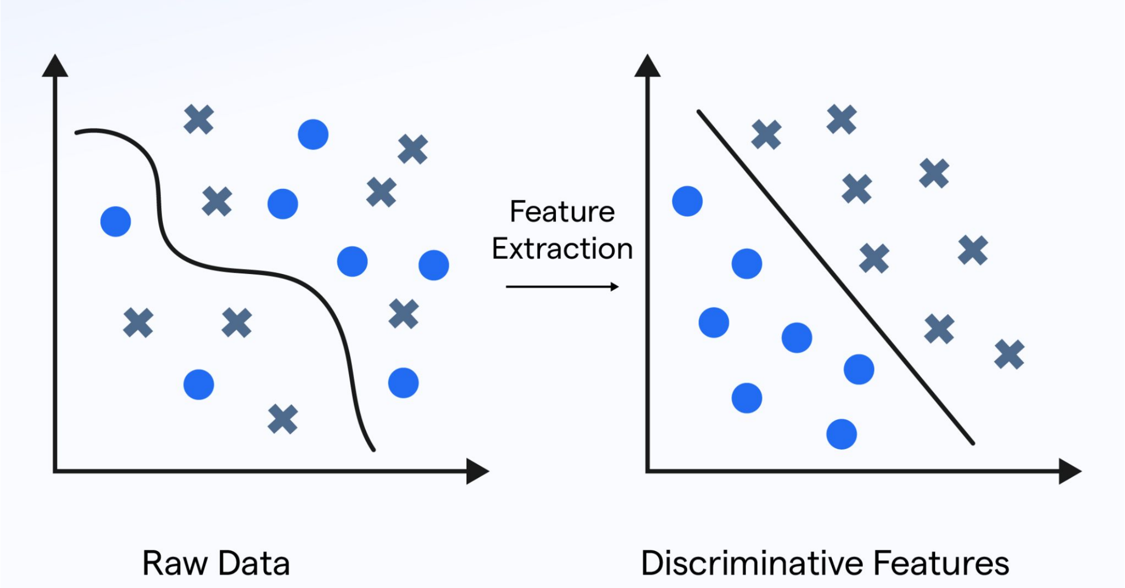
Supervised learning

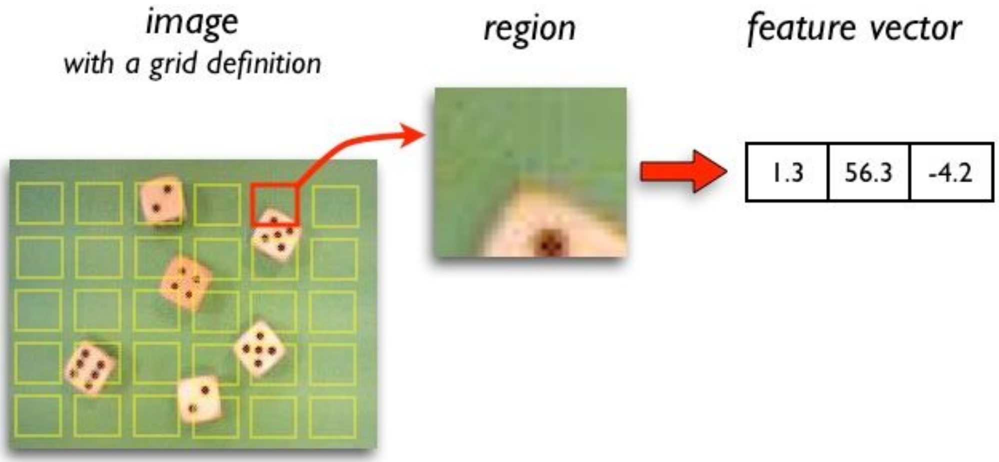
Semi-supervised learning

Multiple Instance Learning

Transfer Learning

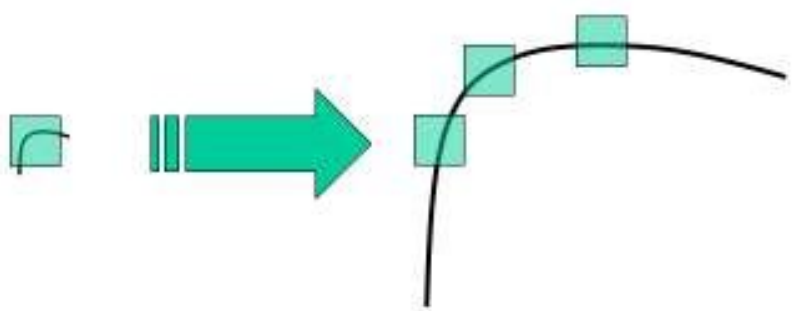
Feature Extraction: Raw Data -> Discriminative Data





Traditional Computer Vision

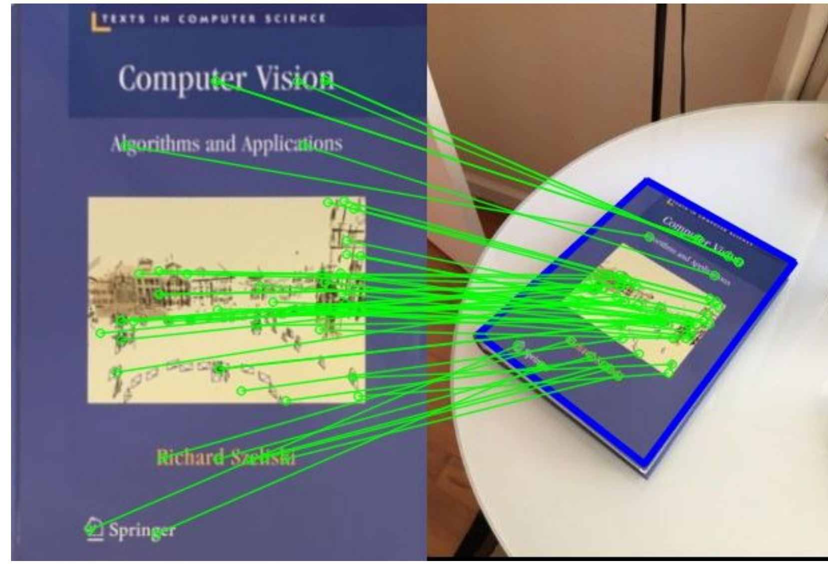
- Thresholding: pixels >= certain set to max
- Edge detection: changes in brightness
- Segmenting: grouping thresholded areas enclosed by edges
- Curve detection: edges adjacent to one another
- Optical flow: detection of movement



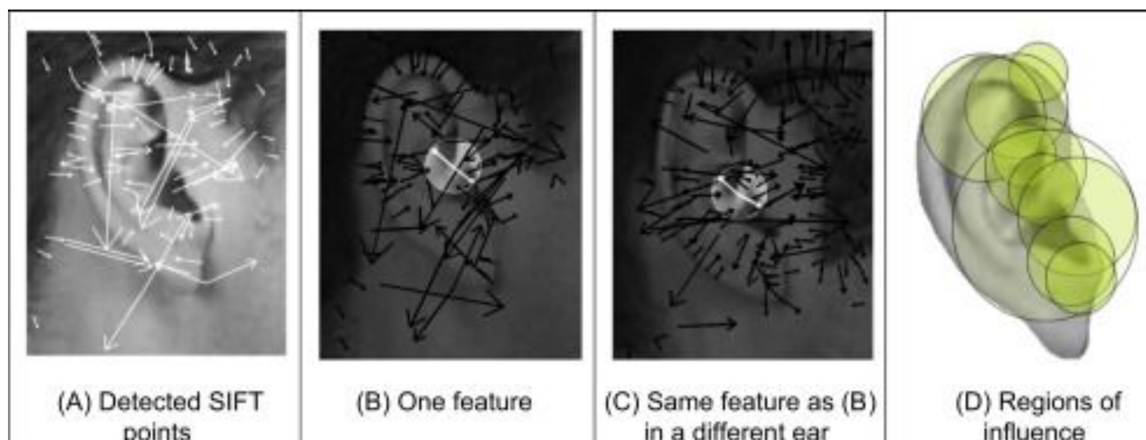
https://pyimagesearch.com/2021/04/28/opencv-thresholding-cv2-threshold/

More complex traditional methods

- Scale-Invariant Feature Transform (SIFT)
- Speeded Up Robust Features (SURF)
- BUT, manual feature engineering is difficult, time-consuming, and often doesn't generalise well



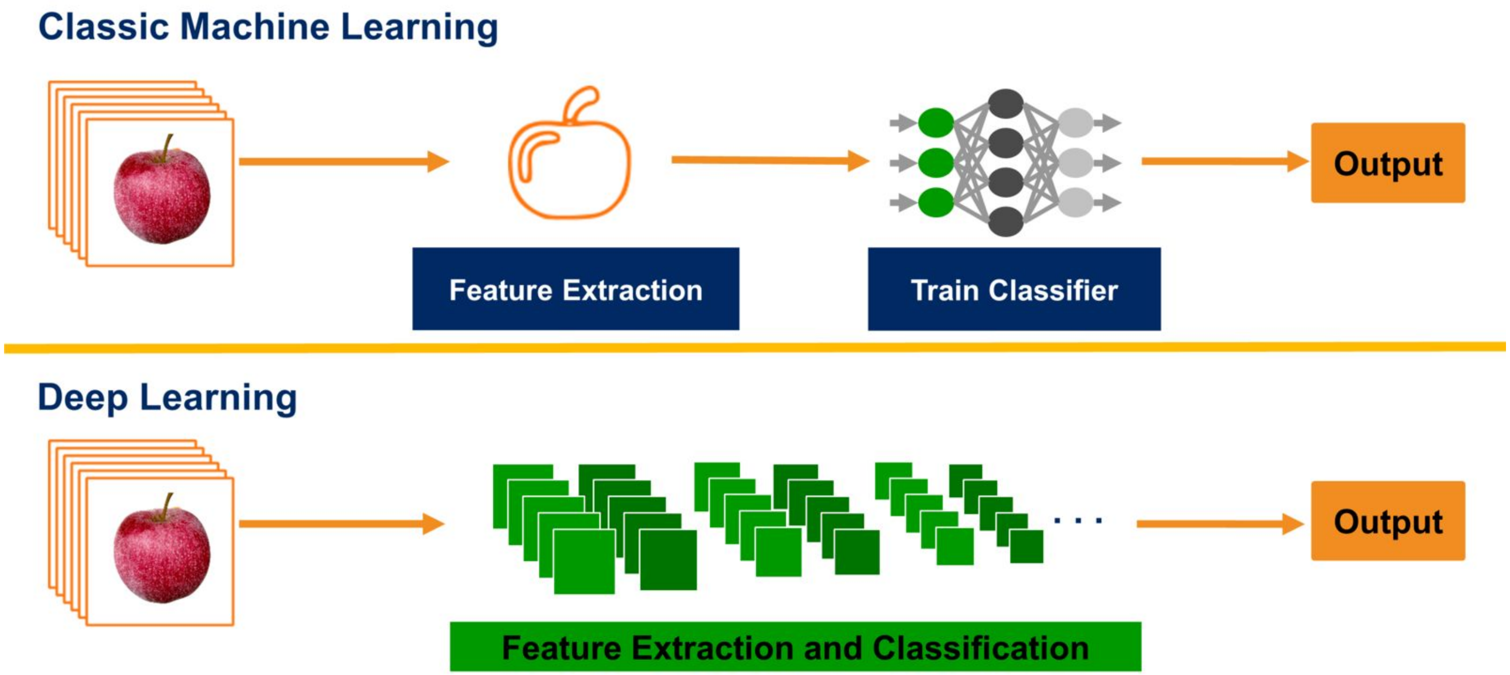
https://stackoverflow.com/questions/51693427/sift-object-matching-in-python



10.1016/B978-0-12-814976-8.00005-1

How do we do analyse images without feature engineering?

