

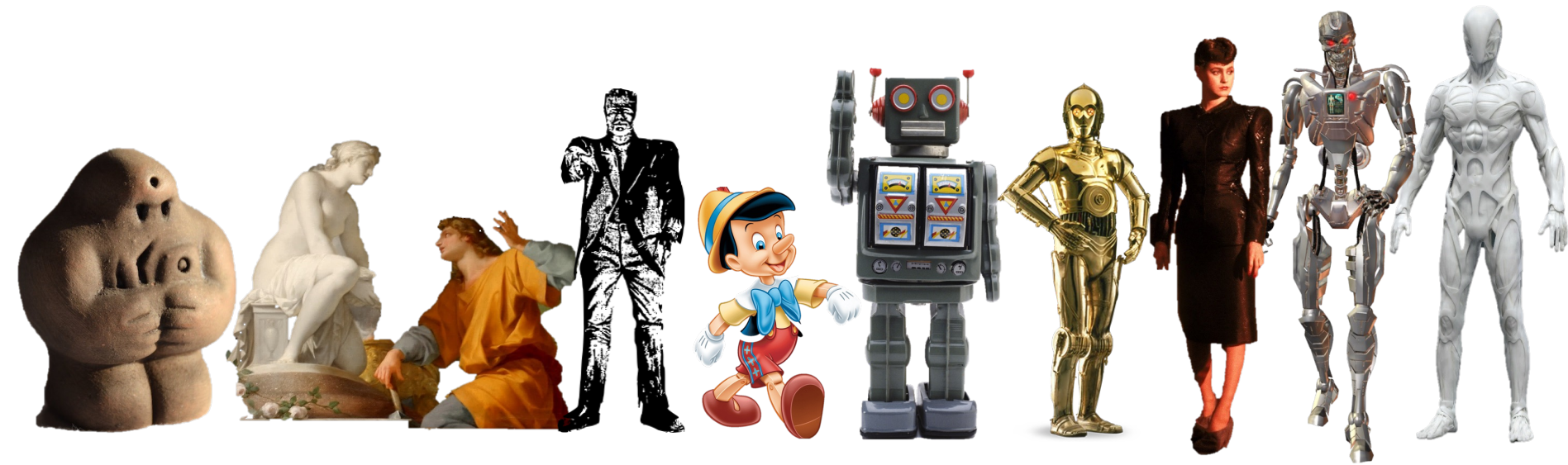


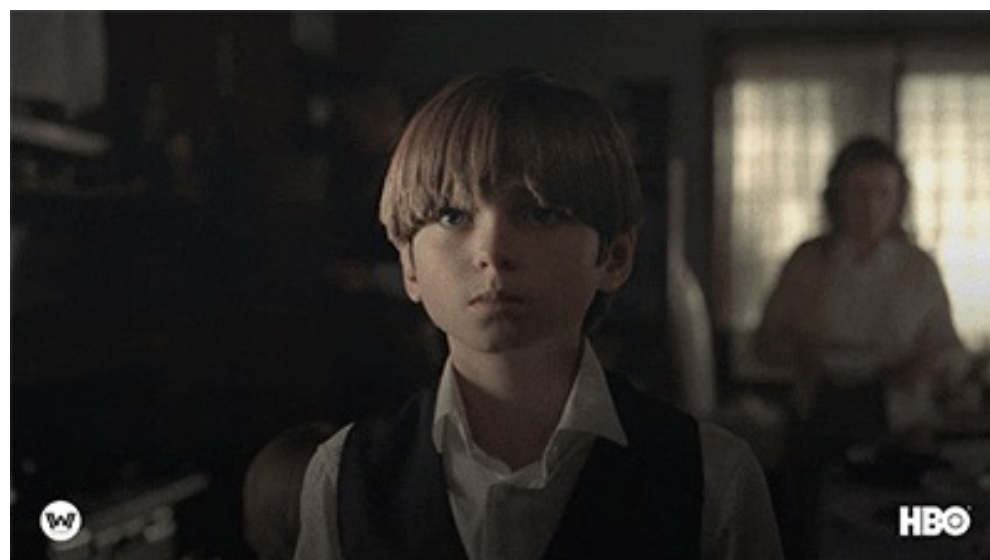
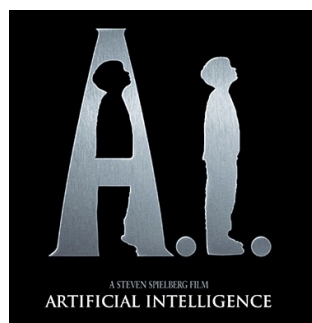
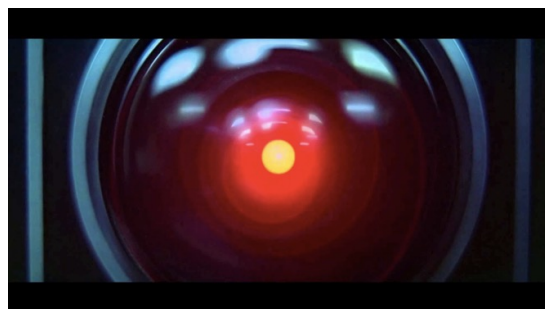
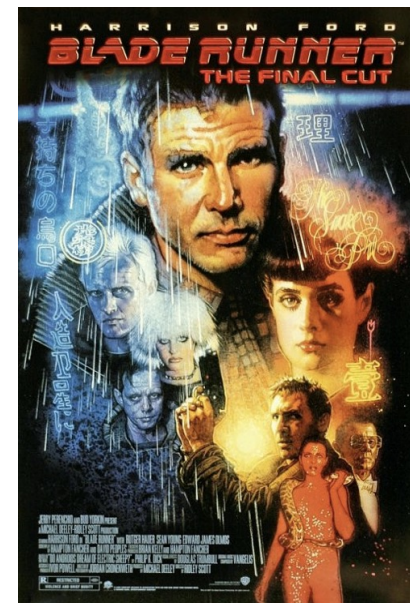
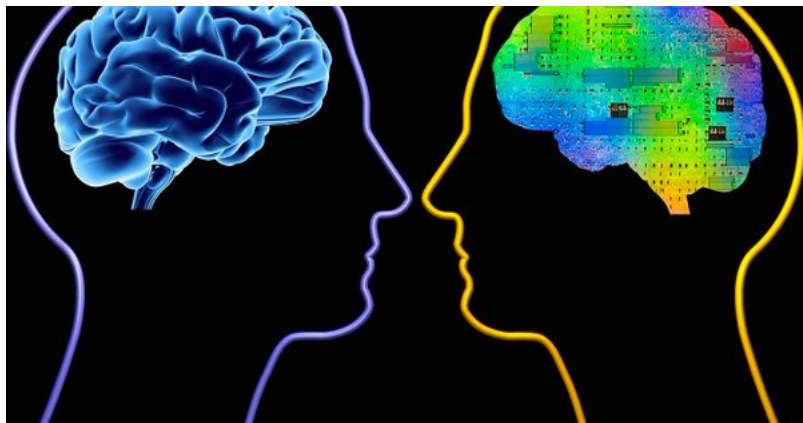
DEEP LEARNING DANS LE DOMAINE MÉDICAL

Master 1 Recherche Biomédicale

E. Reynaud, Labo. EMC, Univ. Lyon 2

Intelligence Artificielle : de l'histoire ancienne...



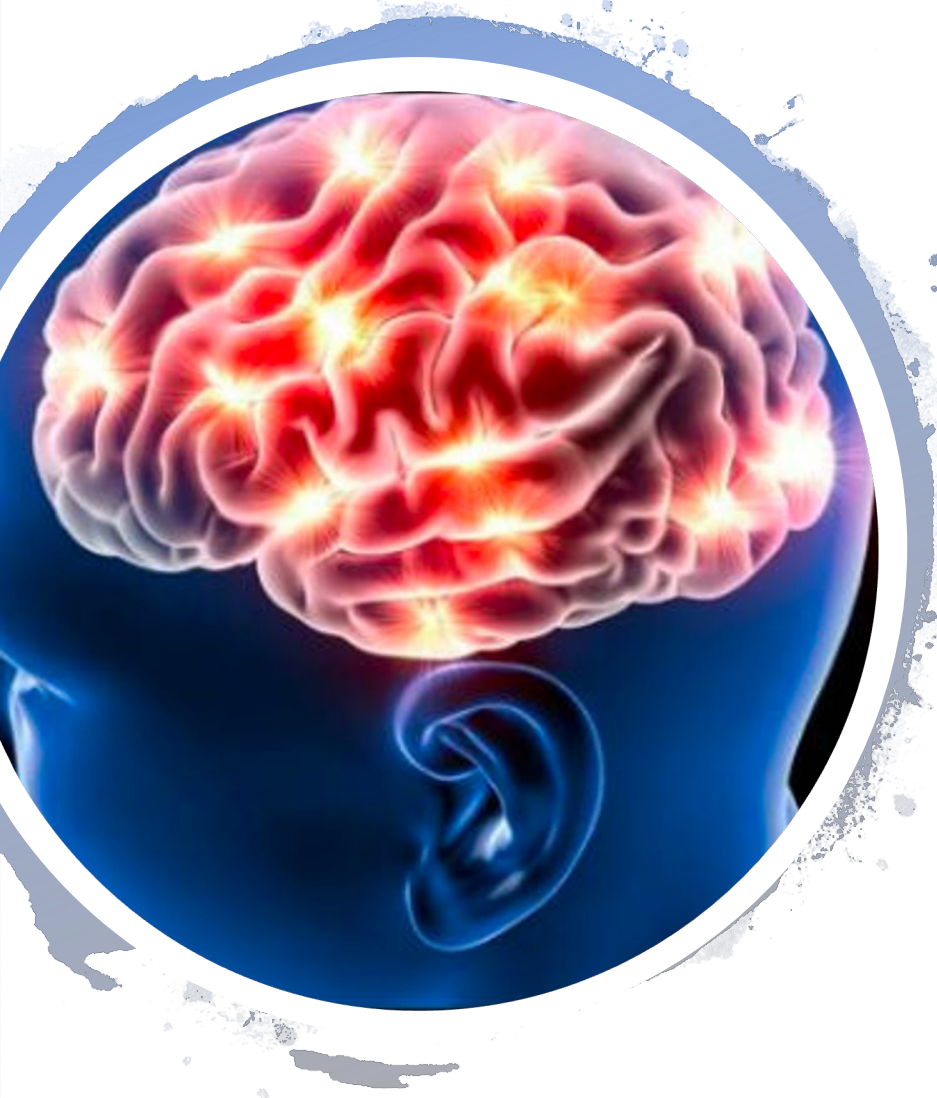


Artificial intelligence

AI wants to build powerful machines
than can think / act intelligently

What is the most intelligent machine ?





Powerful
Portable
Low cost in energy
Adaptative
Can learn
Small
Self-repair
etc



**Artificial intelligence may take your job,
so political leaders need to start
doing theirs**

10 avril 2019, 22:10 CEST

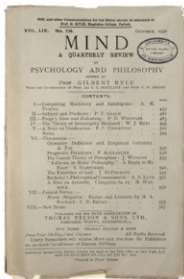
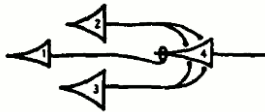
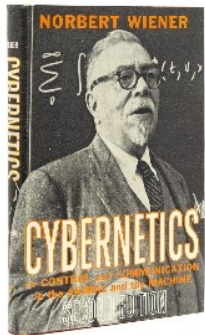
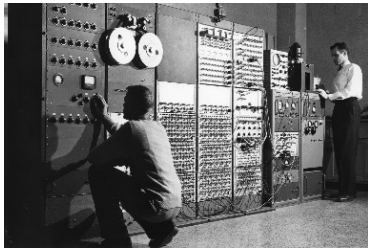
**Elon Musk WARNING: Artificial
Intelligence could be an
'IMMORTAL DICTATOR'**

**Hinton, LeCun, Bengio : la
« conspiracy » du deep learning**

ial intelligence (AI) could bring with it an
, that they do not "have to be evil to destroy

LES REBELLES DE LA SCIENCE 9/10. Au début des années 2000, contre l'opinion de leurs pairs,
trois chercheurs en intelligence artificielle ont remis au goût du jour une voie jugée sans avenir :
les réseaux de neurones.

Artificial Intelligence



Nathaniel Rochester Marvin L. Minsky John McCarthy
Oliver G. Selfridge Ray Solomonoff Trenchard More Claude E. Shannon

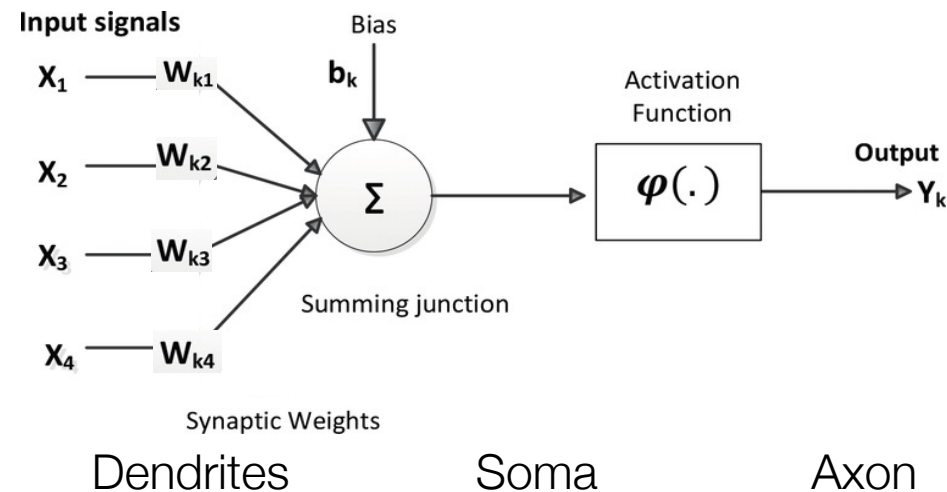
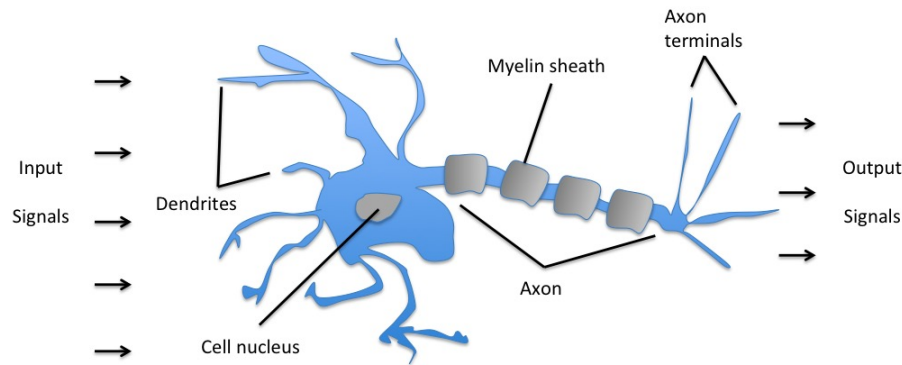
August 1956

"Every aspect of learning or any other feature of intelligence can be so precisely described that a machine can be made to simulate it. » (Dartmouth Proposal, 1956)



Artificial Neural Networks aka ANN

- McCulloch & Pitts, 1943 : MCP neuron model

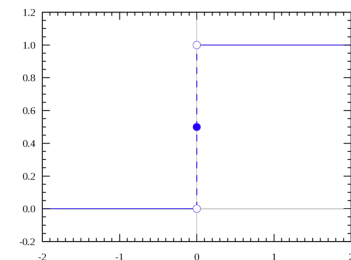


MCP neuron k :

$$\text{Activation } A = x_1 \cdot w_{k1} + x_2 \cdot w_{k2} + x_3 \cdot w_{k3} + x_4 \cdot w_{k4} - b_k$$

$$\text{Output } Y_k = \varphi(A)$$

$$\text{Transfer function : } \varphi(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$



Artificial Neural Networks aka ANN

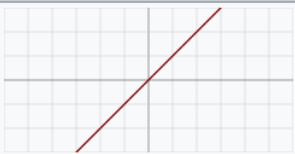
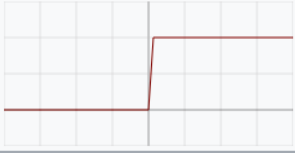
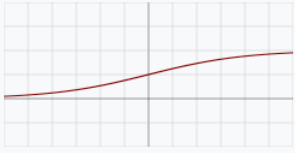
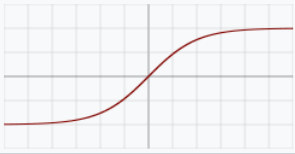
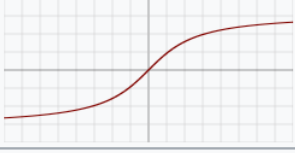

For neuron k with n pre-synaptic inputs :

$$y_k = \varphi \left(\sum_{i=1}^n x_i w_{ki} - \theta_k \right)$$

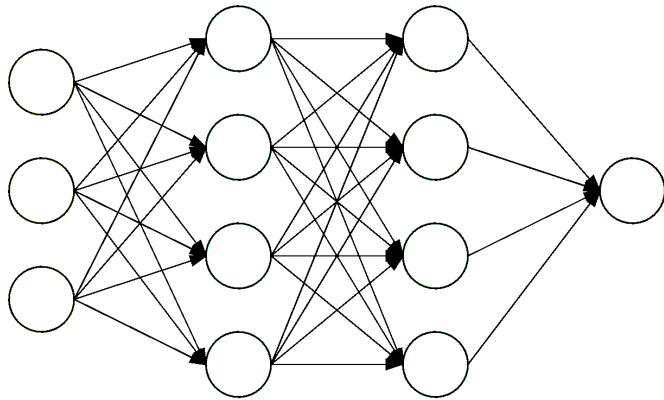
In code :

```
for i = 1 to n :  
    sum = sum + xi * wki  
sum = sum - θ  
out = activation_function(sum)
```

