Co-financed by the Connecting Europe Facility of the European Union

eXplainable Artificial Intelligence in healthcare Management
2020-EU-IA-0098

UNIVERSITÀ DEGLI STUDI DI BRESCIA

Realizing Trustworthy AI solutions for diagnosis and prognosis support

Alberto Signoroni, PhD -- University of Brescia
Associate Professor of Computer Science (ING-INF/05)
Department of Medical and Surgical Specialties, Radiological Sciences, and Public Health (DSMC)
alberto.signoroni@unibs.it

April-May 2025

bio

copyright & disclaimer

Affiliations at University of Brescia

- Department of Medical and Surgical Specialties, Radiological Sciences, and Public Health (DSMC)
- MediMint : Medical Imaging and Multimodal Intelligence Laboratory
- TRAIL : Trustworthy AI Laboratory (multidepartmental)

Education

- 1997 MSc Electronics Engineering, University of Brescia
- 2001 PhD Information Engineering, University of Brescia
- 2019 Master in Management of Research, Innovation and Technology, Polytechnic of Milan

Main Courses taught at University of Brescia

- Image Data Analysis
- Law and Regulations for ICT
- Advanced Methods for Information Representation
- Artificial intelligence to aid medical diagnosis

Research interests

- Medical/Industrial image analysis
- Computer Vision/Graphics
- Machine/Deep Learning
- Signal/Image processing

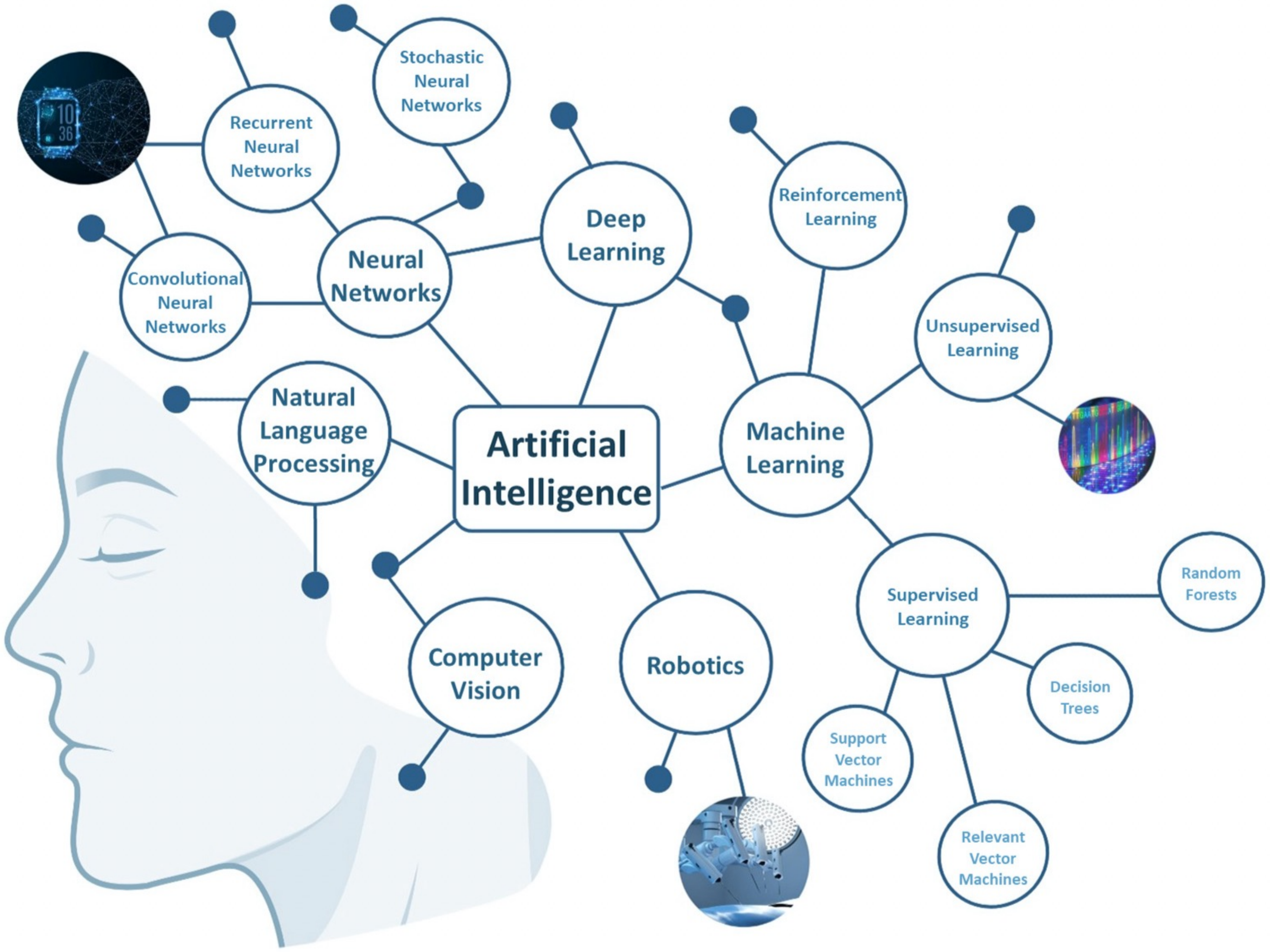
Publications (Google Scholar page)

Course Outline: AI systems in Health and Medicine

- **THE BASES**: definitions, lexicon, taxonomy, core aspects
- **THE WORKFLOW**: the key steps of the AI pipeline
- **THE DATA CURATION**: how to collect, prepare and feed data to AI systems
- **THE EXPERIMENTAL PHASES**: how to lead or participate to the development and testing of new or existing AI-based solutions
- **THE DEPLOYMENT**: ethical and regulatory aspects for reliable, explainable, trustworthy AI in health and medicine

The course in 5 pictures

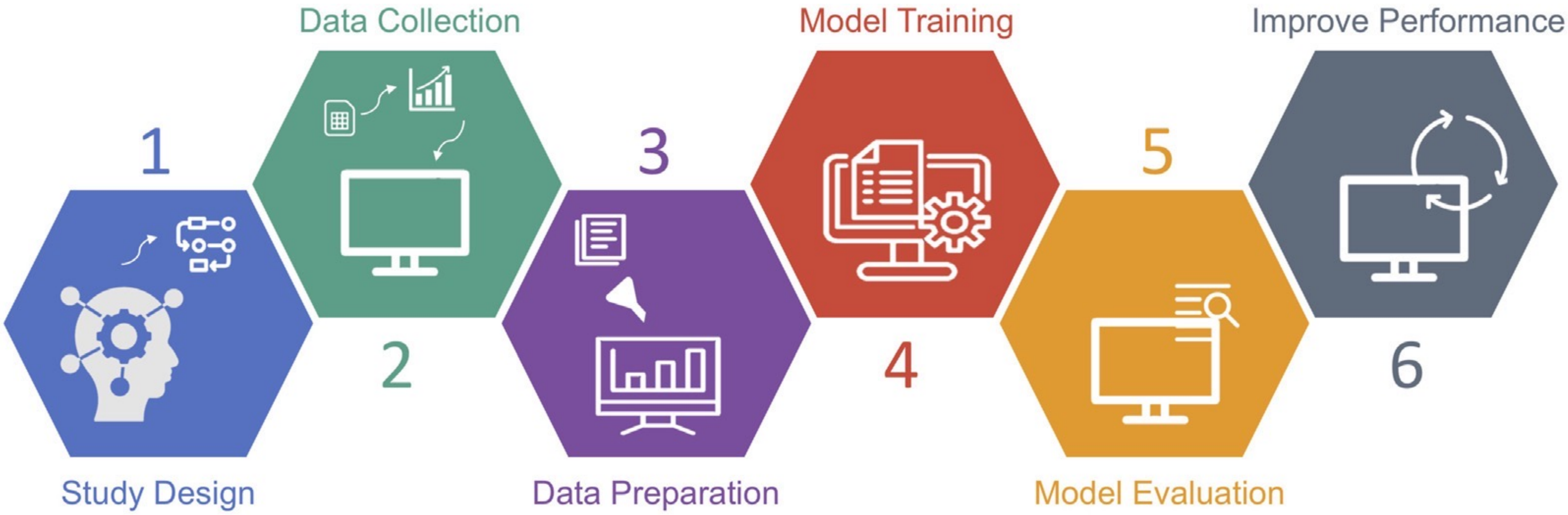
THE BASES: definitions, lexicon, taxonomy, key theory



https://doi.org/10.3390/biomedinformatics2040049

The course in 5 pictures

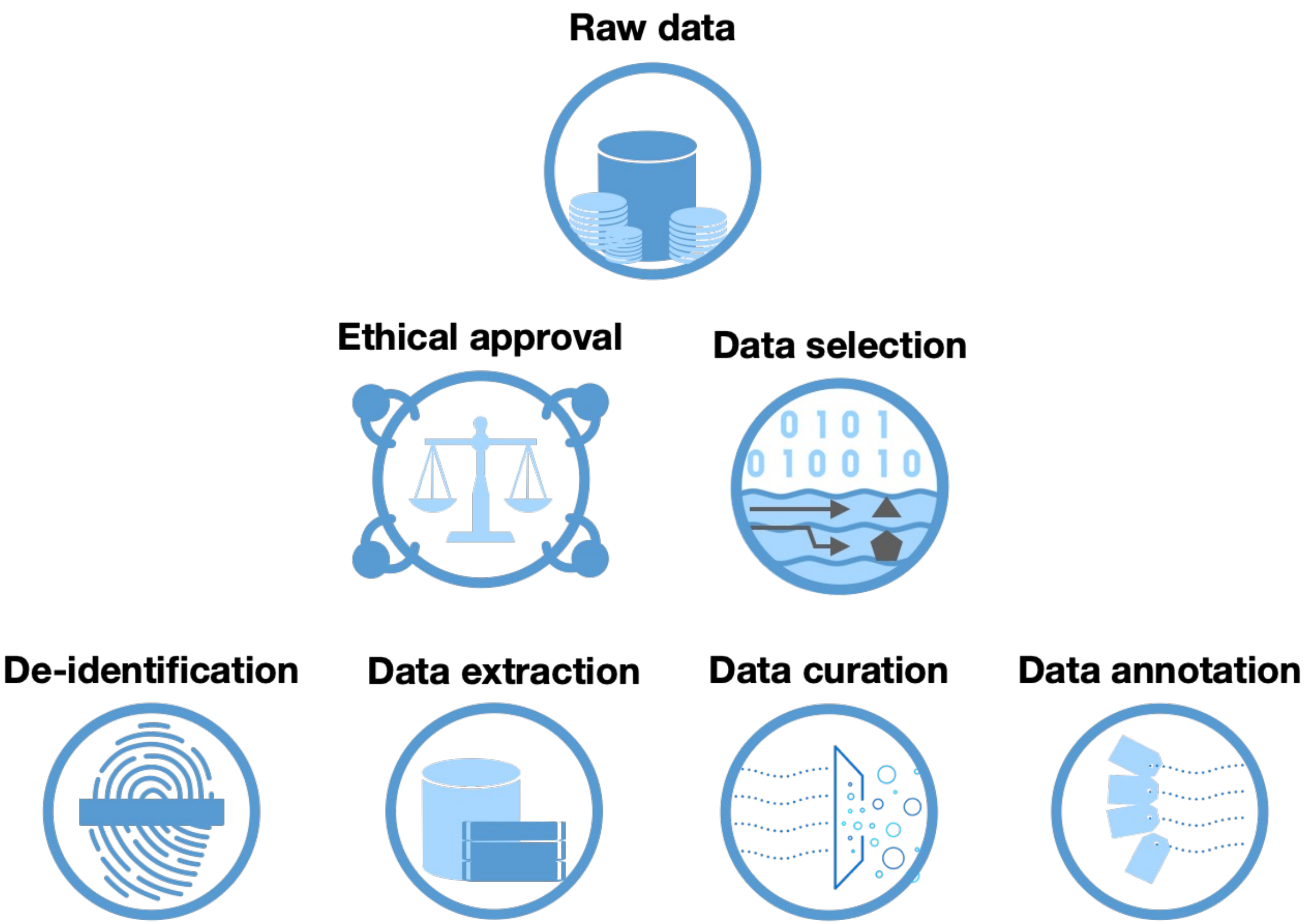
THE WORKFLOW: the key steps of the AI pipeline



https://doi.org/10.3389/fphar.2021.720694

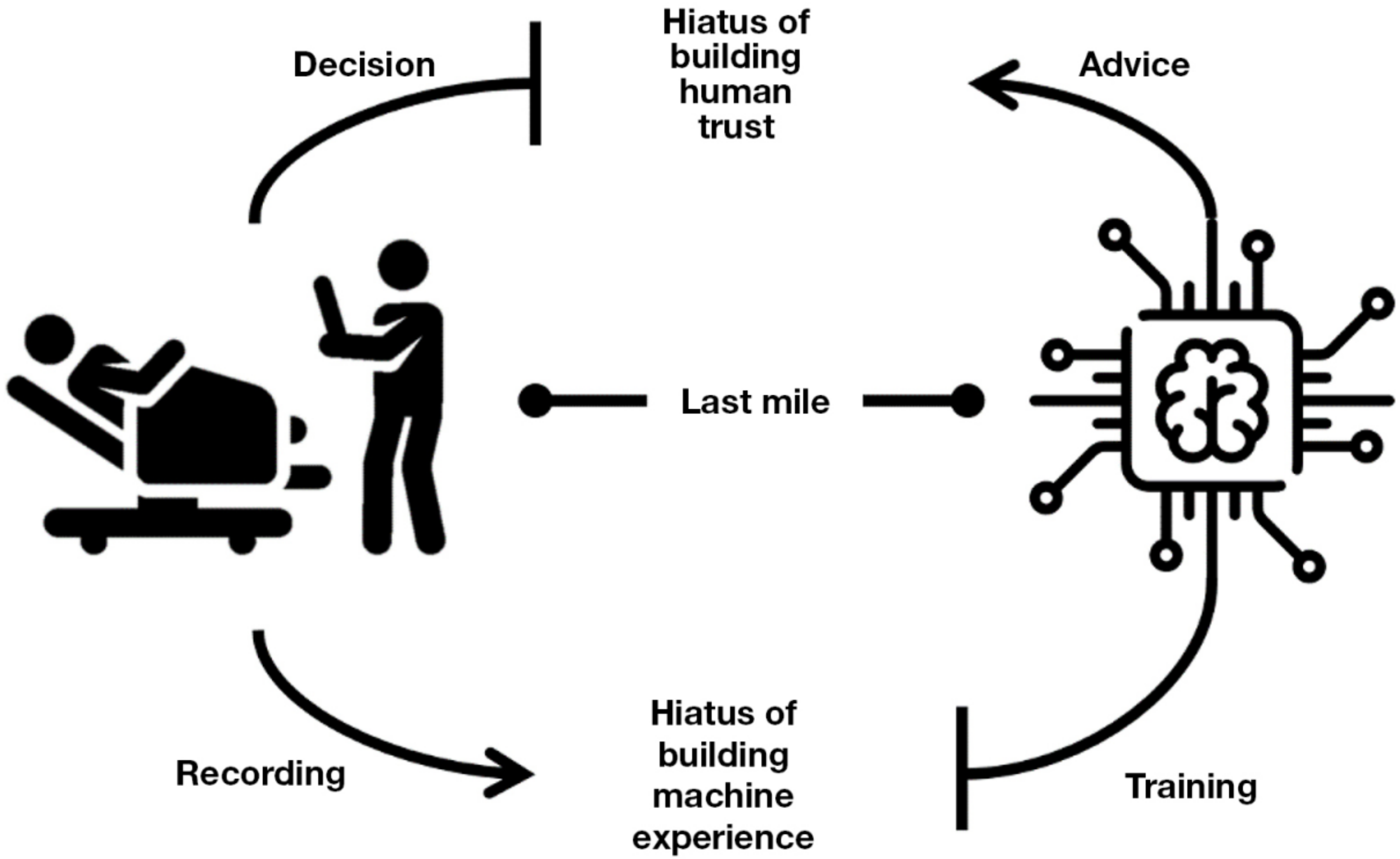
The course in 5 pictures

THE DATA CURATION: how to collect, prepare and feed data to AI systems



The course in 5 pictures

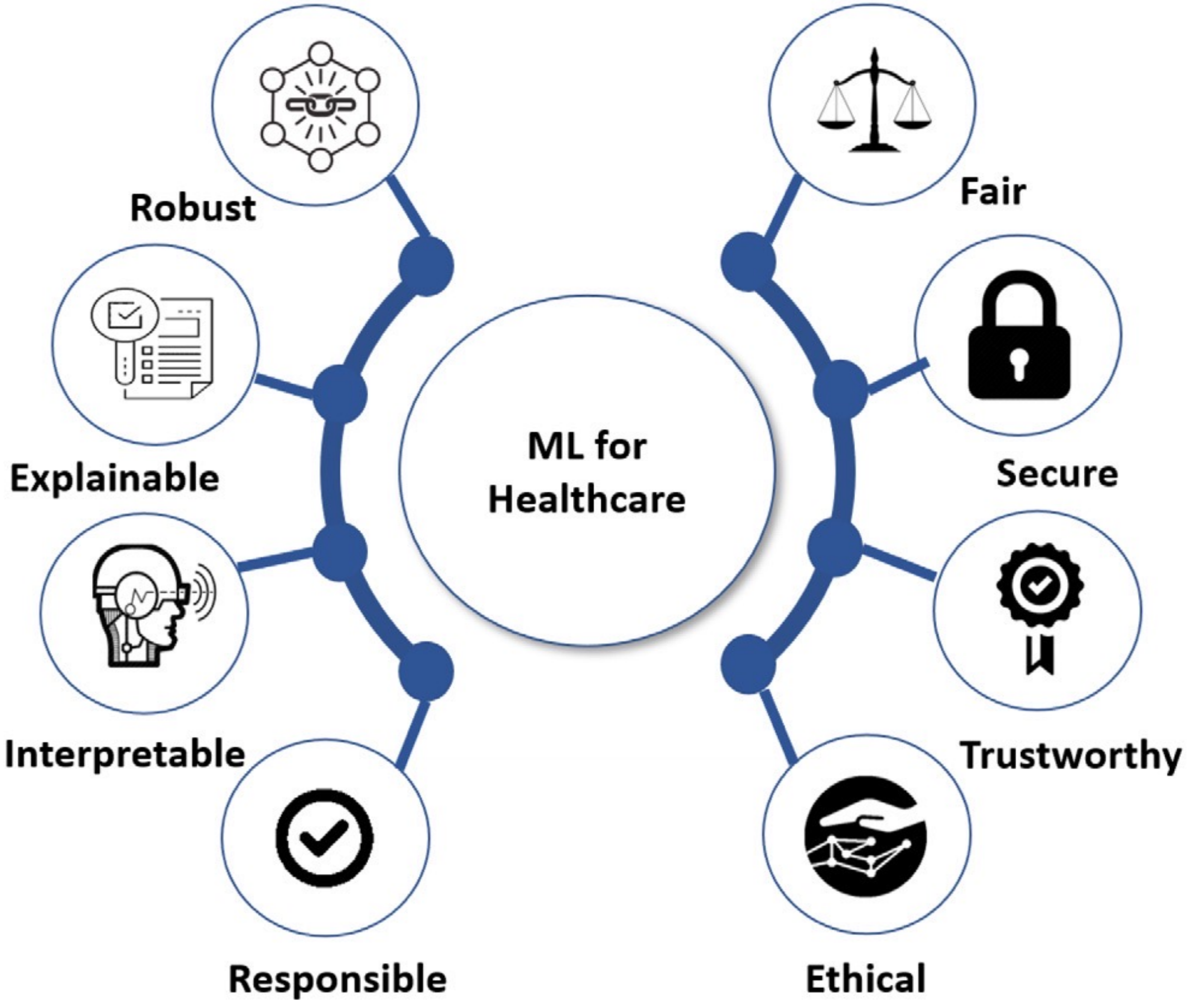
THE EXPERIMENTAL PHASES: how to lead or participate to the development and testing of new or existing AI-based solutions



http://dx.doi.org/10.21037/atm.2020.03.63

The course in 5 pictures

THE DEPLOYMENT: ethical and regulatory aspects for reliable, explainable, trustworthy AI in health and medicine



https://doi.org/10.1016/j.compbiomed.2022.106043

Course Outline: AI systems in Health and Medicine

THE BASES: definitions, lexicon, taxonomy, core aspects

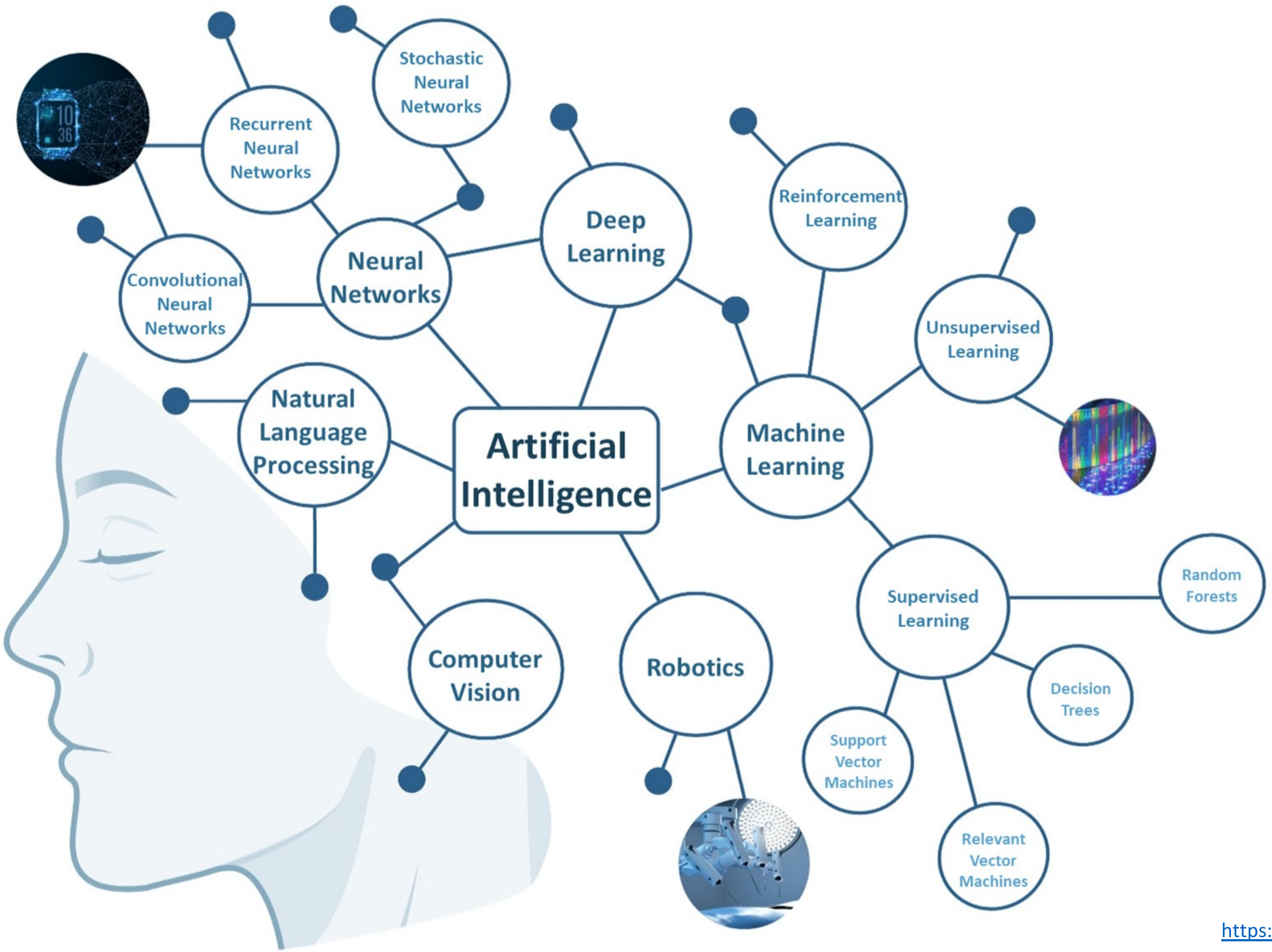
THE WORKFLOW: the key steps of the AI pipeline

THE DATA CURATION: how to collect, prepare and feed data to AI systems

THE EXPERIMENTAL PHASES: how to lead or participate to the development and testing of new or existing AI-based solutions

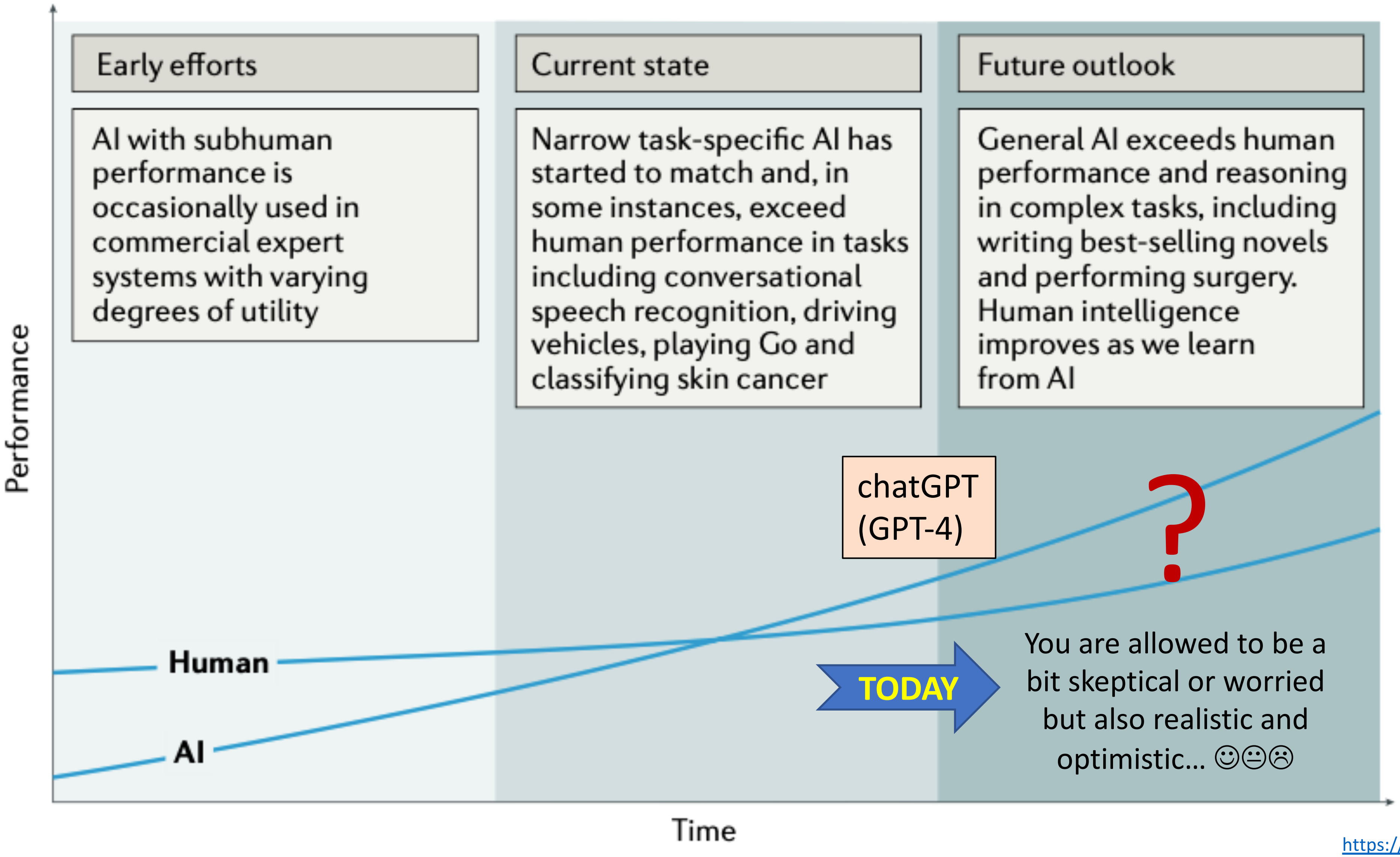
THE DEPLOYMENT: ethical and regulatory aspects for reliable, explainable, trustworthy AI in health and medicine