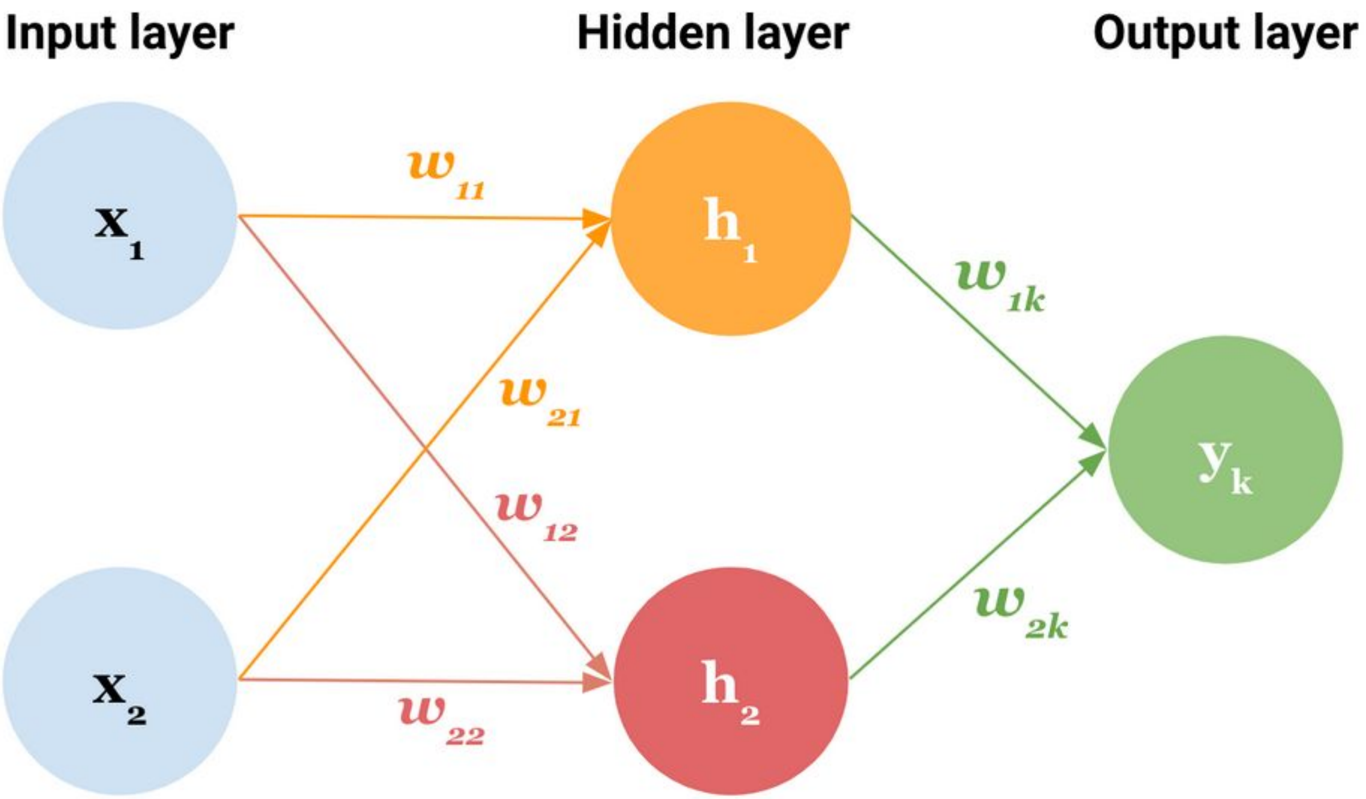
Deep Learning discovers feature representations



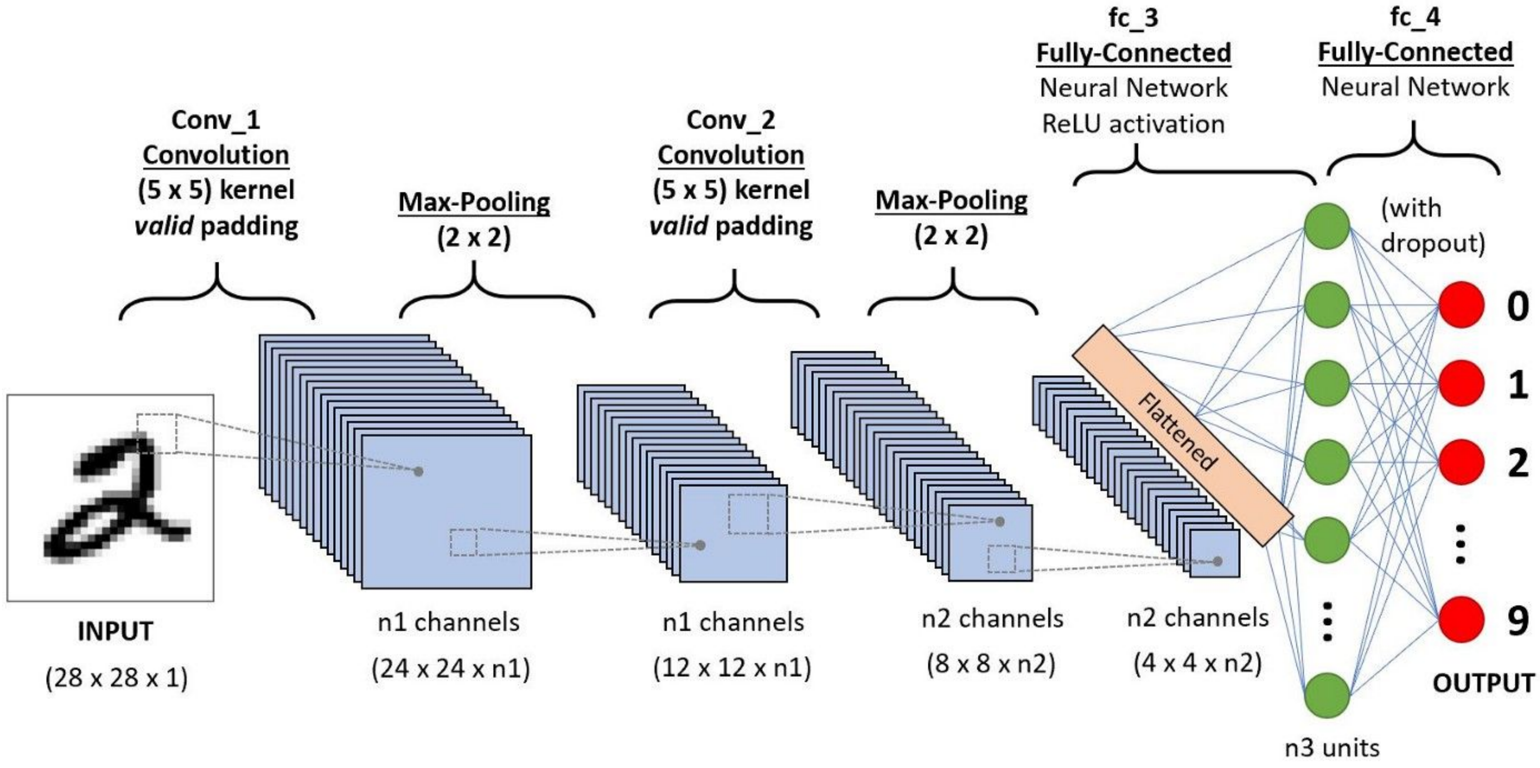
https://www.mvtec.com/technologies/deep-learning/classic-machine-vision-vs-deep-learning

Neural Networks

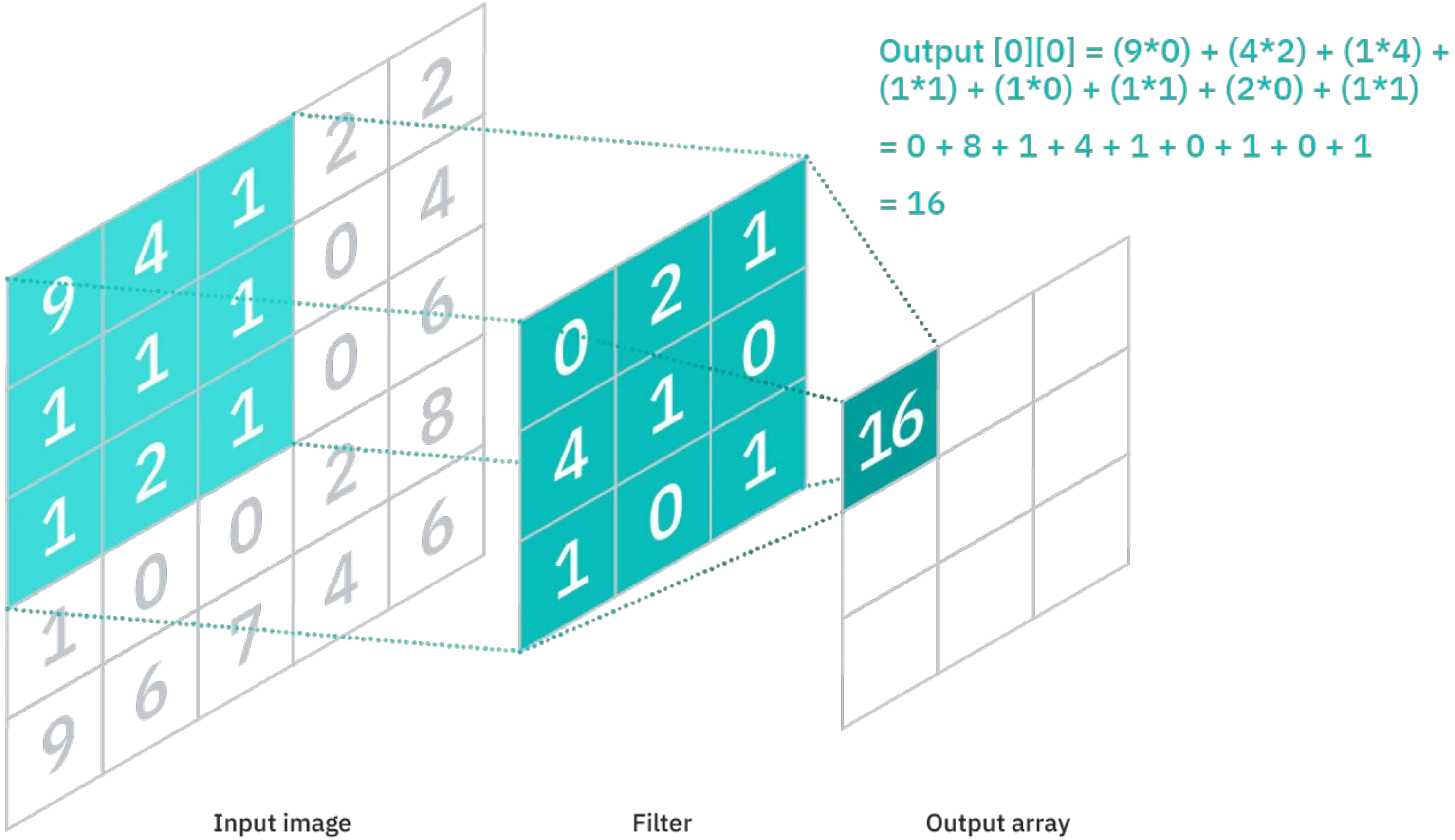
- Need: efficient models that capture high-dimensional non-linear relationships
- Solution: stack many simple models with a non-linearity (e.g., logistic / ReLU)
- **Neural Networks:**