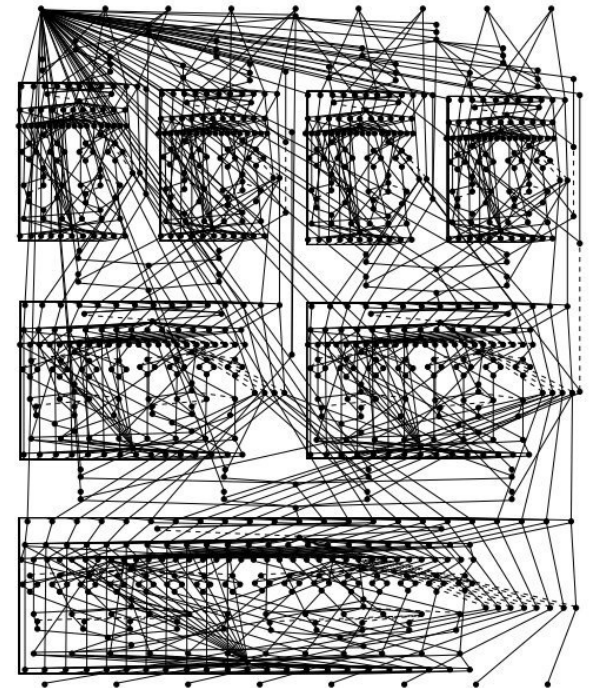
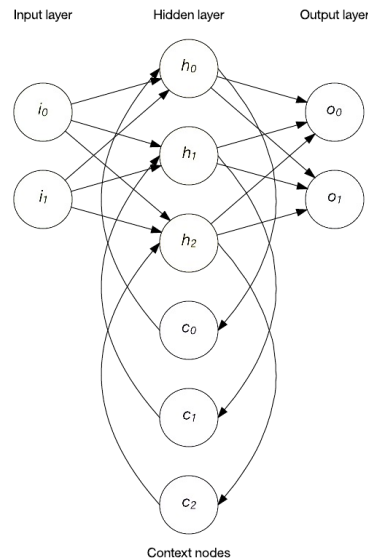
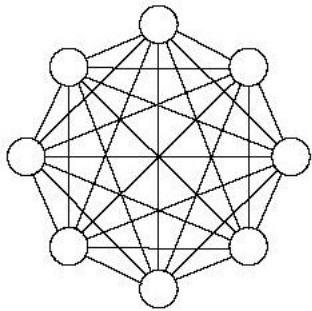
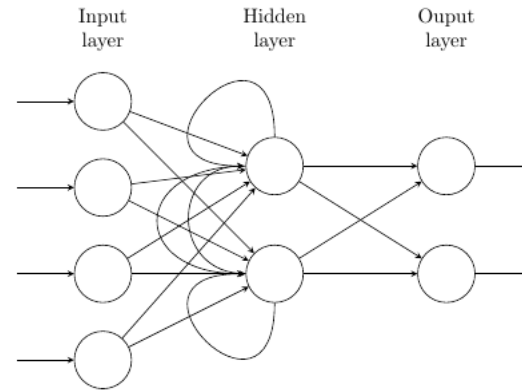
Activation function

Name ◆	Plot ◆	Equation ◆
Identity		$f(x) = x$
Binary step		$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
Logistic (a.k.a. Sigmoid or Soft step)		$f(x) = \sigma(x) = \frac{1}{1 + e^{-x}} \text{ [1]}$
TanH		$f(x) = \tanh(x) = \frac{(e^x - e^{-x})}{(e^x + e^{-x})}$
ArcTan		$f(x) = \tan^{-1}(x)$
Rectified linear unit (ReLU) ^[15]		$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$

Artificial Neural Networks aka ANN



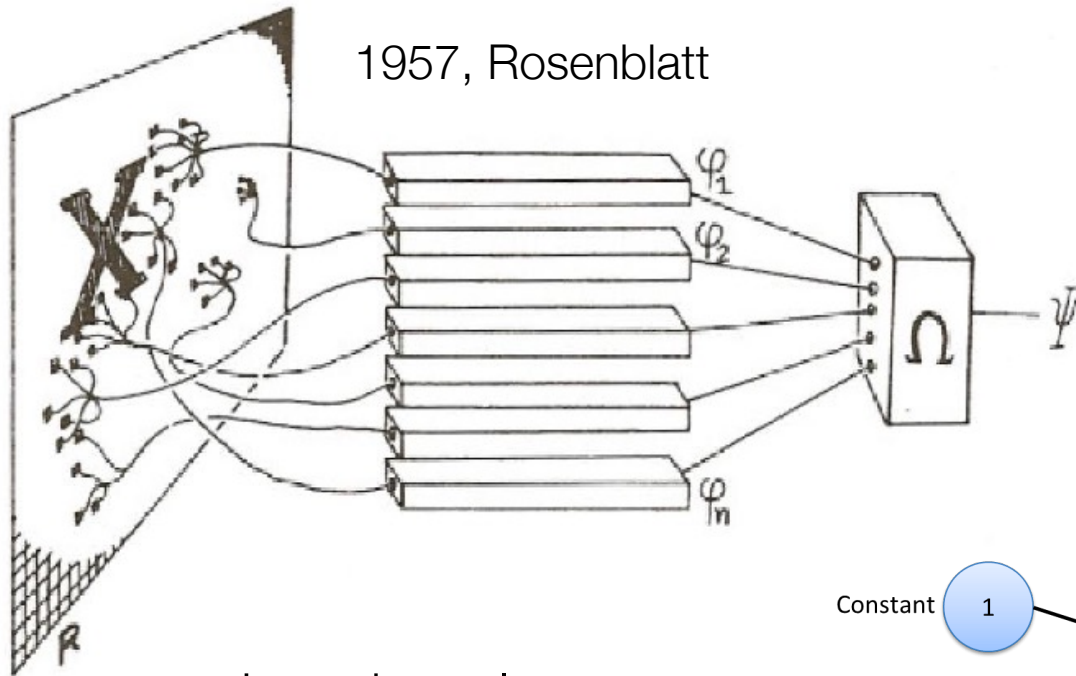
input layer hidden layer 1 hidden layer 2 output layer



Artificial Neural Networks aka ANN

- Simple units
 - Parallel Distributed Processing
 - Units in a network : « knowledge » can emerge
 - « Knowledge » lies in the synaptic weights
-
- ANN need to be trained in most cases : supervised learning
 - Training = weights adjustments
 - Error = difference between the network outputs and the desired outputs for a set of stimuli

The father of all ANN^w: The Perceptron



Function to learn :

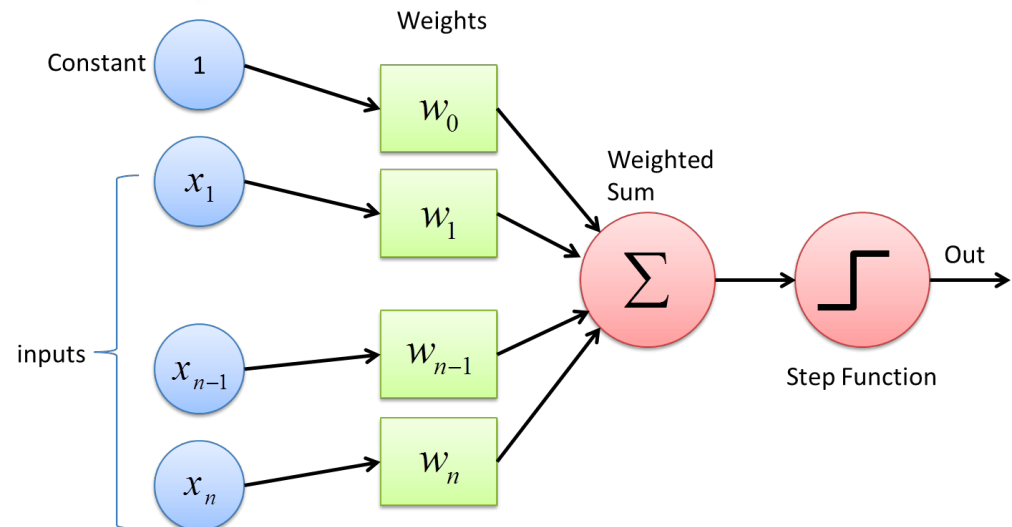
Letter	d :
X	1
T	0

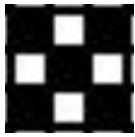
Learning rule :

$$w_j(t+1) = w_j(t) + \eta (d - y) x$$

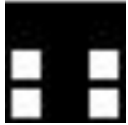
with :

- d : desired output
- y : network output
- η : learning rate





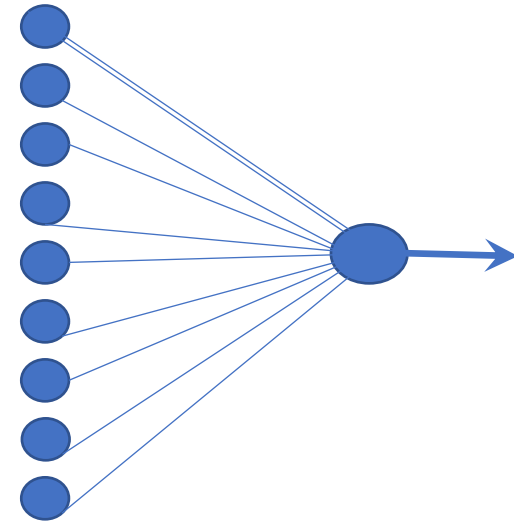
$X1 = [1\ 0\ 1\ 0\ 1\ 0\ 1\ 0\ 1]$



$X2 = [1\ 1\ 1\ 0\ 1\ 0\ 0\ 1\ 0]$

$W_{t=0} = [0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0]$

$\eta = 0.5$



```
X1 = [1 0 1 0 1 0 1 0 1];
X2 = [1 1 1 0 1 0 0 1 0];
```

```
W=[0 0 0 0 0 0 0 0 0];
%W = randn(9,1);
disp(W);
```

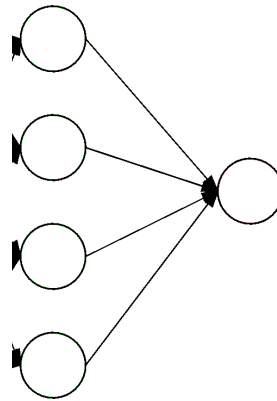
```
eta = 0.5;
d1 = 1;
d2 = 0;
```

```
i=1;
while i < 10
    iter = sprintf('EPOQUE n. %d',i);
    disp(iter)
    A=sum(X1.*W);
    if A>=0 y = 1 ;
    else y = 0;
    end
    fprintf('Presentation X1 lettre X.,
    W=W + (eta * (d1-y))*X1;
    disp(W');

    A=sum(X2.*W);
    if A>=0 y = 1 ;
    else y = 0;
    end
    fprintf('Presnetation X2 lettre T., output reseau: %d\n',y);
    W=W + (eta * (d2-y))*X2;
    disp(W');
    i=i+1;
end
```

	d	Σ	y	d-y	eta*(d-y)	nouveaux poids
X	1	0	1	0	0	[0 0 0 0 0 0 0 0 0]
T	0	0	1	-1	-0.5	[-0.5 -0.5 -0.5 0 -0.5 0 0 -0.5 0]
X	1	-1.5	0	1	0.5	[0 -0.5 0 0 0 0 0.5 -0.5 0.5]
T	0	-1	0	0	0	[0 -0.5 0 0 0 0 0.5 -0.5 0.5]
X	1	1	1	0	0	[0 -0.5 0 0 0 0 0.5 -0.5 0.5]

The Perceptron complexified : MLPs



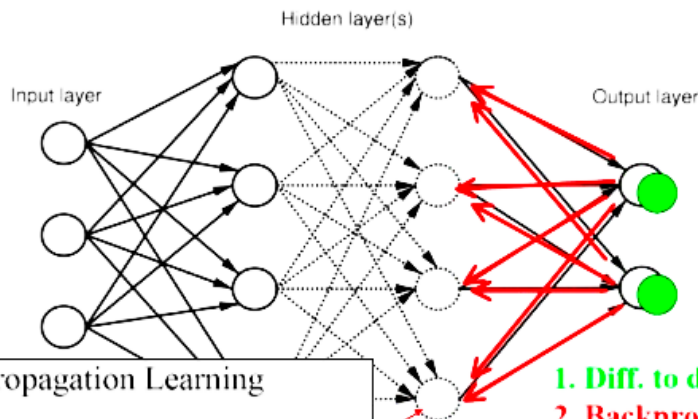
MLP = perceptron

- + hidden layers to complexify decision regions
- + activation function : sigmoïd / tanh
- + backpropagation 😊

Backpropagation simplified

- 1 – From Layer 1 to Output layer : input is propagated
- 2 – Network output is computed
- 3 – Error is computed (from desired -output)
- 4 – Weights are updated from Output Layer to previous layers, Error is distributed backwards according to which weights contributed the most to this error, then to the error in the next layers etc until Layer 1

Backpropagation Learning



Backpropagation Learning

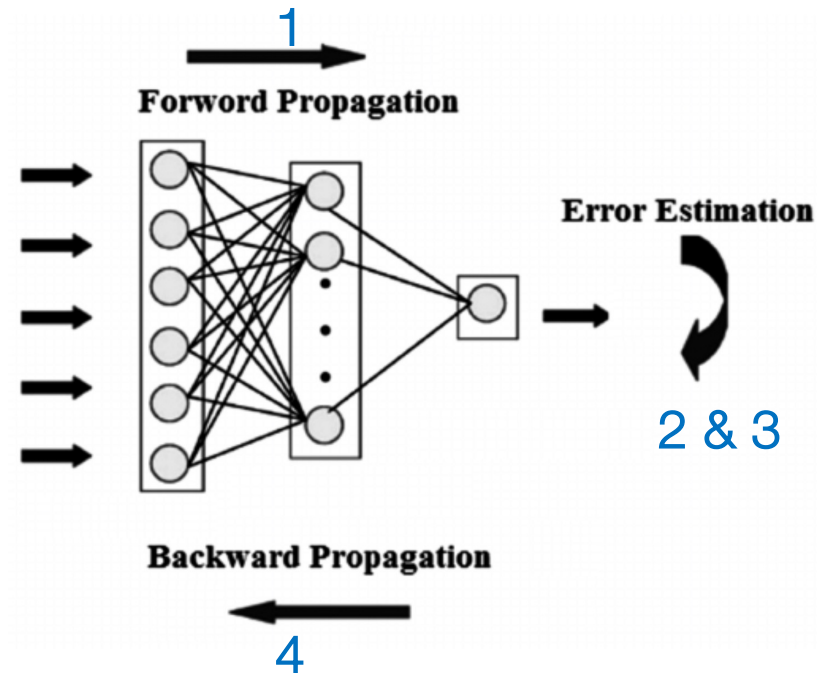
$$E_{out i} = d_{out i} - out_i$$

$$E_{total} = \sum_{i=1}^{num(n_{out})} E_{out i}^2$$

$$E_{hid i} = \sum_{k=1}^{num(n_{out})} E_{out k} \cdot w_{out i k}$$

$$diff_{hid i} = E_{hid i} \cdot (1 - out_{hid i}) \cdot out_{hid i}$$

1. Diff. to desired values
2. Backprop output layer



Deep Learning

- « Brain » / « Cognition » inspired
- MLP with many layers + modified architecture
- Hierarchical representation of information through the layers (see visual cortex for example)
- Modified neurons (eg. : ReLU units)
- Less weights, more learning !
- Krizhevsky et al. (2012): won the ImageNet challenge with AlexNet, similar to **LeNet (LeCun 1998)**



Yann LeCun est avec Yoshua Bengio.

27 mars, 12:27 · 🌐

Official citation from the Association for Computing Machinery for the 2018 Turing Award.