First: we need a shared language



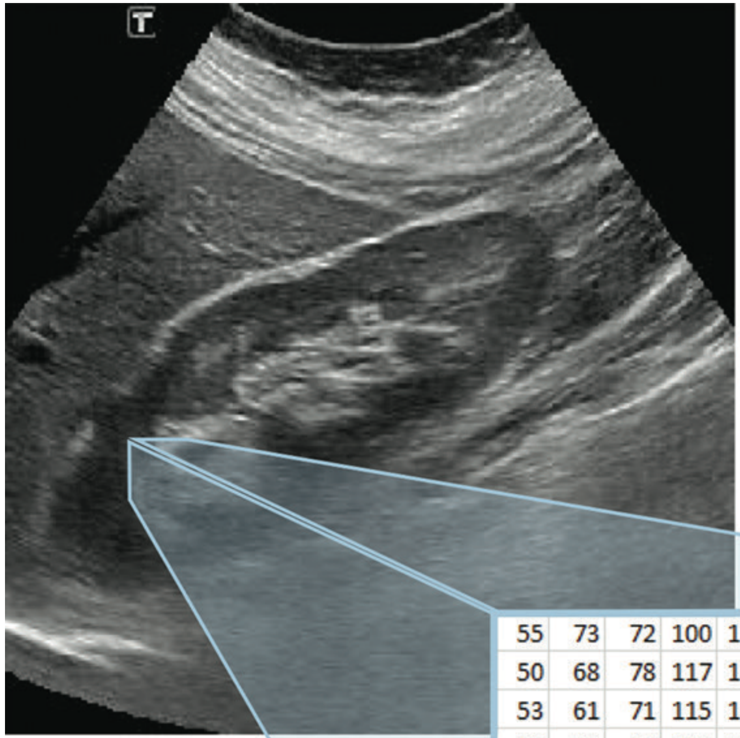
https://doi.org/10.3390/biomedinformatics2040049

Artificial Intelligence vs Human intelligence



https://doi.org/10.1038/s41568-018-0016-5

Computer Vision vs Human Vision



Humans
do not see numbers

55	73	72	100	115	137	115	121	134	145	147	143
50	68	78	117	134	149	117	123	134	142	138	127
53	61	71	115	128	133	139	141	146	146	136	119
61	70	84	138	156	162	145	140	133	121	106	92
62	68	77	127	148	158	144	136	118	95	76	66
54	57	52	86	101	113	134	131	118	94	72	60
52	62	56	80	89	103	103	116	122	109	87	68
58	65	51	61	59	69	93	120	144	143	123	99
67	81	88	84	81	83	73	82	95	109	119	117
73	88	100	105	109	112	75	89	104	112	113	110
67	69	72	78	79	77	102	95	93	108	134	146
76	72	75	85	89	84	82	95	107	114	120	130
74	69	72	81	83	78	92	79	69	81	109	127

Computers
only see numbers

A human expert easily classifies this image as an image of the right kidney.

Why is (**or was**) this task difficult for a computer?

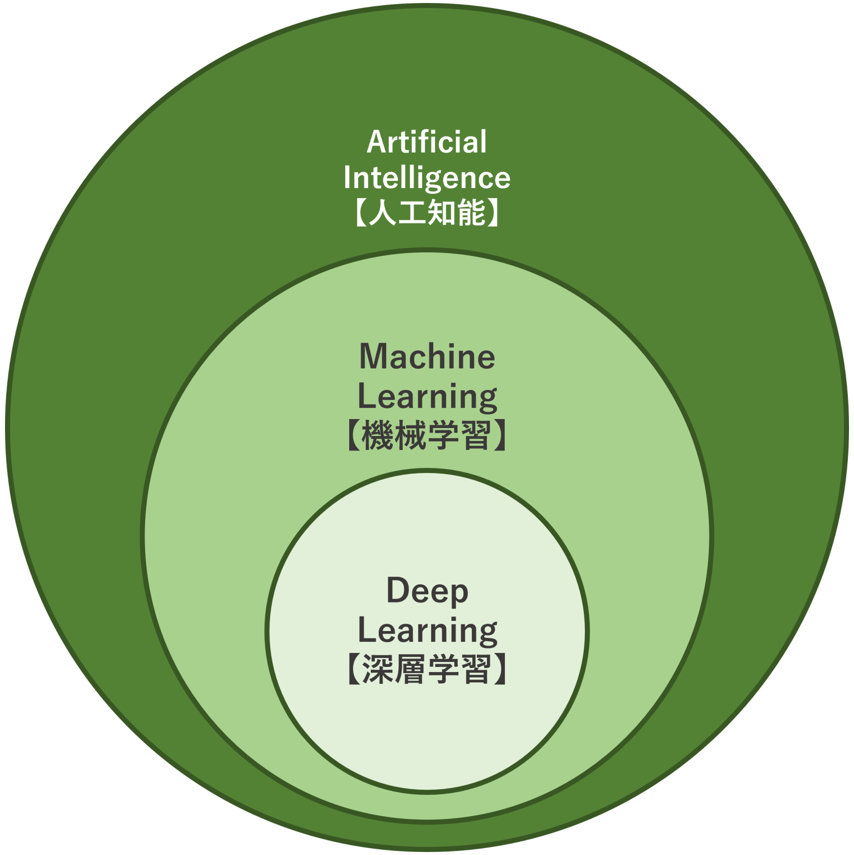
Instead of shades of gray, a computer "sees" a matrix of numbers representing pixel brightness.

Computer vision typically involves

1) computing the presence of **numerical patterns** (called ***features***) in this matrix,
2) applying **model-based** or **machine learning** algorithms to analyze images (local or global understanding) on the basis of these features.

https://doi.org/10.1148/rg.2017170077

What is AI?
main definitions



Artificial Intelligence (AI) *computer systems perform tasks that ordinarily require human intelligence.*

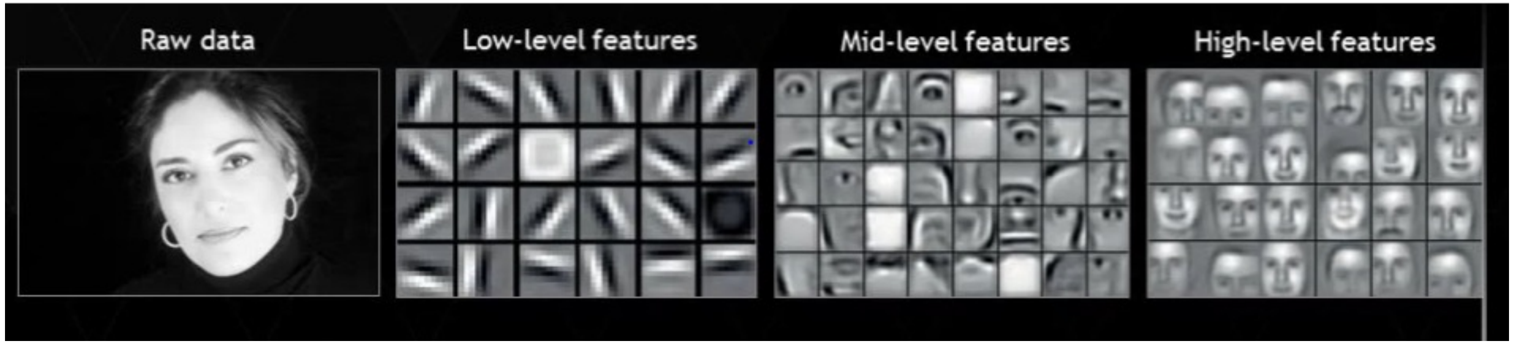
Machine Learning (ML) *subfield of AI where algorithms can learn patterns from data, trough predetermined data **features***

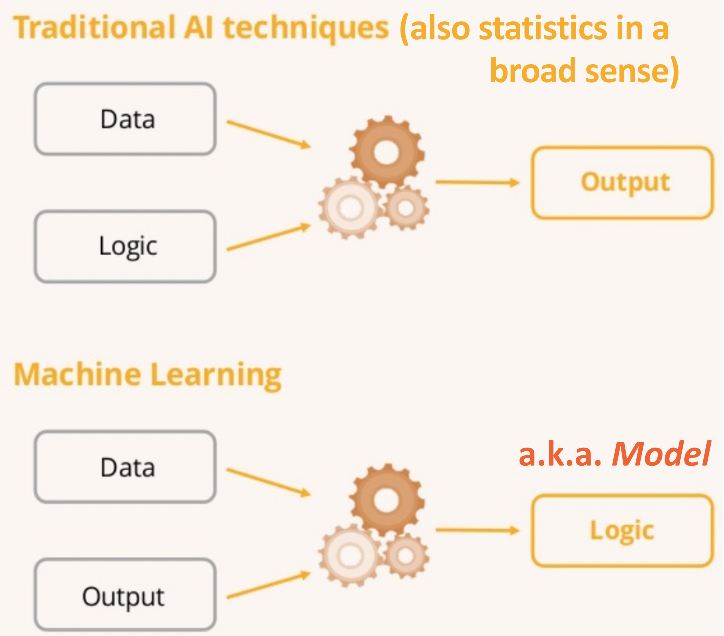
Representation Learning (RL) *a type of ML in which no feature engineering is used, but the algorithm learns on its own the best features to interpret data*

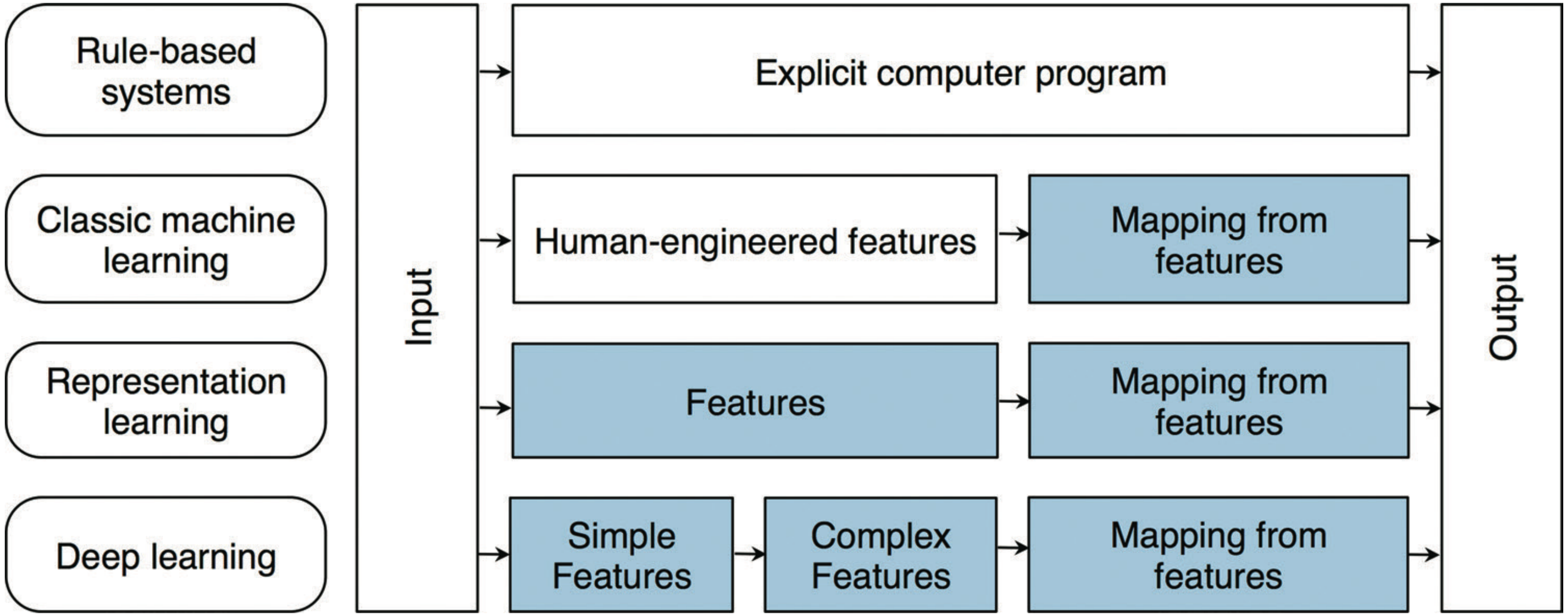
Deep Learning (DL) *a type of RL in which the algorithm learns a composition of features that reflect a hierarchy in the data interpretation structure.*

Foundation Models (FM) *large scale DL models trained on vast amount of data to serve as a base (foundation) of multiple applications*

Main definitions:
The role of data and features in AI



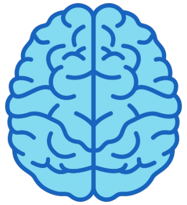


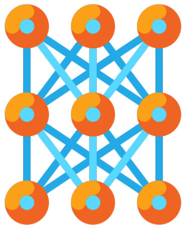


- **Classic ML** depends on carefully designed features, requiring human expertise and complicated task-specific optimization.
- **DL systems** propose an end-to-end approach by learning simple features (such as signal intensity, edges, and textures) as components of more complex features such as shapes, lesions, or organs, therefore leveraging the compositional nature of images

https://doi.org/10.1148/rg.2017170077

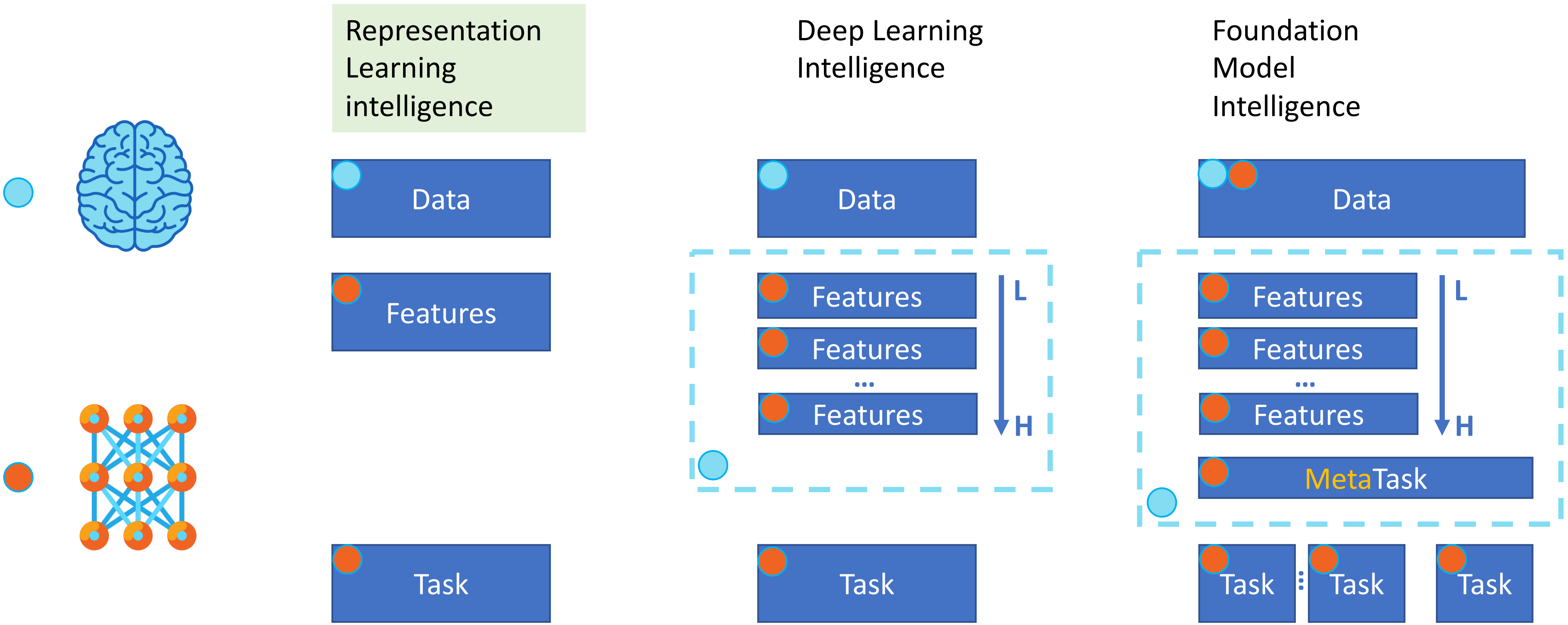
From rule-based to representation learning intelligence





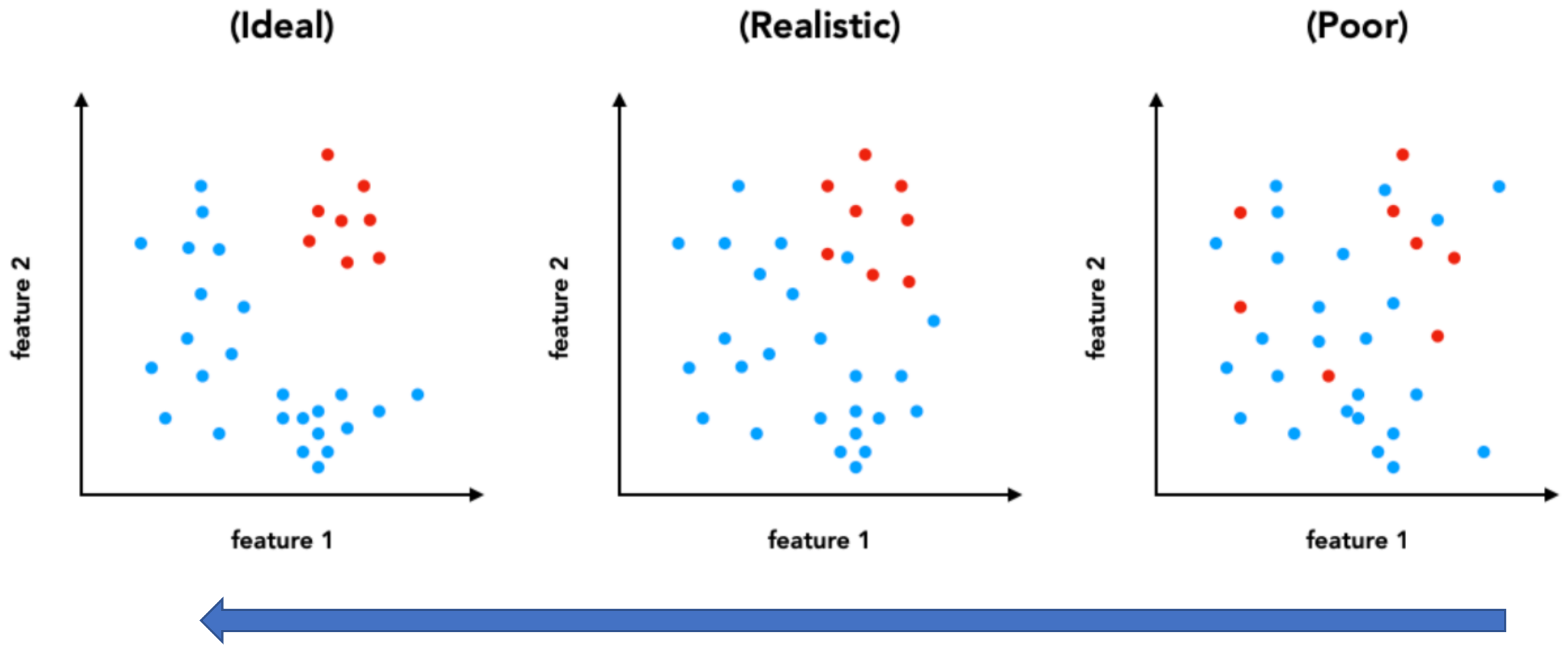
Programmed Intelligence (rule-based systems)	Machine Learning Intelligence	Representation Learning Intelligence
Data	Data	Data
Features	Features	Features
Task	Task	Task

From representation learning to foundation models



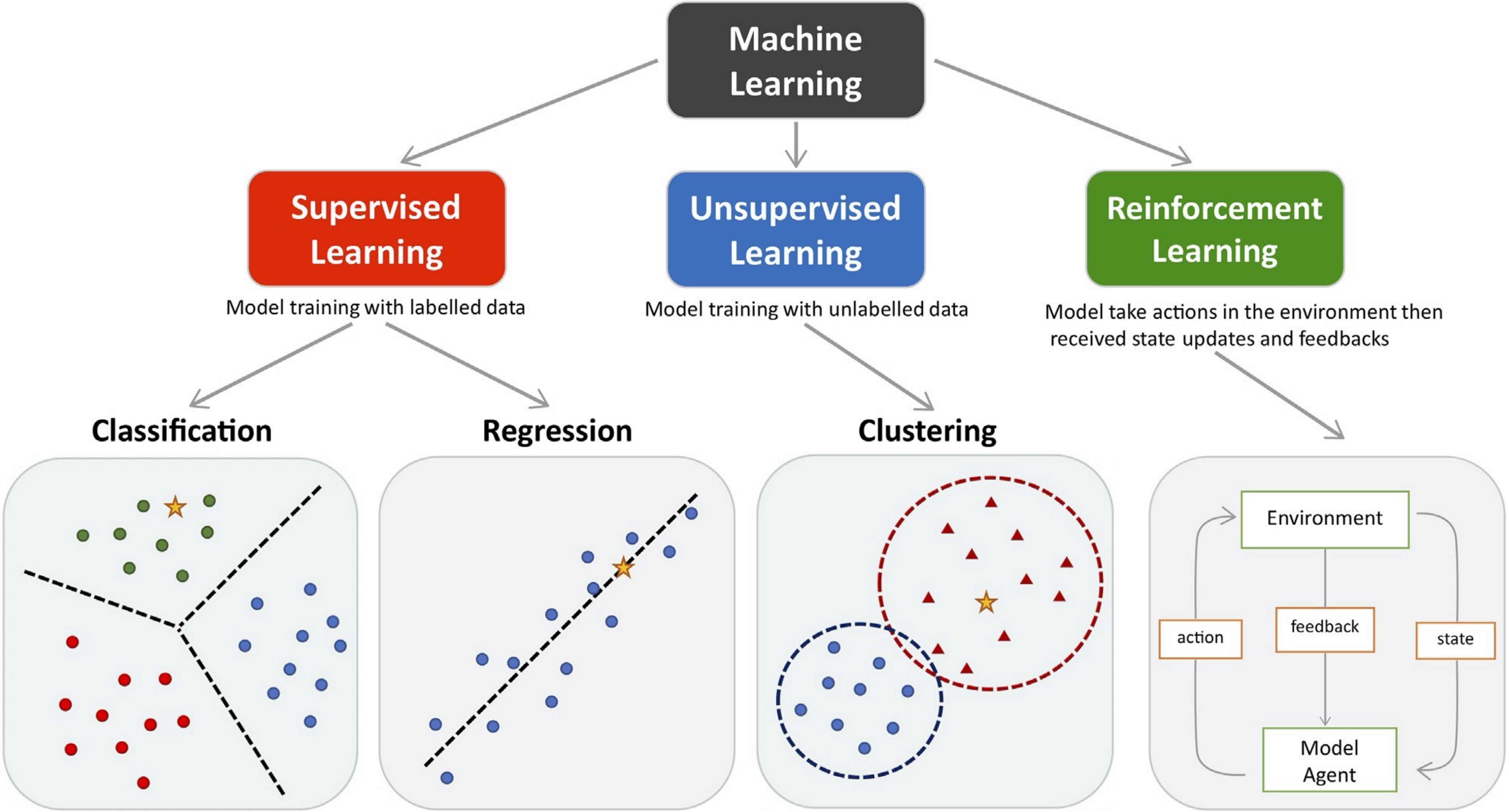
Pattern (or feature) space

The pattern space dimensions are data channels/components or features extracted from data. Which features?