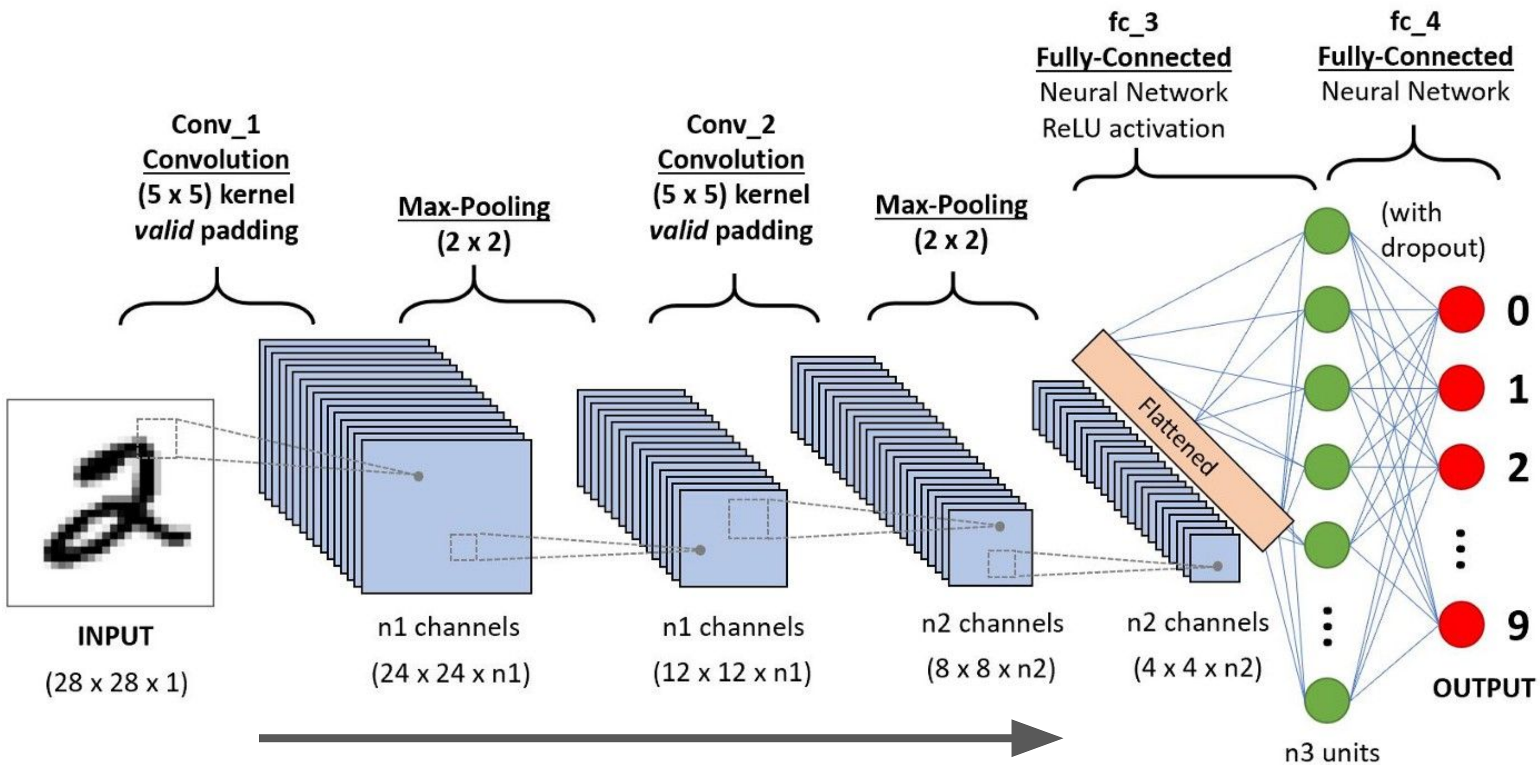
Convolutions capture spatial relations



Convolutions capture spatial relations

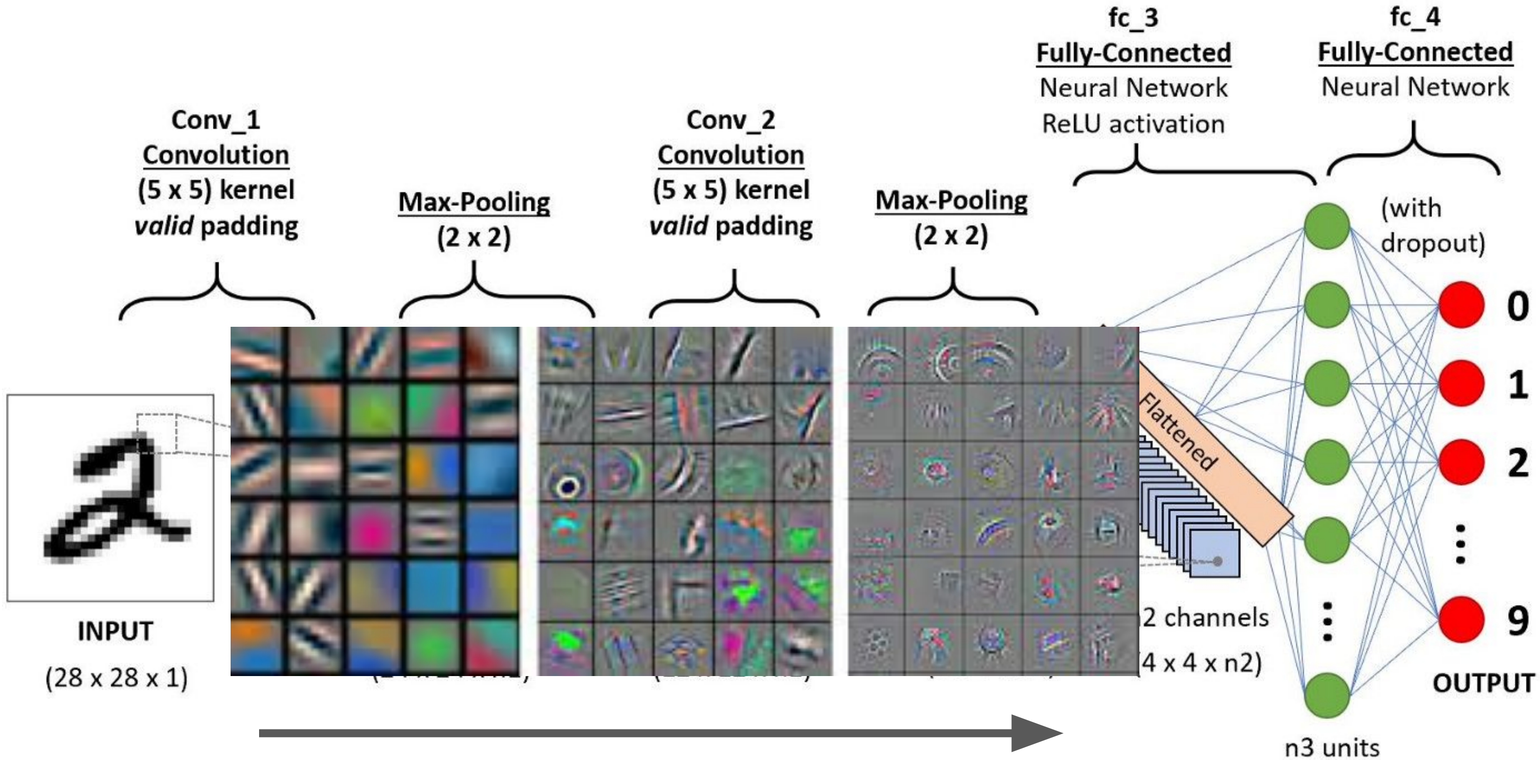


Convolutions capture spatial relations



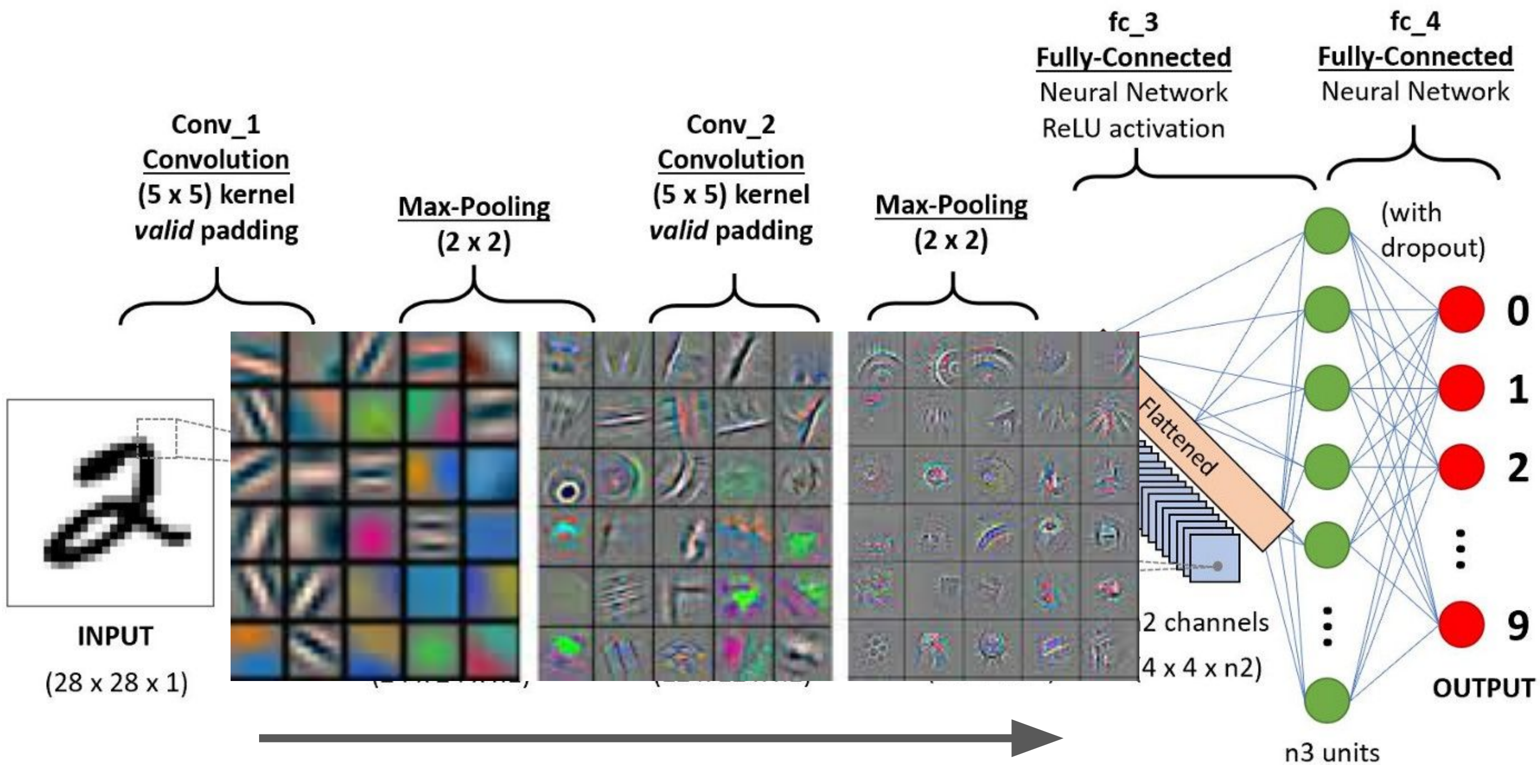
Increasingly higher order learnt representations

Convolutions capture spatial relations



Increasingly higher order learnt representations

Convolutions capture spatial relations

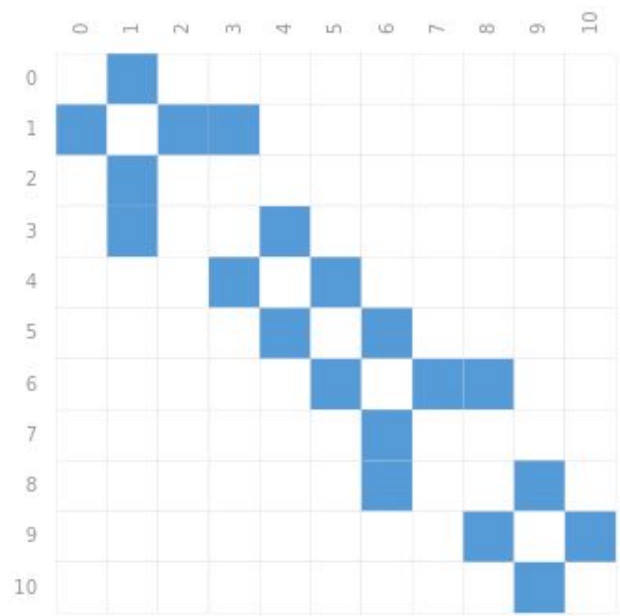


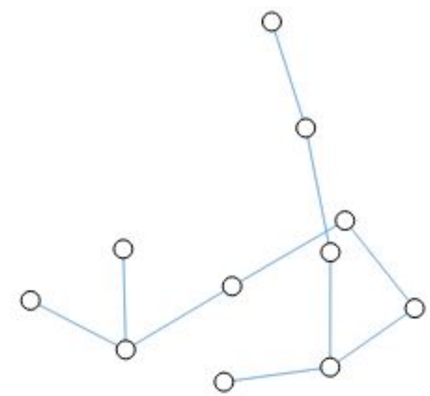
Increasingly higher order learnt representations

- Can use convolutions in more complex architectures (attention, resnets, transformers etc).

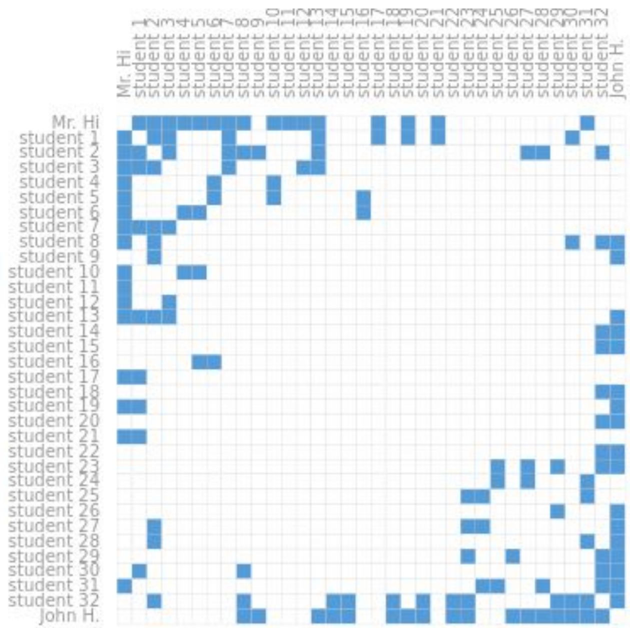
Aside: CNNs can be used for non-image spatial data