Yoshua Bengio



Geoffrey Hinton



Yann LeCun

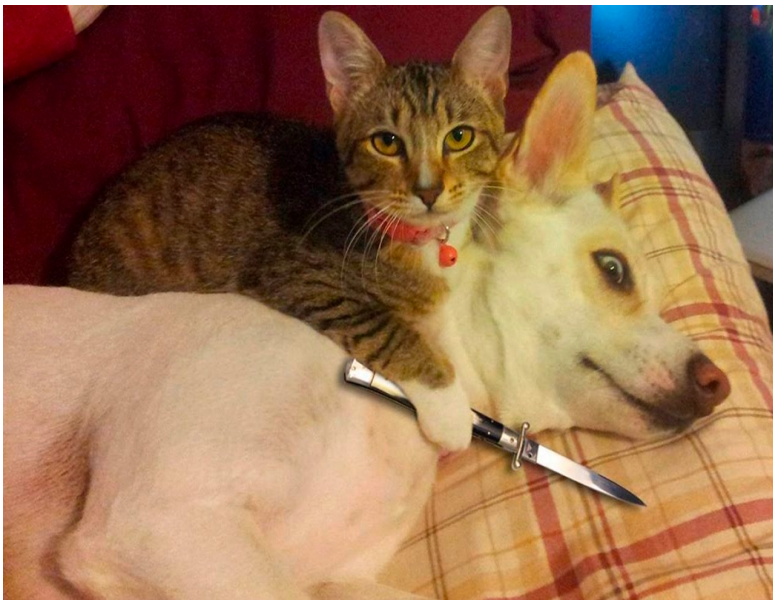
AWARDS.ACM.ORG

Fathers of the Deep Learning Revolution Receive ACM A.M. Turing Award

👍❤️👏 2,7 K

132 commentaires 468 partages

Computer and images



What we see

Image of dimension $N \times M$

Variables - red

picture x red x green x blue x

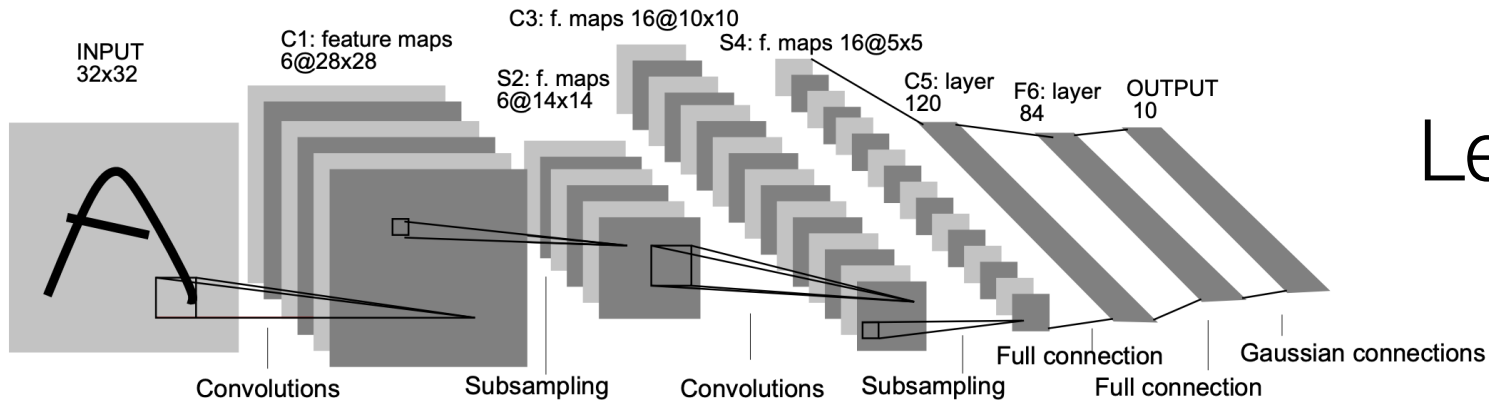
227x227 uint8

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1	125	127	126	125	122	123	122	120	119	120	118	118	119	120	122	128	152	164	163	163	163	162
2	125	125	124	124	122	122	122	119	116	116	117	118	118	118	119	130	158	164	165	165	162	162
3	128	123	122	121	121	123	122	118	113	114	115	116	117	117	118	129	159	165	167	165	162	163
4	128	122	120	120	119	120	120	118	115	116	115	114	116	117	116	137	164	170	168	167	162	164
5	129	124	122	121	119	121	120	119	117	118	117	114	113	114	116	142	168	171	170	168	166	167
6	128	125	123	122	119	118	118	118	116	118	117	120	116	115	115	146	172	172	170	171	171	170
7	125	125	123	122	121	118	118	116	114	116	114	115	113	113	119	153	168	175	174	173	173	173
8	125	124	125	122	118	118	116	114	112	113	114	113	111	112	128	157	167	170	173	177	175	177
9	126	122	122	119	117	119	114	114	115	112	115	113	110	111	130	157	165	166	169	176	181	183
10	128	125	122	121	119	118	114	114	115	114	113	111	111	111	134	157	163	166	168	170	176	182
11	129	127	125	121	119	118	117	115	111	112	111	110	112	111	138	155	160	163	165	167	172	176
12	128	123	124	122	120	118	117	116	110	110	110	111	111	113	139	155	158	162	164	165	167	171
13	133	125	126	123	119	118	115	111	107	109	110	110	110	113	139	154	156	160	164	165	168	168
14	134	128	125	121	118	119	114	110	109	107	108	110	109	109	135	152	154	158	162	164	164	167
15	131	127	126	122	118	117	116	113	112	109	106	108	108	107	133	150	152	158	159	162	163	166
16	131	127	127	122	121	119	117	112	112	110	109	108	107	105	127	149	150	155	156	161	163	163
17	134	129	127	123	121	118	115	112	110	110	110	107	108	105	121	144	148	152	155	159	161	163
18	134	130	127	124	121	116	113	111	109	108	106	105	109	108	115	138	146	148	150	149	147	151
19	133	131	128	127	121	114	113	112	107	107	106	107	106	105	111	134	142	144	143	143	145	145
20	134	132	127	126	121	115	110	110	108	108	106	107	105	105	107	128	141	140	139	147	150	151
21	135	133	127	124	118	115	111	109	109	107	105	105	104	105	105	112	135	138	141	147	150	150
22	137	132	126	123	117	115	111	111	109	107	105	103	104	104	105	105	118	133	141	147	146	149
23	135	129	127	124	120	115	108	109	109	107	105	102	104	104	103	106	111	116	133	143	140	141
24	136	132	130	127	119	115	110	107	107	107	105	102	102	104	102	105	111	113	115	132	140	135
25	137	133	129	124	118	116	112	108	106	106	104	102	101	101	102	104	111	112	111	116	136	141

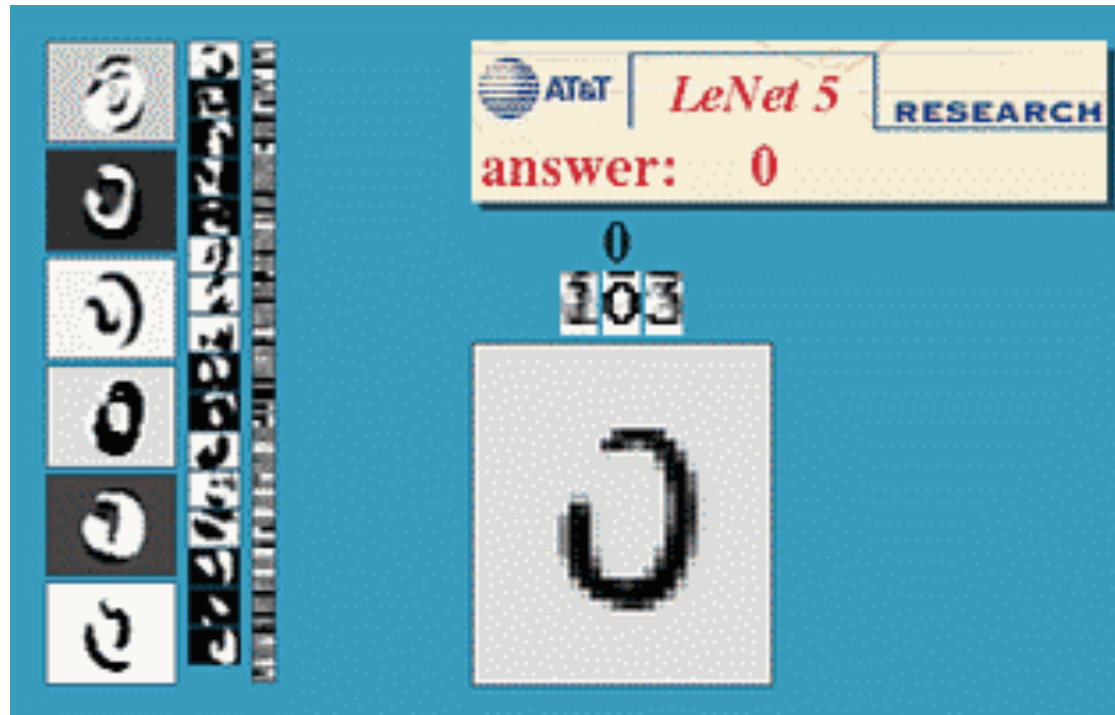
What computers see

tensor of $N \times M \times 3$ integers between 0 and 255
(for 8-bits RGB images)

Convolutional Neural Networks aka CNN

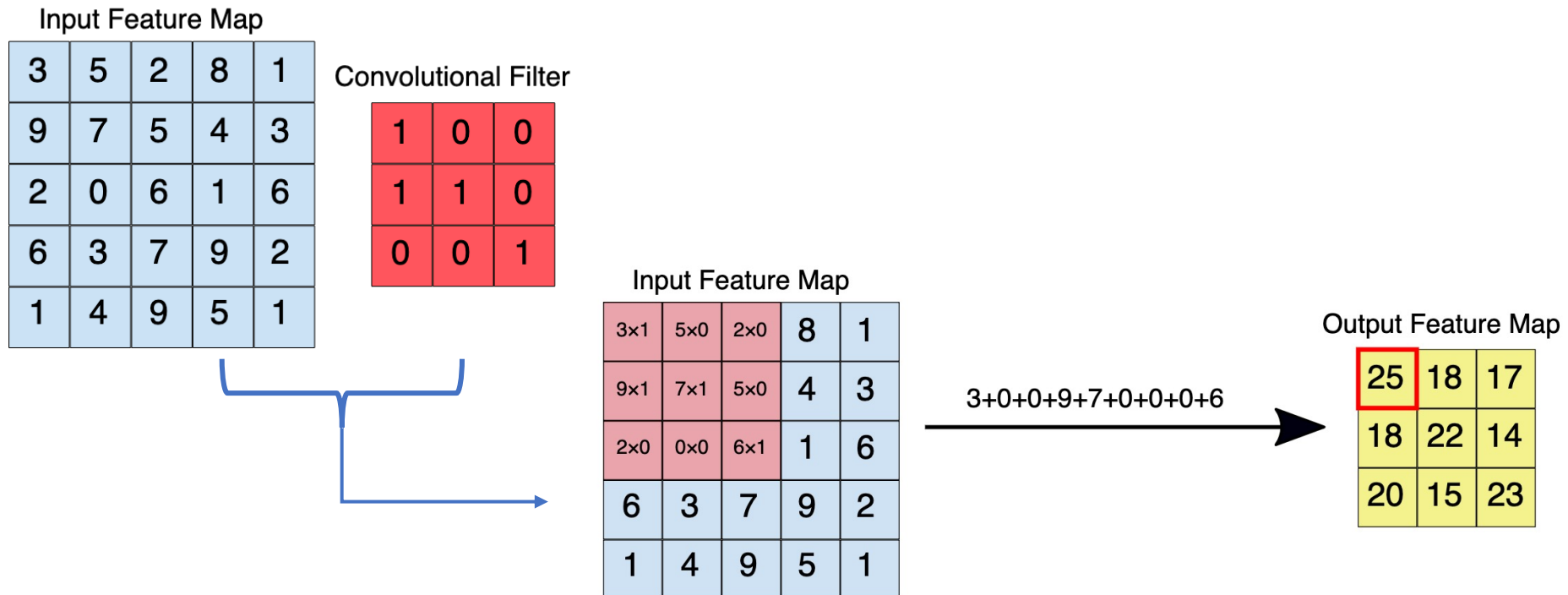
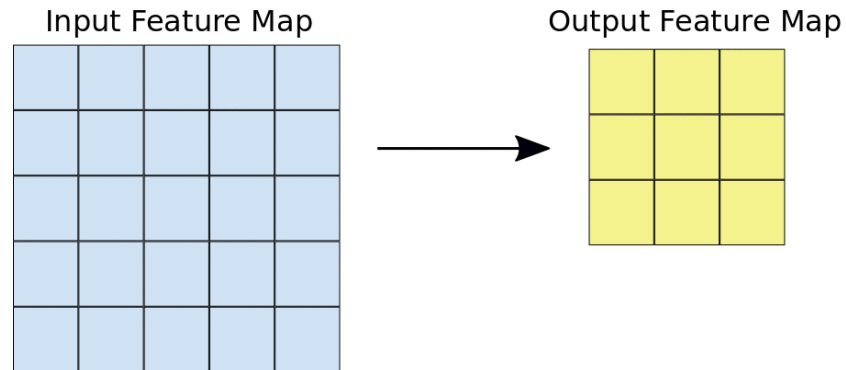


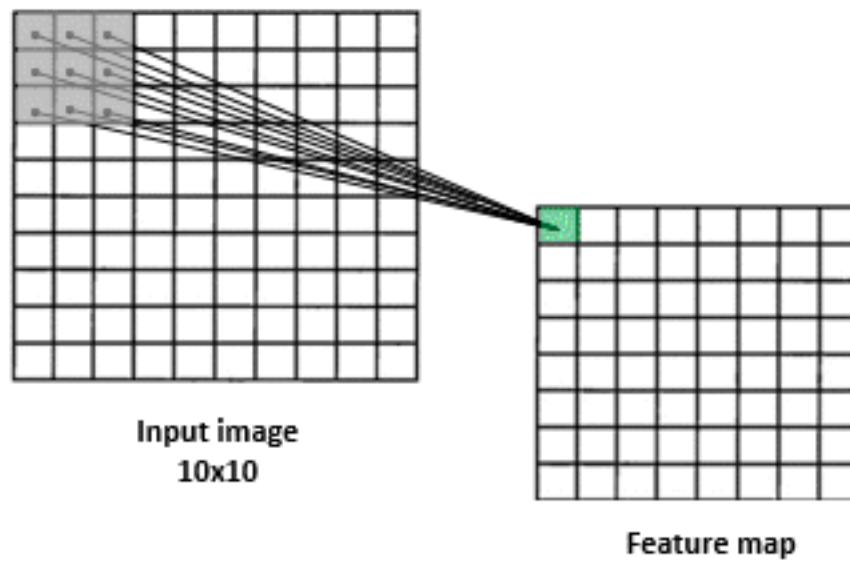
LeNet-5



Convolution Layer

« receptive field »





Kernel



Hidden neuron



1x1	1x0	1x1	0	0
0x0	1x1	1x0	1	0
0x1	0x0	1x1	1	1
0	0	1	1	0
0	1	1	0	0

4		

Convolution Layer : the feature detector

Input

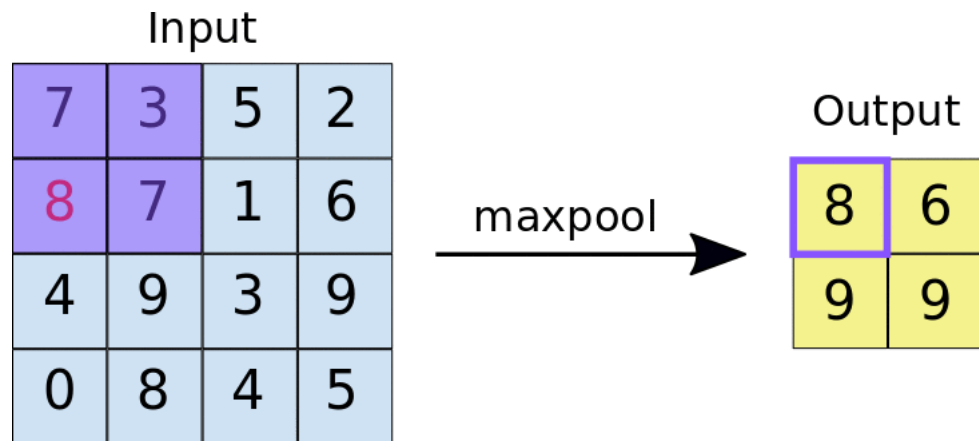
*	0	1	0			3	0	0
	0	1	0	→		3	0	0
	0	1	0			3	0	0
*	0	0	0			1	1	0
	1	1	1	→		1	1	0
	0	0	0			1	1	0

filter (weights)

feature map

Pooling Layer

- Downsampling the feature maps by associating with a smaller filter a single value
(eg : max-pooling with 2×2 filter)
- Saves computation time, reduces dimensions, but preserves critical features



Convolution Layer + Maxpool layer

Input

*

0	1	0
0	1	0
0	1	0

*

0	0	0
1	1	1
0	0	0

filter (weights)



3	0	0
3	0	0
3	0	0

feature map

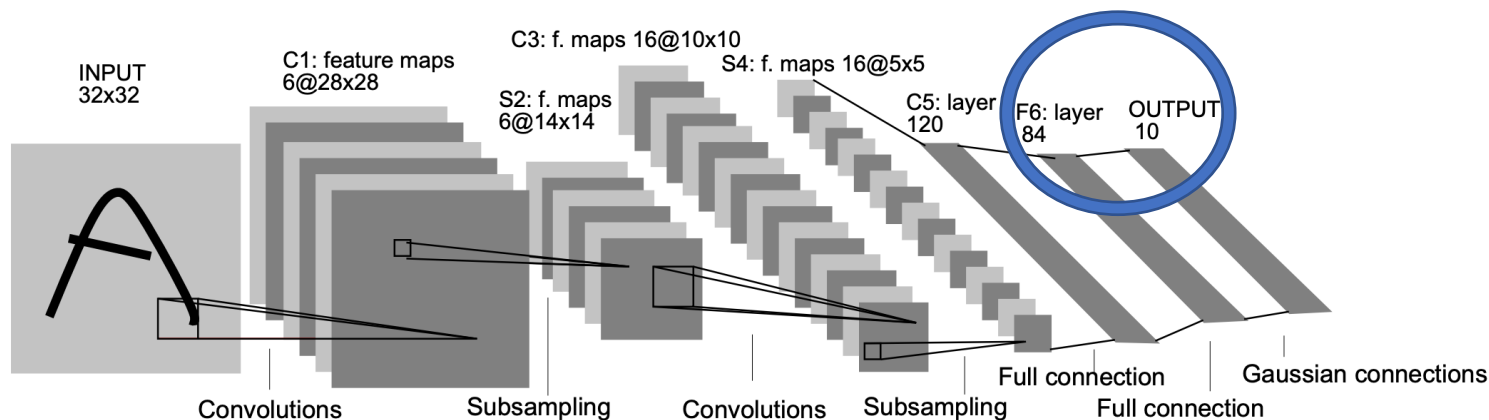


3	0
3	0

maxpool

Fully connected layers

- End of CNN : one or more fully connected layers
- Performs classification based on the features extracted by the convolutions.
- Outputs a probability value from 0 to 1 for each of the classification labels the CNN is trying to predict
- One output neuron / category to include