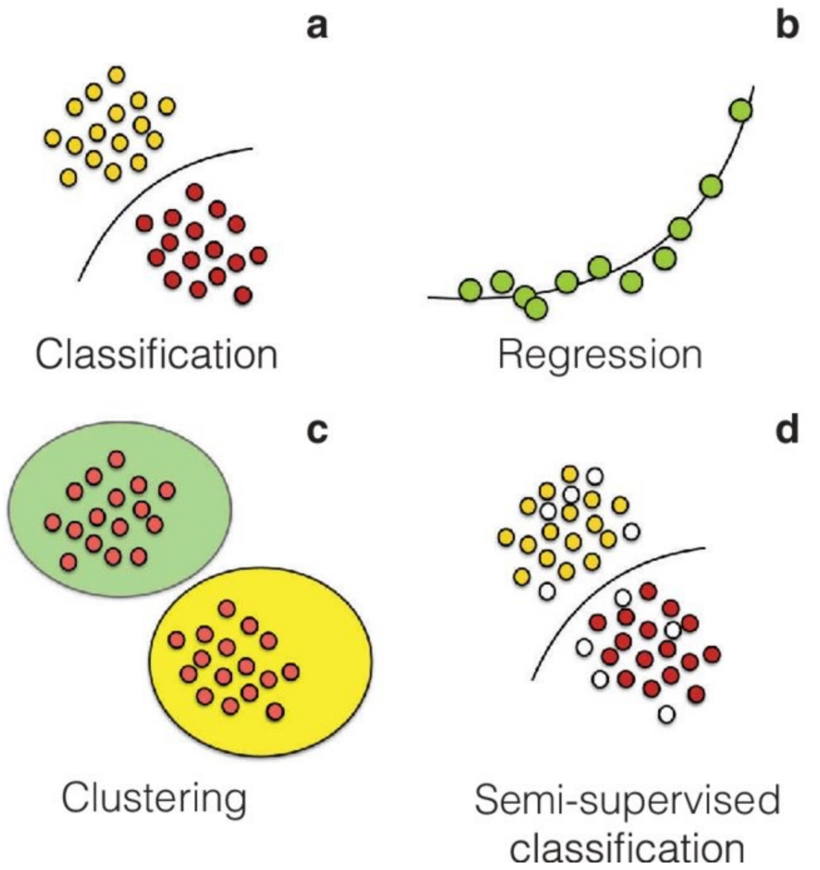
With **representation learning** the expectations is to learn good discriminative features for a given task in order to guarantee easy and accurate enough discriminations

Machine learning: kind of

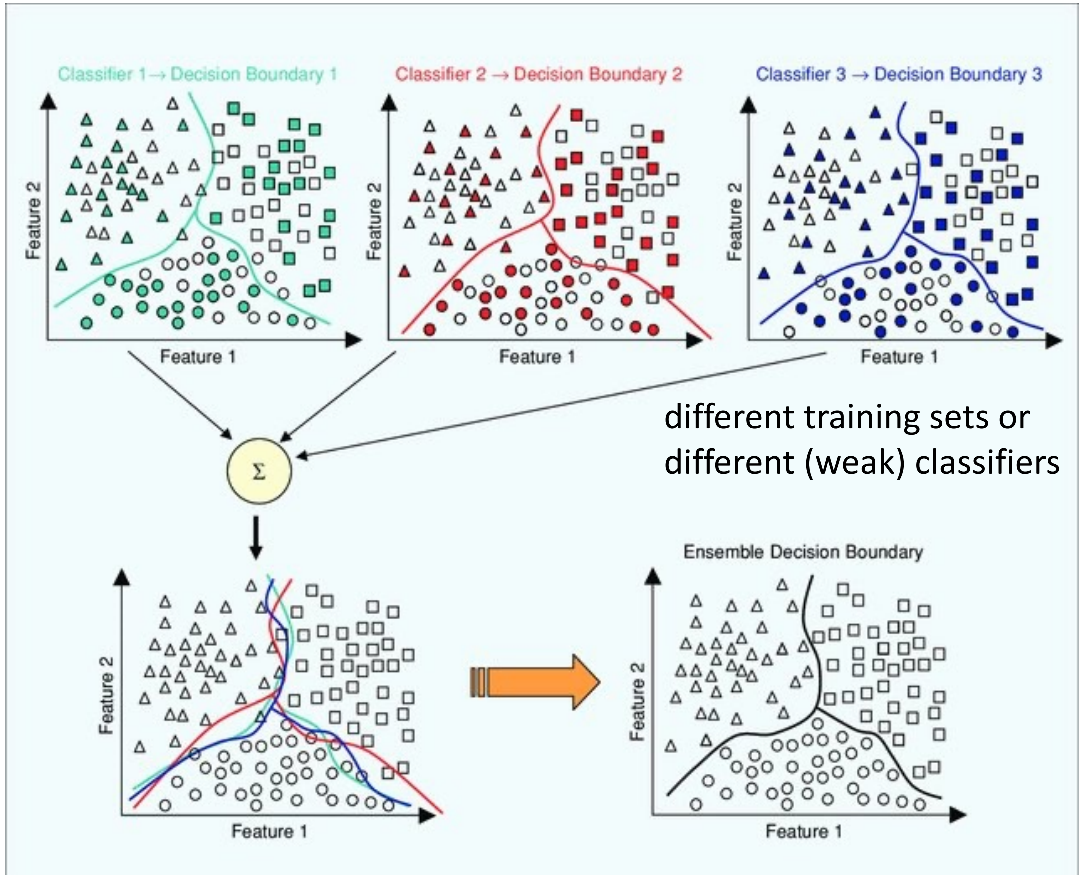


https://doi.org/10.3389/fphar.2021.720694

Machine learning: kind of



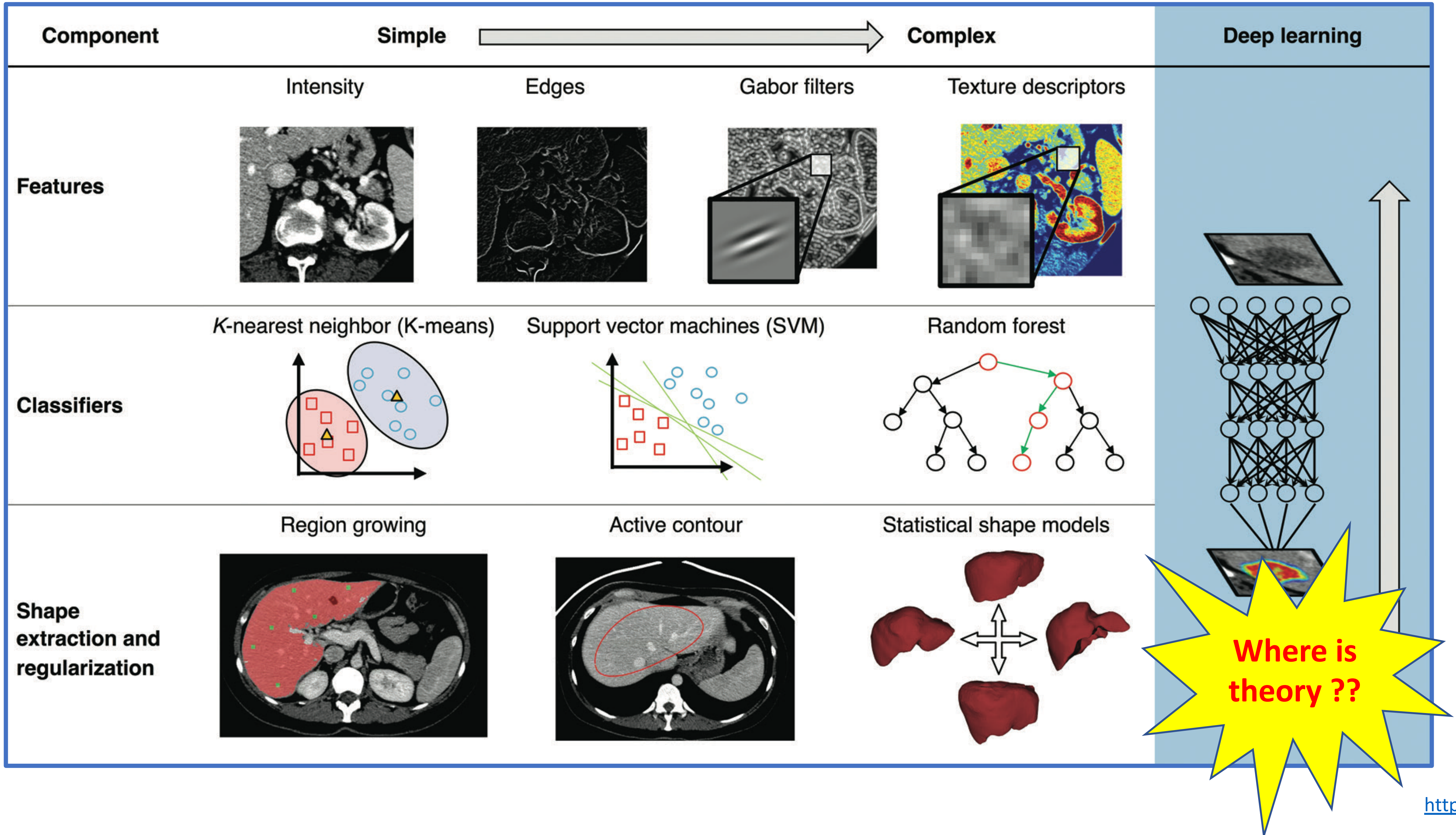
Fitting data according to the kind of data/knowledge



Ensemble learning

http://dx.doi.org/10.1109/MCAS.2006.1688199

Machine learning: methodological notes



Features describe the appearance of organs and points of interest in medical images.
Classifiers integrate features to output a decision (eg. pixel-wise).
Shape extraction and regularization recover a consistent shape despite classification noise.
Deep Learning proposes an end-to-end approach where features are learned to maximize the classifier's performance.
Shape extraction can become implicit (regularized pixel-wise info easy to obtain).

https://doi.org/10.1148/rg.2017170077

ML (learning from data to interpret new ones) vs Statistics (fitting to models to explain given data)

VS

	Machine learning	Statistics
1	Machine learning is a subset of artificial intelligence.	Statistics is a field of mathematics that studies data through various techniques.
2	Predicting accurate outcomes is the strength of machine learning algorithms.	The statistical models are intended for interference about the connections between the variables.
3	The models in machine learning are designed to conclude the most accurate predictions possible.	Many statistical models make predictions, but they are not accurate enough.
4	Machine learning is all about outcomes.	Statistics is all about finding relationships between variables and their significance.

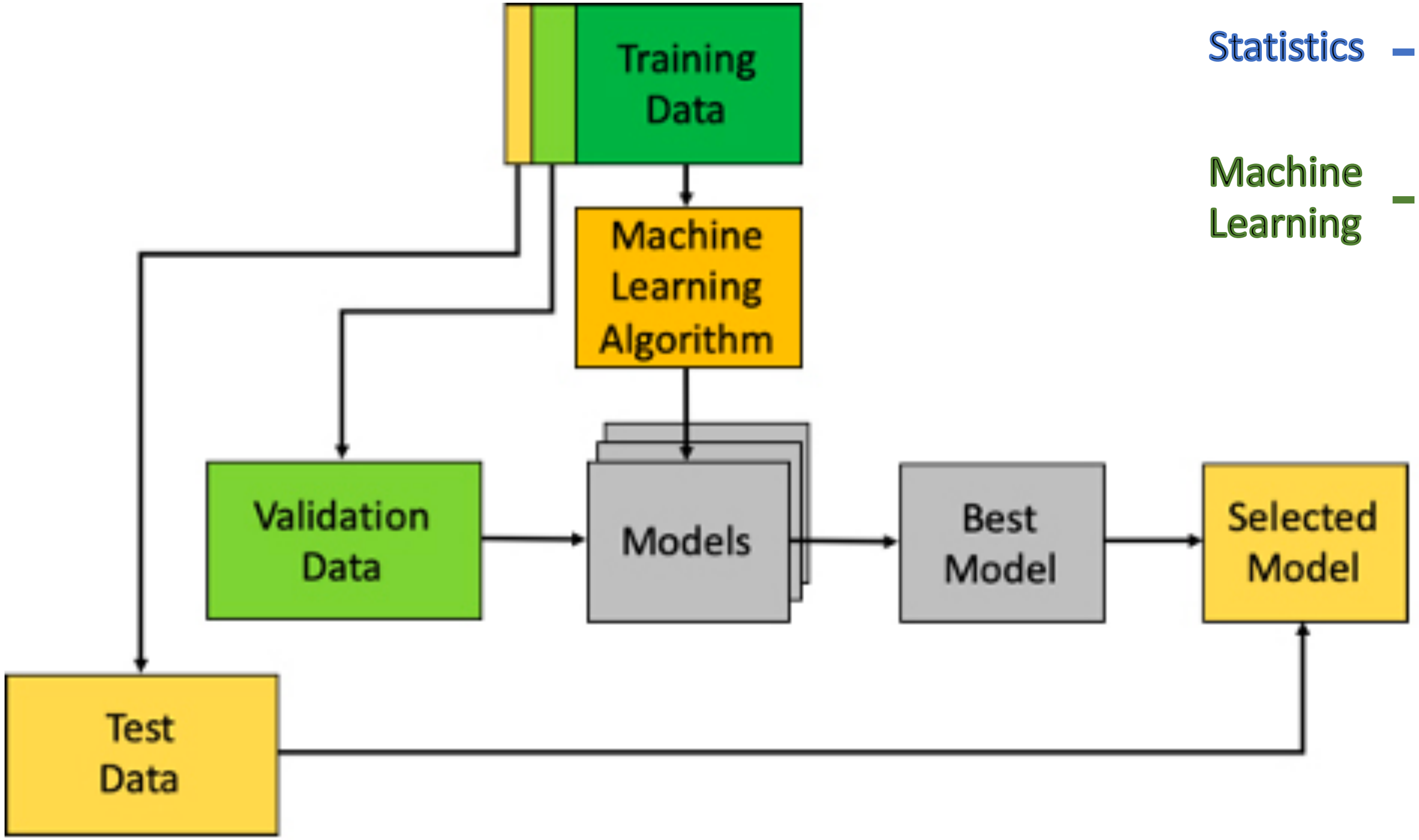
Turing

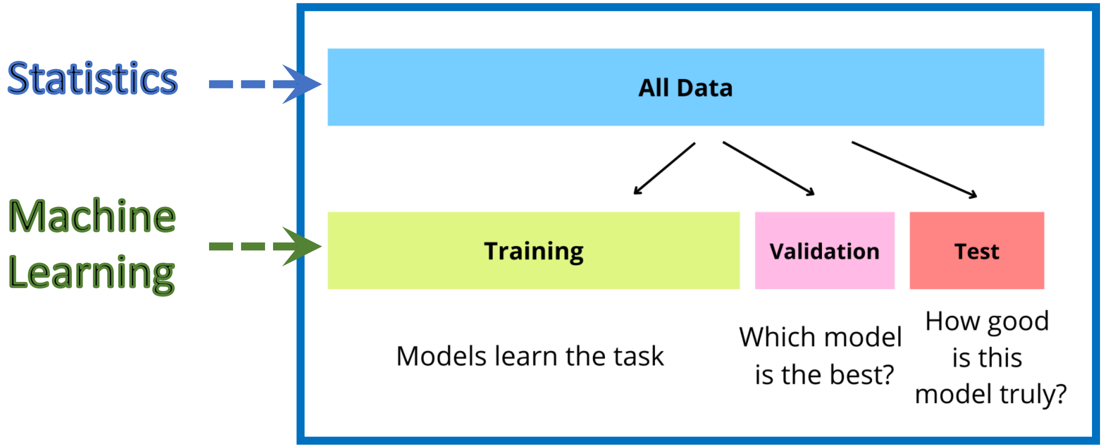
	MACHINE LEARNERS	STATISTICIANS
Network/Graphs vs. Models	Network/Graphs to train and test data	Models to create predictive power
Weights vs. Parameters	Weights used to maximize accuracy scoring and hand tuning	Parameters used to interpret real-world phenomena - stress on magnitude
Confidence Interval	There is no notion of uncertainty	Capturing the variability and uncertainty of parameters
Assumptions	No prior assumption (we learn from the data)	Explicit a-priori assumptions
Distribution	Unknown a priori	A-priori well-defined distribution
Fit	Best fit to learning models (generalization)	Fit to the distribution

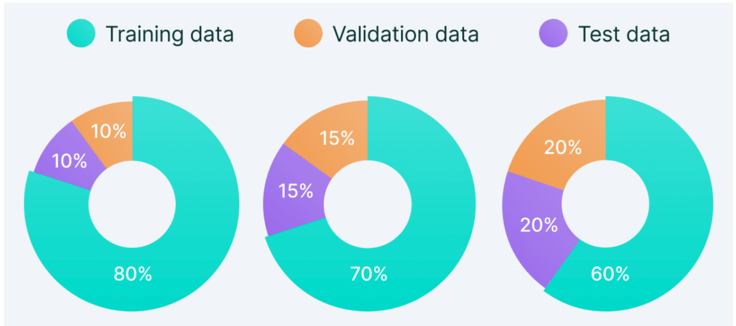
https://www.turing.com/kb/introduction-to-statistics-for-machine-learning

NB Good statistics knowledge is highly beneficial for ML (not the vice versa). For more insights read at this link and this other one

The *supervised* Machine Learning paradigm





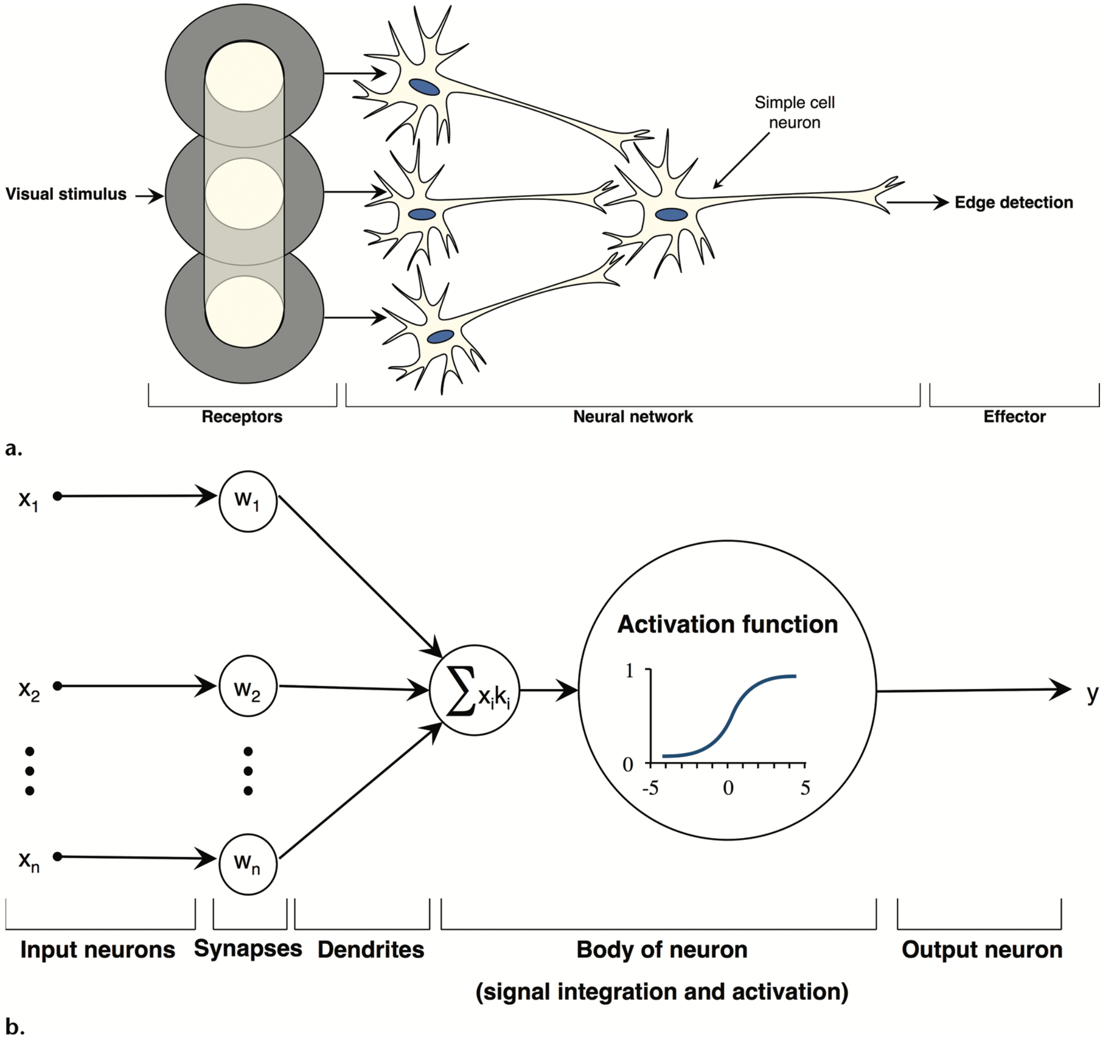


https://odsc.com/blog/the-comprehensive-guide-to-model-validation-framework-what-is-a-robust-machine-learning-model/

https://www.v7labs.com/blog/train-validation-test-set

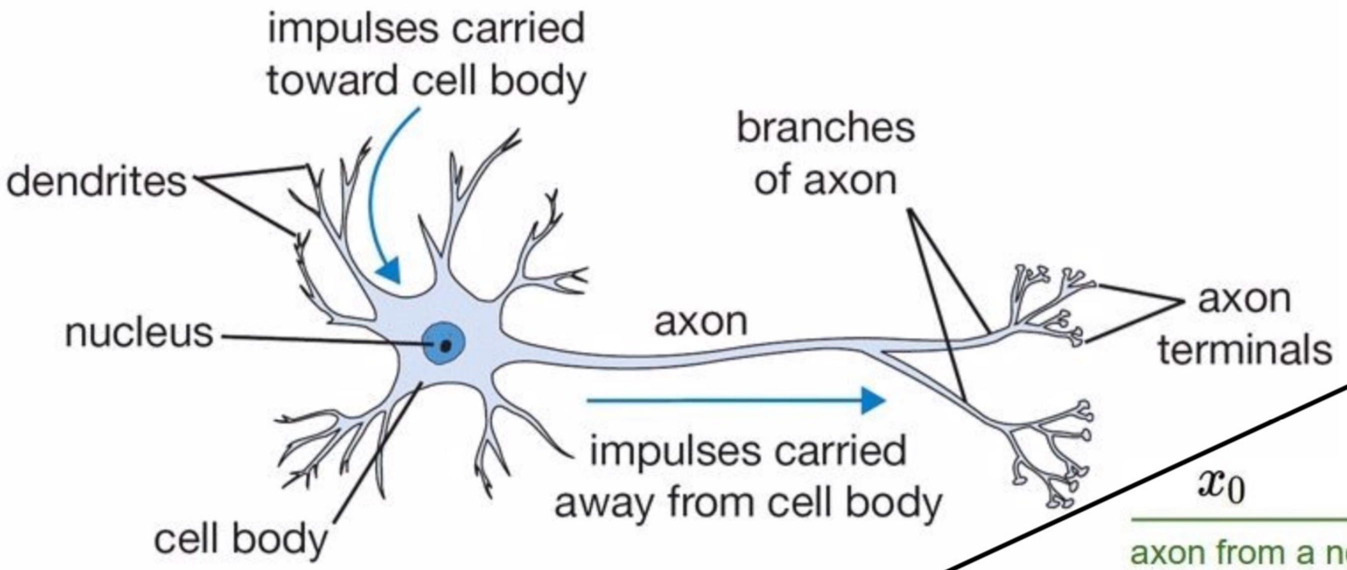
(Deep) Learning through artificial Neural Networks

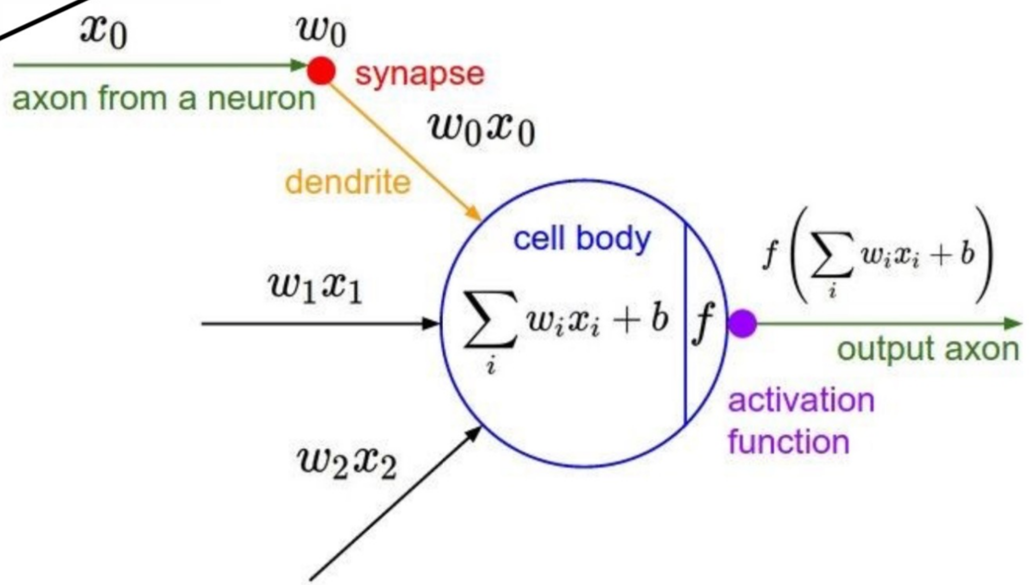
- Although neural networks were known and used for decades, in recent years three **key factors** have enabled the training of large neural networks:
  i. **the** possible **availability of large quantities of labeled data**,
  ii. cost-effective ~~inexpensive~~ **and powerful parallel computing hardware**,
  iii. **improvements in training techniques and architectures**.
- Artificial neural networks **are inspired by biologic neuron activation** process and from what we know about the structure of the **visual cortex** (a).
- The artificial neuron (b) takes as an input a set of values representing features, each multiplied by a corresponding weight. The **weighted features are summed and passed through a non-linear activation function**. In this way, an artificial neuron can be viewed as producing an **activation decision** by weighing a set of evidence.