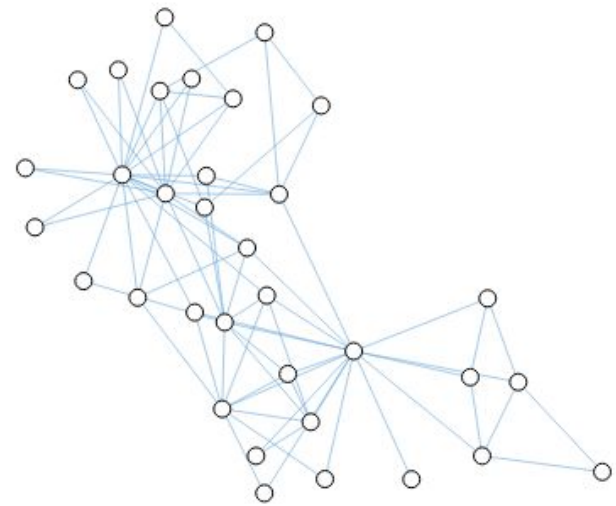
https://distill.pub/2021/gnn-intro





Aside: CNNs can be used for non-image spatial data

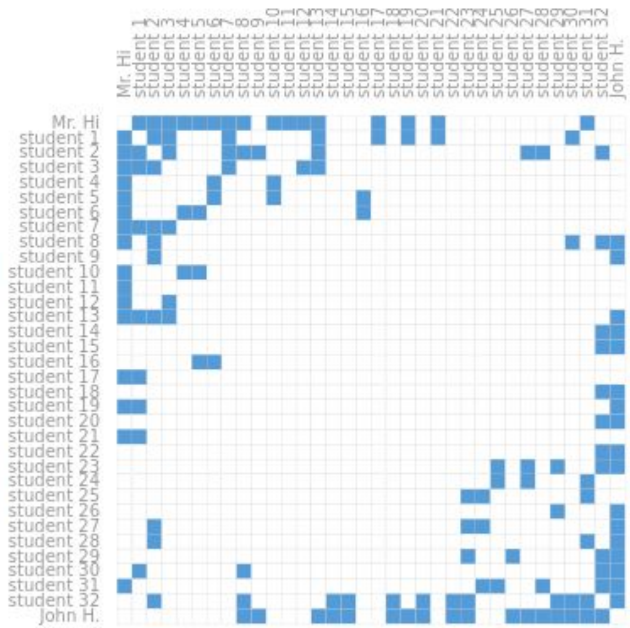


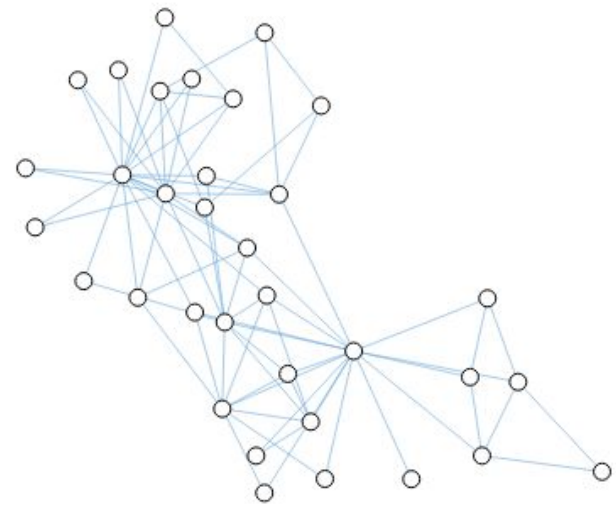


https://distill.pub/2021/gnn-intro

- Graph neural networks

Aside: CNNs can be used for non-image spatial data

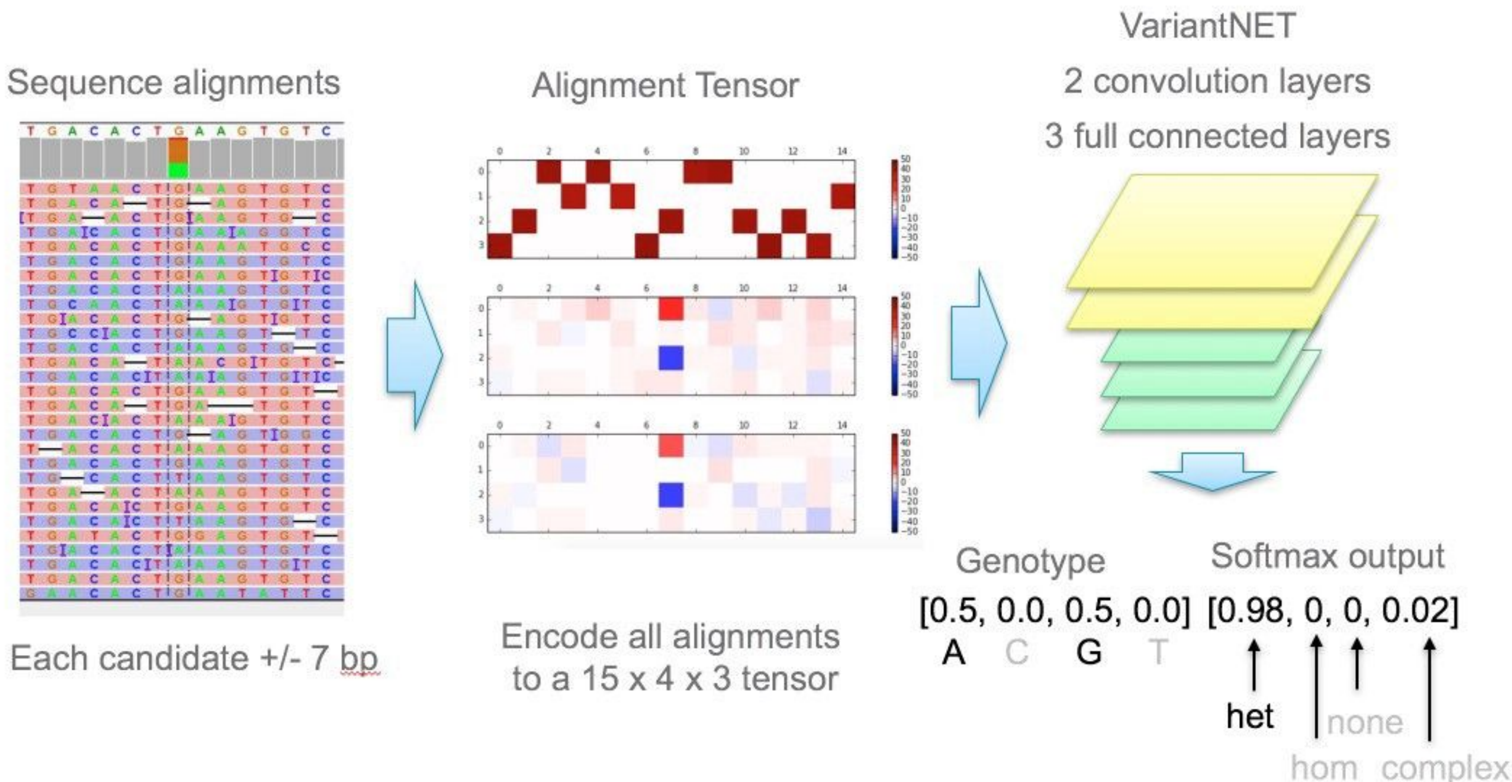




https://distill.pub/2021/gnn-intro

- Graph neural networks
- Text data (semantic networks)

Aside: CNNs can be used for non-image spatial data



- Graph neural networks
- Text data (semantic networks)
- Mutation calling…

U-Net forms basis of most SOTA approaches

- Encoder-decoder architecture
- Down and up-convolutions
- Convolutions get features but lose spatial relations (Down)