ConvNets for image recognition

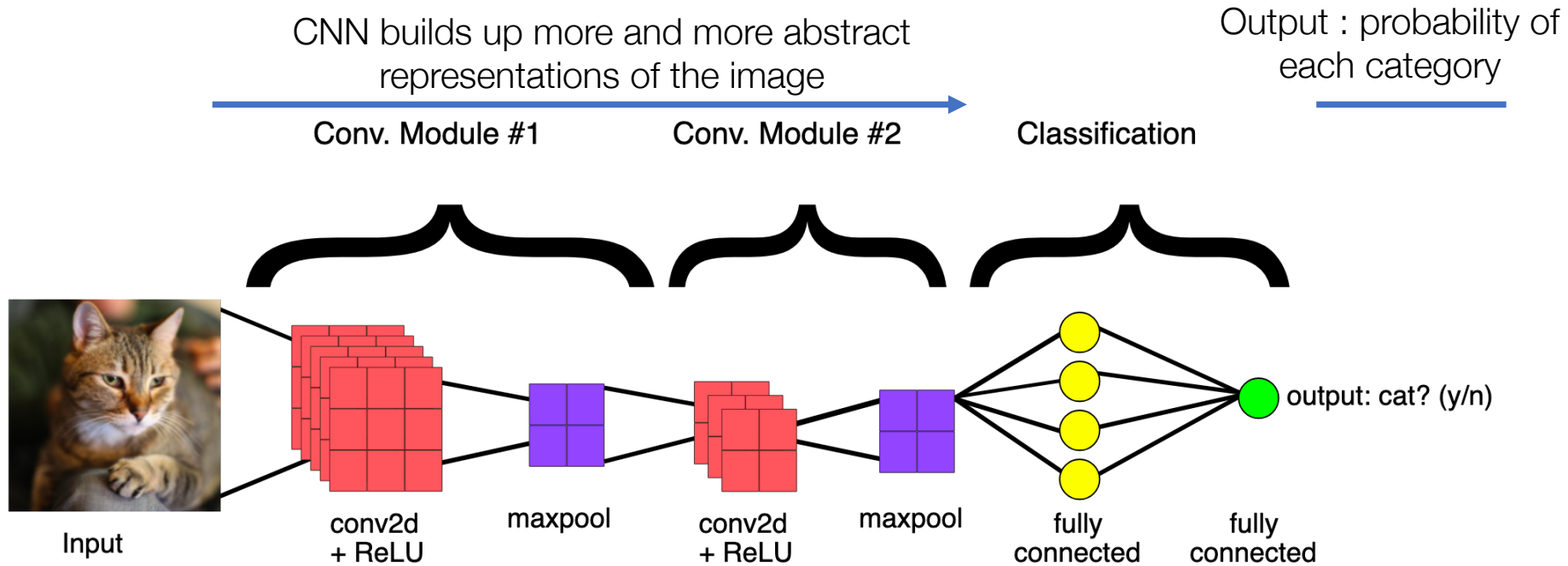
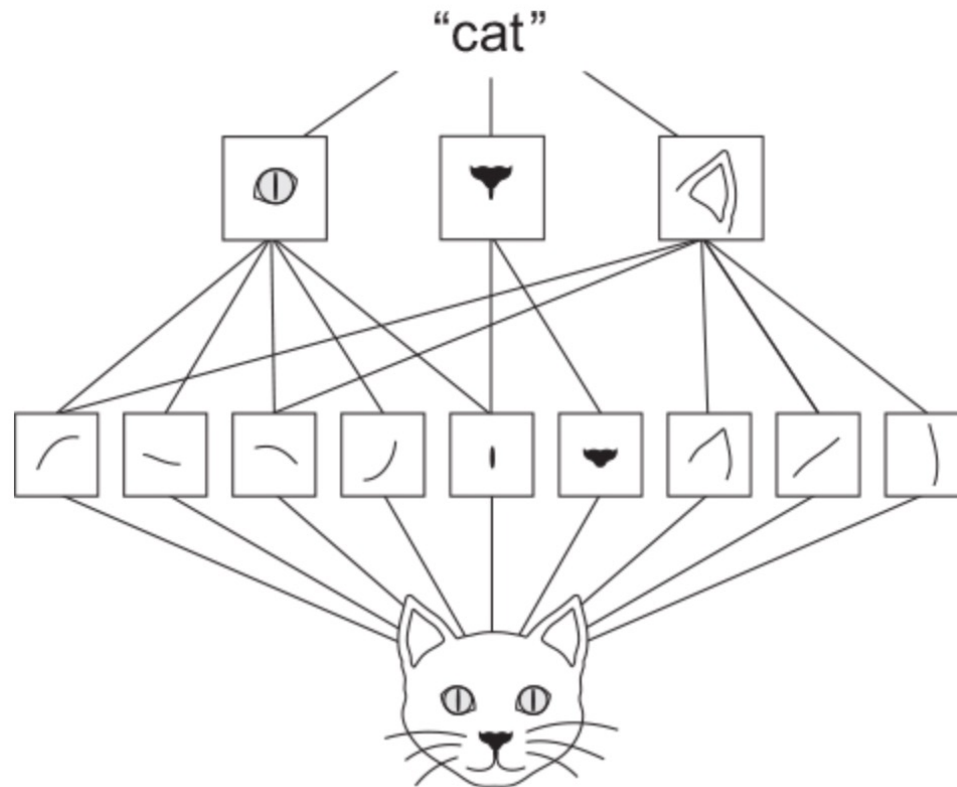
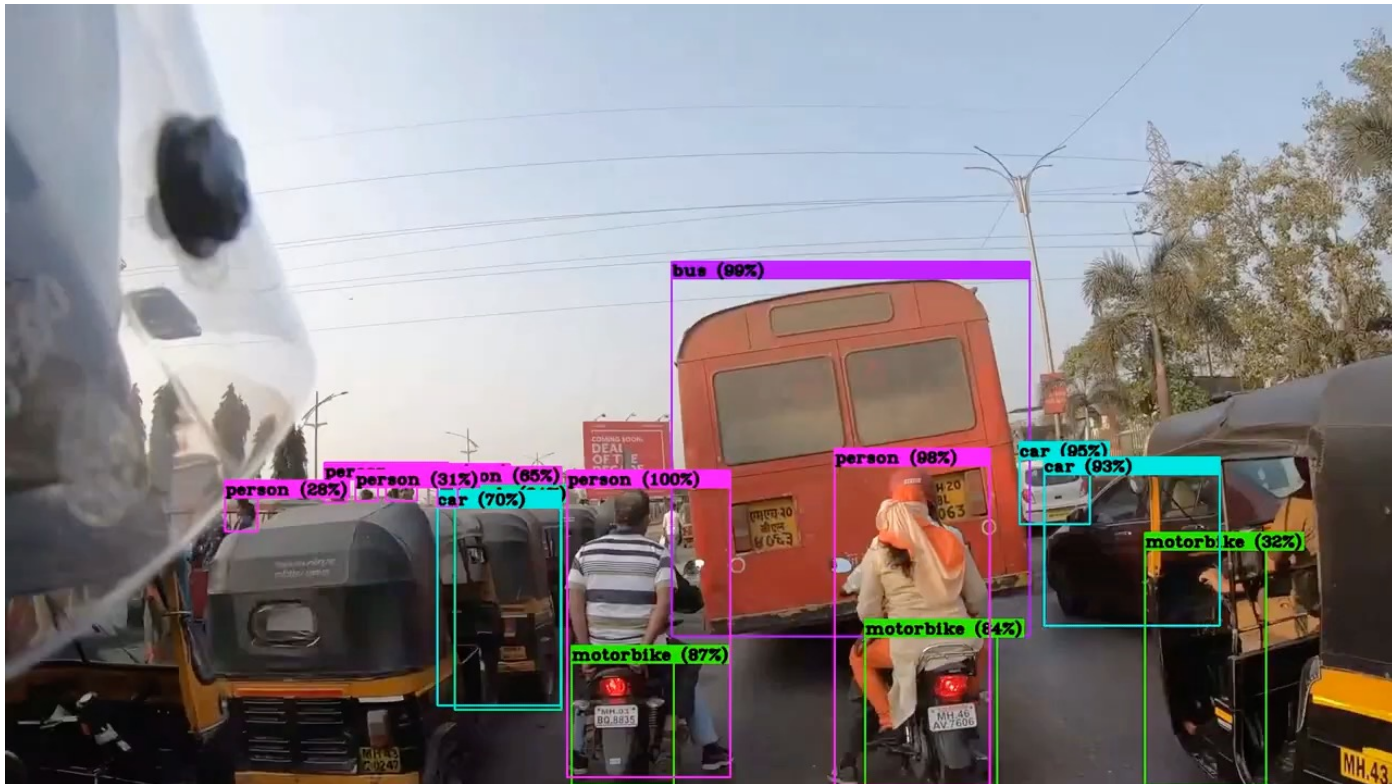


Figure 5.2. The visual world forms a spatial hierarchy of visual modules: hyperlocal edges combine into local objects such as eyes or ears, which combine into high-level concepts such as “cat.”



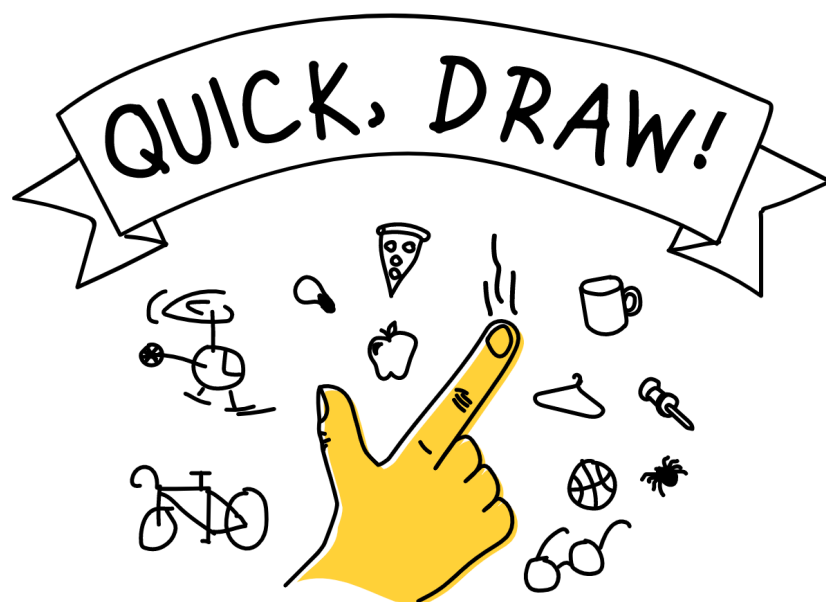
Démo : détection d'objets, YOLOv4



Source : <https://www.youtube.com/watch?v=h08N0HX16I8> , GrahamML

Démo : Quickdraw (AI Experiments,
Google)

<https://quickdraw.withgoogle.com/>



Demo : ChatGPT



Please write a dog food advertisement in a magazine read by dogs



"Woof! Are you tired of the same old kibble day after day? Switch things up with our delicious and nutritious dog food! Made with high-quality ingredients, our food will keep your tail wagging and your taste buds satisfied. Plus, our specially-formulated recipes cater to dogs of all sizes and breeds, so every pup can enjoy a hearty meal. Try it today and see the difference it makes in your dog's health and happiness. Order now and receive a free bag of treats with your purchase! Woof woof!"



What is the diameter of an apple divided by the thickness of an apple peel?

The diameter of an apple can vary depending on the variety, but on average it's about 7-10 cm. The thickness of an apple peel is around 1-2 mm.

Dividing the diameter of an apple by the thickness of the peel, we get:

$$7 \text{ cm} / 0.1 \text{ cm} = 70$$

Note: The above numbers are estimates and the exact numbers may vary depending on the apple being considered.



To provide some perspective, the diameter of an apple divided by the thickness of the apple peel is around 70, as calculated earlier. However, the diameter of the Earth divided by the deepest ocean depth is approximately 1162.8. This indicates that the ocean depths would be much thicker than the apple peel, if the Earth was scaled down to the size of an apple.

You are actually wrong, it is the other way around. We are talking about fractions, and $1/1162$ is actually smaller than $1/70$. But you made good improvement with the latest update. good job!



Thank you for your feedback. I apologize for the mistake in my previous answer. Yes, you are correct, if the Earth was the size of an apple, the ocean depths would be much thinner than the apple peel. The fraction $1/1162$ is indeed smaller than $1/70$. I appreciate the opportunity to correct my mistake. Thank you for your understanding.



Getting someone to like you can be difficult, as it depends on many factors such as compatibility and timing. However, there are some things you can do to increase the chances of your crush developing feelings for you:



- Be confident: Confidence is attractive, so try to be comfortable in your own skin and believe in yourself.
- Show interest: Let your crush know that you are interested in them by flirting, making eye contact, and showing genuine interest in their life and interests.
- Be a good listener: Show your crush that you value what they have to say by actively listening to them and asking questions.
- Be kind and respectful: Treat your crush and others with kindness and respect, as people are more likely to be attracted to those who display good character.
- Be yourself: Don't try to be someone you're not in order to impress your crush, as ultimately, they will appreciate the real you.

my creativity and ability to come up with new ideas?"



Sure! Here are some sample behavioral interview questions about your ability to create and execute successful marketing campaigns:

1. Can you describe a time when you had to develop and implement a marketing campaign from start to finish? What was the campaign for, and what were the results?



continue this guitar pattern

$e||-1-----1-x-6-----||$
 $B||-1-----1-----1-x-8-----||$
 $G||-3-3-----3-----2-x-10^*-----||$
 $D||-3-3-----3-3-----5-x-8-----||$
 $A||-1-1-----1-0-1-----1-3-x-6-----||$
 $E||-----3-3-3-----3-0-1-1-x-----||$



Out of these four shoes, which would you say works best with outfit #3: light brown Timberlands, white Air Force Ones, black Air Max, dark red Converse



Of the four shoe options you mentioned, the white Air Force Ones would work best with Outfit #3 (white T-shirt, black leather jacket, and dark blue jeans). The white Air Force Ones are a classic and versatile sneaker that can add a clean and casual look to the outfit.

The black Air Max may be a bit too sporty for the outfit, as the black leather jacket and dark blue jeans have a more edgy and casual vibe. The dark red Converse may clash with the black leather jacket and look out of place with the rest of the outfit. The light brown Timberlands may work, but they could be a bit too bulky and heavy for the outfit, as the outfit already has a lot of texture and weight with the leather jacket and jeans.

Remember, your personal style and preference should ultimately guide your choice of shoes, so if you feel that one of the other options complements your outfit better, then go for it!

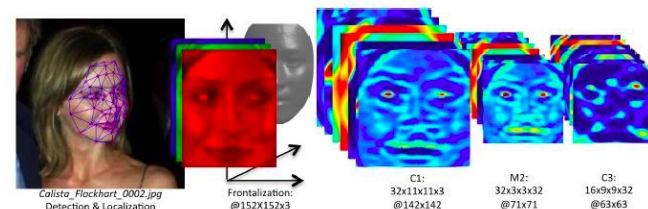
Identifier son interlocuteur



<https://www.visagetechnologies.com/HTML5/latest/Samples/ShowcaseDemo/ShowcaseDemo.html>

Deep Learning for AI

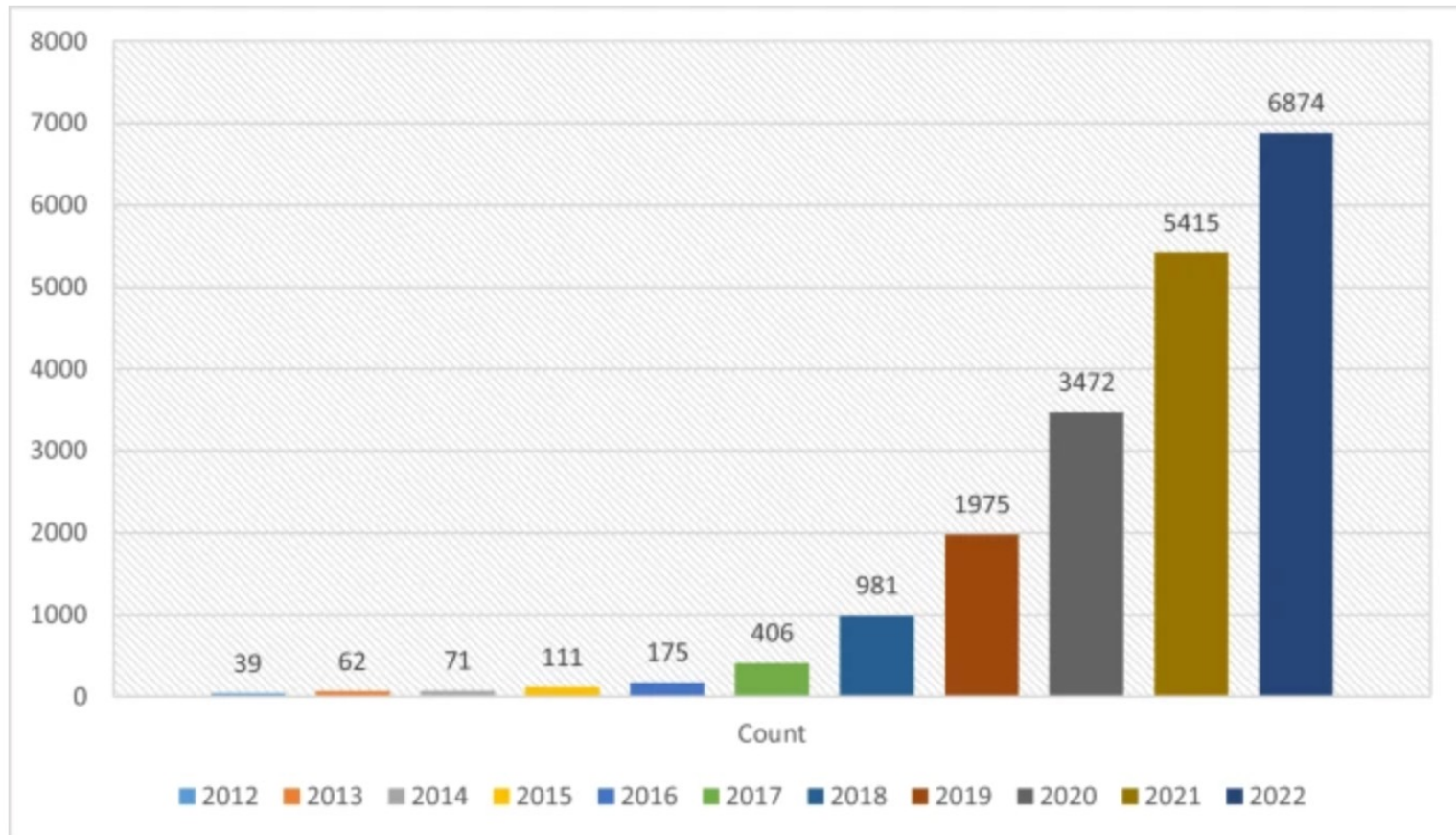
- GAFAM :
DeepMind, AlphaGo, Google Car, Cortana, Siri, Watson etc
- Image recognition, automatic labelling, voice recognition, buying recommendations, automatic translation, art creation, fraud detection, spam filtering, customer profiling, personalization of contents, facial identification, emotion recognition, video creation, chatbots, game playing, natural language processing, etc ...



Deep Learning for ... medical applications?