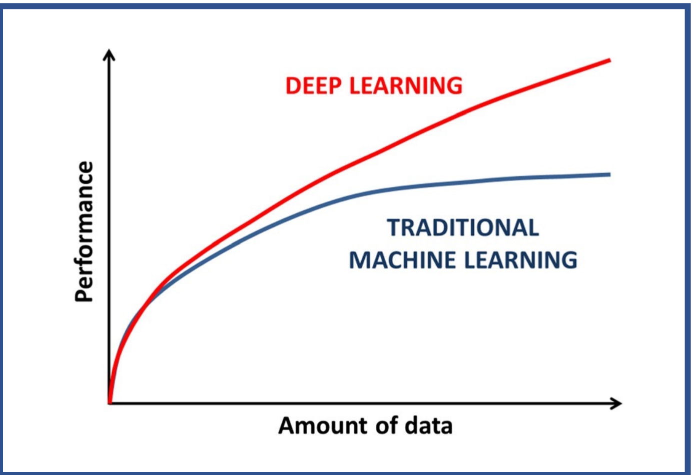


https://doi.org/10.1148/rg.2017170077

(Deep) Learning through artificial Neural Networks



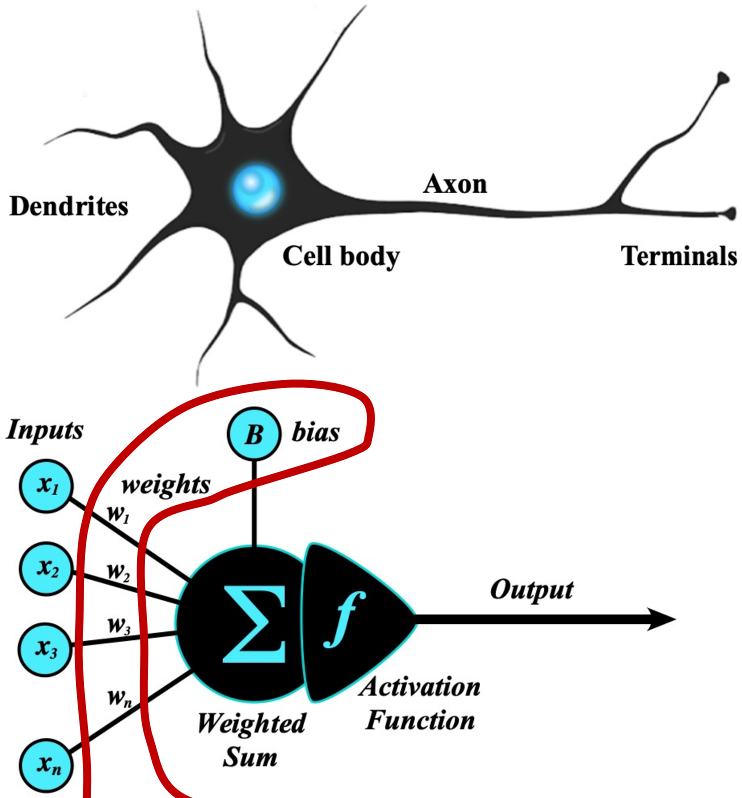


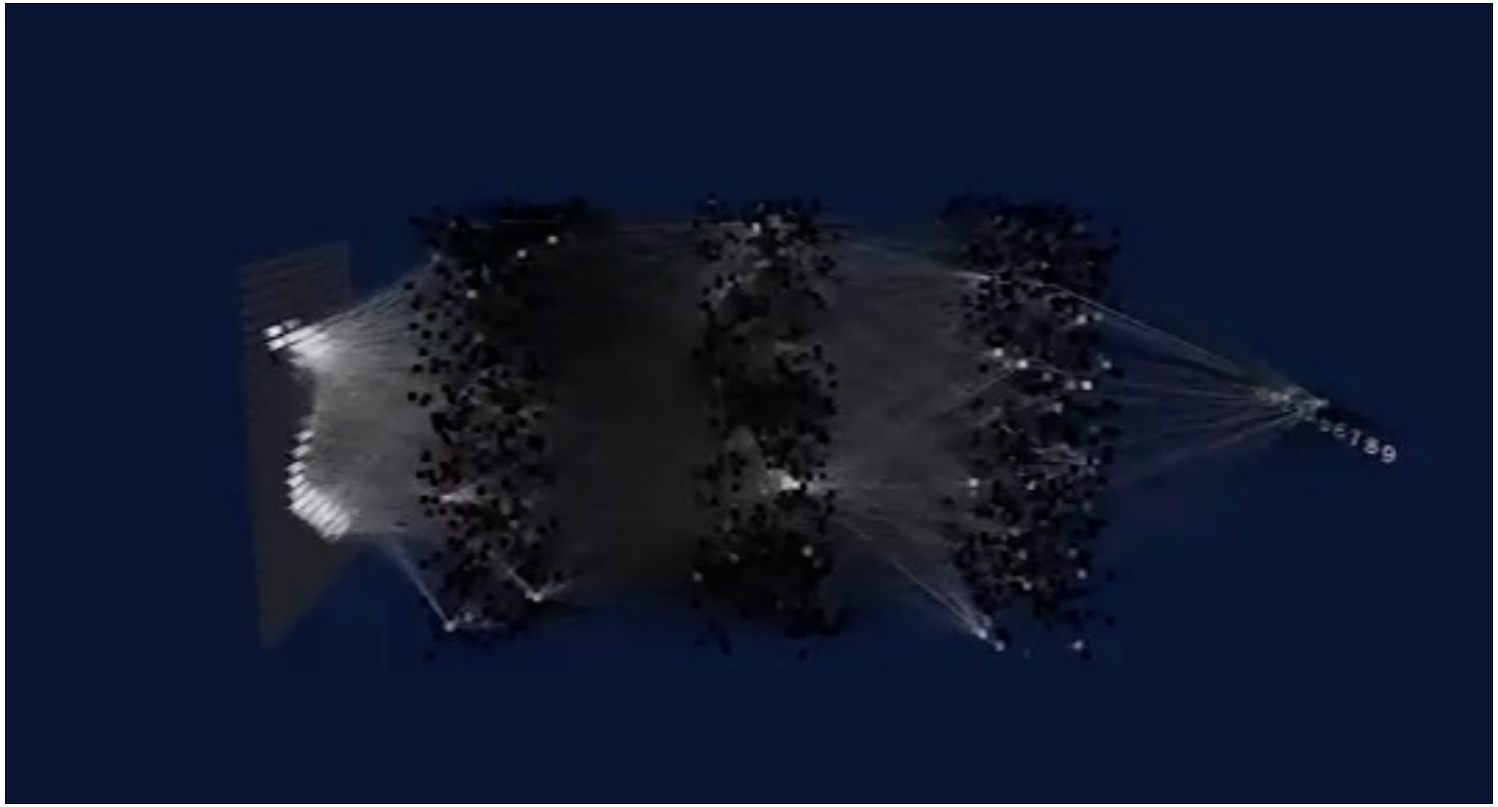


(Deep) Learning through artificial Neural Networks

Artificial Neural Network

Structure (layers→deep)
Number of parameters





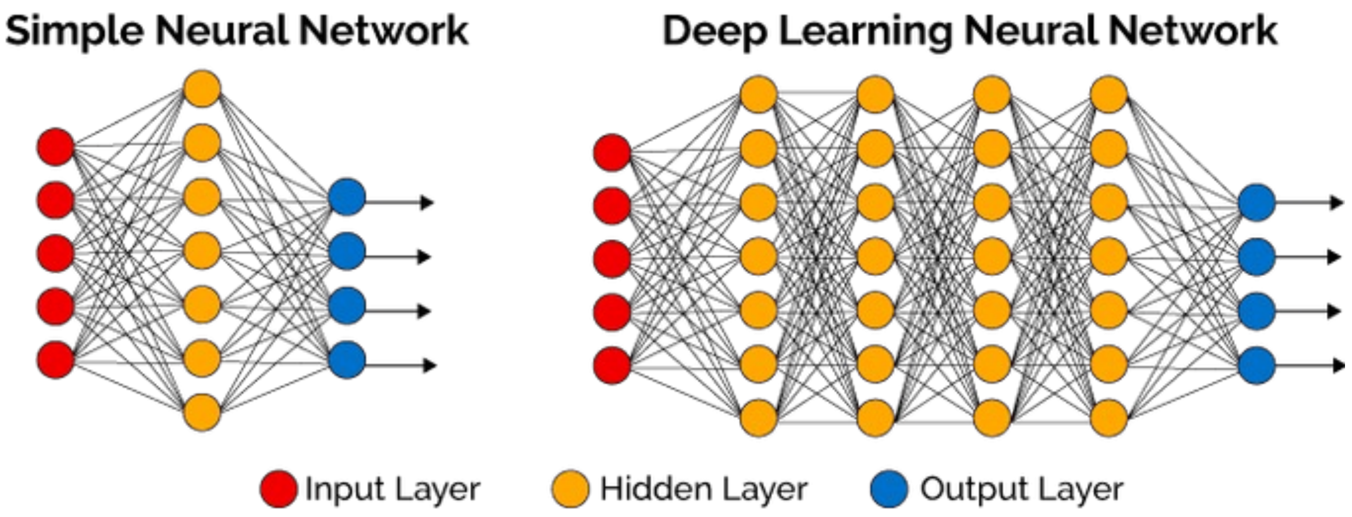
WEIGHTS & BIAS

parameters to be estimated for every neuron

https://www.youtube.com/watch?v=Tsvxx-GGlTg

(Deep) Learning through artificial Neural Networks

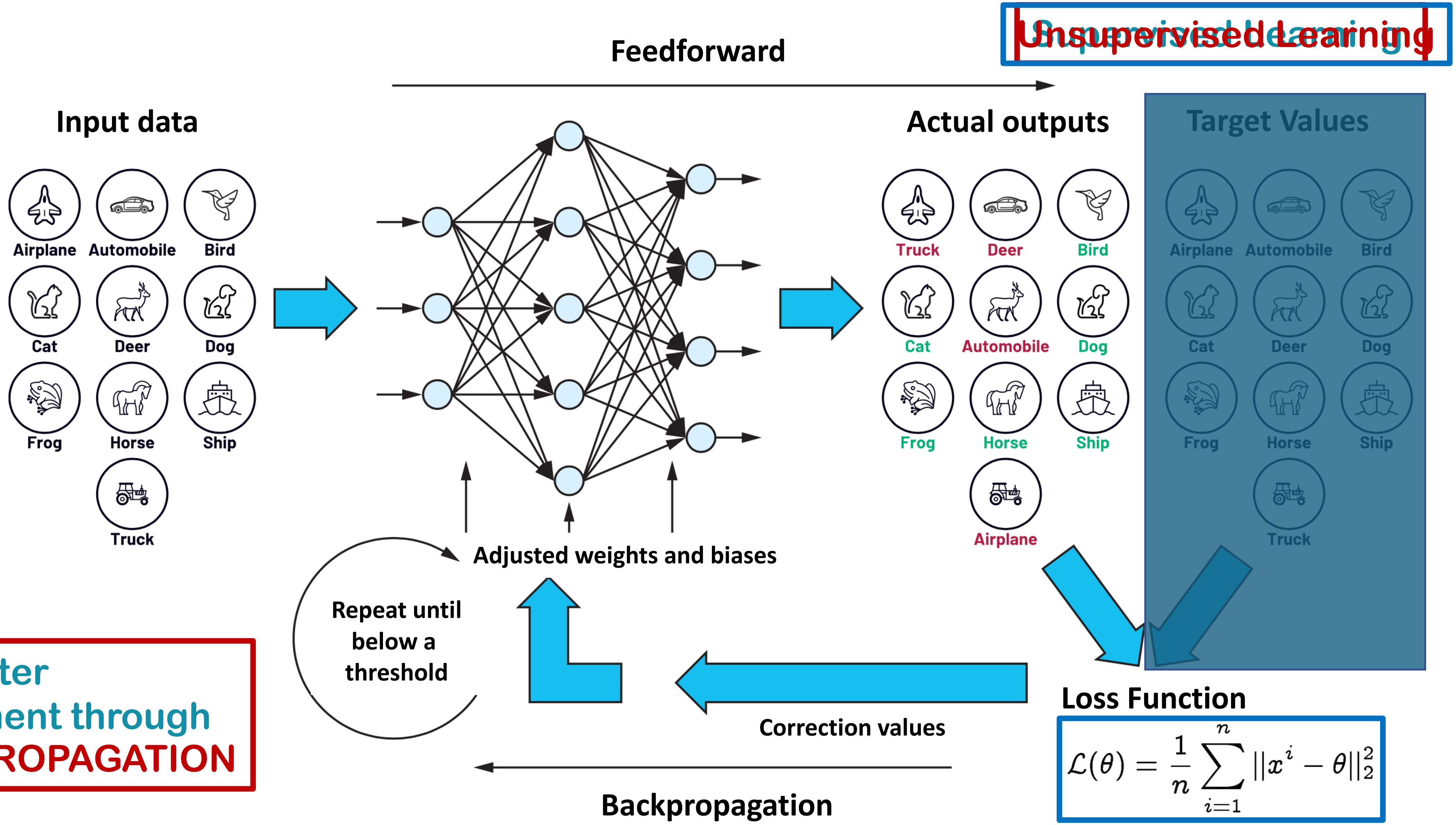
- Although an individual artificial neuron is simple, neural network architectures called **multilayer perceptrons** (MLP) that consist of thousands of neurons can represent very complex nonlinear functions.
- These multilayer perceptrons are typically constructed by assembling multiple neurons to form **layers** and by stacking these layers connecting the output of one layer to the input of the following layer. This produces a **hierarchy of features** that are an increasingly complex composition of low-level input features, thereby modeling higher levels of abstractions in the data.
- **MLPs perform poorly on images** in which the object of interest tends to vary in shape, orientation, and position because they must encode redundant representations for the many feature arrangements that this results in.



- A neural network is **trained** by adjusting the parameters, which consist of the weights of each node. Modern neural networks contain millions of such parameters.
- Starting from a random initial configuration, the parameters are adjusted via an optimization algorithm called **gradient descent**, which attempts to find a set of parameters that performs well on a **training dataset**.

https://doi.org/10.1148/rg.2017170077

(Deep) Learning through artificial Neural Networks

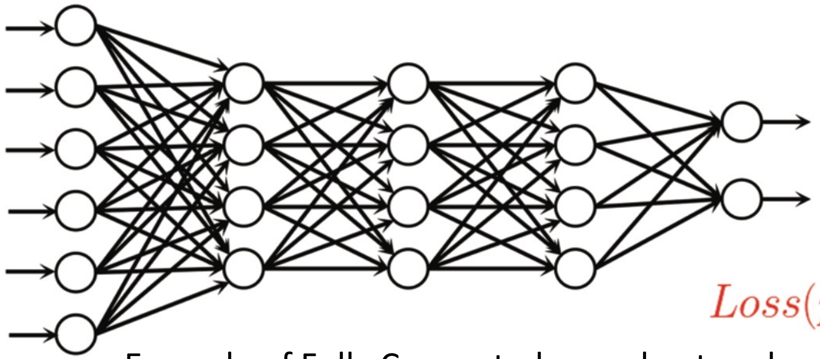


Parameter adjustment through BACKPROPAGATION

https://www.embedded.com/training-convolutional-neural-networks/

(Deep) Learning through artificial Neural Networks

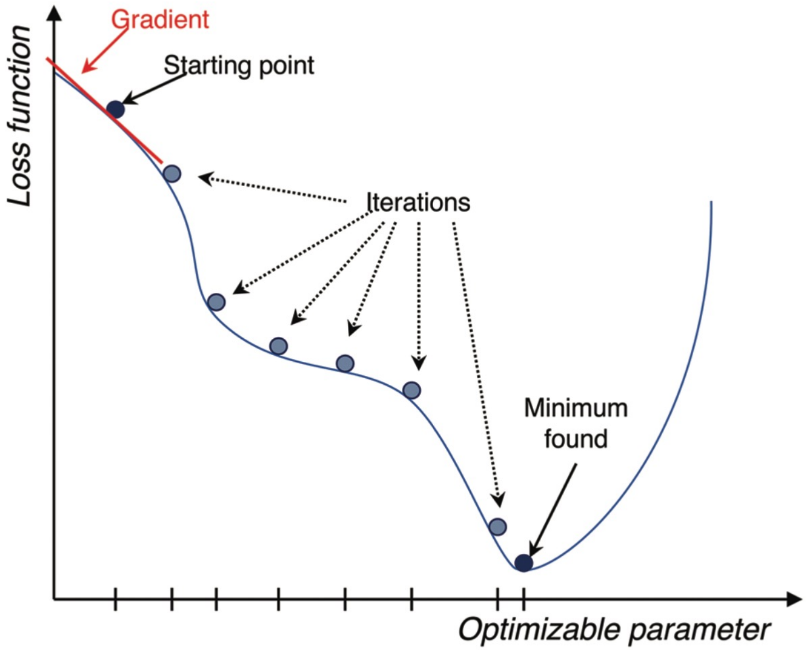
Learning process. Weights used by artificial neurons can be billions within a deep neural network. These parameters, are randomly initialized, are progressively adjusted via an optimization algorithm called ***gradient descent***. When presenting a series of training samples to the network, a ***loss function*** measures quantitatively how far the prediction is to the target class or regression value. All parameters are then slightly updated in the direction that will favor minimization of the loss function.



Prediction q	Target p
0.89	1.0
0.11	0.0

$$Loss(p, q) = -\sum_{x} p(x) log(q(x))$$

Example of Fully Connected neural network with 3 hidden layers: each connection carries a weight. and each neuron an additional one.

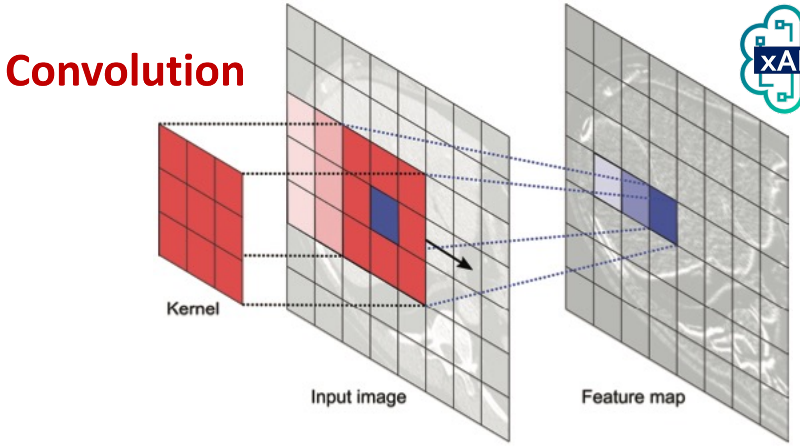


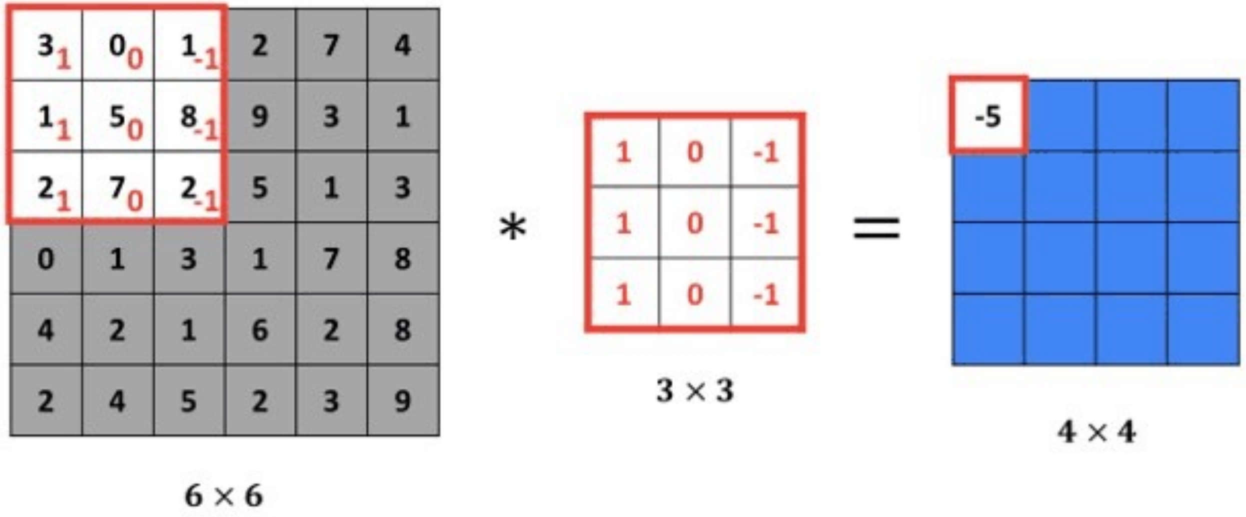
Learning algorithm

1. Present a batch of training samples to the network to evaluate a prediction based on its current configuration.
2. Evaluate the loss function by comparing the output prediction with the target values or classes.
3. Compute the gradient of the loss function with respect to every parameter of the model.
4. Update the weights of the model.
5. Repeat steps 1 to 4 until the loss function reaches a minimum.

https://doi.org/10.1148/rg.2017170077

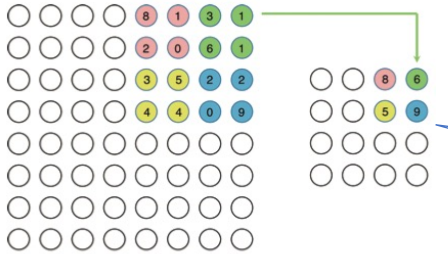
Convolutional Neural Networks

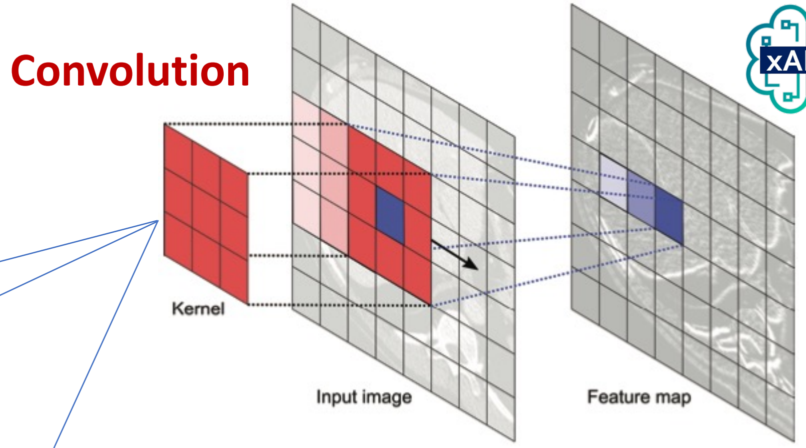




$$3 \times 1 + 1 \times 1 + 2 \times 1 + 0 \times 0 + 5 \times 0 + 7 \times 0 + 1 \times -1 + 8 \times -1 + 2 \times -1 = -5$$

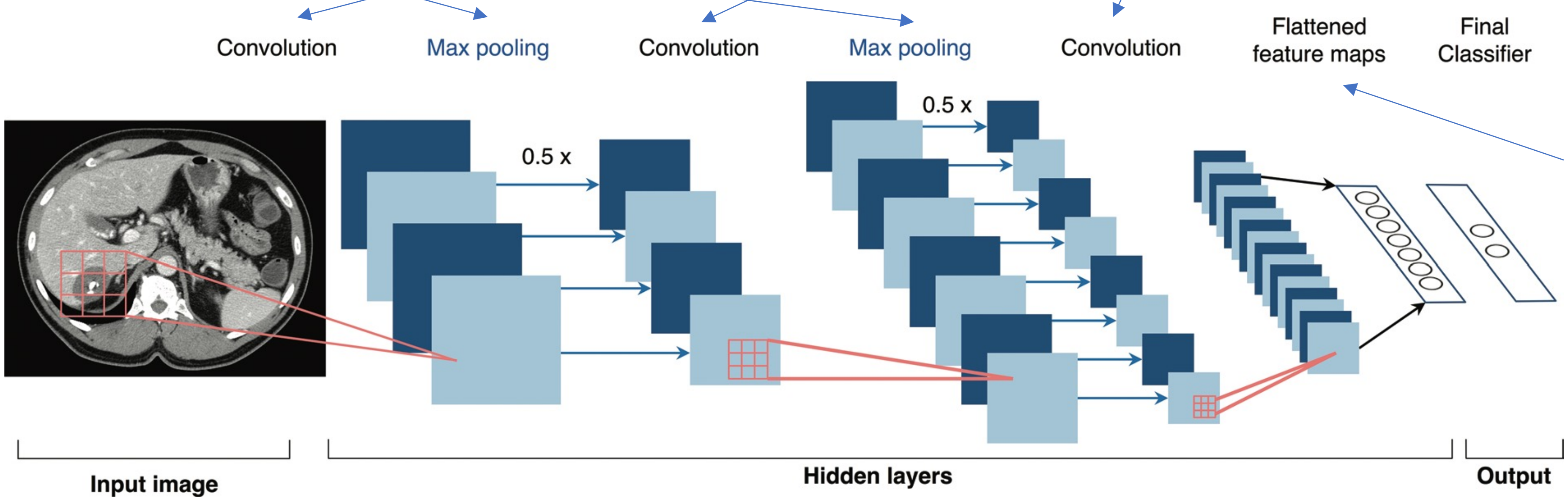
Convolutional Neural Networks





Convolutional layers and activation functions transform the feature maps

Down-sampling/ **pooling layers** reduce the spatial resolution



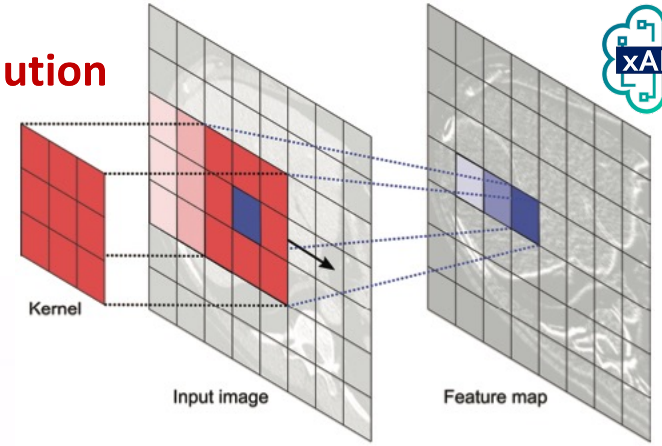
Coarser feature maps are transformed into a vector form through **fully connected layers**. allowing reasoning about the entire content of the image.