- Rebuild spatial relations (Up)

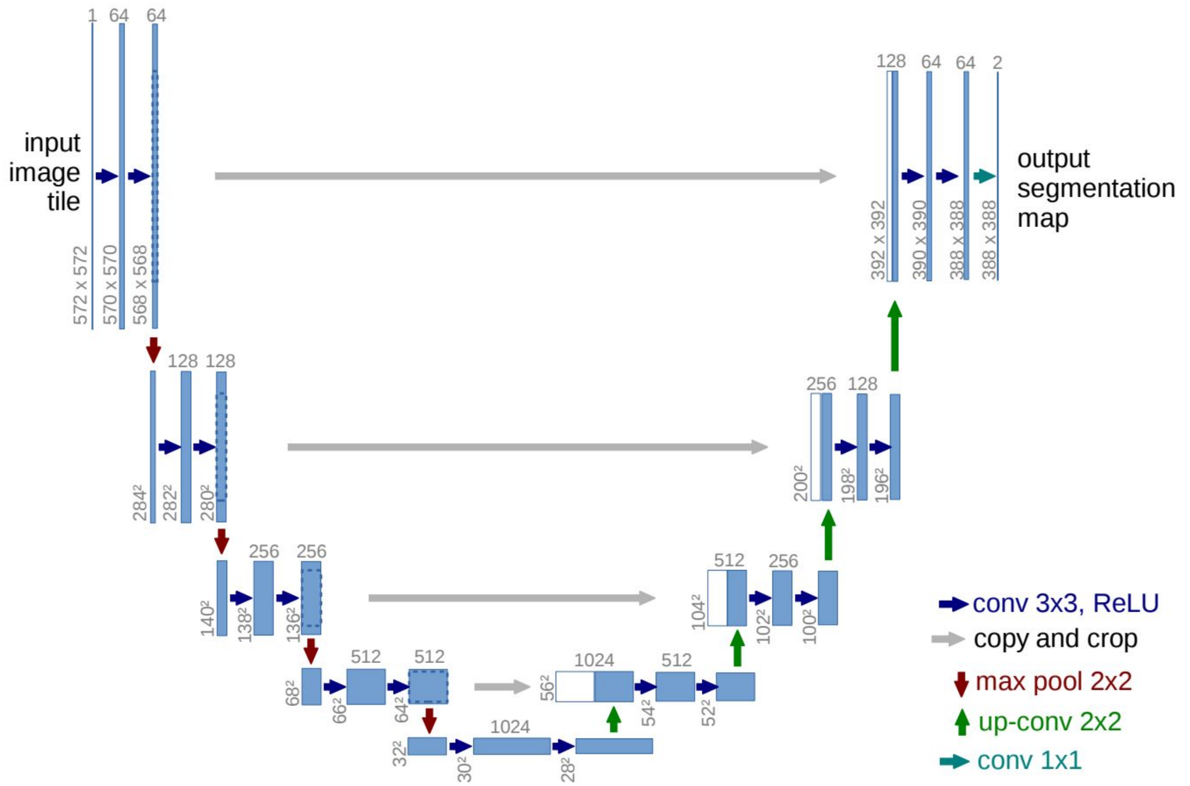


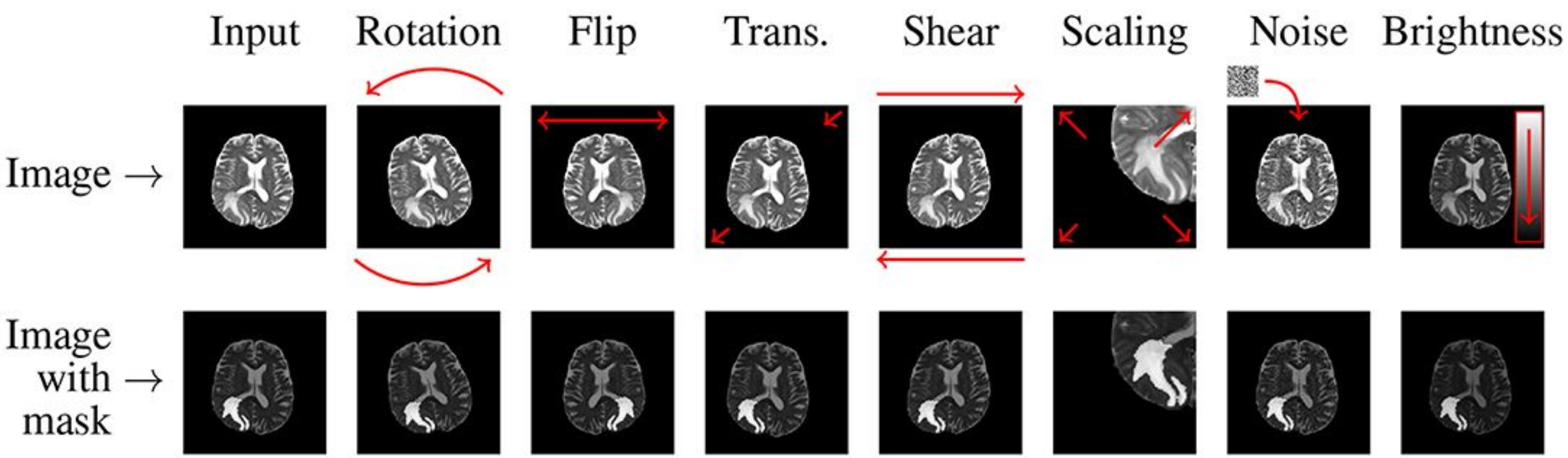
Fig. 1. U-net architecture (example for 32x32 pixels in the lowest resolution). Each blue box corresponds to a multi-channel feature map. The number of channels is denoted on top of the box. The x-y-size is provided at the lower left edge of the box. White boxes represent copied feature maps. The arrows denote the different operations.

[arXiv:1505.04597 [cs.CV]](arXiv:1505.04597 [cs.CV])

Didn’t you say training data is hard to get?

Making your data go further: augmentation

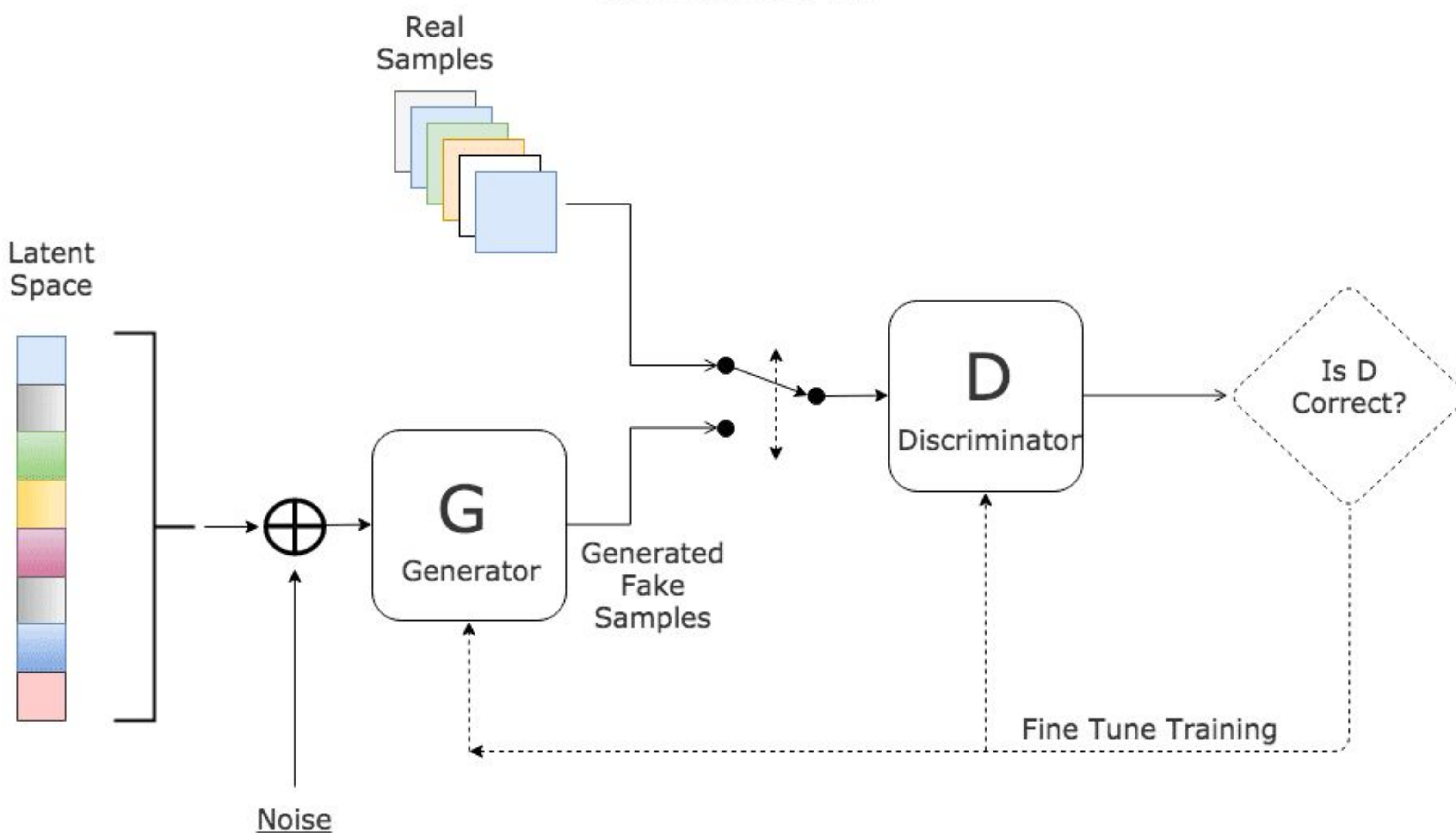
- Apply affine and pixel transformations to your data -> more training samples