- DL is a great tool for image recognition and classification
- All medical diagnostics that are based on medical images and that require image classification / object detection can benefit from DL :
 - Dermatology
 - Radiology
 - Ophtalmology
 - Pathology...
- DL can help physicians to detect relevant areas in images
- DL can help to predict the outcomes
- DL can offer a second opinion
- DL can be « the physician's third eye » (a well-trained eye)

Fig. 1

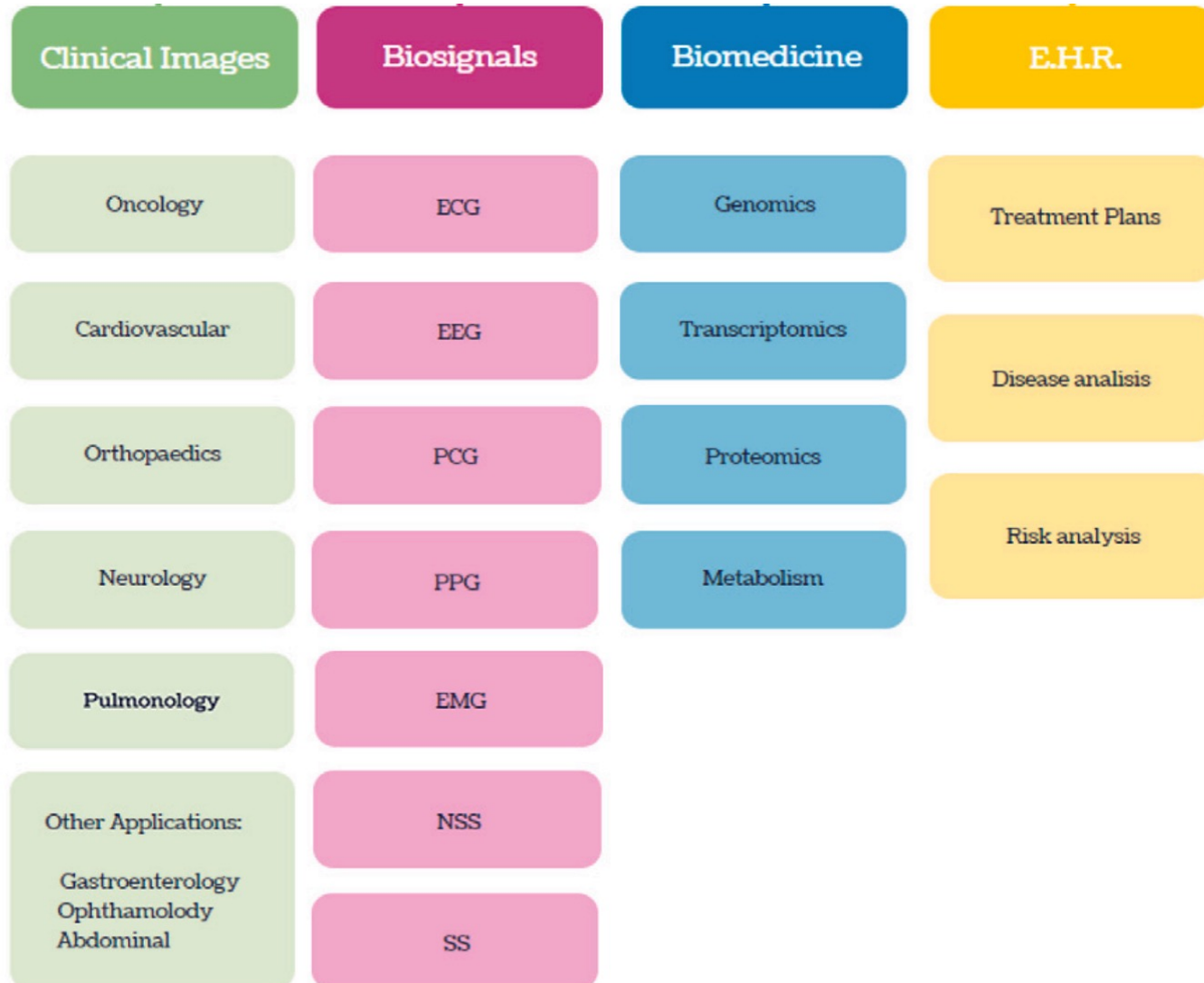


The yearly distribution of DL techniques in HCS in the PubMed database for the last decade

AI in Health : long story, many questions

- Old story, beginning with Expert Systems - GOFAI
- Problem : can we make « experience » explicit ?
- Can we replace humans with machines in the medical domain?
- Can we make patients accept a machine's opinion ?
- What if there is no agreement ?

Champs d'application du DL en médecine



MÉDECINE PRÉDICTIVE



PRÉDICTION D'UNE MALADIE
ET/OU DE SON ÉVOLUTION

MÉDECINE DE PRÉCISION



RECOMMANDATION DE
TRAITEMENT PERSONNALISÉ

AIDE À LA DÉCISION



DIAGNOSTIQUE
ET THÉRAPEUTIQUE

ROBOTS COMPAGNONS



NOTAMMENT POUR LES
PERSONNES ÂGÉES
OU FRAGILES

CHIRURGIE ASSISTÉE PAR ORDINATEUR



PRÉVENTION

en population générale

- ANTICIPATION
D'UNE ÉPIDÉMIE
- PHARMACOVIGILANCE



Les domaines d'application de l'IA en médecine

Le deep learning dans la pratique :

transcription de consultations

résumé de dossiers médicaux

assistance aux tâches médico-administratives et cliniques

génération de contenus pédagogiques destinés à la formation.

Mais aussi ... :

analyse d'images médicales

aide au diagnostic

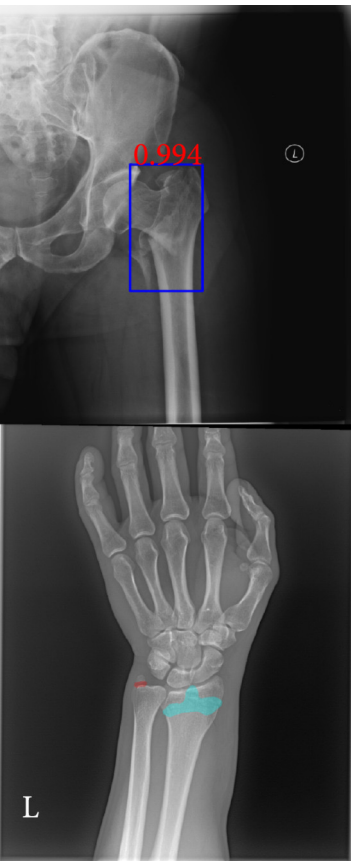
découverte de molécules pour médicaments

personnalisation des traitements

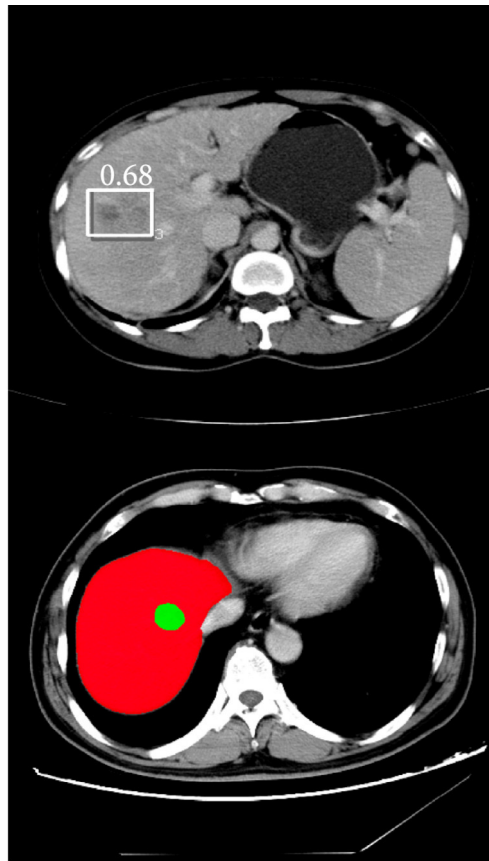
aiguillage aux urgences

etc etc

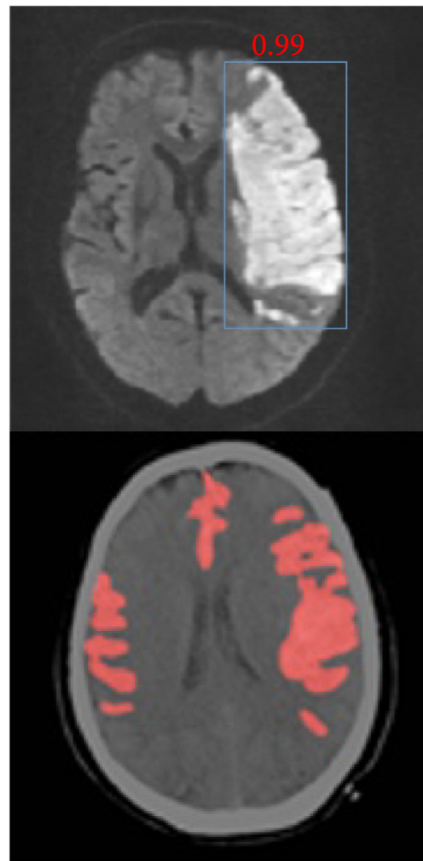
1 – DL and clinical images



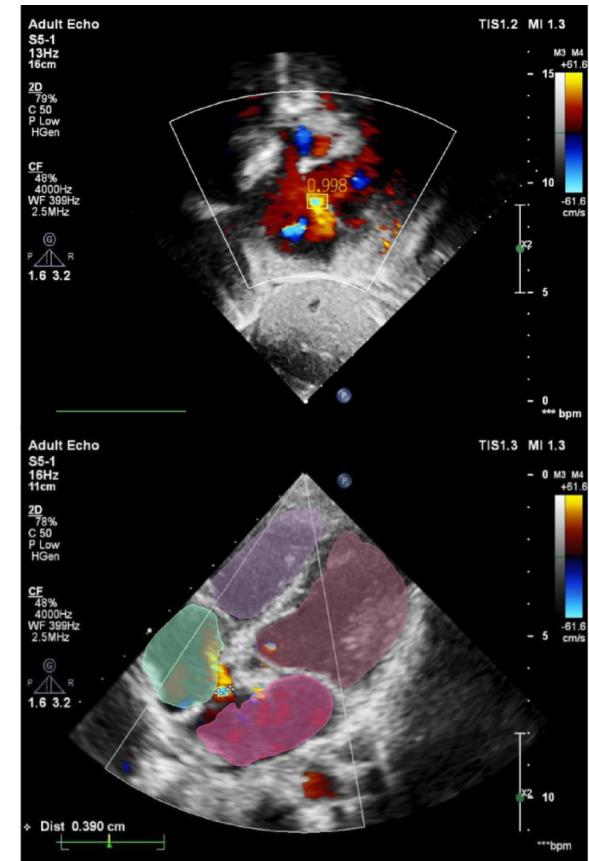
Bone X-ray



Liver CT



Brain MRI



Cardiac ultrasound

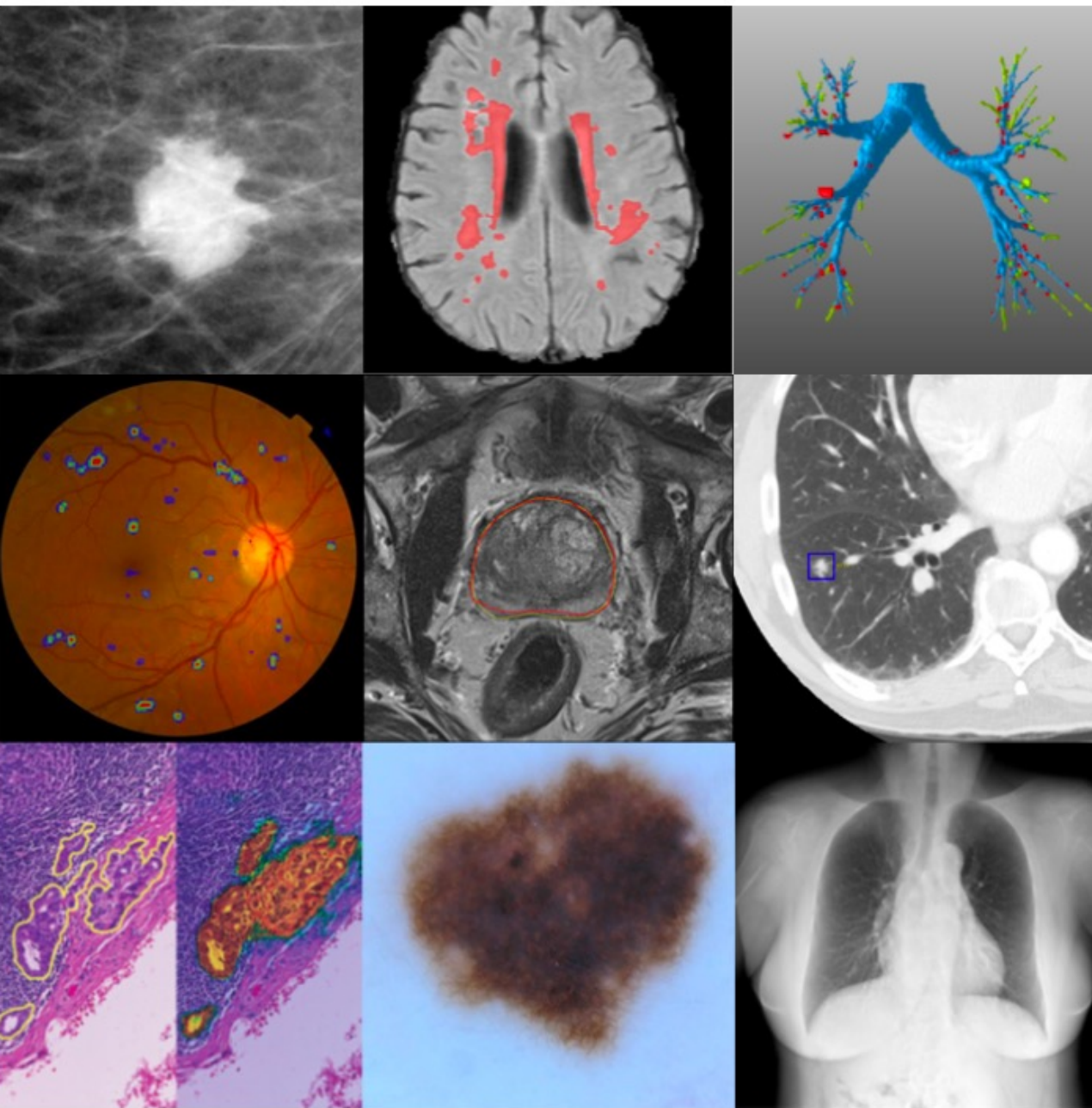
With MRI images, a CNN network can do these things :

- Tissue/anatomy/lesion/tumor **segmentation** → 1
- Image (re)construction/**enhancement** → 2
- Disorder **classification** (eg. AD, MCI, Schizophrenia) → 3
- Lesion/tumor **detection** and classification
- Survival/disease activity/development **prediction**
- Other

All image-based tasks can benefit from DL

From images, DL can help to :

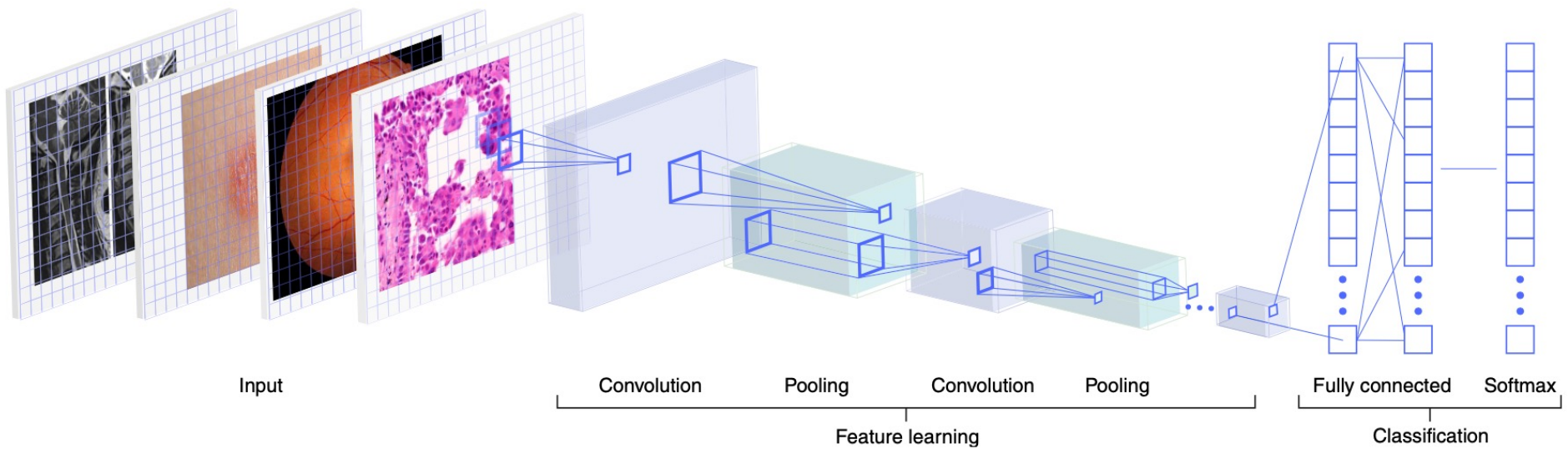
- Identifying moles vs melanomas
- Diabetic retinopathy
- Cardiovascular risks
- Breast lesion detection in mammograms
- Flagging large artery occlusions
- Predicting survival probabilities
- Combining imaging modalities



What DL can do :

Mammographic mass classification,
 Segmenting brain lesions
 Detecting leaks in airway tree,
 Classifying diabetic retinopathy
 Segmenting prostate,
 Classifying nodules
 Detecting breast cancer metastases
 in lymph nodes
 Classifying skin lesions,
 Suppressing bones in chest X-Rays

How does it work ?