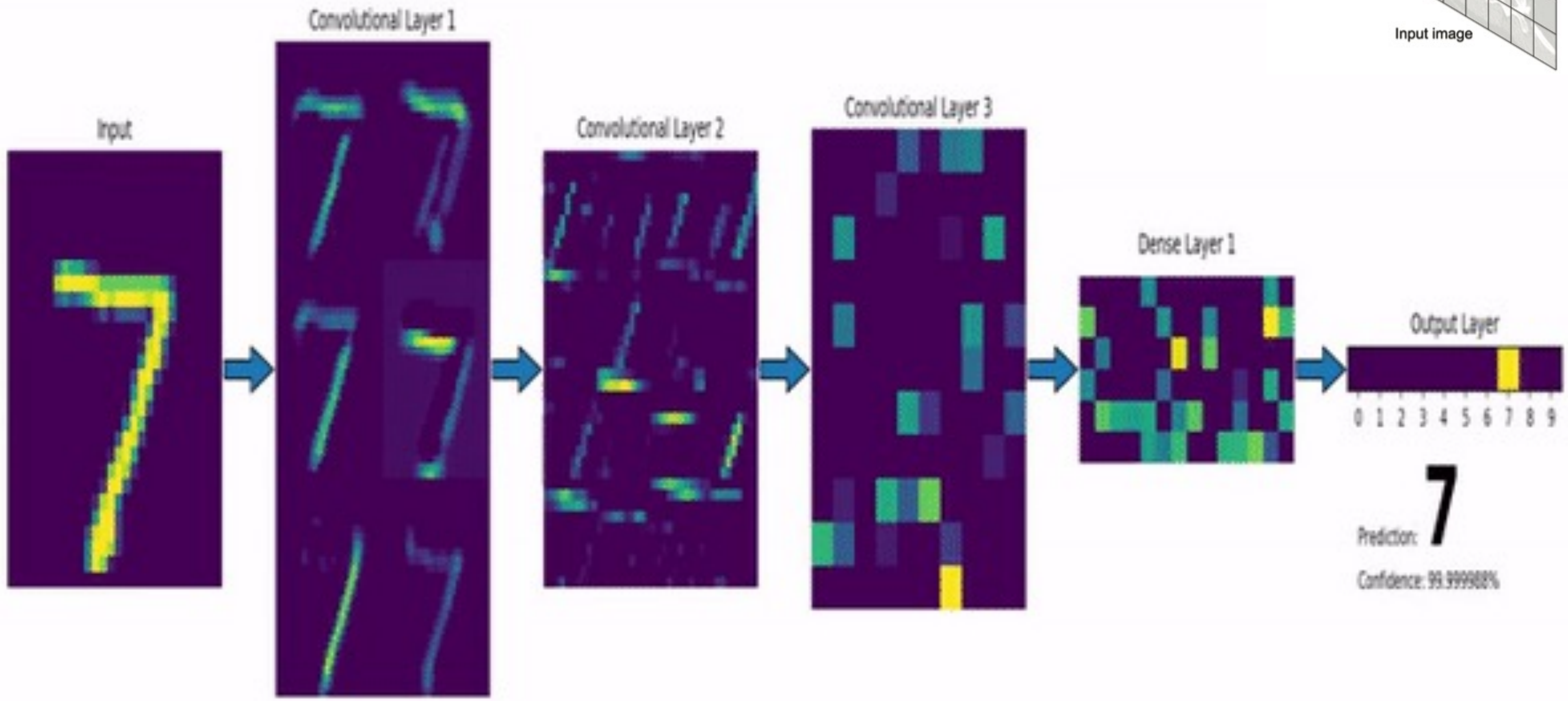
- For processing images, a deep learning architecture known as the ***convolutional neural network*** (CNN) has become dominant (from 2012 CNNs of increasing complexity kept winning popular image classification competitions even exceeding human performance).
- **CNNs introduces some robustness to image variations** by passing each feature detector over every part of the image in a **convolution operation**. Each feature detector is limited to detecting local features in its immediate input, which is acceptable for natural images.
- Since a feature may occur anywhere in the image, the filters' weights are shared across all the image positions. Thus, image features can be modeled with fewer parameters (shared kernel weights for all spatial positions), increasing model efficiency.

https://doi.org/10.1148/rg.2017170077

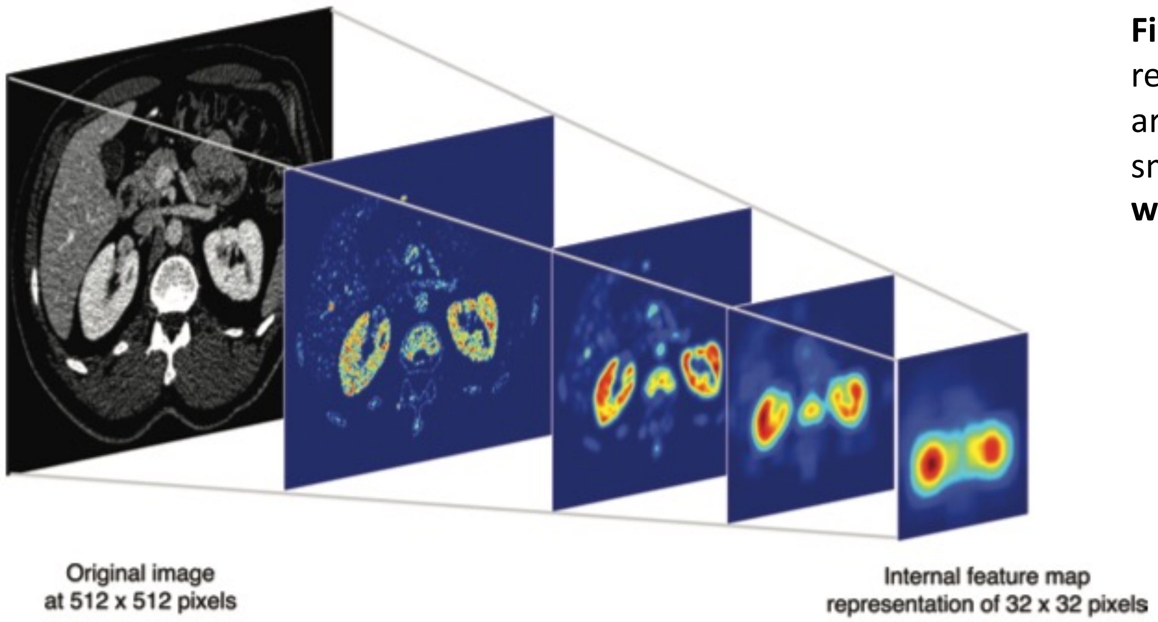
Convolutional Neural Networks

Convolution

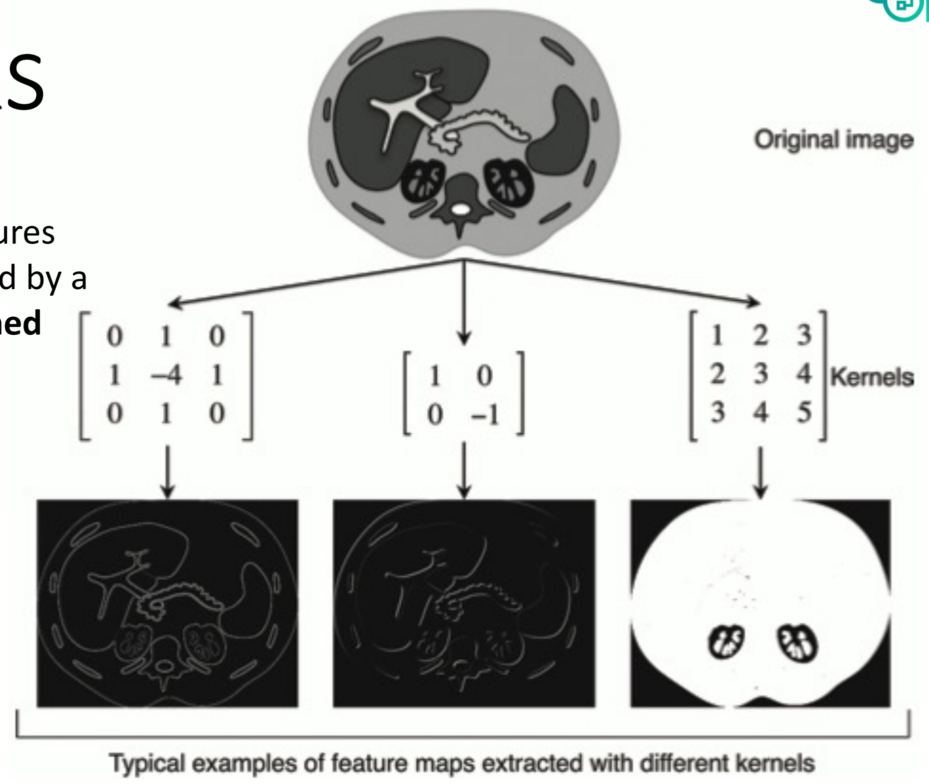




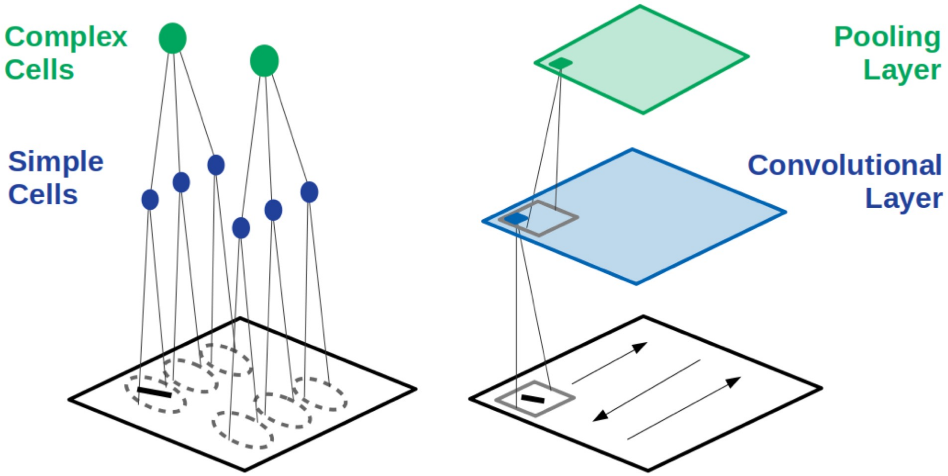
Convolutional Neural Networks



Filter kernels representing features are usually defined by a small grid of **learned weights** (eg, 3x3).



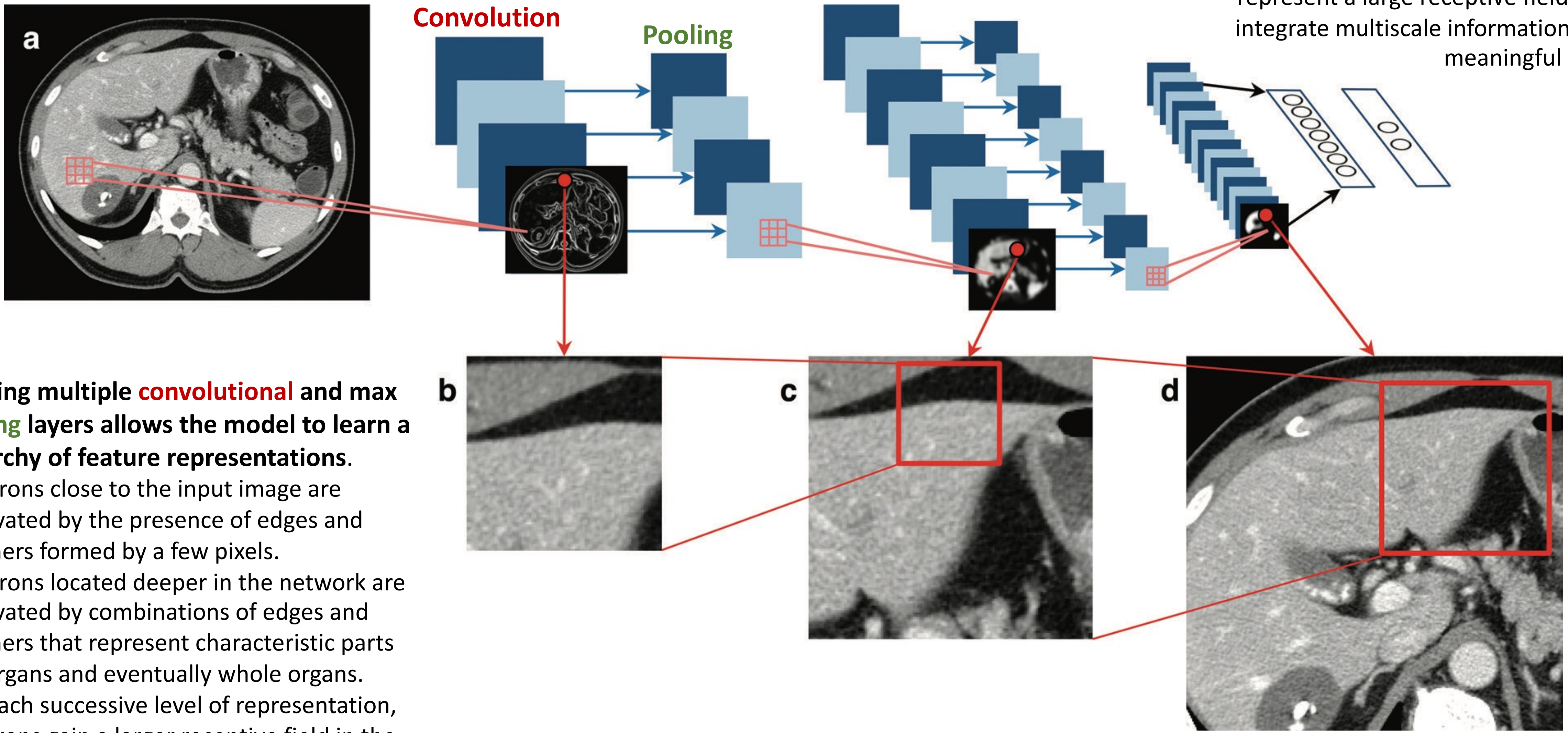
- CNNs were inspired by **early findings on biological vision** by Hubel & Wiesel in 1962 and from an early artificial NN called Neocognitron by Fukushima in 1979.
- In CNNs, **multiple different convolutional filters are learned for each layer**, yielding many different feature maps, each highlighting where different characteristics of the input image or of the previous hidden layer have been detected.
- In a CNN, the deeper the layer of representation, the coarser the characterization of the feature's spatial position (due to downsampling/pooling); thus, kernels in these deeper layers consider features over increasingly larger spatial scales.



https://doi.org/10.1148/rg.2017170077
https://doi.org/10.1162/jocn_a_01544

Convolutional Neural Networks

The **final classification** task relies on a rich set of hidden features that represent a large receptive field and integrate multiscale information in a meaningful way.

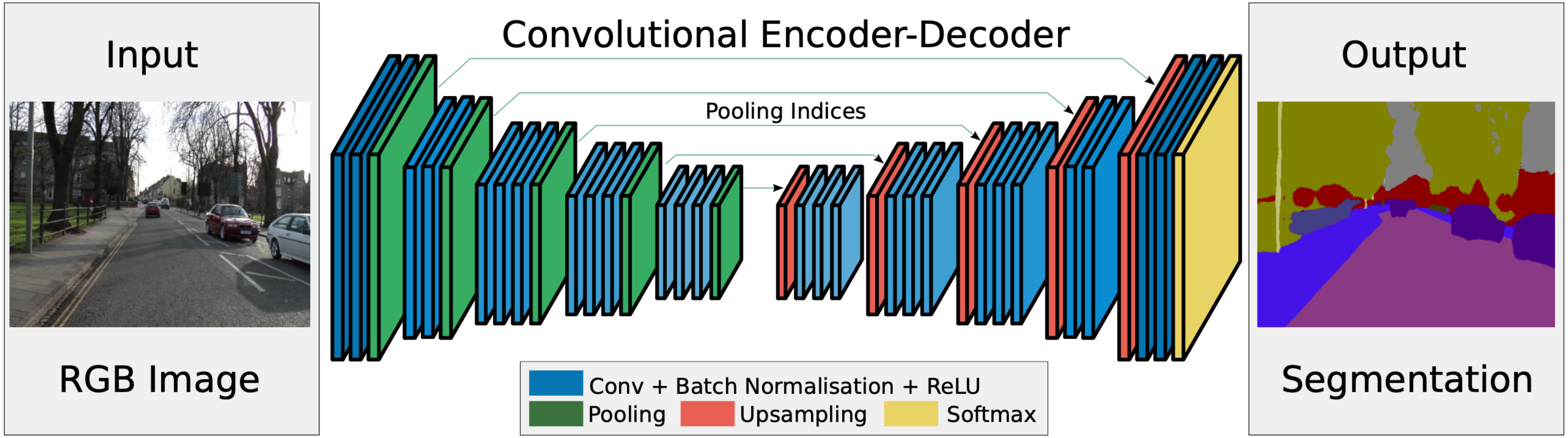


Stacking multiple convolutional and max pooling layers allows the model to learn a hierarchy of feature representations.

- Neurons close to the input image are activated by the presence of edges and corners formed by a few pixels.
- Neurons located deeper in the network are activated by combinations of edges and corners that represent characteristic parts of organs and eventually whole organs.
- At each successive level of representation, neurons gain a larger receptive field in the input image.

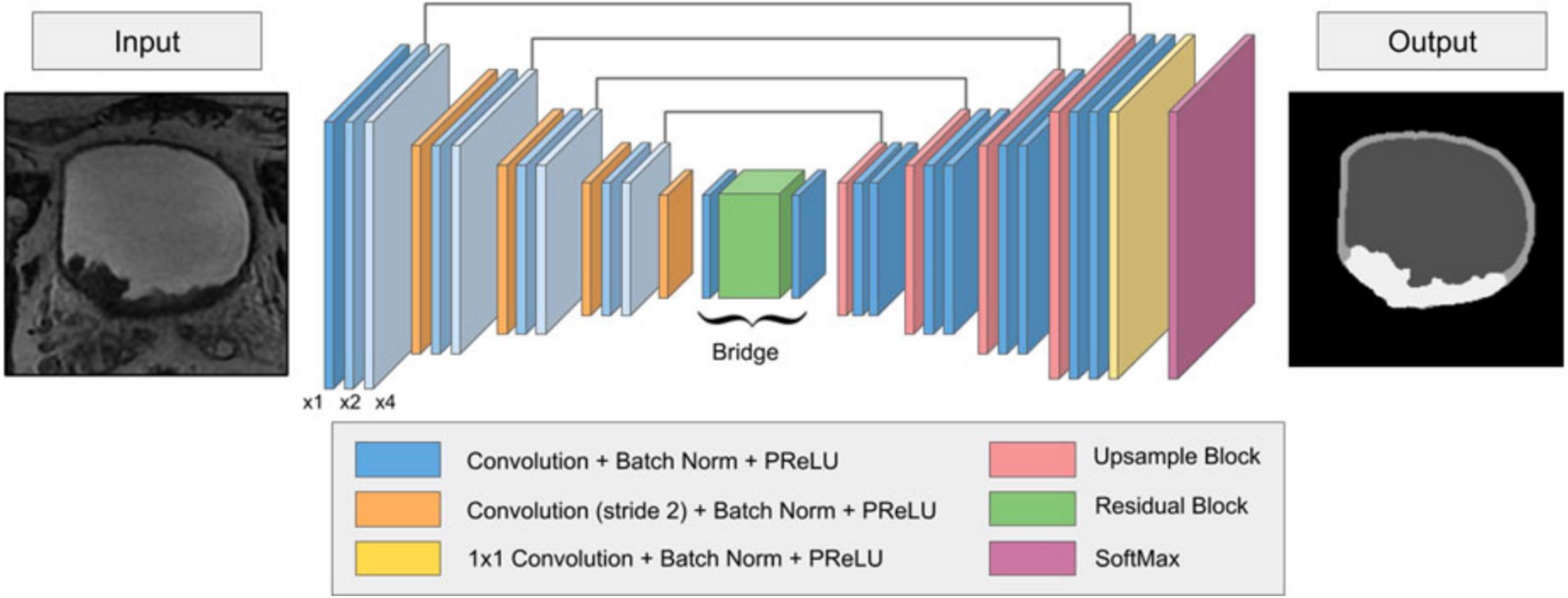
https://doi.org/10.1148/rg.2017170077

CNNs for segmentation (coding+decoding)



- If **pixel-level outputs** are desired, we have to **upsample** the features again
  - e.g. to obtain semantic level maps, typically pixel-wise semantic classification (or segmentation, or contouring)
- Why we need downsampling (coding) first and then upsampling (decoding)?
  - Downsampling provides strong features with large receptive field (quick information aggregation with limited number of network parameters)
  - Upsampling yields output at the same resolution as input
  - Usually skip connections are used, like in ResNet, but longer, as we will see in U-Nets
    - Skip connections allow maintaining high level of accuracy related e.g. to object boundaries

CNNs for segmentation (coding+decoding)



Multiregion segmentation of bladder cancer structures in MRI with progressive dilated convolutional networks, J.Dolz et al., Medical Physics, 2018
https://doi.org/10.1002/mp.13240

The AI “epochs” categorization

JAMA | Special Communication | AI IN MEDICINE

Three Epochs of Artificial Intelligence in Health Care

Michael D. Howell, MD, MPH; Greg S. Corrado, PhD; Karen B. DeSalvo, MD, MPH, MSc

Jan 2024

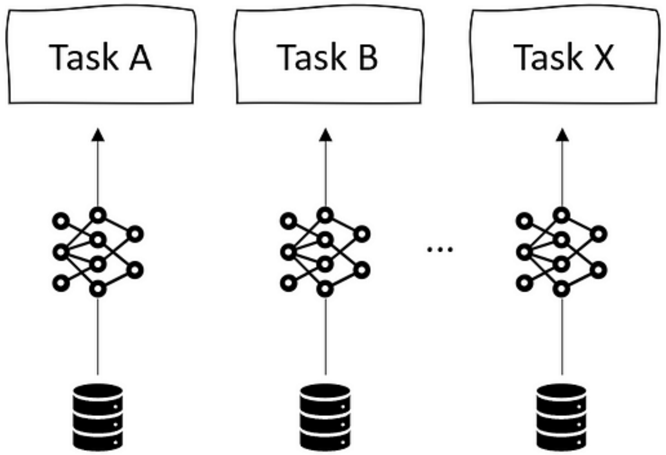
Approximate beginning year	*1950s*	*2011*	*2018-2022*
	AI 1.0 Symbolic AI and probabilistic models	**AI 2.0** Deep learning	**AI 3.0** Foundation models
Core functionality and key features	Follows directly encoded rules (if-then rules or decision trees)	Predicts and/or classifies information Task-specific (1 task at a time); requires new data and retraining to perform new tasks	Generates new content (text, sound, images) Performs different types of tasks without new data or retraining; prompt creates new model behaviors
Training method	Rules based on expert knowledge are hand-encoded in traditional programming	Learning patterns based on examples labeled as ground truth	Self-supervised learning from large datasets to predict the next word or sentence in a sequence
Good for	Follow rules or decision paths	Classify/detect based on training	Interpret and respond/assist
Issues and risks	Underfit real-world complexity, Errors in the model/rules	Unrepresentative/uncomplete data Bias in the training data	Hallucinations (plausible but incorrect) Bias in the training data
Healthcare applications	Rule-based clinical decision supporting systems	Diabetic retinopathy detection, lung cancer screening, skin condition classification, predictions based on electronic health records	Medically tuned large language models improve patient/clinician communication NB Foundation Models not only LLM! (*ndr*)

JAMA. 2024;331(3):242-244. doi:10.1001/jama.2023.25057

Single task VS Foundation (generative) AI

Single Task DL Models

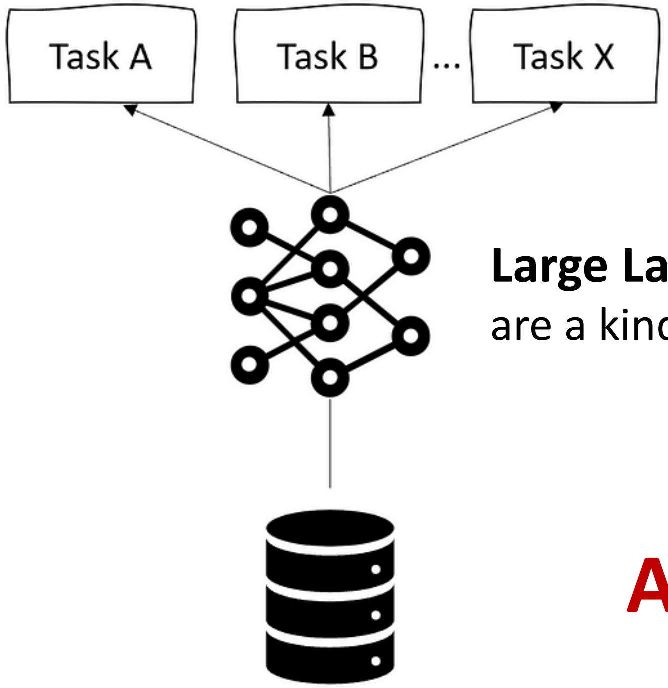
Typical technology
Convolutional Neural Networks
(are a kind of Deep Neural Network)



AI 2.0

Foundation Models

Typical technology
Transformers (are a kind of Deep Neural Network)



Large Language Models
are a kind of Foundation Models

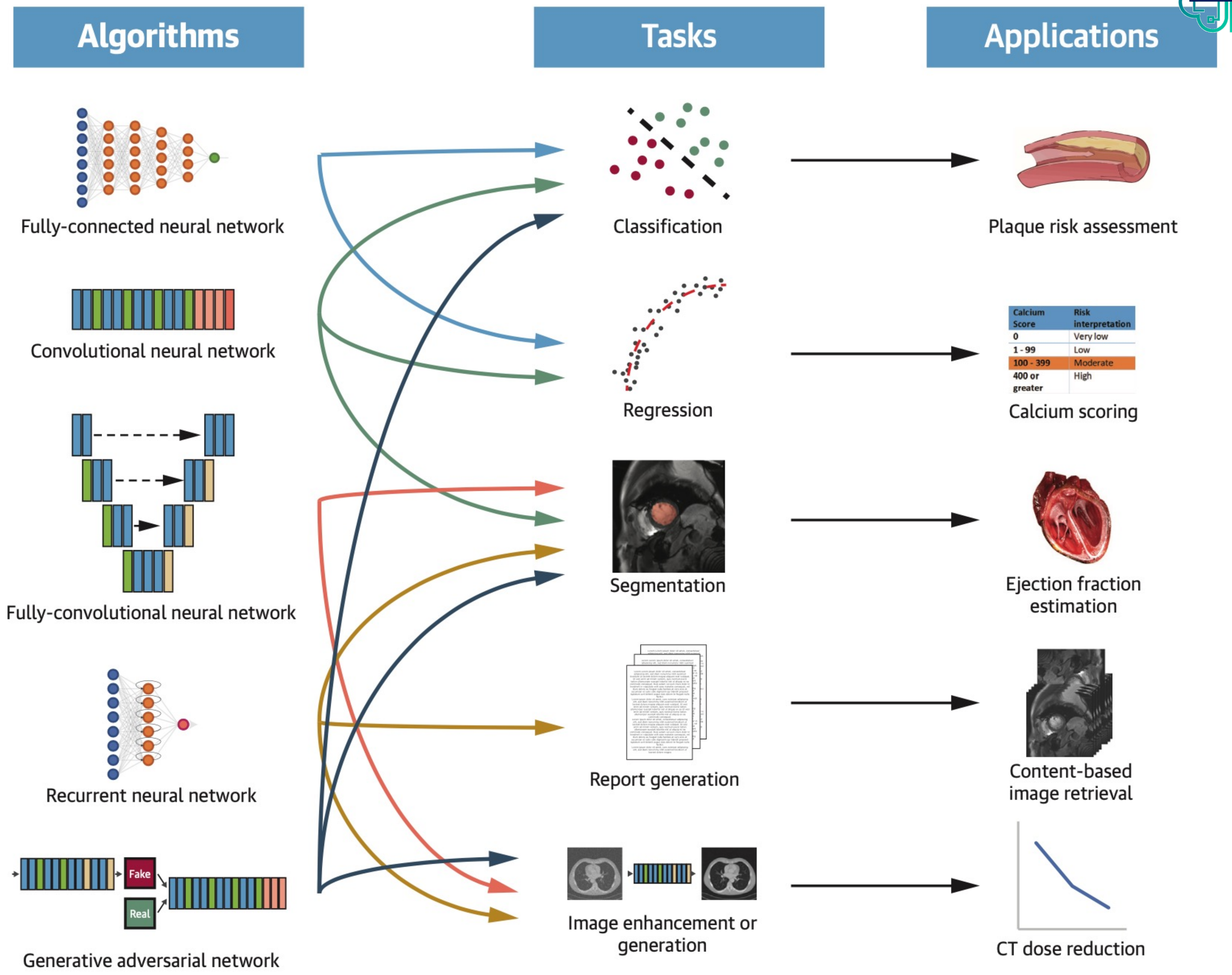
AI 3.0

The “big-mix” of AI 2.0 AI in medicine

Countless combinations among
- Algorithms
- Tasks
- Applications

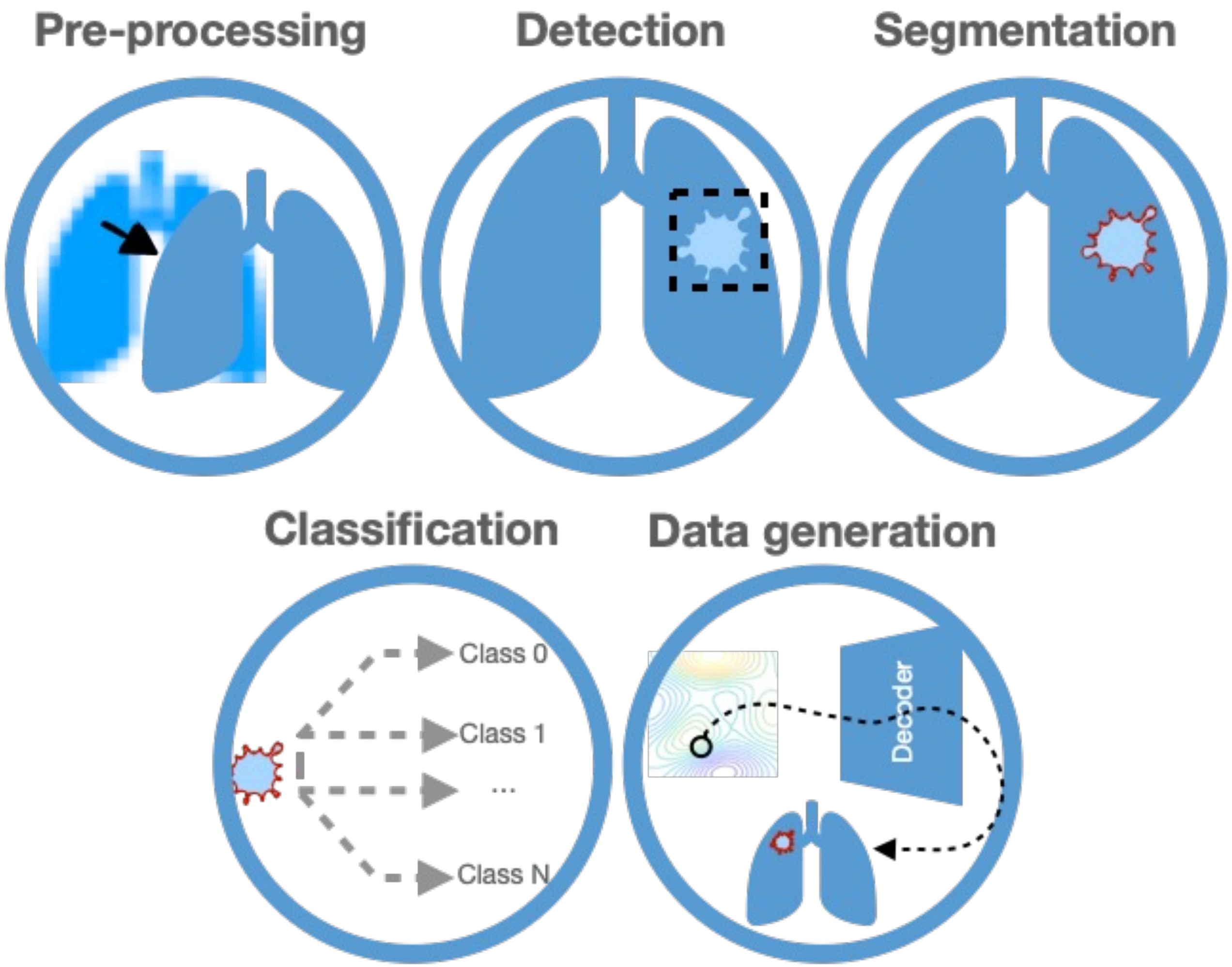
are possible.