(*Aside*: how can data generated from existing images create meaningfully novel signal? Think of augmentation as smoothing the discrete distribution of images OR as applying a prior on image variability)

Feels a bit like manual feature engineering, can we automate this?

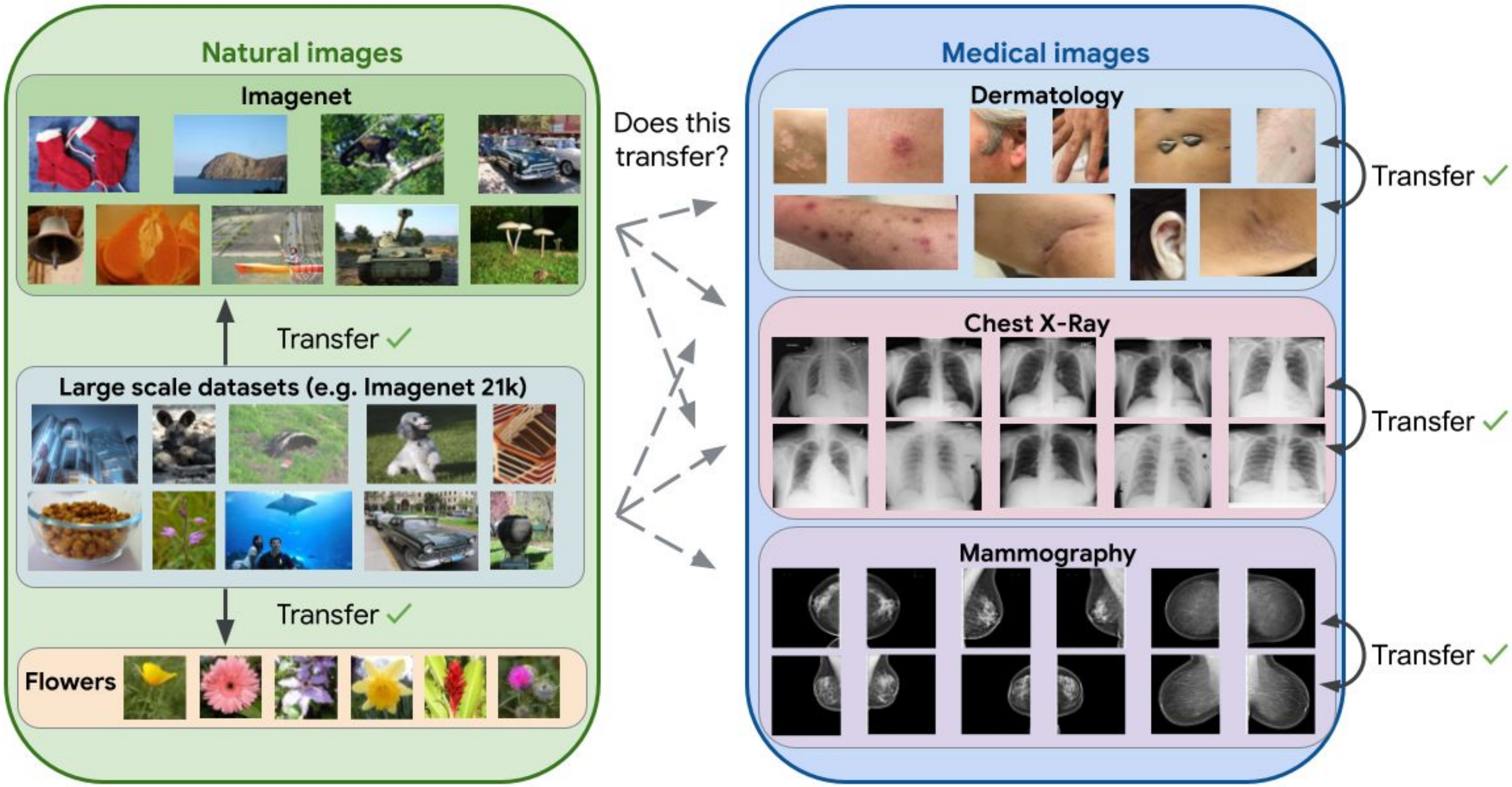
Yes! Generative Adversarial Network



https://www.kdnuggets.com/2017/01/generative-adversarial-networks-hot-topic-machine-learning.html

Can we use different image data then tune?

Transfer learning

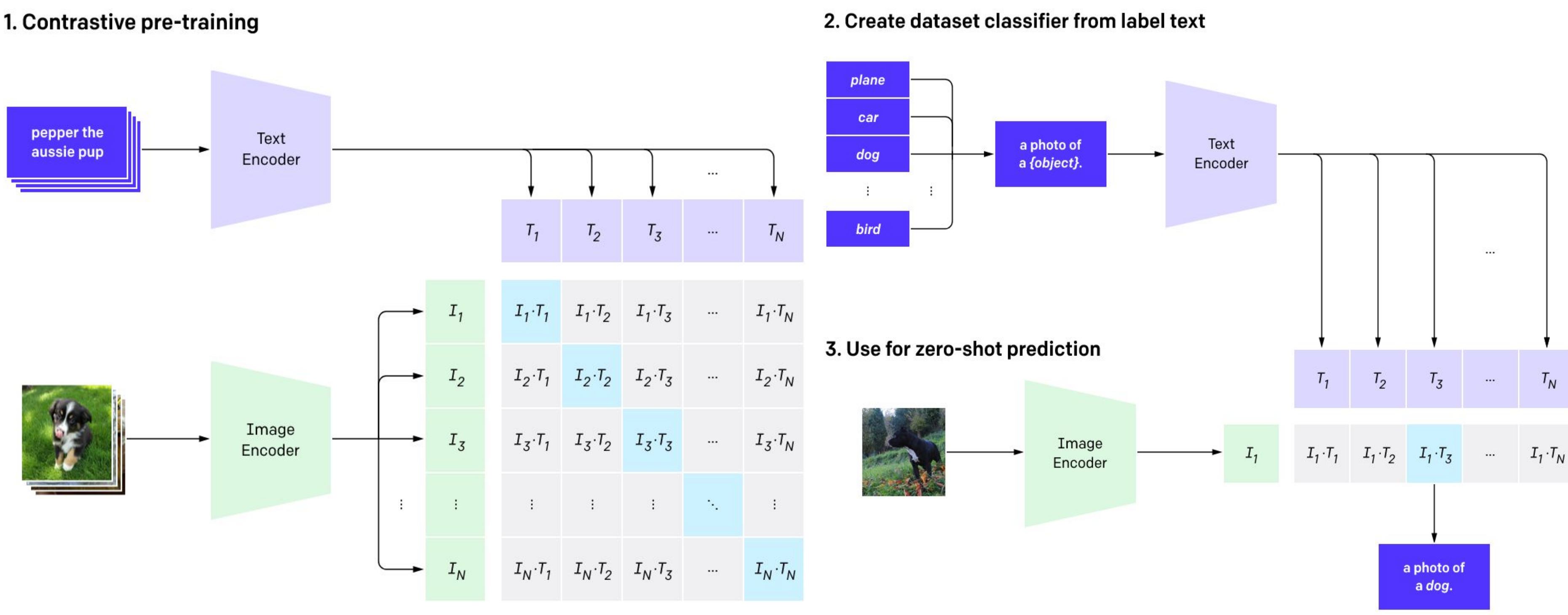


10.2101.05913

Can we make bad labels better?

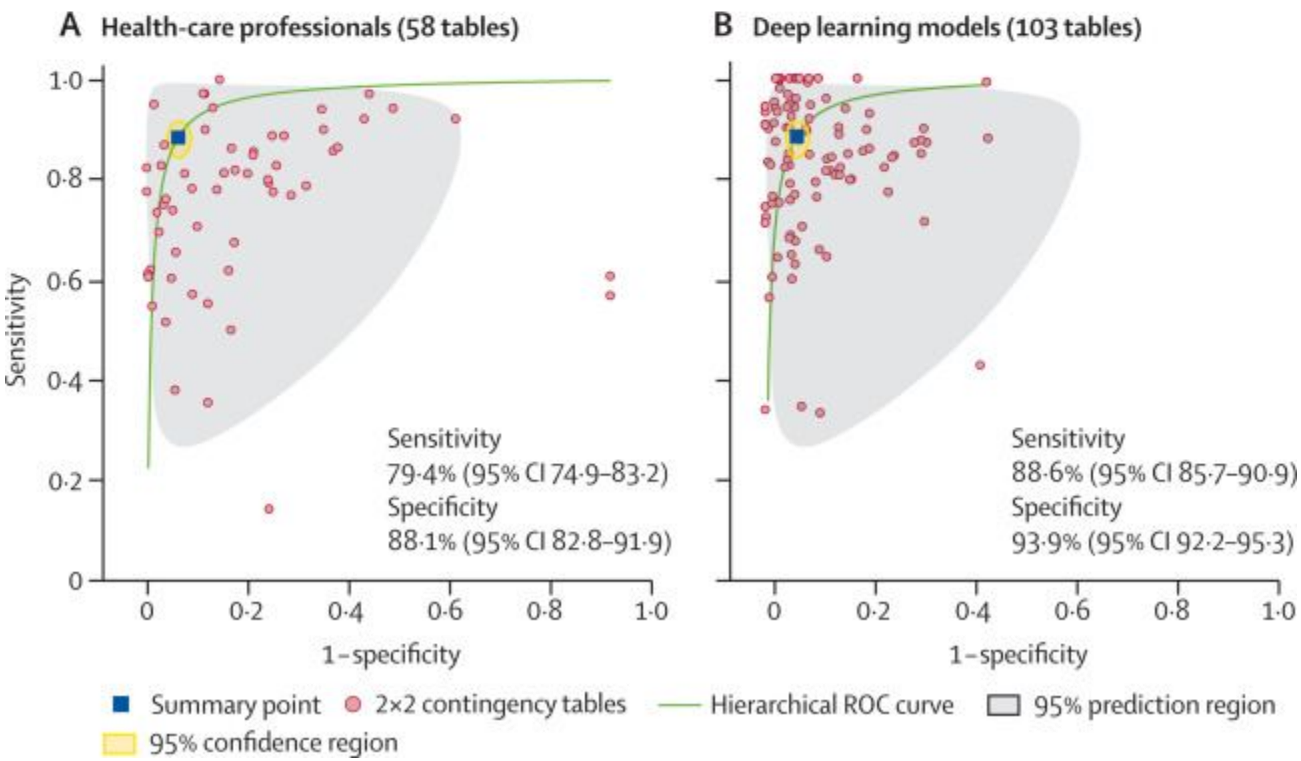
Bad text labels, bad images -> why not embed both?