1.1- DL for segmenting medical images

- Segmenting from patches of images + context
- Segmenting = classifying « in » vs « out »
- Class balance is skewed severely towards « out » class in a training setting.
- Segmenting organs, lesions, surgical objects (stents)

Segmenting the pancreas

- Ex. : pancreas segmentation with CNN from CT scans
- Pancreas has a very high anatomical variability : challenging

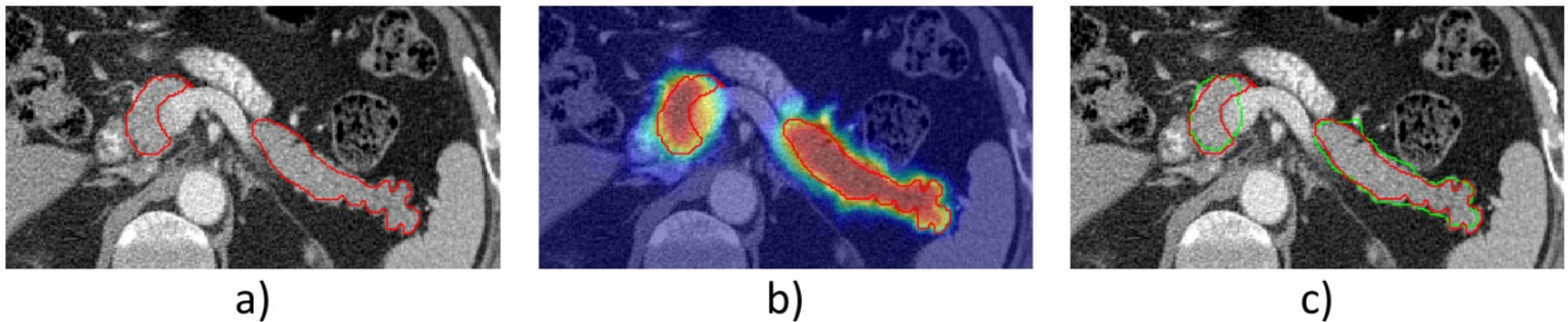


Figure 6: Example of pancreas segmentation using the proposed R_2 -ConvNet approach in testing. a) The manual ground truth annotation (in red outline); b) the $G(P_2(x))$ probability map; c) the final segmentation (in green outline) at $p_2 = 0.6$ (DSC=82.7%).

Segmenting the brain

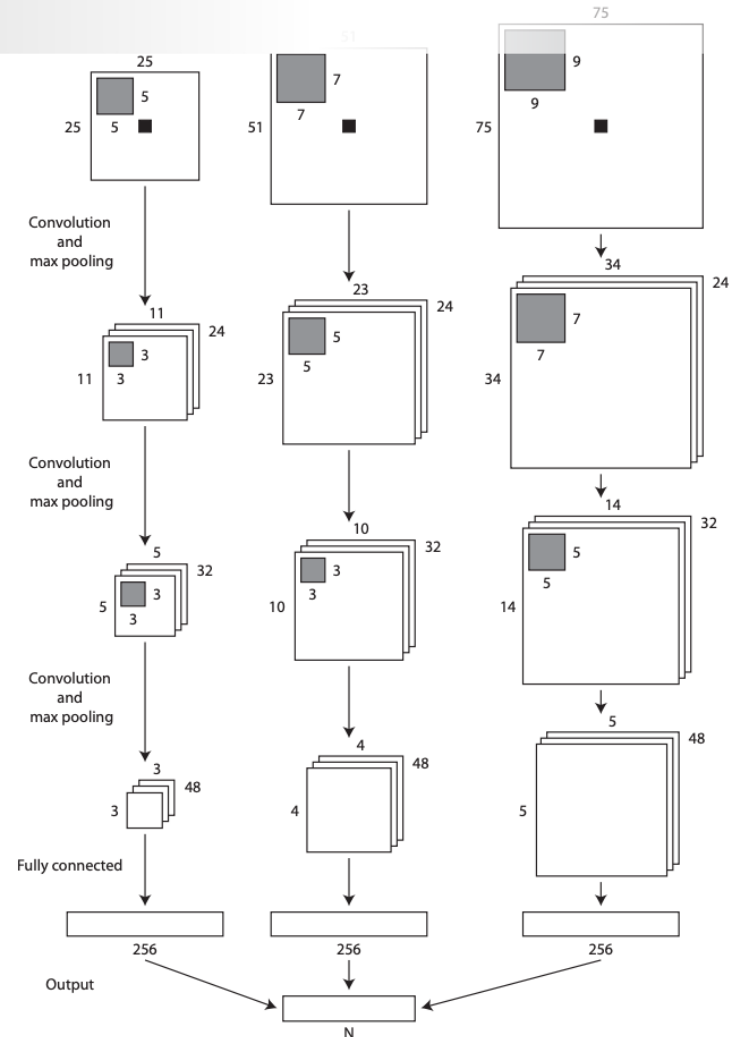


Fig. 1: Schematic overview of the convolutional neural network. The number of output classes, N , was set to 9 (8 tissue classes and background) for the neonatal images, to 8 (7 tissue classes and background) for the ageing adult images, and to 7 (6 tissue classes and background) for the young adult images. After the third convolution layer, max-pooling is only

Inside the CNNs...

TABLE I: Acquisition parameters for the images used in this paper.

	Cor. 30 wks	Cor. 40 wks	Ax. 40 wks	Ageing adults	Young adults
Age	30 weeks PMA	40 weeks PMA	40 weeks PMA	70 years	23 years
Acquisition protocol	Coronal T ₂ -weighted	Coronal T ₂ -weighted	Axial T ₂ -weighted	Axial T ₁ -weighted	Sagittal T ₁ -weighted
Number of images	10	5	7	20	15
Reconstruction matrix	$384 \times 384 \times 50$	$512 \times 512 \times 110$	$512 \times 512 \times 50$	$240 \times 240 \times 48$	$256 \times 256 \times (261-334)$
Reconstructed voxel sizes [mm ³]	$0.34 \times 0.34 \times 2.0$	$0.35 \times 0.35 \times 1.2$	$0.35 \times 0.35 \times 2.0$	$0.96 \times 0.96 \times 3.0$	$1.0 \times 1.0 \times 1.0$

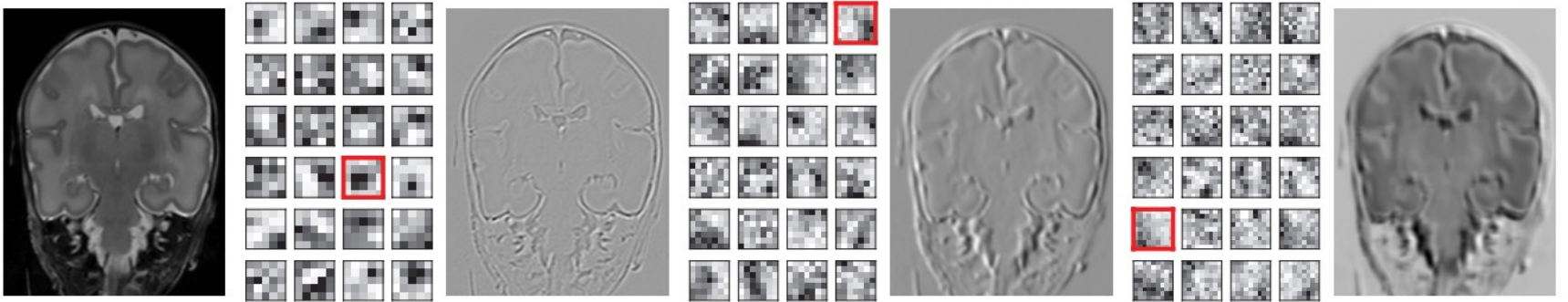


Fig. 2: Trained convolution kernels in the first layer after 10 epochs using the 5 training images acquired at 30 weeks PMA, and the kernels indicated in red applied to a test image. From left to right: the T₂-weighted test image, the kernels of 5×5 voxels, the image convolved with the indicated 5×5 kernel, the kernels of 7×7 voxels, the image convolved with the indicated 7×7 kernel, the kernels of 9×9 voxels, and the image convolved with the indicated 9×9 kernel.

30 - 40 weeks PMA

aging

young

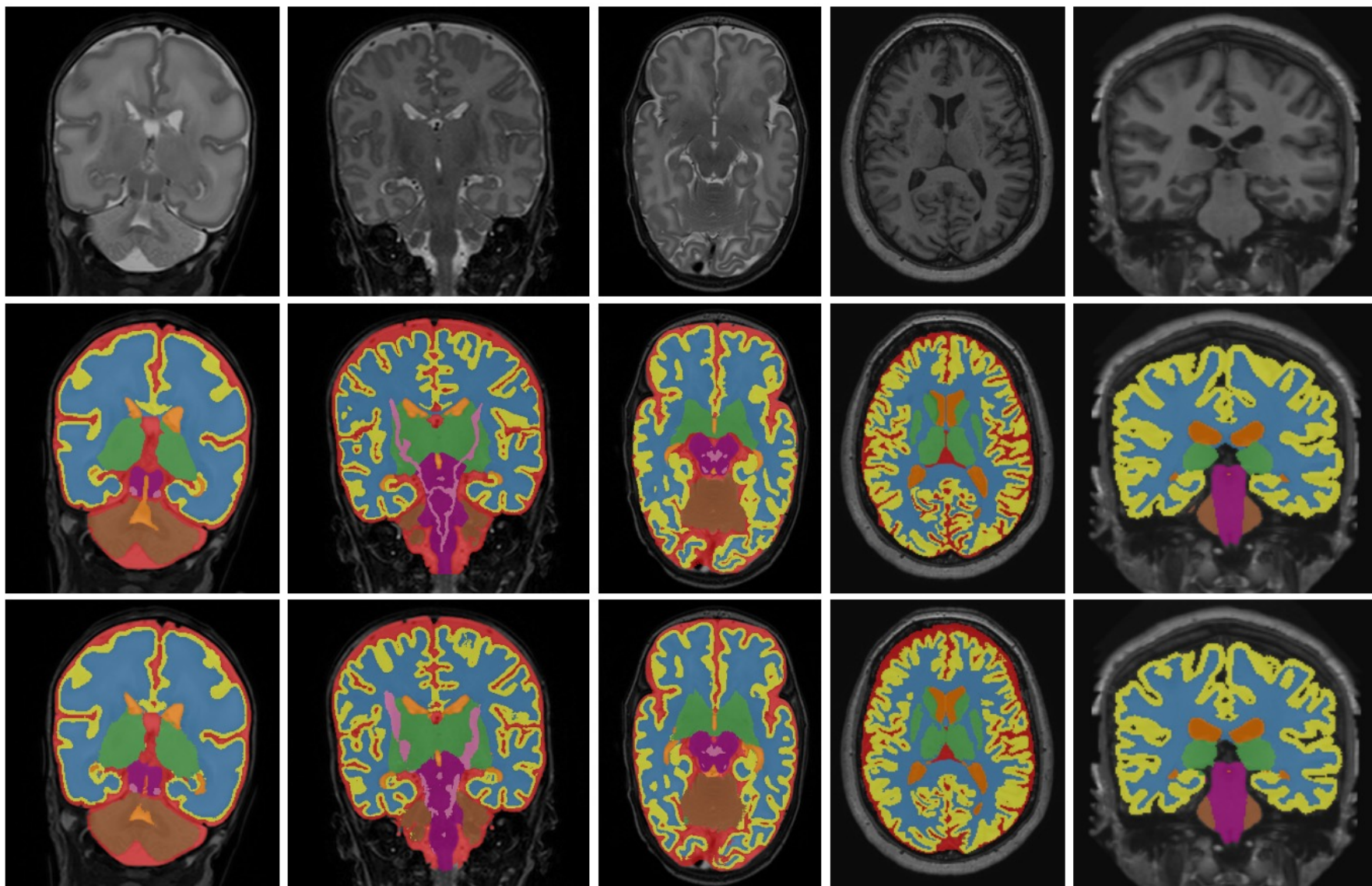


Fig. 4: Segmentation results for CB (brown), mWM (pink), BGT (green), vCSF (orange), (u)WM (blue), BS (purple), cGM (yellow), eCSF (red) in coronal images acquired at 30 weeks PMA (first column), coronal images acquired at 40 weeks PMA (second column), axial images acquired at 40 weeks PMA (third column), axial images of ageing adults (fourth column), and

Inside a sulcus of a 30 weeks PMA brain

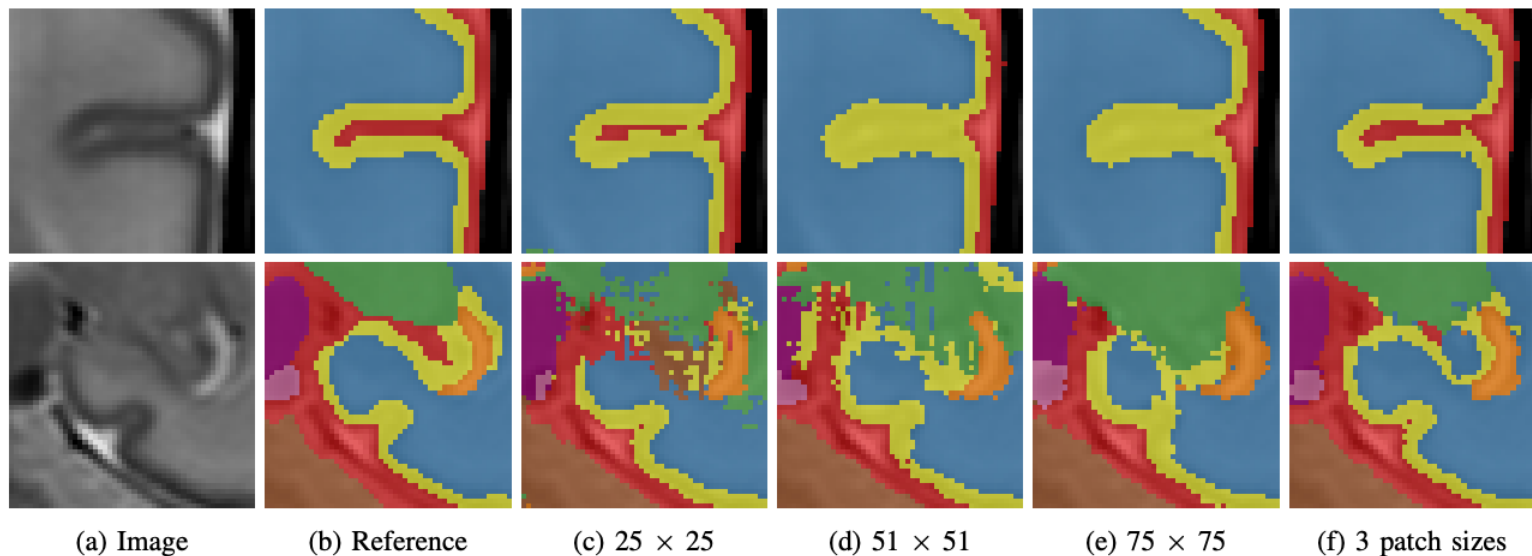


Fig. 5: Segmentation results in a T_2 -weighted image (left) acquired at 30 weeks PMA for the lateral sulcus (top) and the hippocampus (bottom) using (from left to right), manual segmentation, only a patch of 25×25 voxels, only a patch of 51×51 voxels, only a patch of 75×75 voxels, and these 3 patch sizes combined. The tissues are labelled as follows: CB in brown, mWM in pink, BGT in green, vCSF in orange, uWM in blue, BS in purple, cGM in yellow, and eCSF in red.