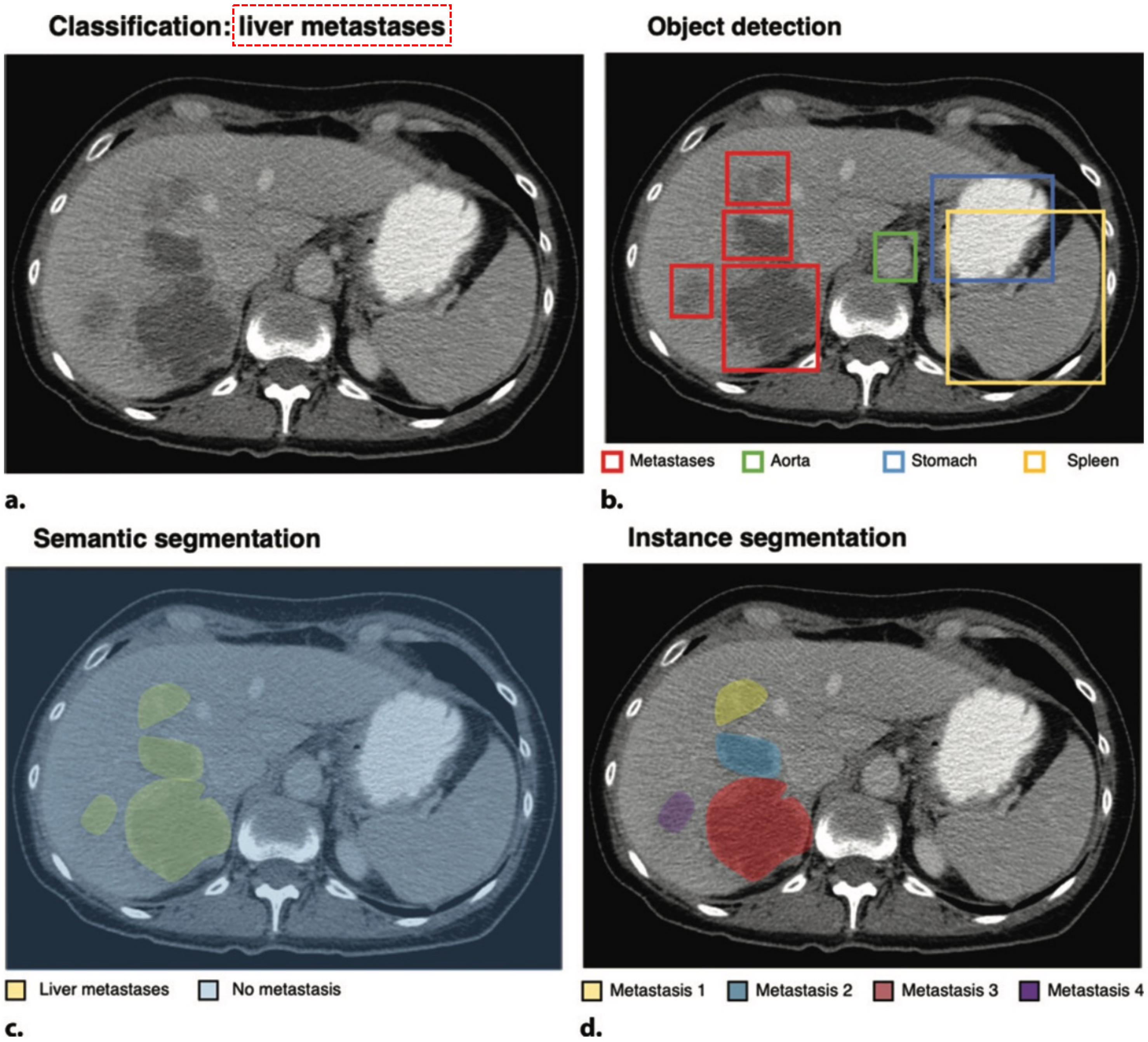
The need of a sound partnership between **technical** (computer scientists, information engineers, data scientists) and **clinical** specialties (physicians, medical technologists, biologists, physicists,…) is evident.



Litjens, G. et al. J Am Coll Cardiol Img. 2019;12(8):1549-65.

Interpretation tasks





Interpretation tasks: diagnostic imaging

- Classification

  allows to associate a “class” (output) to a given data (input). If the choice is among 2 classes we say *binary* classification, otherwise we say *multiclass*

- Detection

  allows to localize a given object or a plurality of them by means of bounding boxes and to associate a class label to each of them.

- Segmentation (semantic)

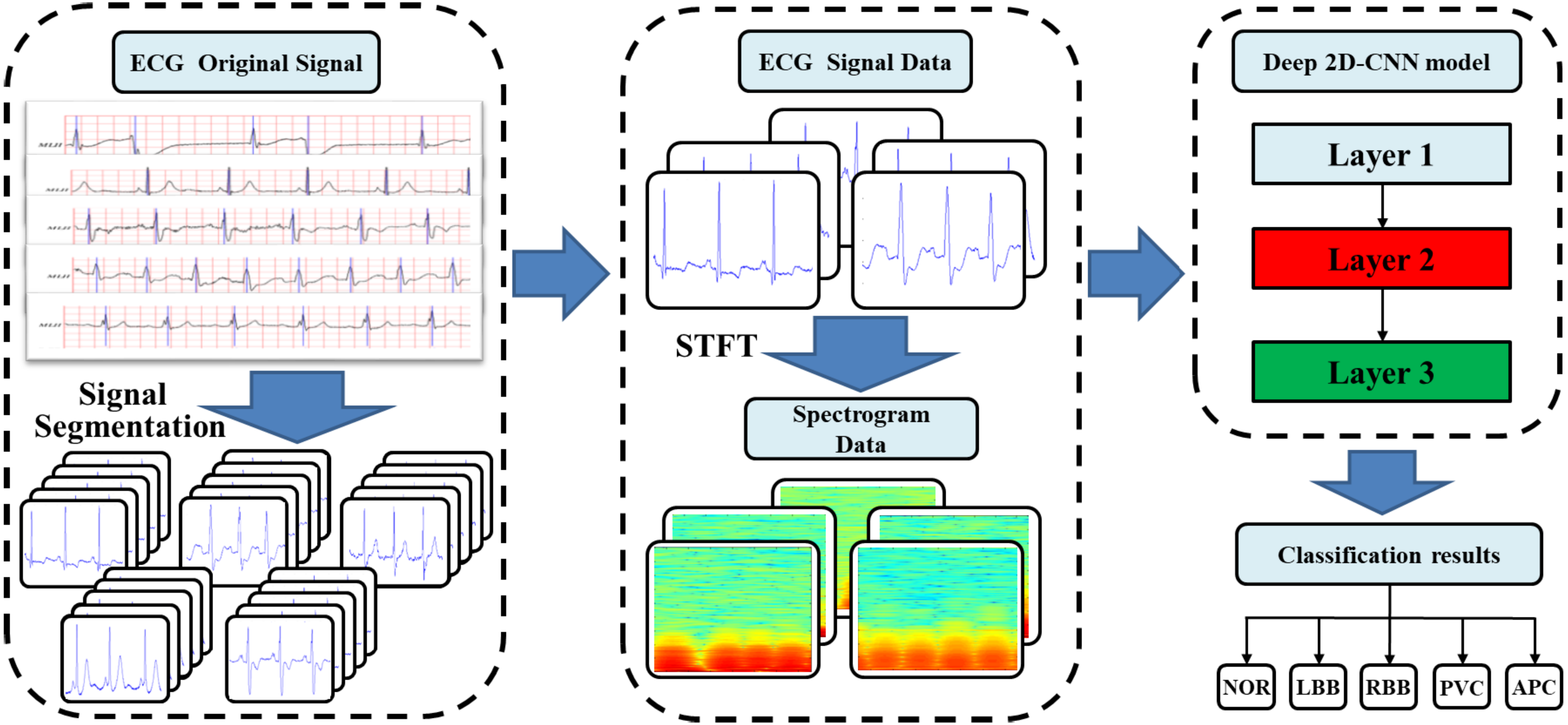
  allows to label every pixel in an image as to belonging to one of the available classes. Is a pixel-wise classification. Semantic segmentation can also be contour-aware (in this case we can say *contouring*)

- Segmentation (instance)

  like semantic segmentation with the possibility to differenciate different instances of the same object (in this case it is necessarily contour-aware).

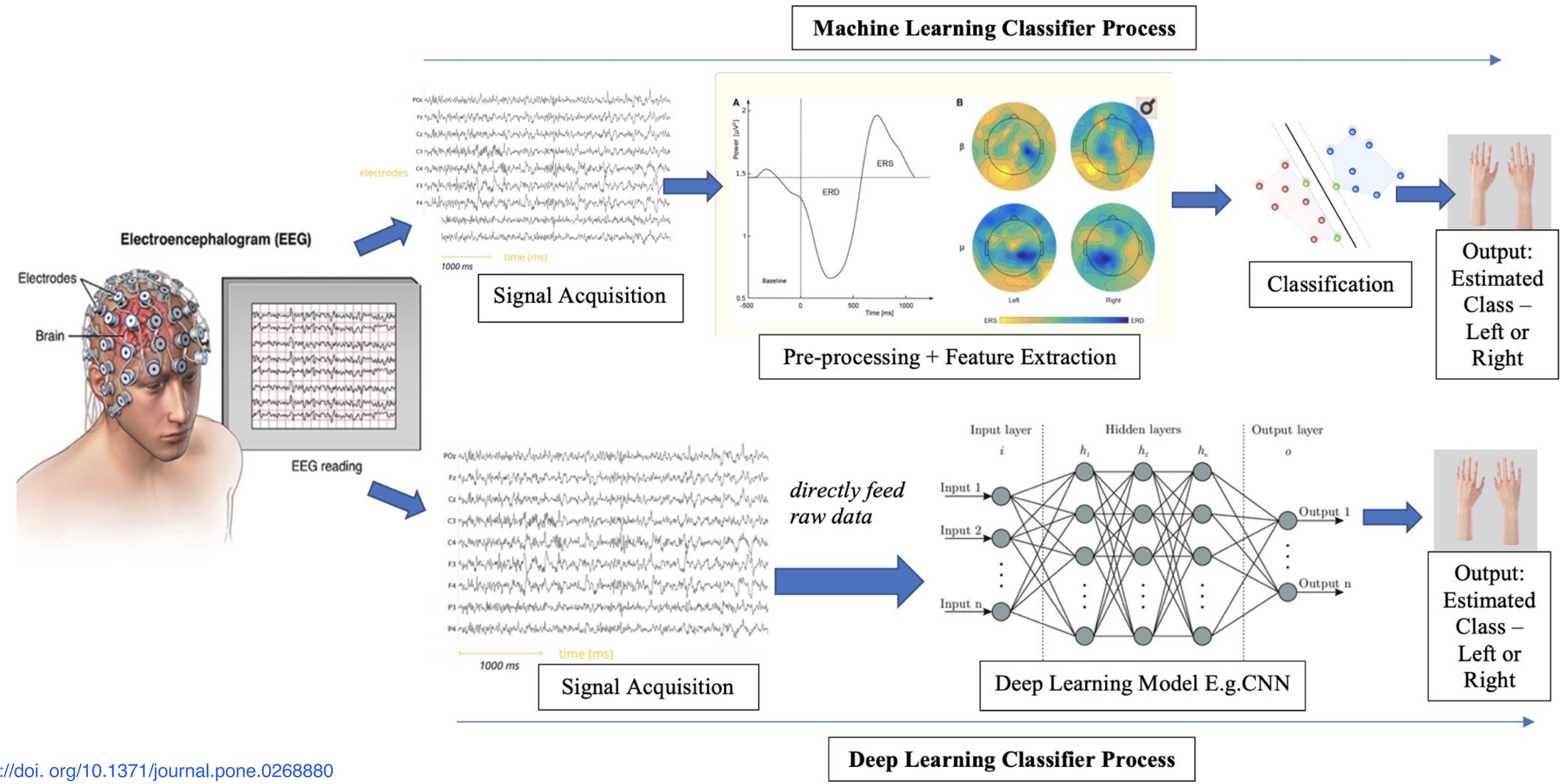
https://doi.org/10.1148/rg.2021200210

Interpretation tasks: non-image data sources



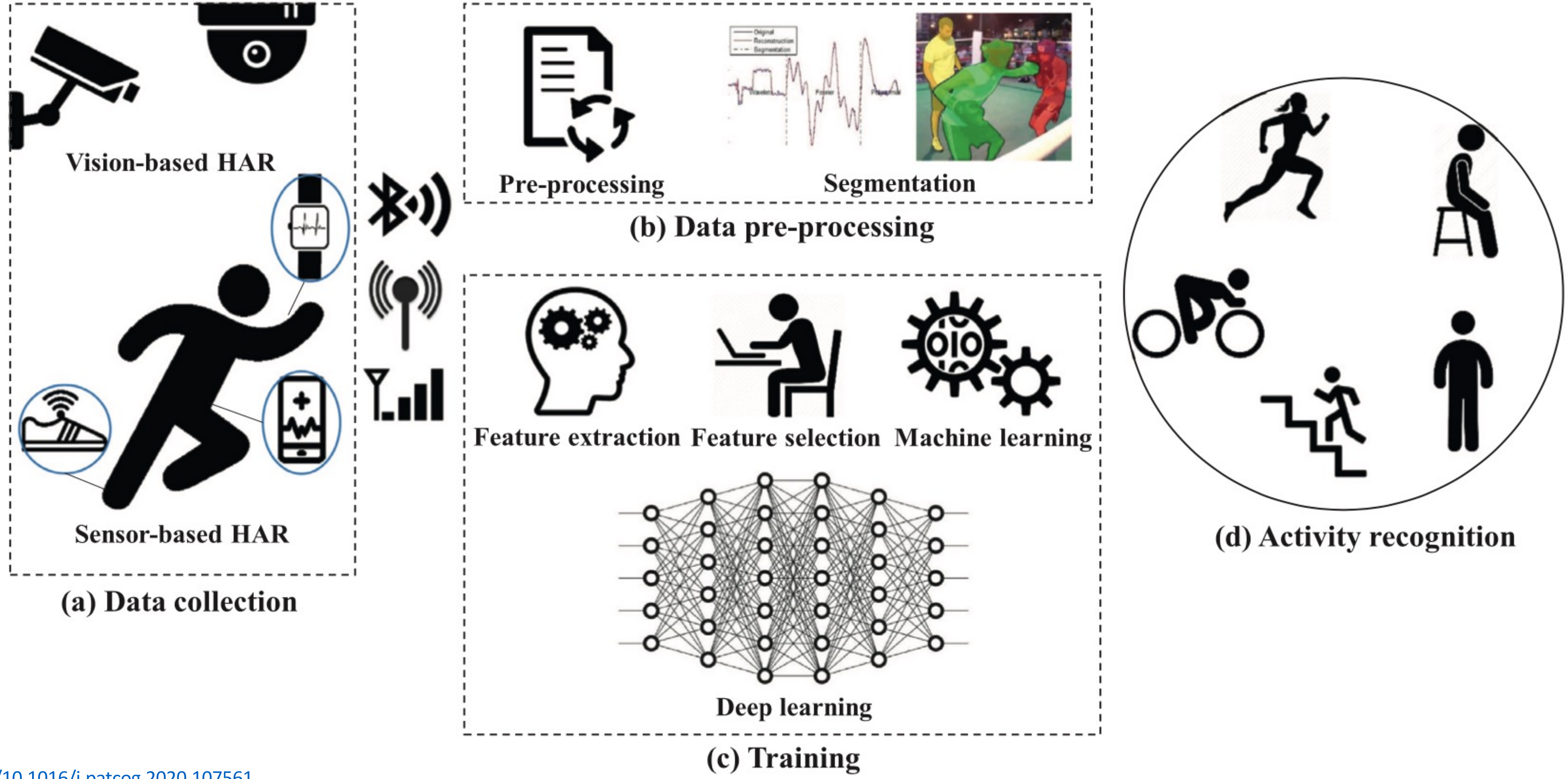
https://doi.org/10.1109/ACCESS.2019.2928017

Interpretation tasks: non-image data sources



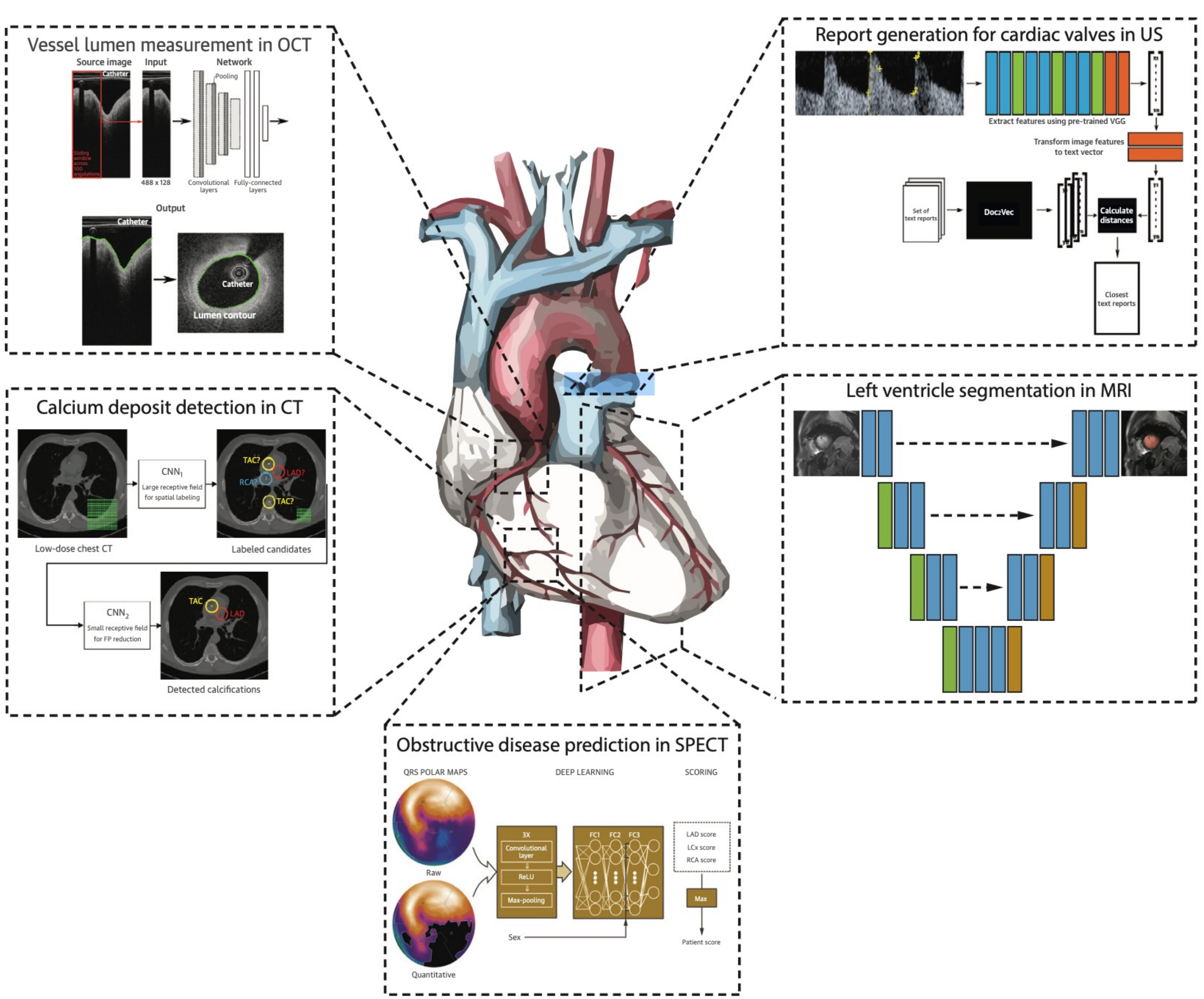
https://doi. org/10.1371/journal.pone.0268880

Interpretation tasks: non-image data sources

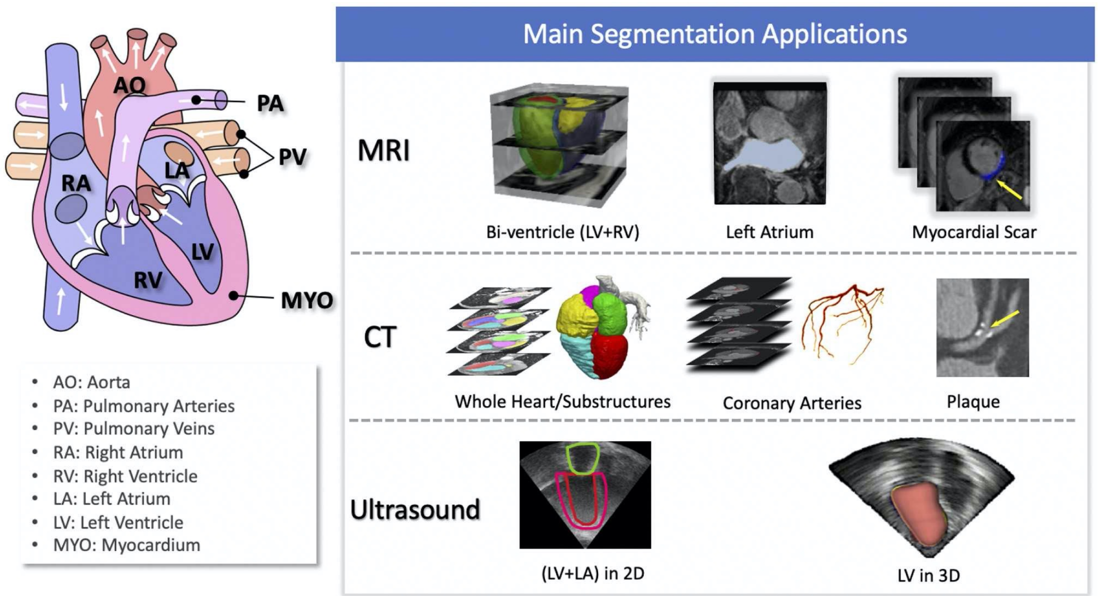


https://doi.org/10.1016/j.patcog.2020.107561

Organ-based/pathology-driven tasks



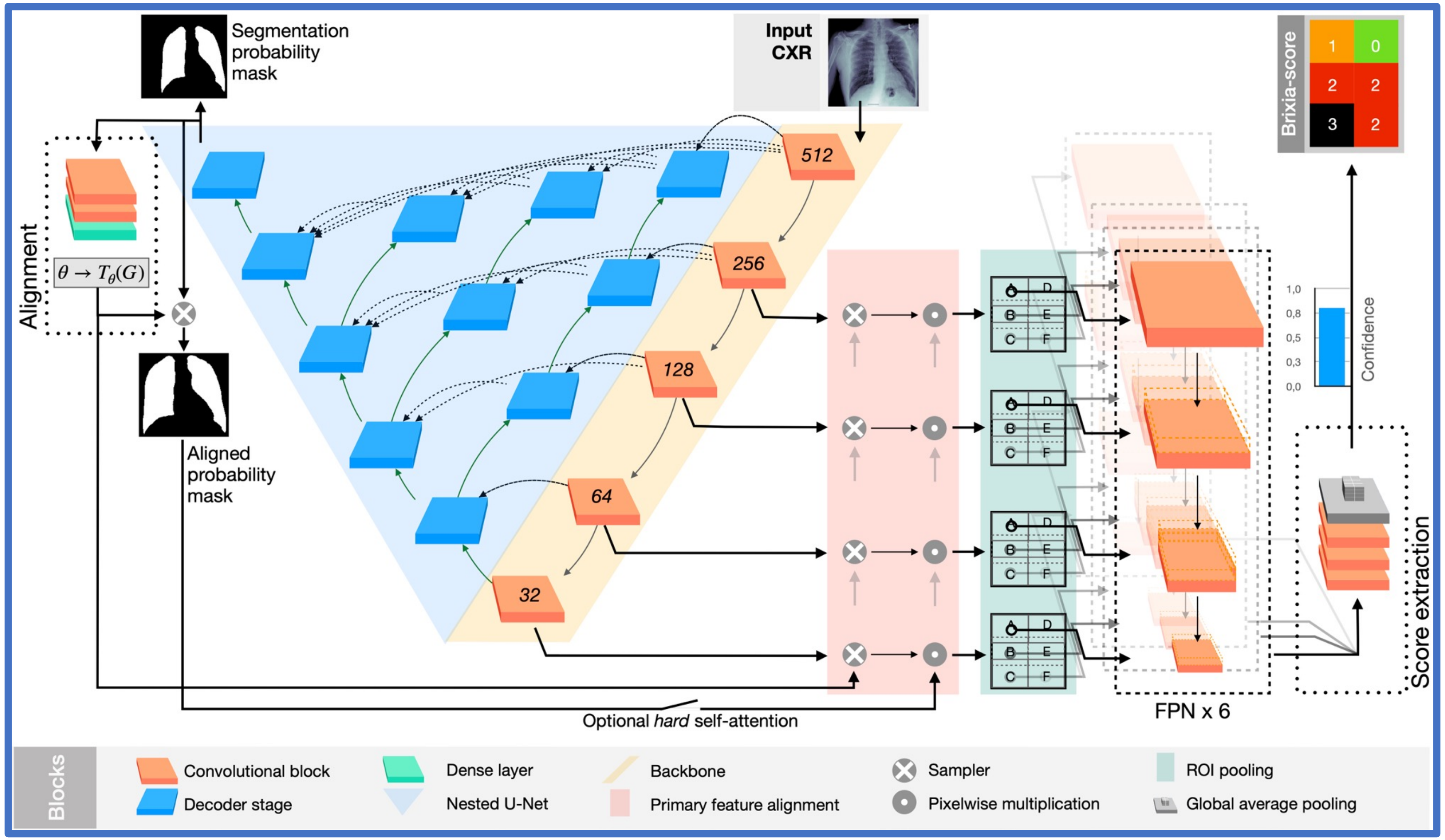
https://doi.org/10.1016/j.jcmg.2019.06.009