- OpenAI's Contrastive Language-Image Pre-training (CLIP)
- Use 400 million images trawled from internet (with variable quality labels)
- Initial training: 30 days 592 GPUs -> $1,000,000 equivalent cost



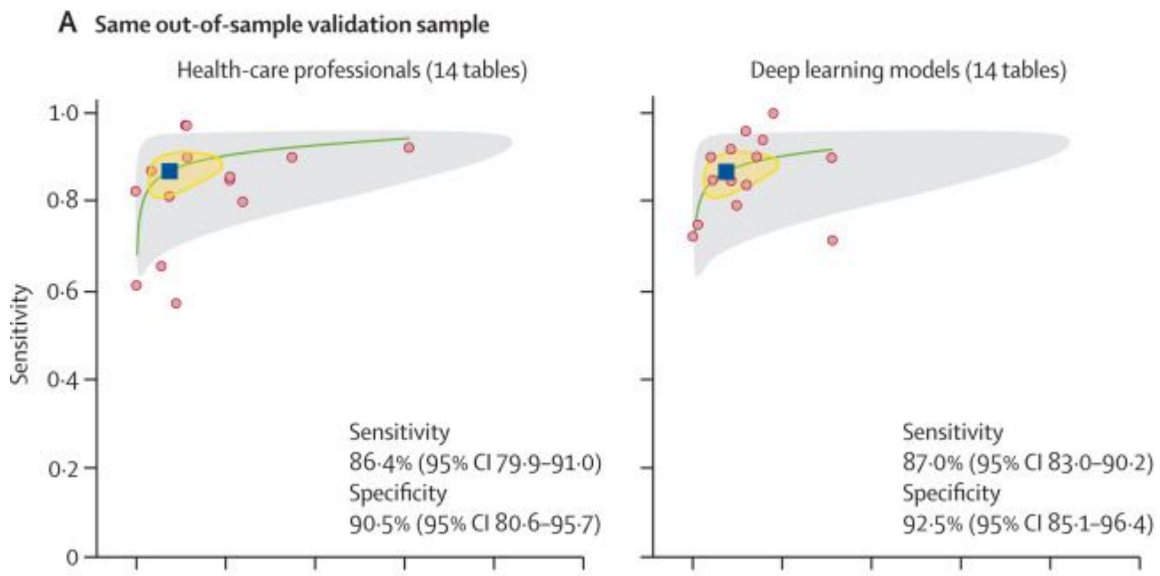
So, can we use these clinically?

Outperforming humans is possible



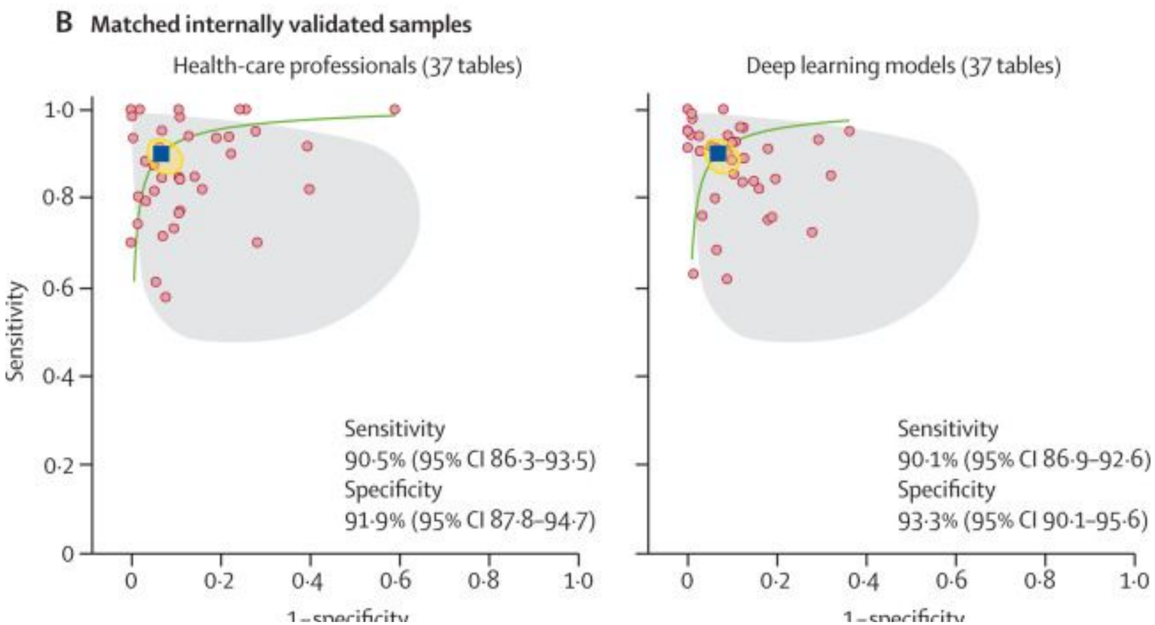
10.1016/S2589-7500(19)30123-2

Outperforming humans is possible



10.1016/S2589-7500(19)30123-2

Outperforming humans is possible

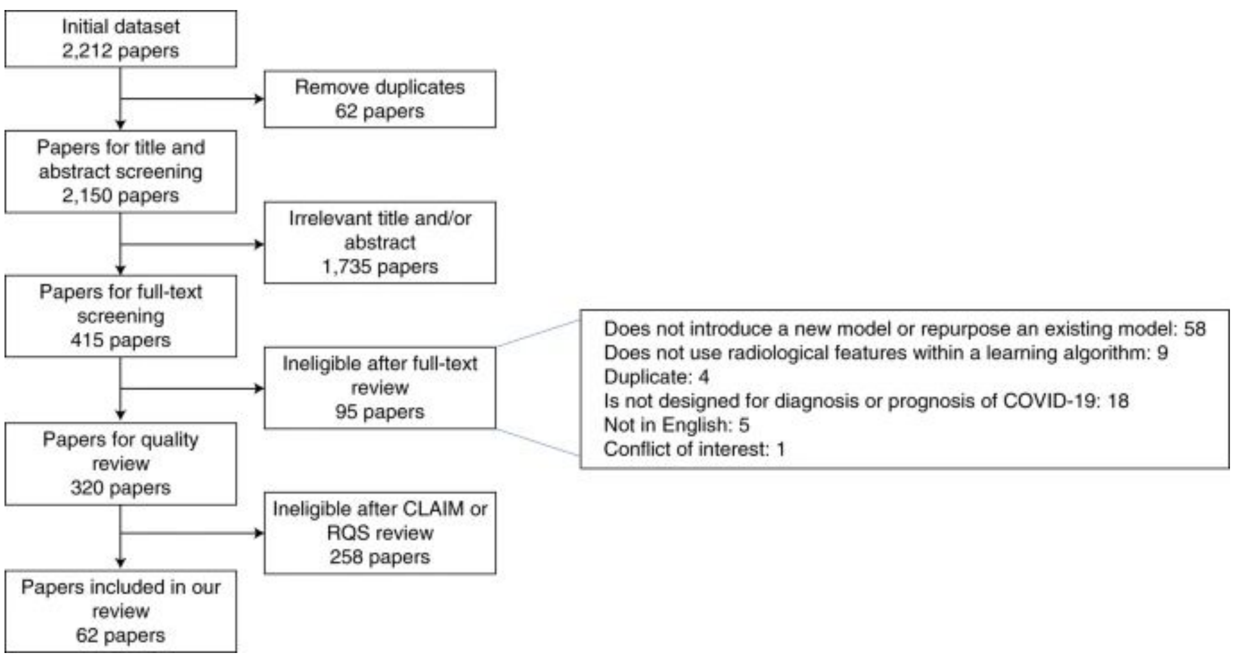


10.1016/S2589-7500(19)30123-2

Great, so why don’t we use these all the time?

COVID-19 Case Study

Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans



https://doi.org/10.1038/s42256-021-00307-0

COVID-19 Case Study

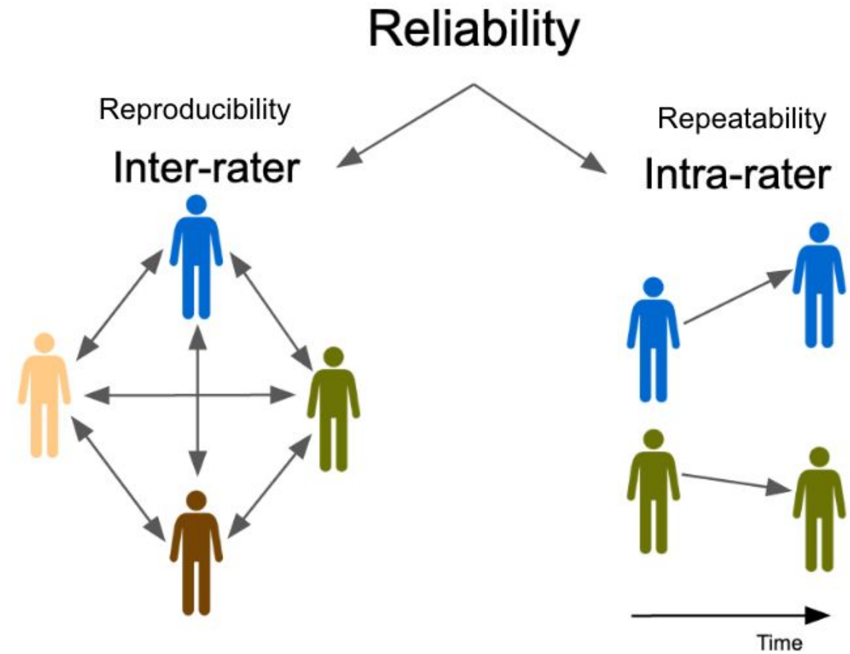
- 254/320 deep-learning papers. 215 failed standard quality checks:
  - 132 (61%) didn't include clear description of how final model was selected
  - 125 (58%) didn't provide documentation of image pre-processing
  - 105 (49%) didn't provide sufficient details of training approach (optimizer, loss function, LR
- 37 passed quality checks:
  - 29 had no external validation
  - 30 had no sensitivity analysis
  - 26 didn't report data demographics
  - 25 didn't assess significance of results (statistics)
  - 23 did not report confidence intervals
  - 22 did not discuss their limitations, biases, generalizability