Left ventricle segmentation - cardiovascular diseases

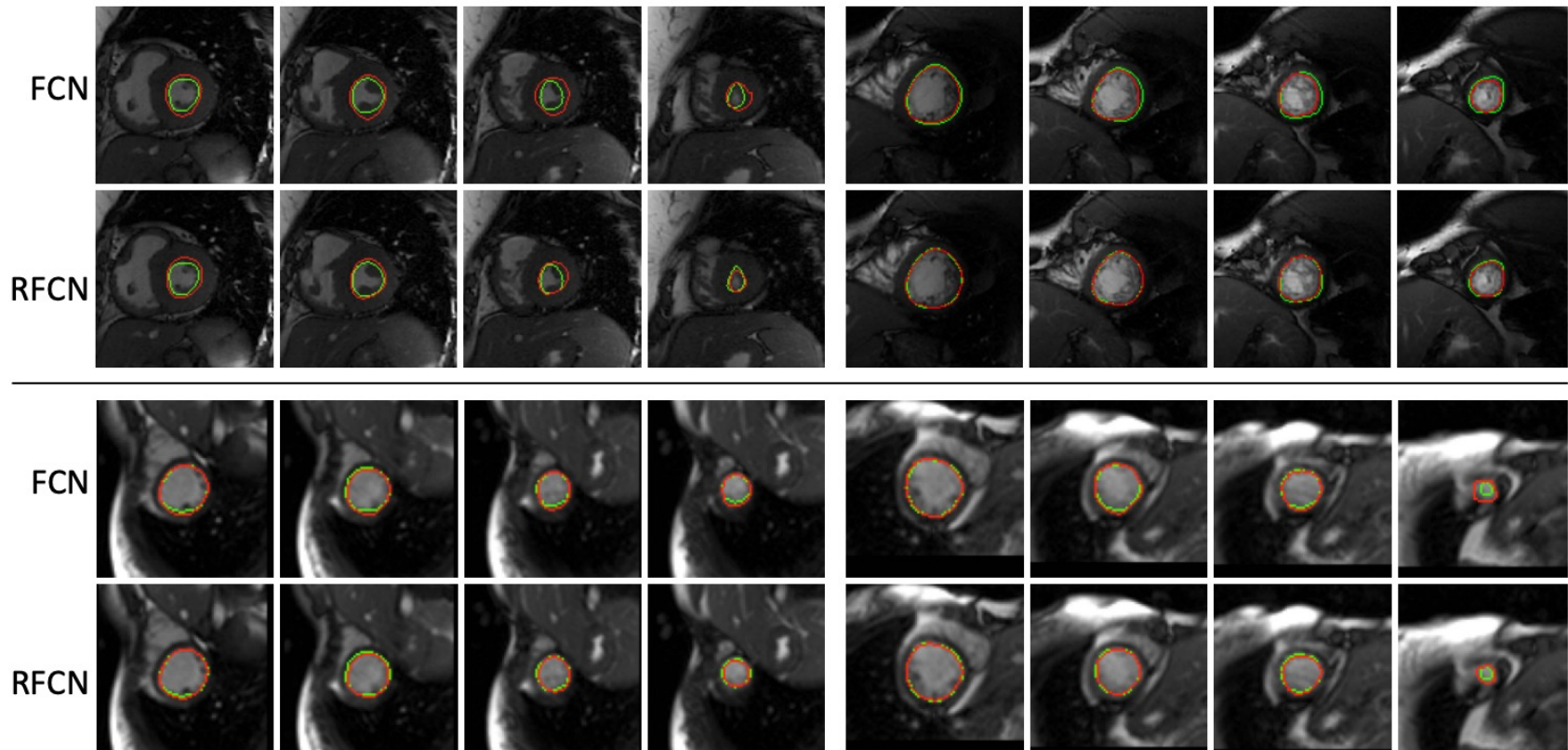
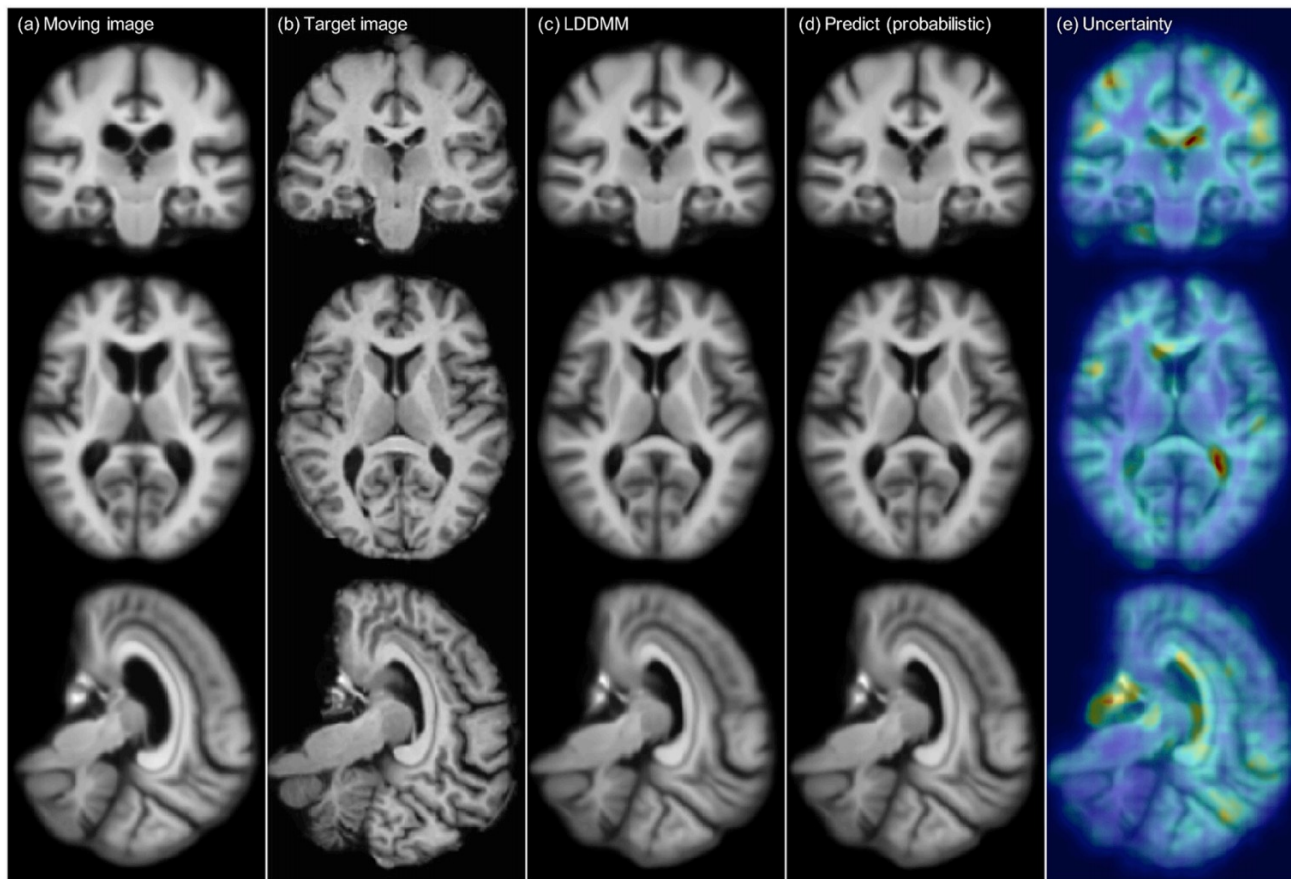


Fig. 3. Example of segmented left ventricle using RFCN and FCN architectures from MICCAI dataset (top two rows) [21] and PRETERM dataset (bottom two rows). Green contours represent the ground truth and red contours are the predicted contours. RFCN is often able to better delineate the left-ventricle contours with weaker boundaries

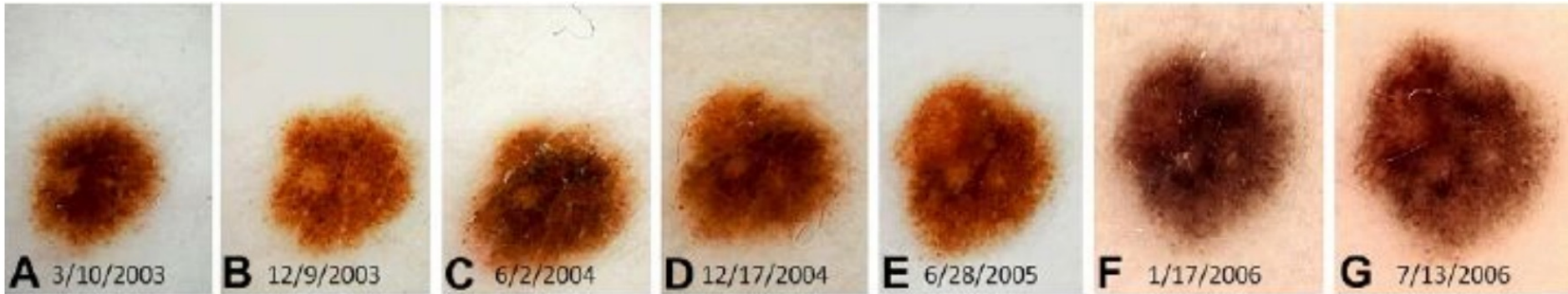
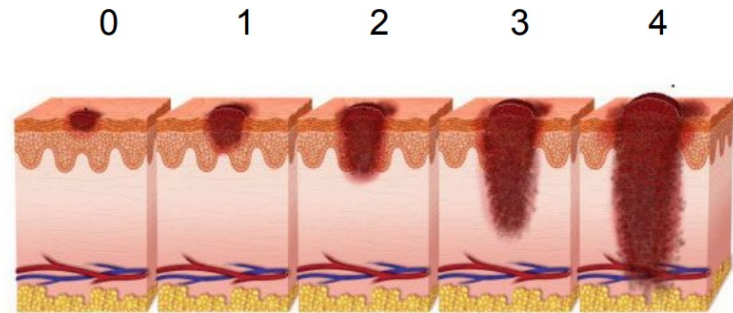
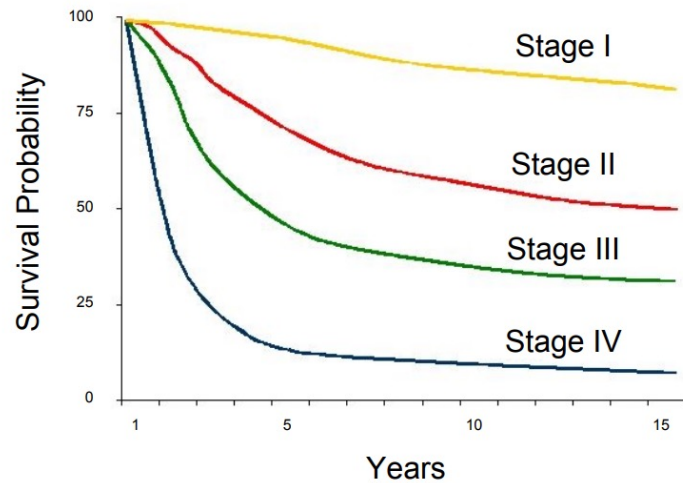
1.2- Image registration : example of fMRI



1.3 - DL for classifying medical images

- Predict the category of an image
- Classify images (group similar images)
- Ex. : cancer, melanoma, AD, depression, schizophrenia etc
- From images (anat and functional) + other data

Skin cancer

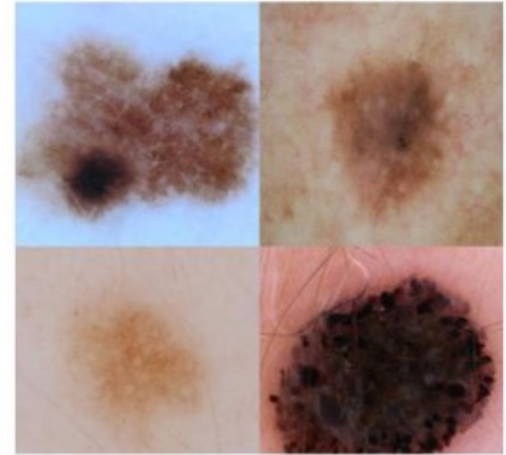


Epidermal Lesions

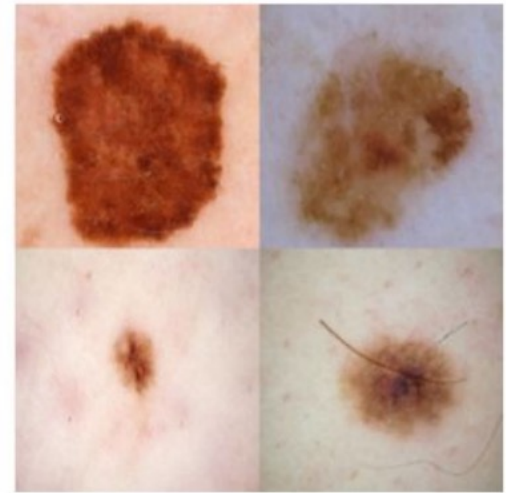
Melanocytic Lesions

Melanocytic Lesions (Dermoscopy)

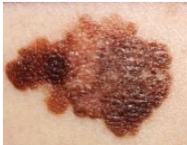
Benign



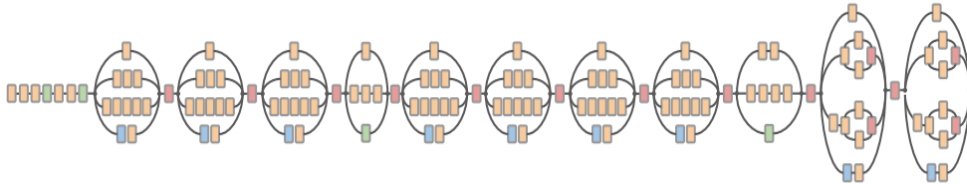
Malignant



Skin Lesion Image



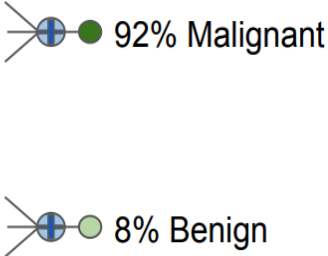
Deep Convolutional Neural Network (Inception-v3)



Training Classes (757)

- Acral-lent. melanoma
- Amelanotic melanoma
- Lentigo melanoma
- ...
- Blue nevus
- Halo nevus
- Mongolian spot
- ...
-
-
-

Inference Classes (varies by task)



Transfer learning

- Application of a process suited for one specific task to a different problem
 - DL model trained to recognize every day color images, such as animals
 - Same model used to classify radiographs.
- All images share similar features such as edges and blobs
 - The model has learnt to « see »
 - Better than random initialization
 - Re-learning on small medical databases

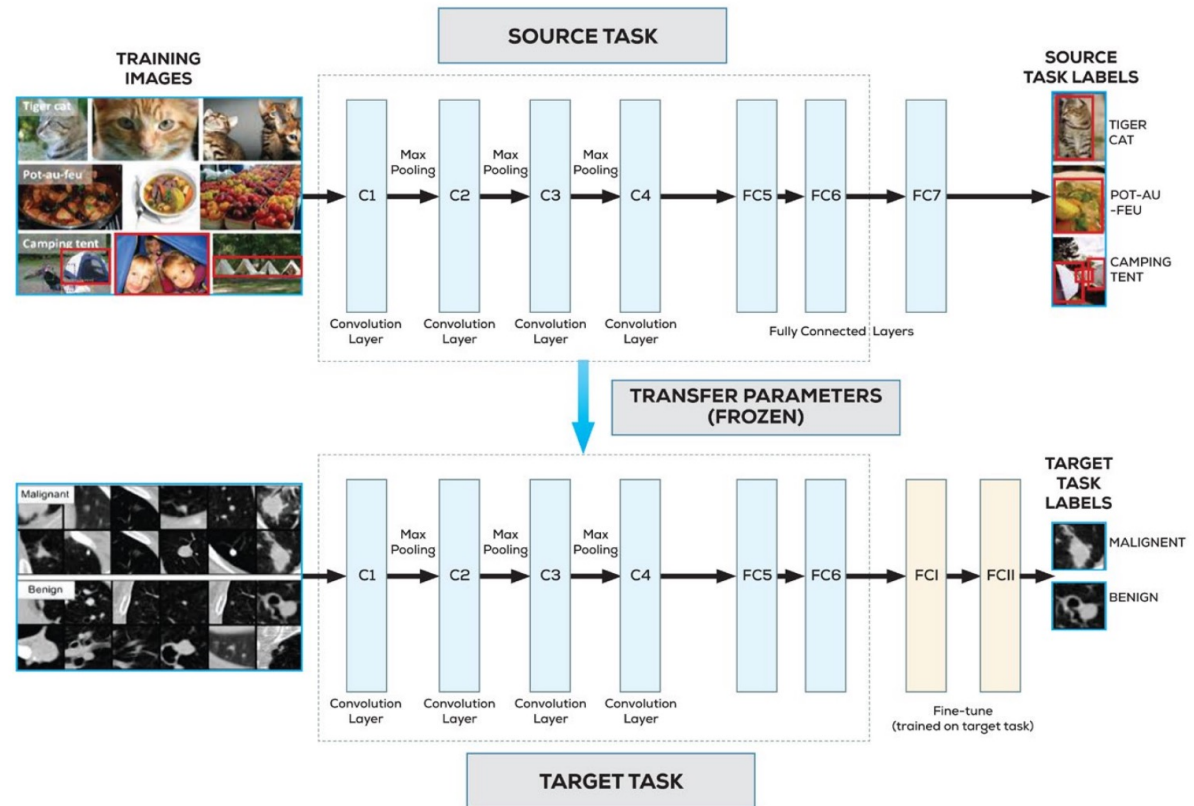
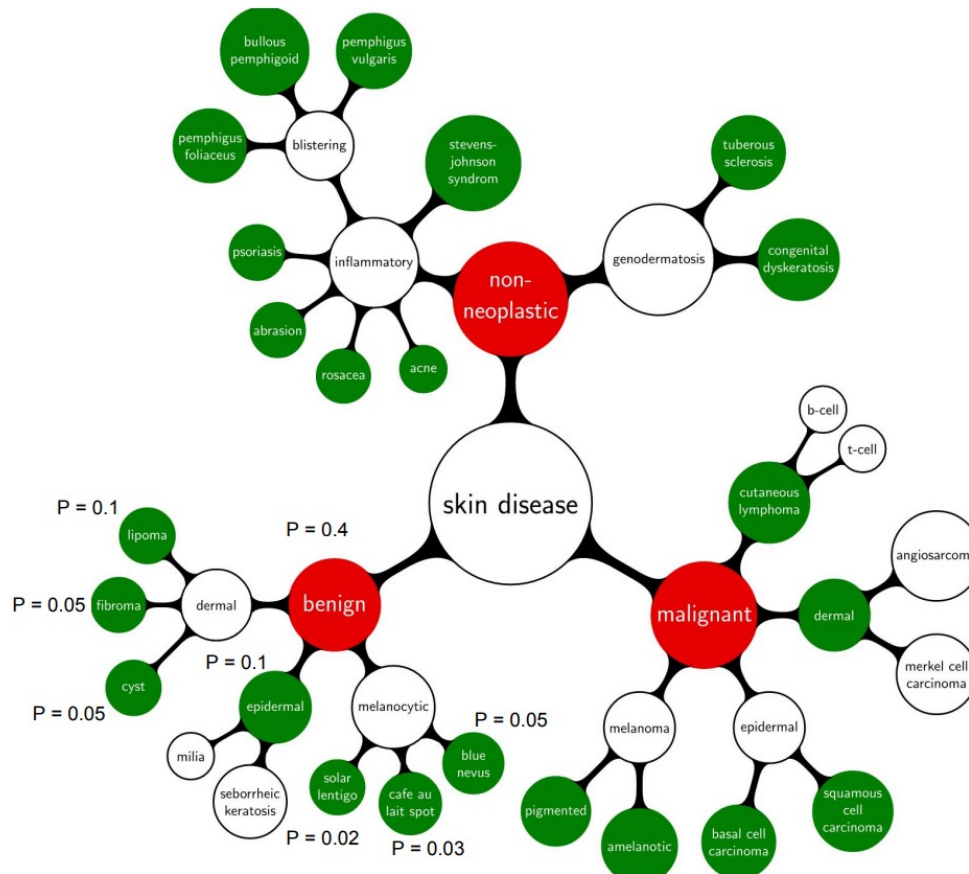


Figure 3. Schematic representation of convolutional neural network (CNN) architecture and the concept of “transfer learning.” Because images from multiple sources have common salient features (borders, shapes, etc), the core of a CNN trained for 1 task (*Top row*: eg, cat from pot-au-feu from camping tent) can be “transferred” (ie, used without modification) for a second task (*Bottom row*: benign from malignant tumors in computed tomography images). Because only the distal layers (fully connected [FC] I and FCII) remain to be trained, much less training data are required for the second task. Max indicates maximum.