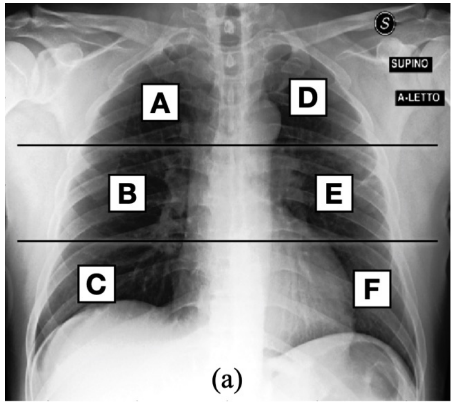


https://doi.org/10.3389/fcvm.2020.00025

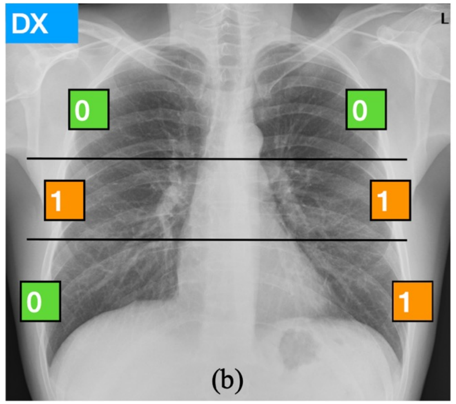
Composite interpretation tasks: multi-network approach

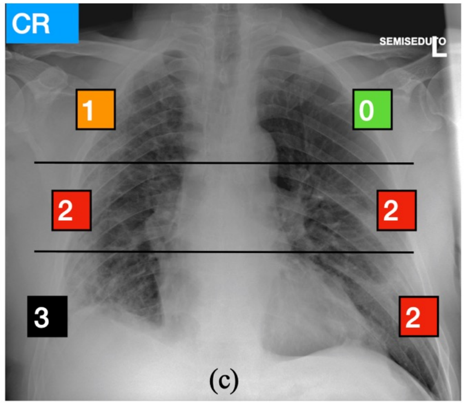
AI-driven evaluation of the «Brixia score» for COVID-19 pneumonia severity assessment
https://brixia.github.io

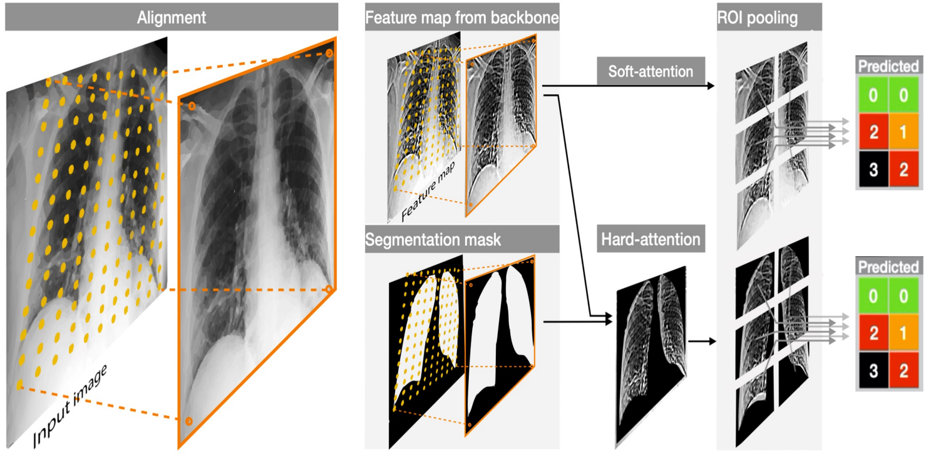




- **Score 0**: no lung abnormalities
- **Score 1**: interstitial infiltrates
- **Score 2**: interstitial and alveolar infiltrates (interstitial predominance)
- **Score 3**: interstitial and alveolar infiltrates (alveolar predominance)







https://doi.org/10.1016/j.media.2021.102046

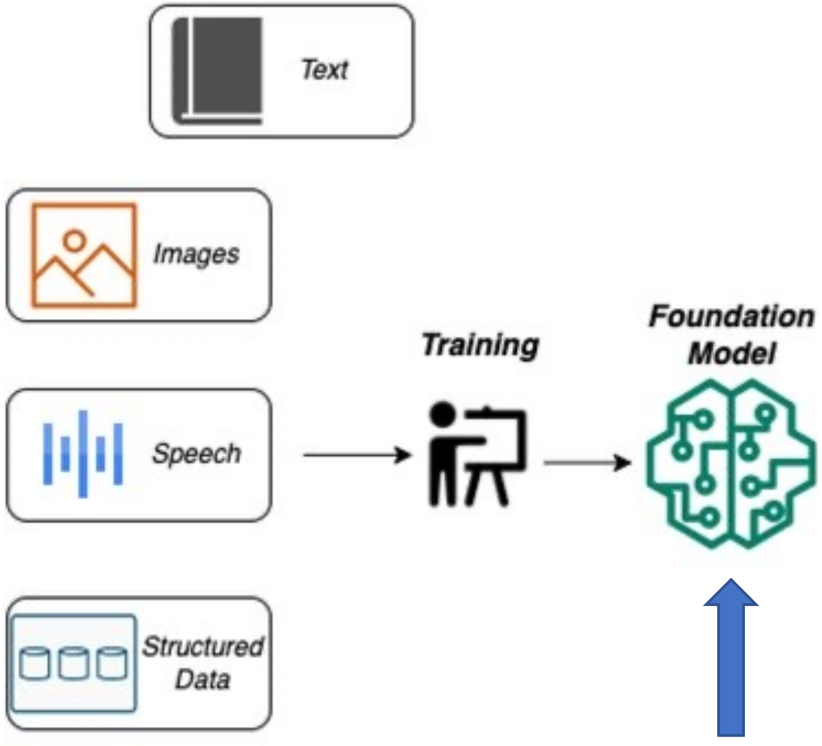
Foundation Models (FM)

Application driven tasks
(not defined in advance)

Data source/sources

AI 3.0

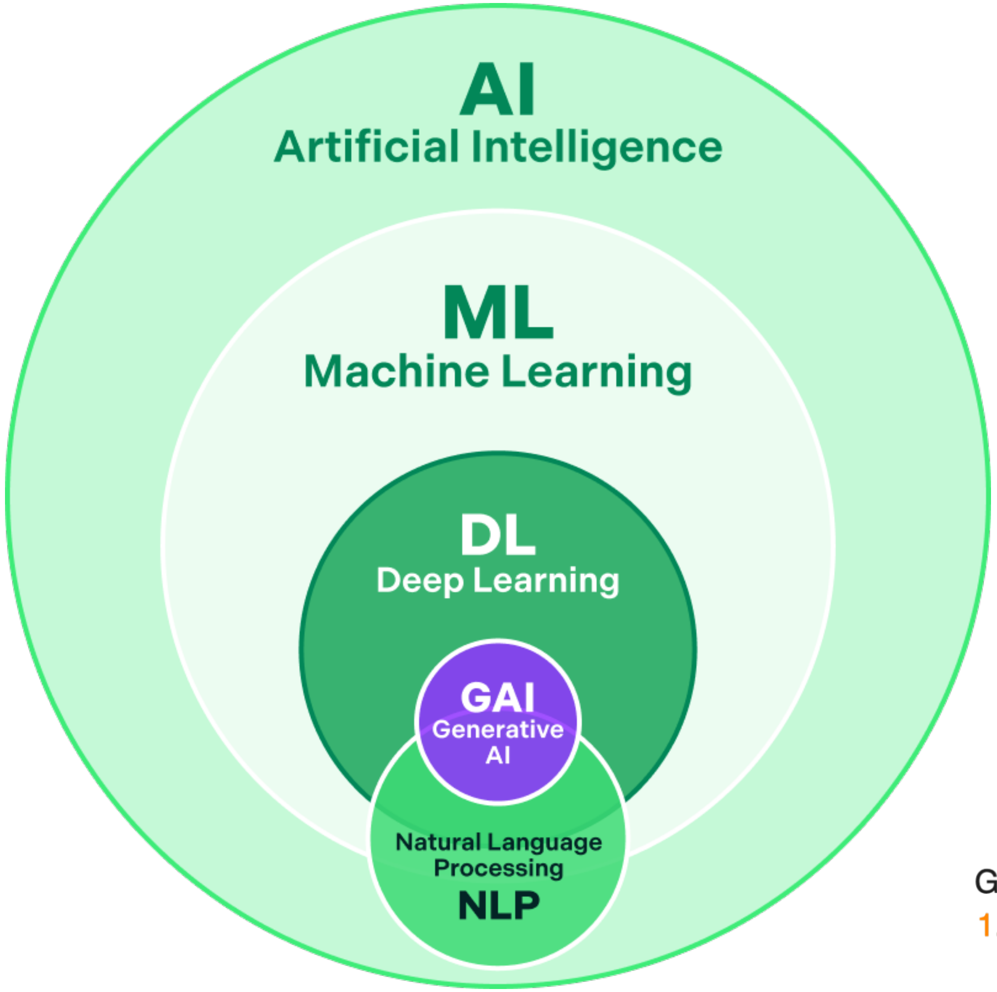
Foundation Models



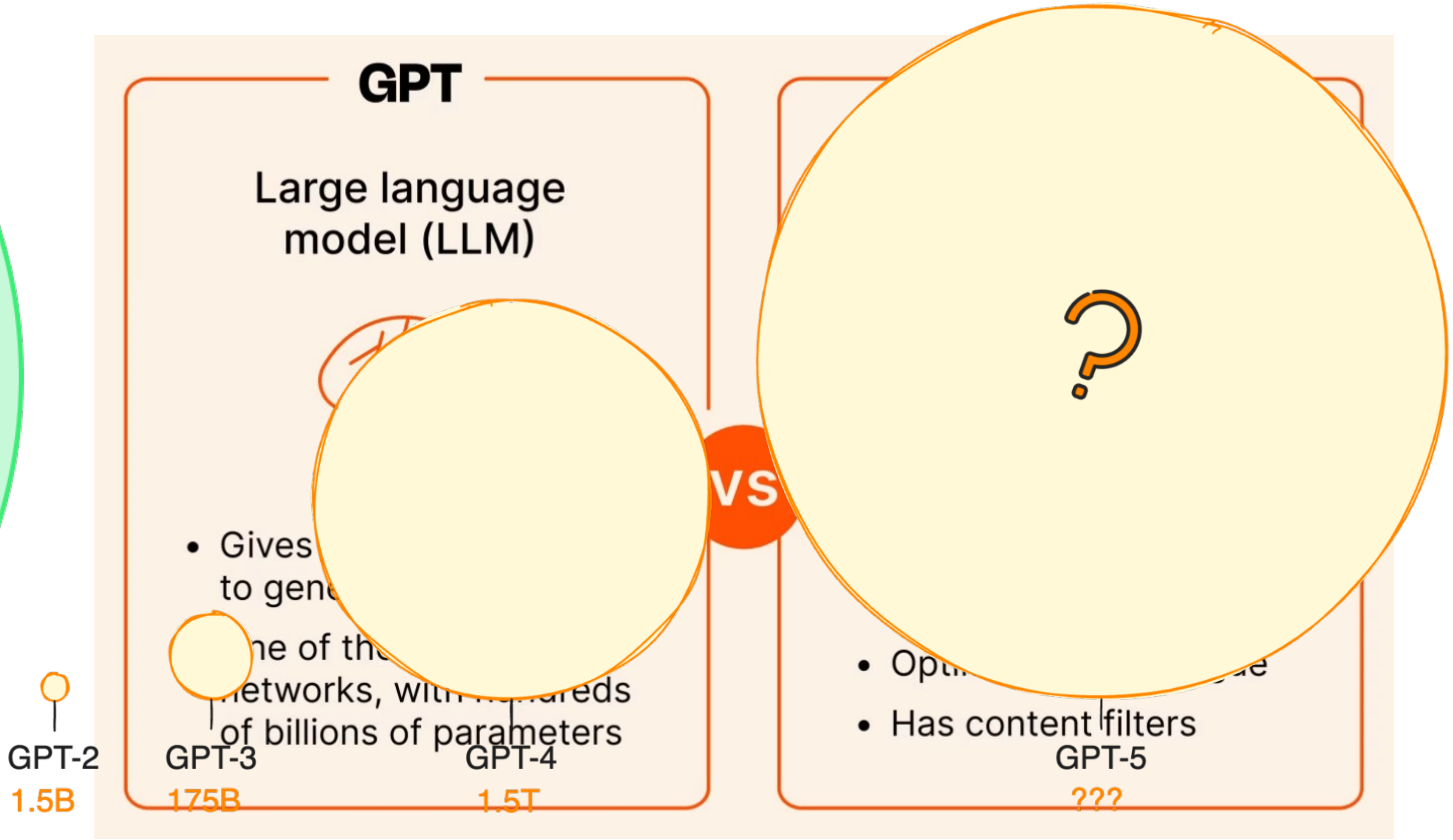
Core (general) task

https://doi.org/10.1007/s10586-023-04203-7

FM: Large Language Models



GPT Generative Pre-trained Transformer



Generalist/Generative tasks in medicine

Perspective

Foundation models for generalist medical artificial intelligence

https://doi.org/10.1038/s41586-023-05881-4

Received: 3 November 2022

Accepted: 22 February 2023

Published online: 12 April 2023

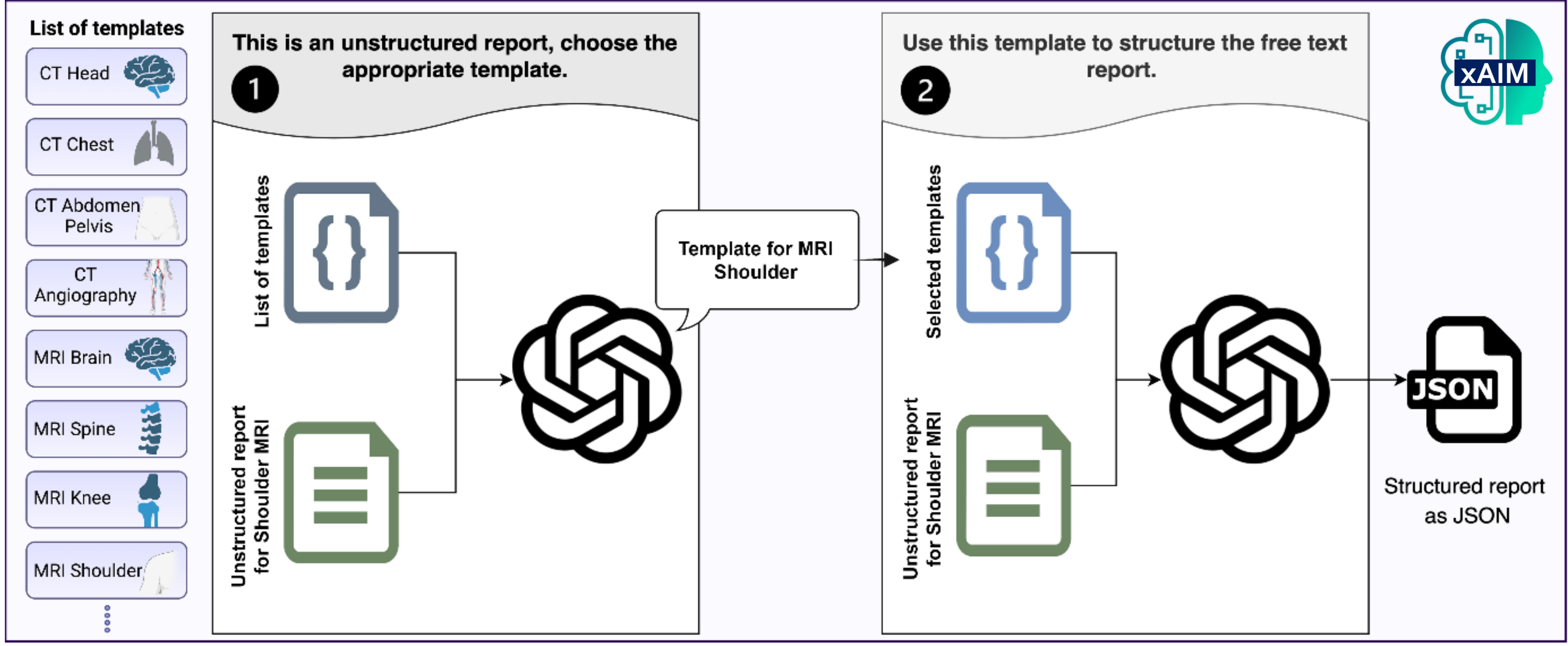
Check for updates

Michael Moor[1,6], Oishi Banerjee[2,6], Zahra Shakeri Hossein Abad[3], Harlan M. Krumholz[4], Jure Leskovec[1], Eric J. Topol[5,7] ✉ & Pranav Rajpurkar[2,7] ✉

The exceptionally rapid development of highly flexible, reusable artificial intelligence (AI) models is likely to usher in newfound capabilities in medicine. We propose a new paradigm for medical AI, which we refer to as generalist medical AI (GMAI). GMAI models will be capable of carrying out a diverse set of tasks using very little or no task-specific labelled data. Built through self-supervision on large, diverse datasets, GMAI will flexibly interpret different combinations of medical modalities, including data from imaging, electronic health records, laboratory results, genomics, graphs or medical text. Models will in turn produce expressive outputs such as free-text explanations, spoken recommendations or image annotations that demonstrate advanced medical reasoning abilities. Here we identify a set of high-impact potential applications for GMAI and lay out specific technical capabilities and training datasets necessary to enable them. We expect that GMAI-enabled applications will challenge current strategies for regulating and validating AI devices for medicine and will shift practices associated with the collection of large medical datasets.

https://www.nature.com/articles/s41586-023-05881-4

Generalist/Generative tasks: based on LLMs



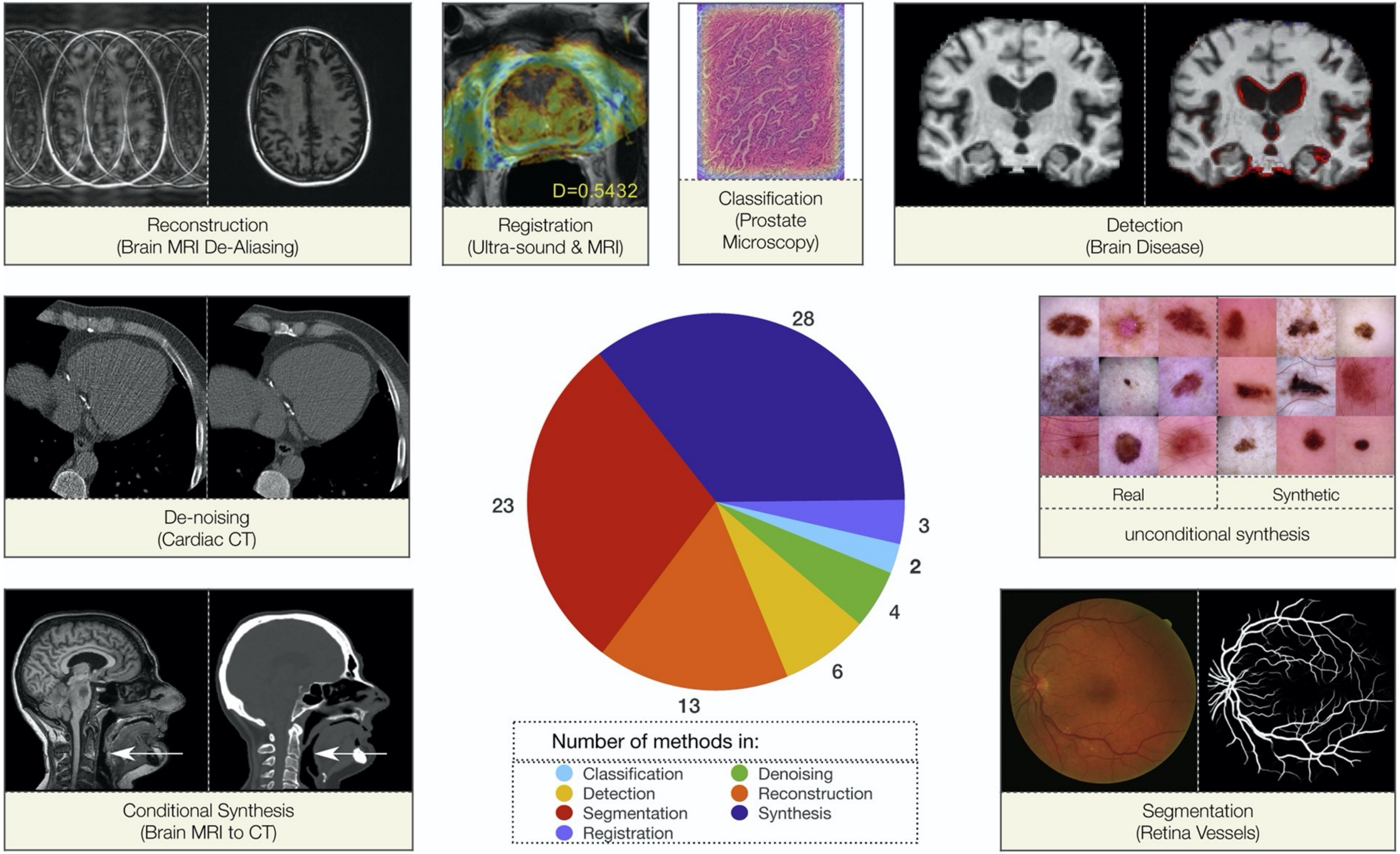
1 Free Text	HISTORY: A 42-year-old female presents with shoulder pain. COMPARISON: None available. TECHNIQUE: A 3T MRI of the right shoulder was conducted, using a variety of sequences, including axial proton density with fat saturation, axial GRE, coronal proton density with fat saturation, coronal T2 with fat saturation, coronal T1, sagittal T2 with fat saturation, and sagittal T1. FINDINGS: The MRI reveals preserved articular cartilage and hematopoietic bone marrow. The acromioclavicular joint displays mild to moderate degenerative changes, with inferior joint capsule hypertrophy and enthesophytes causing mass effect on the underlying supraspinatus tendon. The coracoacromial ligament appears mildly thickened. Bone marrow edema is present within the clavicle, likely due to osteoarthritis, but may also be indicative of early distal clavicular osteolysis. A small amount of fluid and edema can be seen within the subacromial/subdeltoid bursa, which may suggest bursitis. No glenohumeral joint effusion is observed. Mild tendinopathy is present in the supraspinatus, infraspinatus, and subscapularis tendons. The supraspinatus tendon also exhibits mild bursal-sided fraying. The teres minor tendon appears normal. The biceps tendon and anchor are intact. Evaluation of the labrum is limited due to lack of joint fluid. A tear of the superior labrum, posterosuperiorly, is identified. Incidentally noted are breast implants. IMPRESSION: Posterosuperior superior labral tear. Bone marrow edema in the clavicle, likely secondary to osteoarthritis but possibly due to early changes of distal clavicular osteolysis. Mild tendinopathy in the supraspinatus, infraspinatus, and subscapularis tendons. Mild to moderate degenerative changes in the acromioclavicular joint.