2,212 papers on simple clinical problem

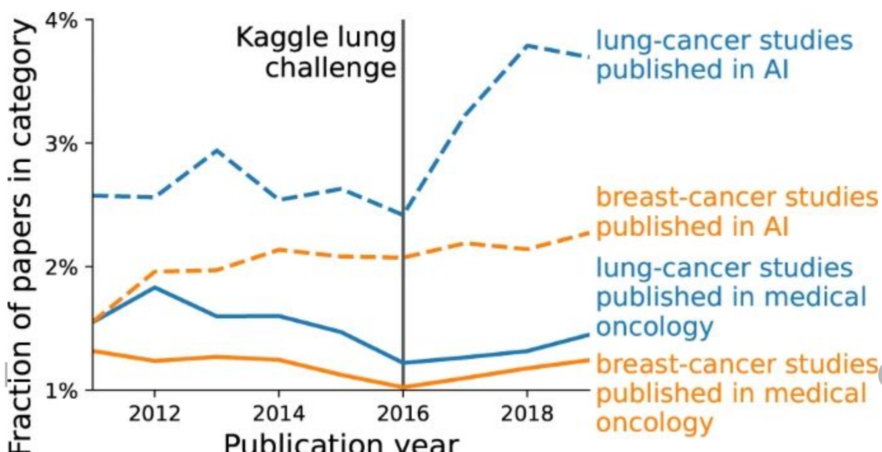
0 evaluated as being usable clinically

Data Challenges

- Reliable and repeatable labelling



- Too much of the wrong data
  - Few clinical questions come as well-posed discrimination tasks
  - Few datasets exist with appropriate data for clinical questions
- Dataset availability distorts research:



https://www.natu es/2

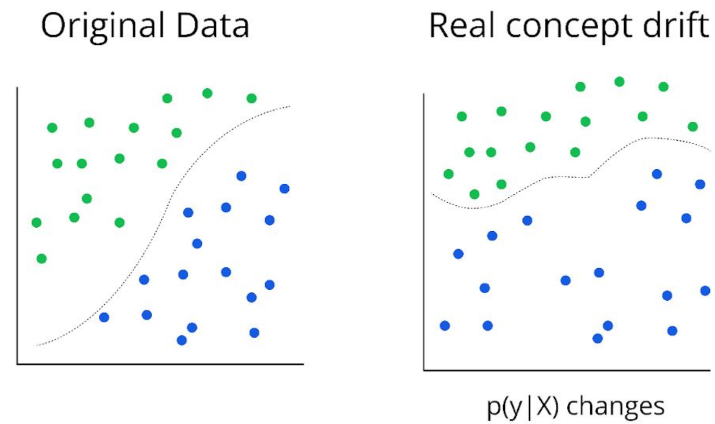
Domain Shift: distributional change from training to deployment

Covariate Shift: images look different but mean the same.

- Imaging equipment (GE vs Siemens)
- Population demographics (Elderly vs Young)
- Imaging Protocol (Settings/Preparations)
- ***Solution***: normalisation & monitor data distributions

Concept Shift: images look similar but mean different things

- Changes in diagnostic criteria
- Changes in treatment
- Changes in disease (e.g., COVID-19 variants)
- ***Solution:*** monitor performance & re-train



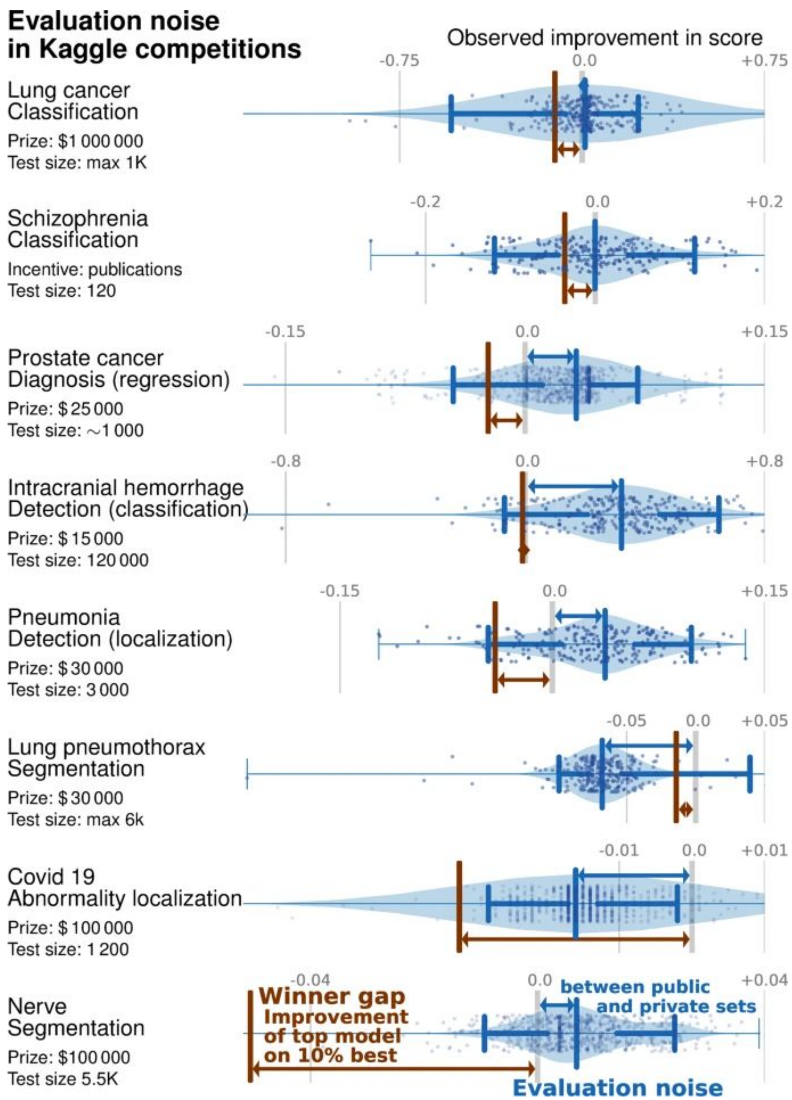
Evaluation Challenges

- **Evaluation error is often larger than algorithmic improvements**

- Evaluation noise*:* public test-set vs private test-set performance on kaggle.

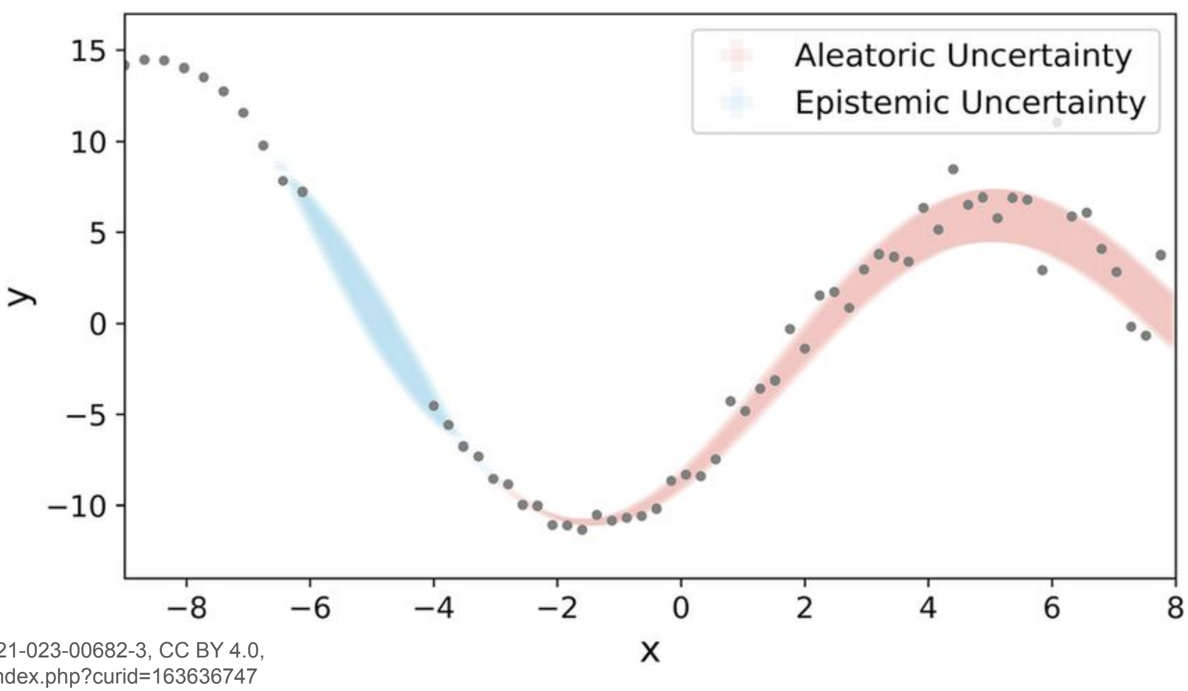
- Positive = public better than private (overfitting)
- Negative = private better than public

- Brown bar = improvement between best model and 10th percentile model



Incorporation of uncertainty

- High accuracy (or sensitive, specificity,...) does mean accurate estimation of uncertainty
- Epistemic Uncertainty: Var(E(ŷ))
- Aleatoric Uncertainty: E(Var(ŷ)
- Evidential Deep Learning: predict parameters of distribution
- Conformal Prediction: predict set of compatible labels up to certainty level
- Bayesian Surrogates: approximate posterior distribution using simpler models



Legal hurdles are daunting but important

- Commercial software for medical images is a medical device
- Regulated as such.
- Regulatory frameworks actively changing/developing
- Requires explanatory power (still a work-in-progress for deep learning)
- Requires acceptance by clinicians
- Can have unpredictable failure modes!

Learning Overview

- Many types of medical imaging data and their respective formats
- DICOM file format is ubiquitous but complex
- Medical imaging data analysis involves many different questions
- Lots of different machine learning paradigms are used to handle challenges of medical image data
- Traditional computer vision approaches
- Deep learning enables learning features/representations
- Convolutions key to capturing spatial relationships
- Augmentation and generative models enable better training with limited data
- Transfer learning and joint-transformer models hugely expand training options
- Machine learning in medical image analysis is very promising but has several major hurdles to broad acceptance