2 Structured

INDICATION	A 42-year-old female presents with shoulder pain.	
TECHNIQUE	A 3T MRI of the right shoulder was conducted, using a variety of sequences, including axial proton density with fat saturation, axial GRE, coronal proton density with fat saturation, coronal T2 with fat saturation, coronal T1, sagittal T2 with fat saturation, and sagittal T1.	
COMPARISON	None available.	
FINDINGS	**SUPRASPINATUS**	Mild tendinopathy with mild bursal-sided fraying.
	INFRASPINATUS	Mild tendinopathy.
	TERES MINOR	Normal.
	SUBSCAPULARIS	Mild tendinopathy.
	ROTATOR CUFF	Not mentioned.
	ACROMIO-CLAVICULAR JOINT	Mild to moderate degenerative changes, with inferior joint capsule hypertrophy and enthesophytes causing mass effect on the underlying supraspinatus tendon.
	BICEPS TENDON	Intact.
	GLENOID LABRUM	A tear of the superior labrum, posterosuperiorly.
	GLENOHUMERAL JOINT	No joint effusion.
	HYALINE CARTILAGE	Preserved.
	BONE MARROW	Bone marrow edema is present within the clavicle, likely due to osteoarthritis, but may also be indicative of early distal clavicular osteolysis.
	SOFT TISSUES	The coracoacromial ligament appears mildly thickened. Small amount of fluid and edema within the subacromial/subdeltoid bursa suggestive of bursitis. Incidentally noted breast implants.
IMPRESSION	Posterosuperior superior labral tear. Bone marrow edema in the clavicle, likely secondary to osteoarthritis but possibly due to early changes of distal clavicular osteolysis. Mild tendinopathy in the supraspinatus, infraspinatus, and subscapularis tendons. Mild to moderate degenerative changes in the acromioclavicular joint.	

https://doi.org/10.1148/radiol.223312

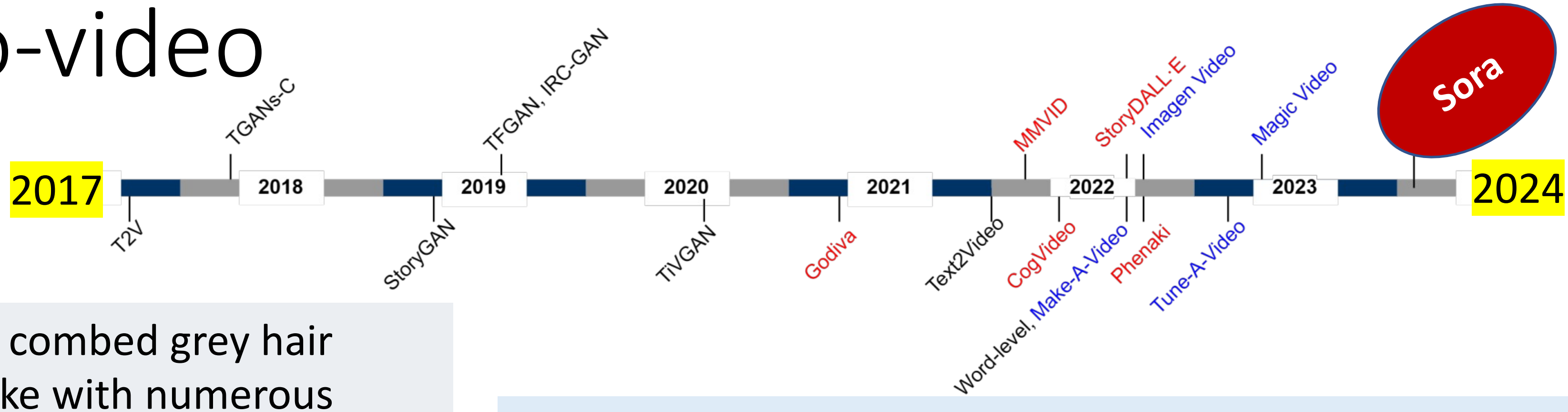
Generative tasks: imaging (both A2.0 and A3.0 approaches)



https://doi.org/10.1016/j.artmed.2020.101938

Current frontiers: generative text-to-video

Impressive… but still "statistics"
(no evidence of "meaning appropriation")



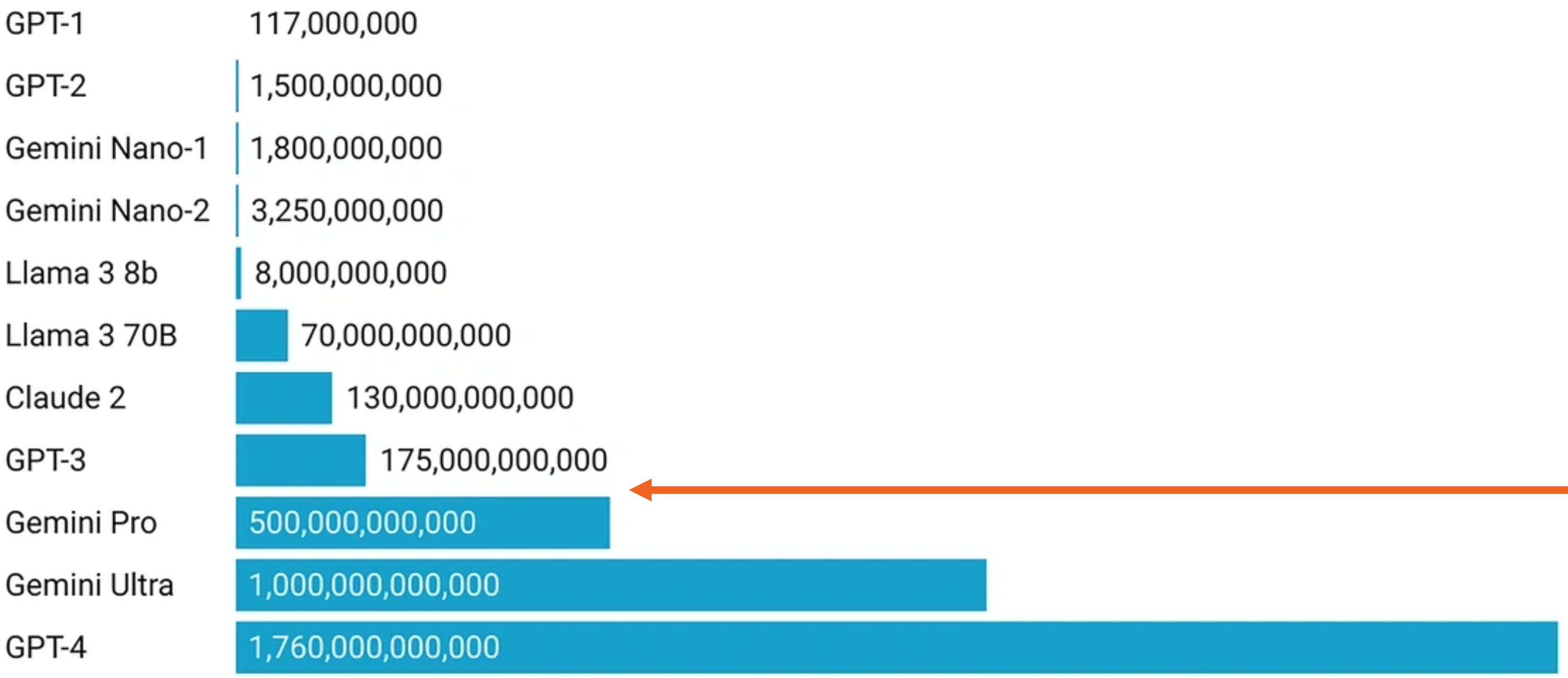
Prompt: A grandmother with neatly combed grey hair stands behind a colorful birthday cake with numerous candles at a wood dining room table, expression is one of pure joy and happiness, with a happy glow in her eye. She leans forward and blows out the candles with a gentle puff, the cake has pink frosting and sprinkles and the candles cease to flicker, the grandmother wears a light blue blouse adorned with floral patterns, several happy friends and family sitting at the table can be seen celebrating, out of focus. The scene is beautifully captured, cinematic, showing a 3/4 view of the grandmother and the dining room. Warm color tones and soft lighting enhance the mood..

Prompt: Archeologists discover a generic plastic chair in the desert, excavating and dusting it with great care.

Parameters in foundation models

Parameters in Selected AI Models

Some of these figures are estimates. Newer models are many times larger than their predecessors.



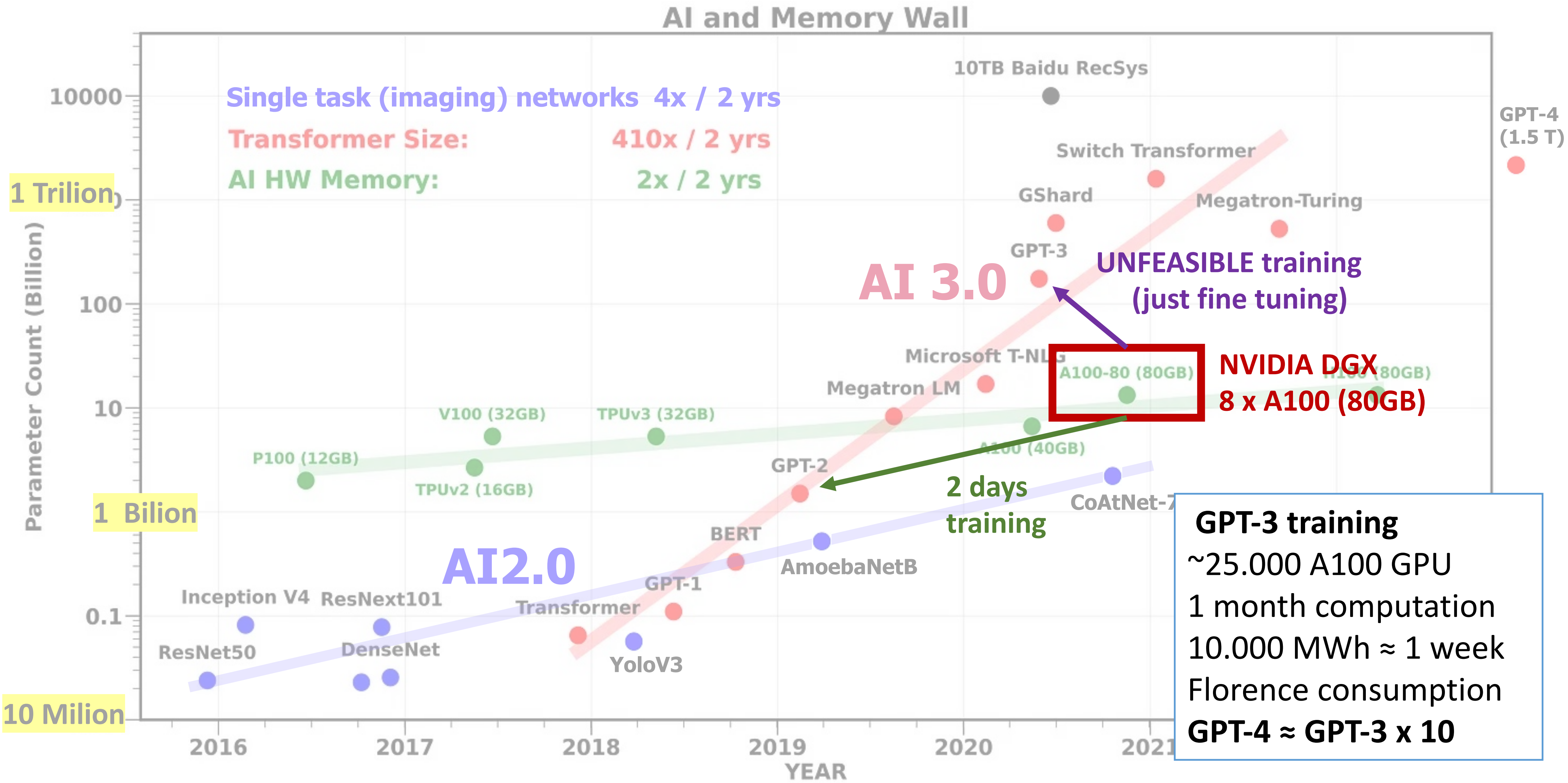
At this point in time, we seem to have reached a plateau.

Today models can use reasoning modules and divide up tasks (Mixture of Experts).

GPT-4o

https://explodingtopics.com/blog/gpt-parameters

Parameter trends



GPT-3 training

~25.000 A100 GPU

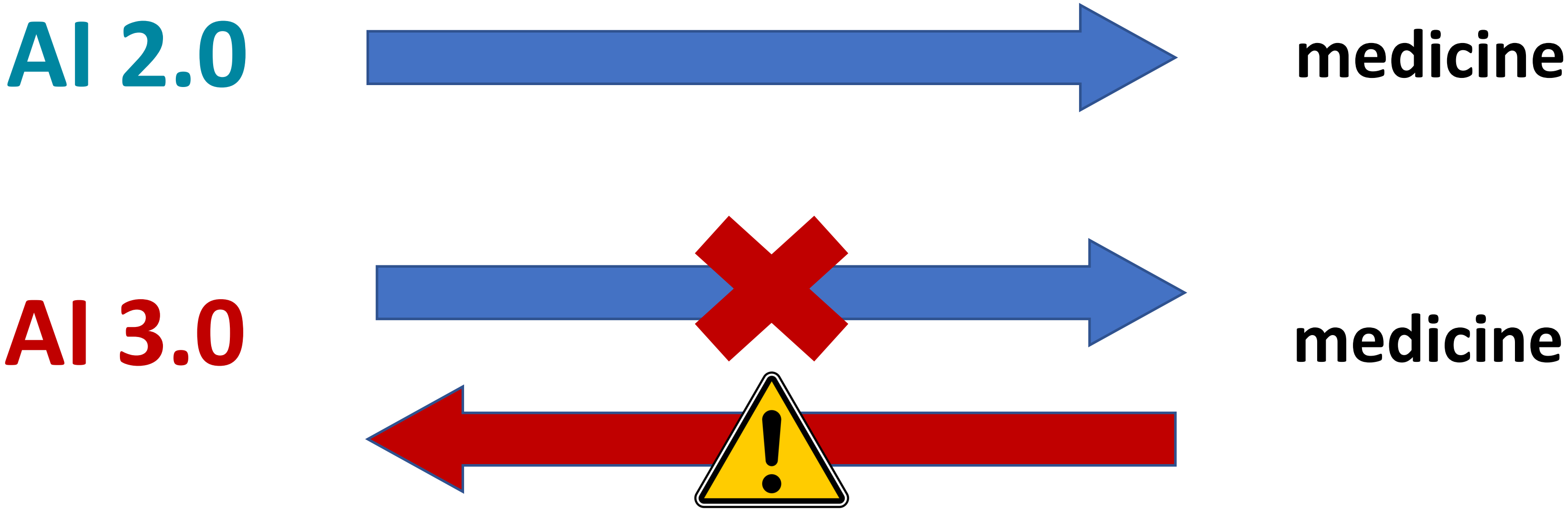
1 month computation

10.000 MWh ≈ 1 week Florence consumption

GPT-4 ≈ GPT-3 x 10

https://hua-hua.org/large-

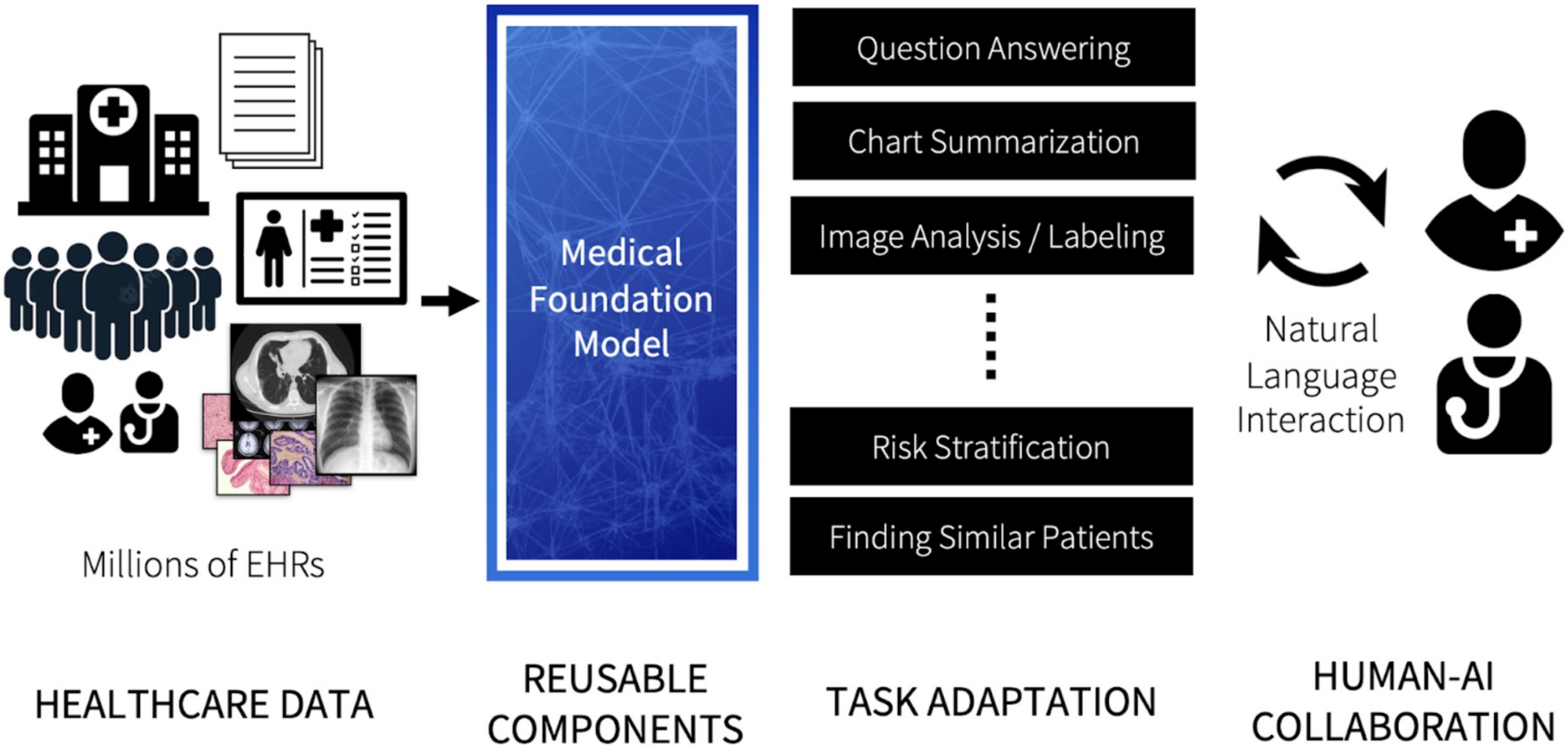
AI in medicine: current “directions”



Other higher complexity dimensions still to be fully explored, e.g.:

- **Multimodality** (what, where, when)
- **Preventive medicine** (patient timeline)
- **Affordable size Foundation Models**

Large language models and foundation models in medicine



https://hai.stanford.edu/news/how-foundation-models-can-advance-ai-healthcare

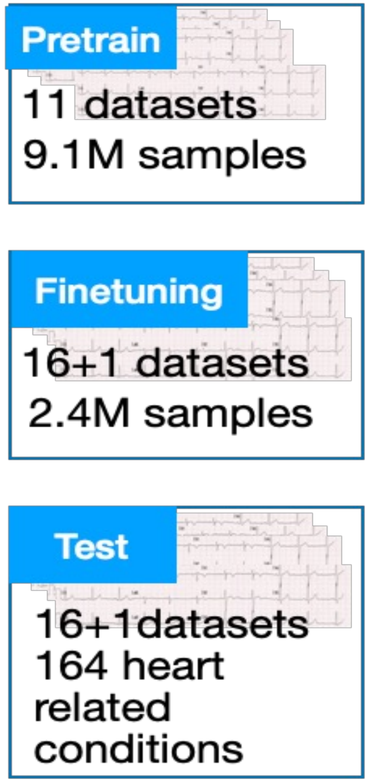
A new cardiology foundation model

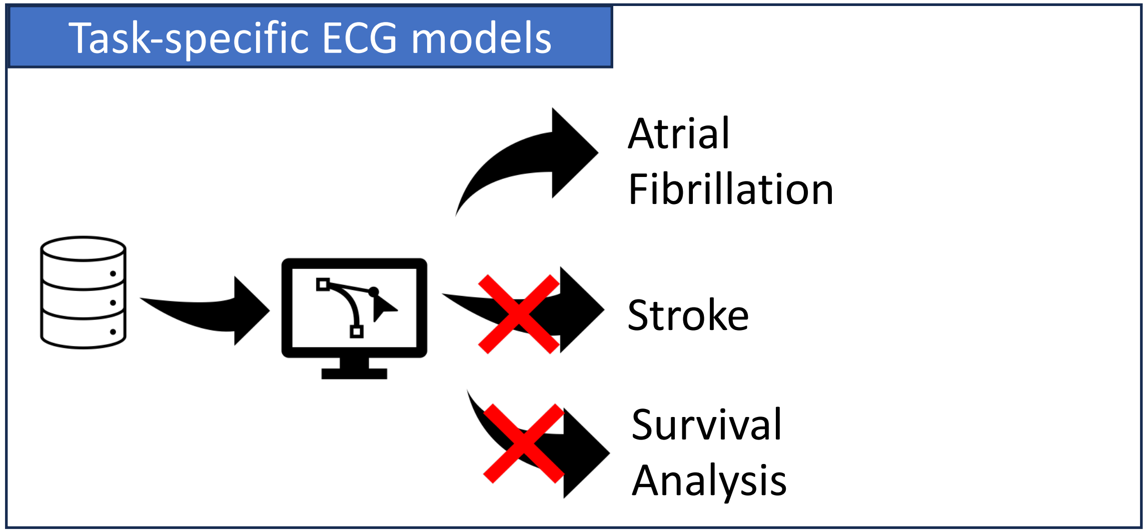
medRχiv
THE PREPRINT SERVER FOR HEALTH SCIENCES

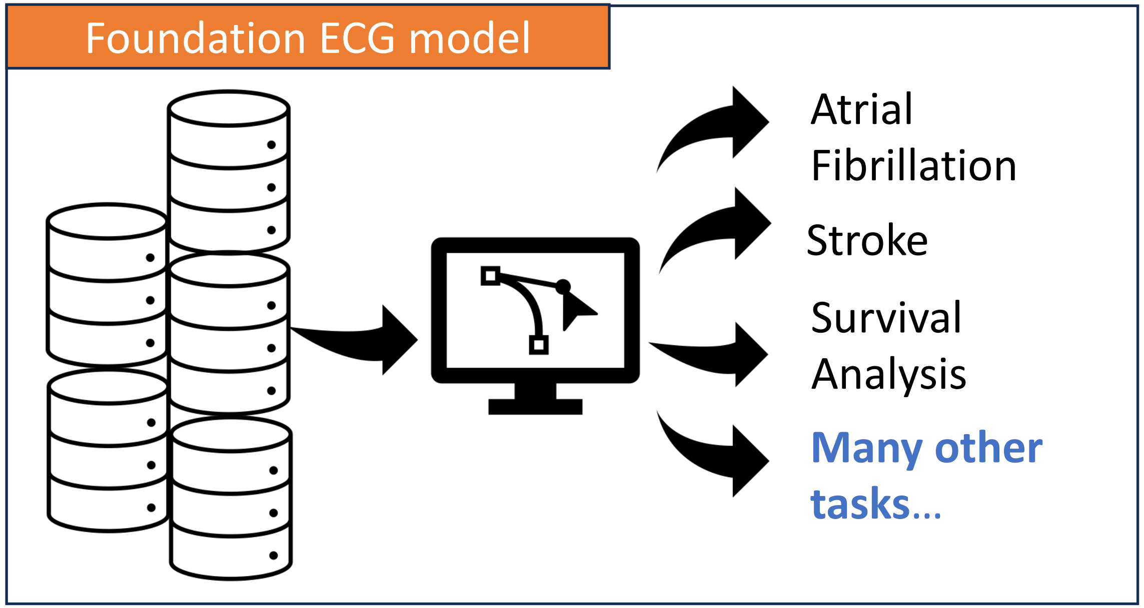
HuBERT-ECG: a self-supervised foundation model for broad and scalable cardiac applications

Edoardo Coppola[1], Mattia Savardi[2], Mauro Massussi[2,3], Marianna Adamo[2,3], Marco Metra[2,3], Alberto Signoroni[2]

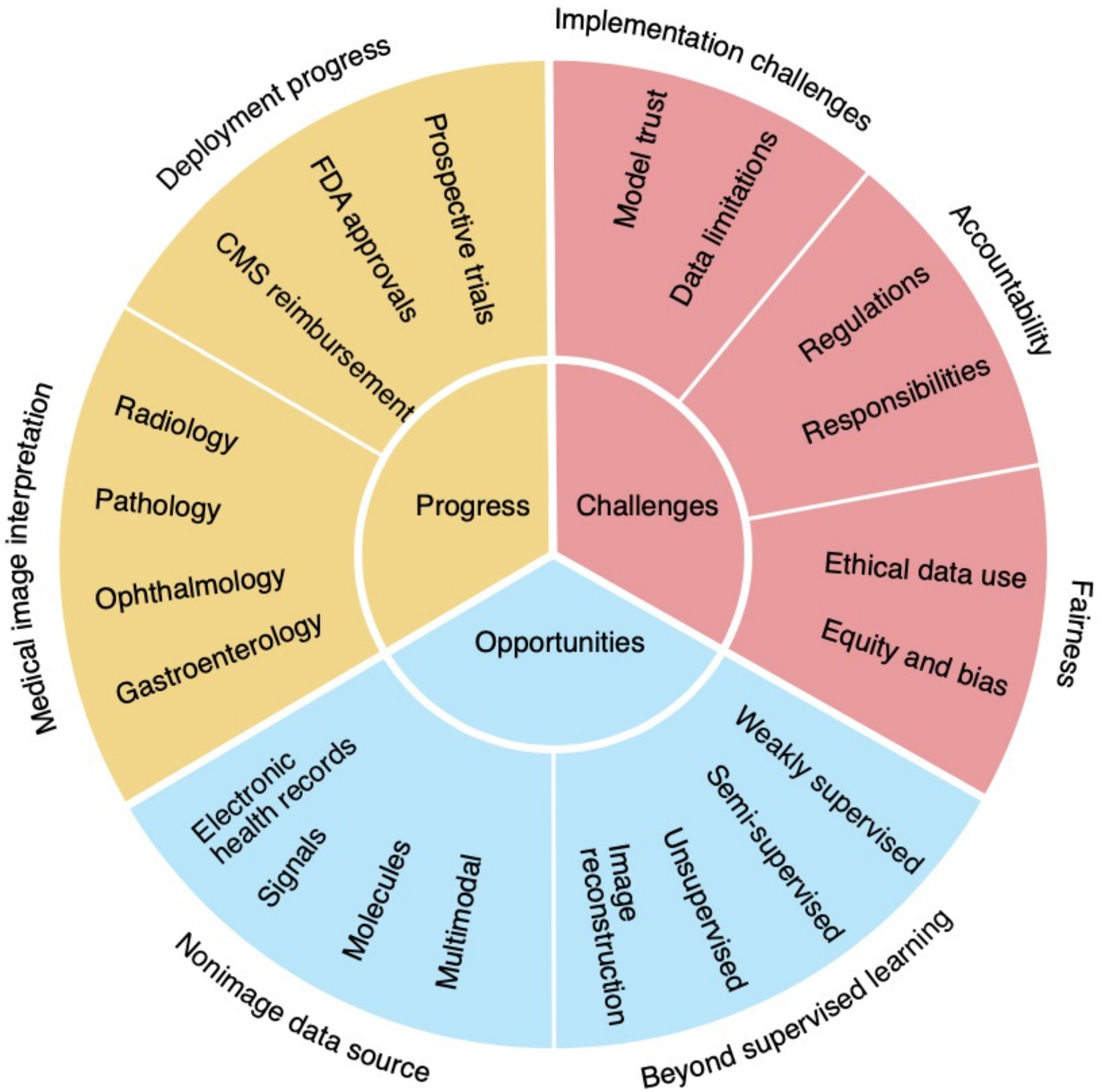
- **9.1 million** unlabeled and labeled ECGs from multiple sources
- **Self-supervised pre-training** on unlabeled ECGs
- Facilitated fine-tuning on **many specific clinical tasks**: testing on many scenarios and **164 clinical conditions**







Progress, challenges and opportunities for AI in health



- Progresses
  - Almost all medical specialties have been impacted by AI: not only diagnosis but also risk prediction and treatment
  - Not only experimental: growing market of deployed (FDA or CE certified) AI products
- Opportunities
  - Multimodal data fusion (not only images)
  - Not only supervised learning: many possible “learning” paradigms (full data labeling can be too expensive/time consuming)
  - Not only human vs machine: collaborative approaches, human-in-the-loop
- Challenges
  - Implementation: data/label acquisition costs, dataset dimensions, data biases
  - Regulation: locked vs continual learning, accountability, explainability
  - Ethics: fairness, privacy, equity, do not harm

https://doi.org/10.1038/s41591-021-01614-0