Skin Cancer Classification



$$P[u] = \sum_{v \in C(u)} P[v]$$

- Training Classes
- Inference Classes

Dermatologist-level Classification of Skin Cancer with Deep Neural Networks

Andre Esteve*, Brett Kuperl*, Rob Novoa, Justin Ko, Susan Swetter, Helen Blau, Sebastian Thrun
Nature, 2017
(Equal contribution authors*)



Classifier	Three-way accuracy
Dermatologist 1	65.6%
Dermatologist 2	66.0%
CNN	69.5%
CNN - PA	72.0%

Disease classes: three-way classification

0. Benign single lesions
1. Malignant single lesions
2. Non-neoplastic lesions

Classifier	Nine-way accuracy
Dermatologist 1	53.3%
Dermatologist 2	55.0%
CNN	48.9%
CNN - PA	55.3%

Disease classes: nine-way classification

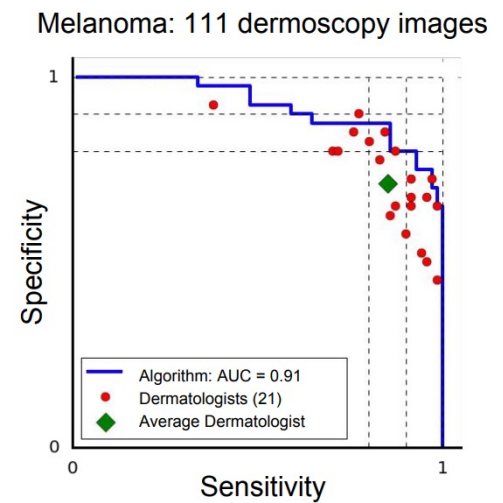
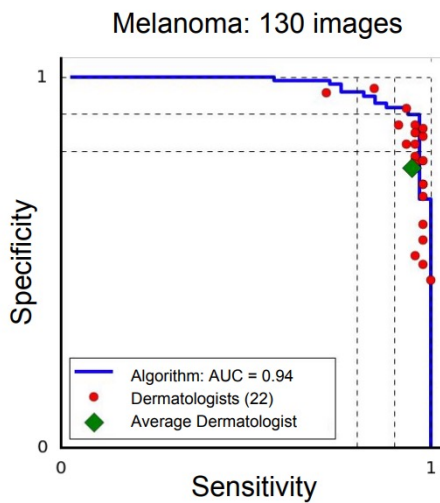
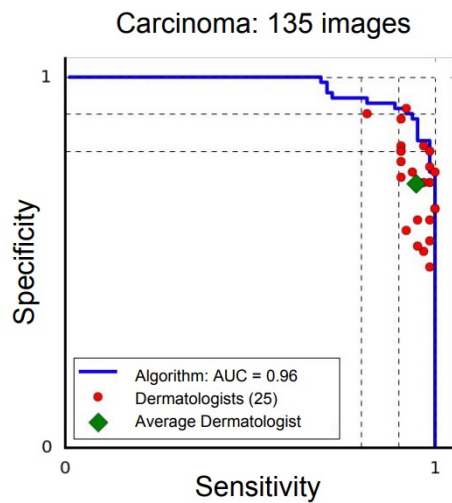
0. Cutaneous lymphoma and lymphoid infiltrates
1. Benign dermal tumors, cysts, sinuses
2. Malignant dermal tumor
3. Benign epidermal tumors, hamartomas, milia, and growths
4. Malignant and premalignant epidermal tumors
5. Genodermatoses and supernumerary growths
6. Inflammatory conditions
7. Benign melanocytic lesions
8. Malignant Melanoma

Signal Detection Theory

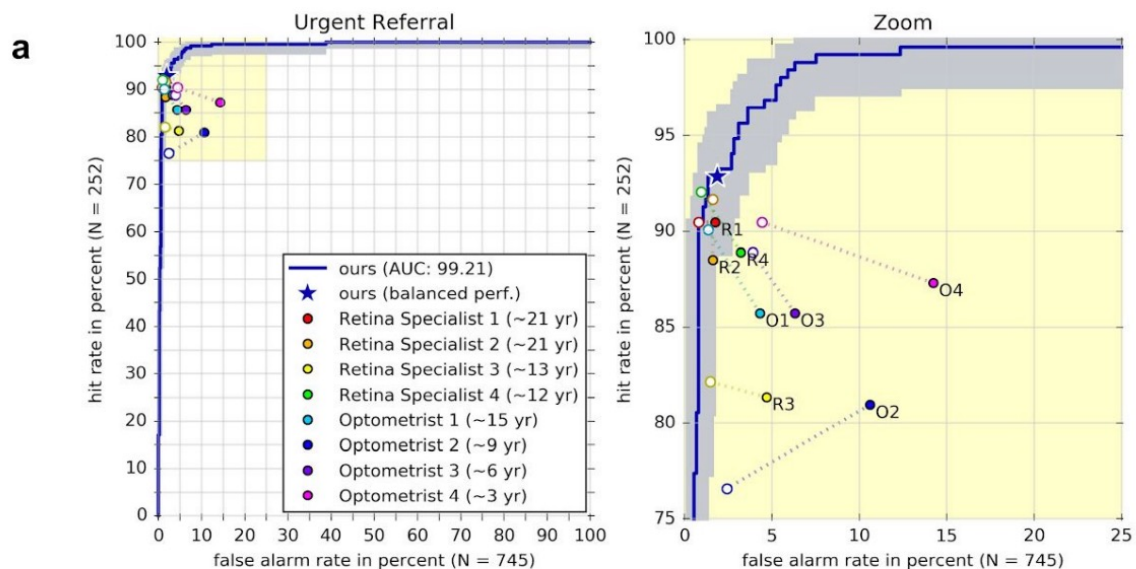
- Accuracy (ACC) = $(TP + TN) / (TP + TN + FP + FN)$
- Sensitivity (SEN) = $TP / (TP + FN)$
- Specificity (SPEC) = $TN / (TN + FP)$

	H₁: signal present	H₀: signal absent
Detection	True Positive	False Positive <i>type I error</i>
Null result	False Negative <i>type II error</i>	True Negative

ROC Curves : sensitivity / specificity



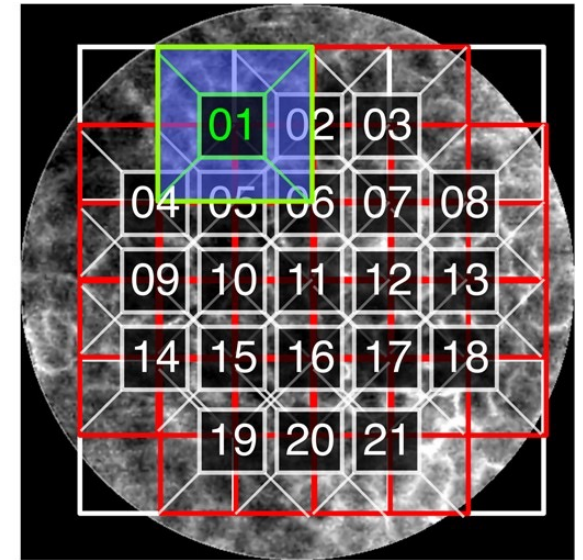
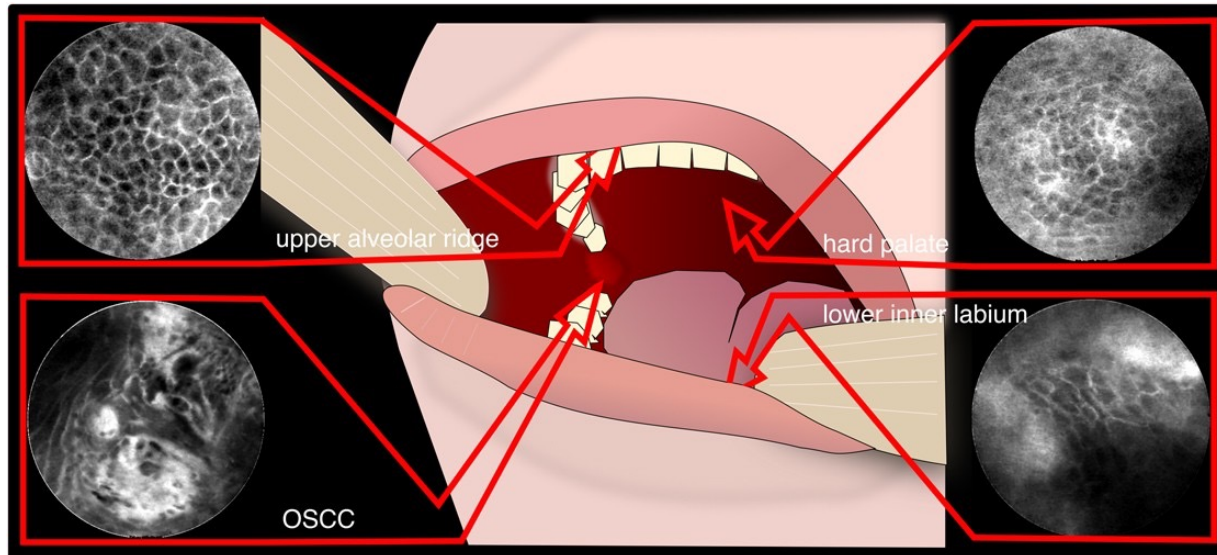
Retina

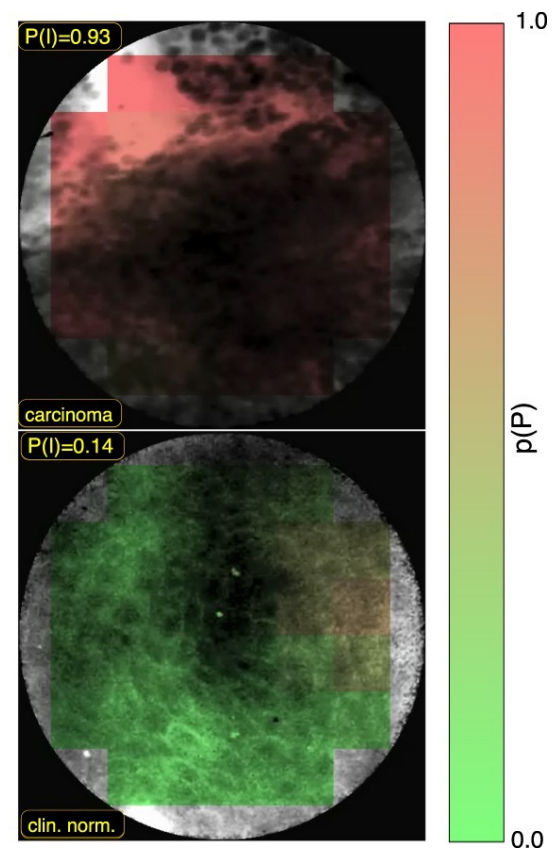
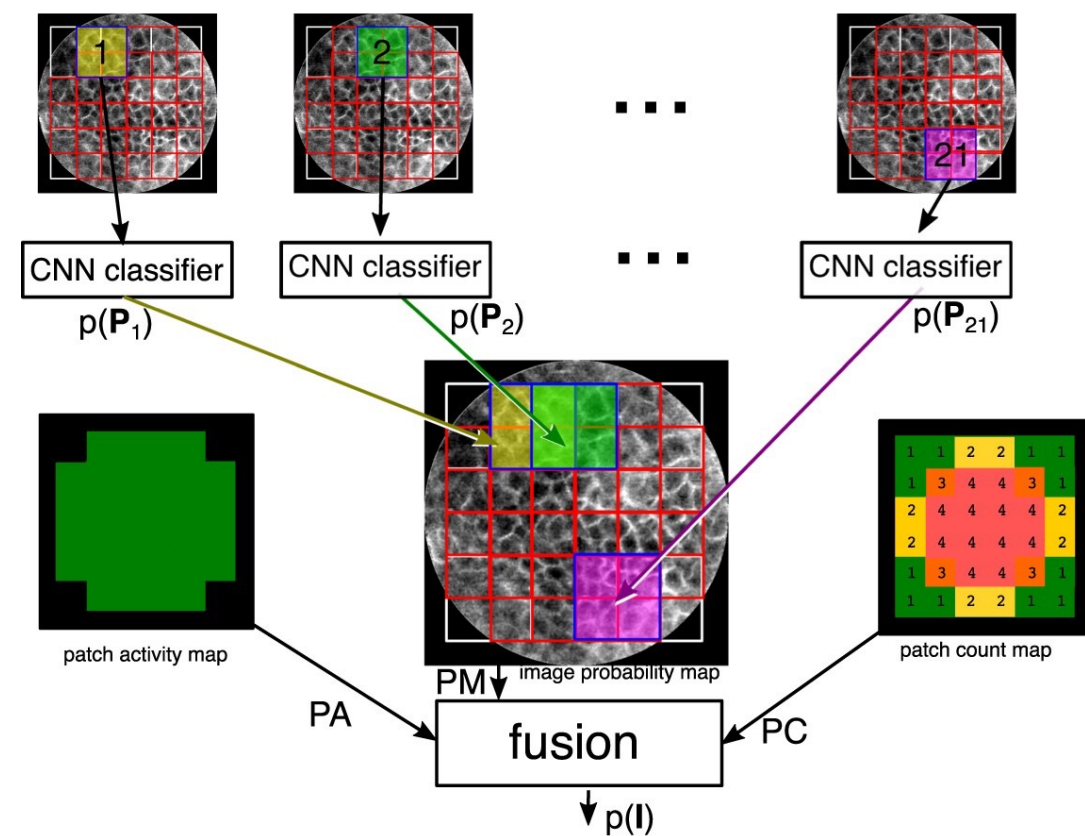
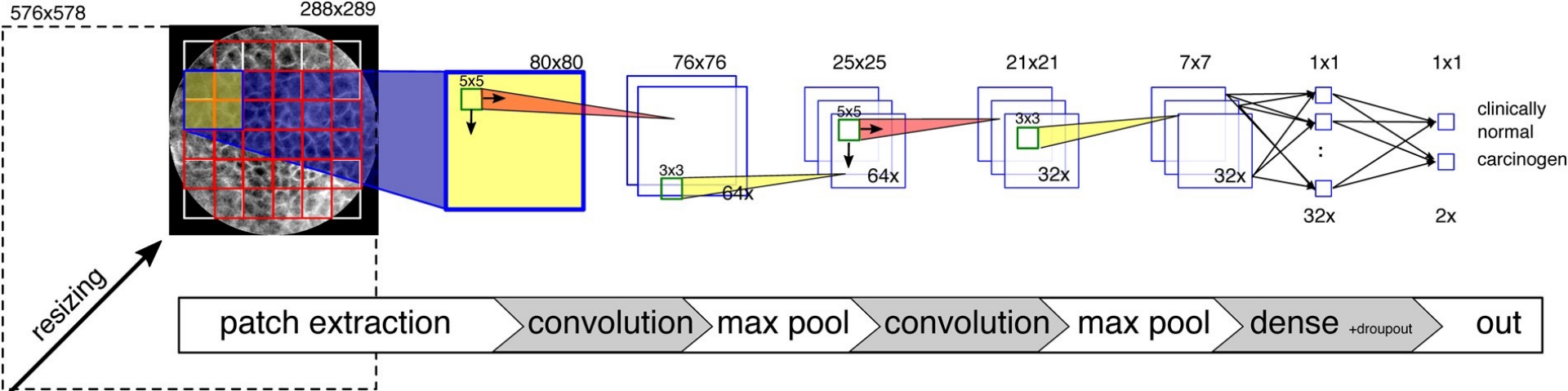


b

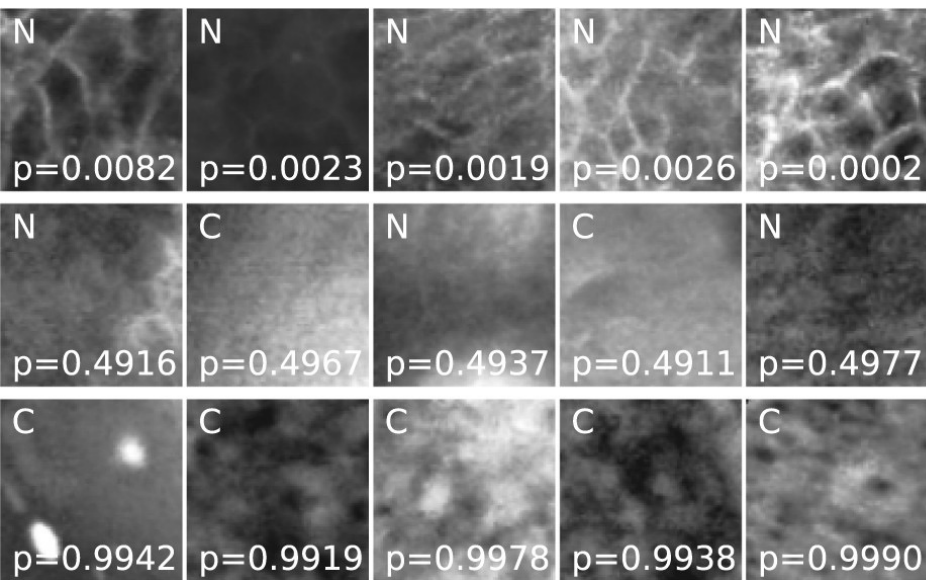
		Our model (OCT only)				Retina Specialist 1 (OCT+fundus+notes)				Retina Specialist 2 (OCT+fundus+notes)			
		Predicted Referral			Observation	Predicted Referral			Observation	Predicted Referral			Observation
		Urgent	Semi-urgent	Routine	Observation	Urgent	Semi-urgent	Routine	Observation	Urgent	Semi-urgent	Routine	Observation
Gold Standard Referral	Urgent	234	5	13	0	228	4	20	0	231	8	13	0
	Semi-urgent	3	225	2	0	3	223	4	0	1	226	3	0
	Routine	10	2	250	4	2	7	254	3	11	1	250	4
	Observation	1	1	14	233	1	1	10	237	0	2	20	227

Classifying laserendomicroscopy images of the oral cavity for cancer (Oral Squamous Cell Carcinoma (OSCC))



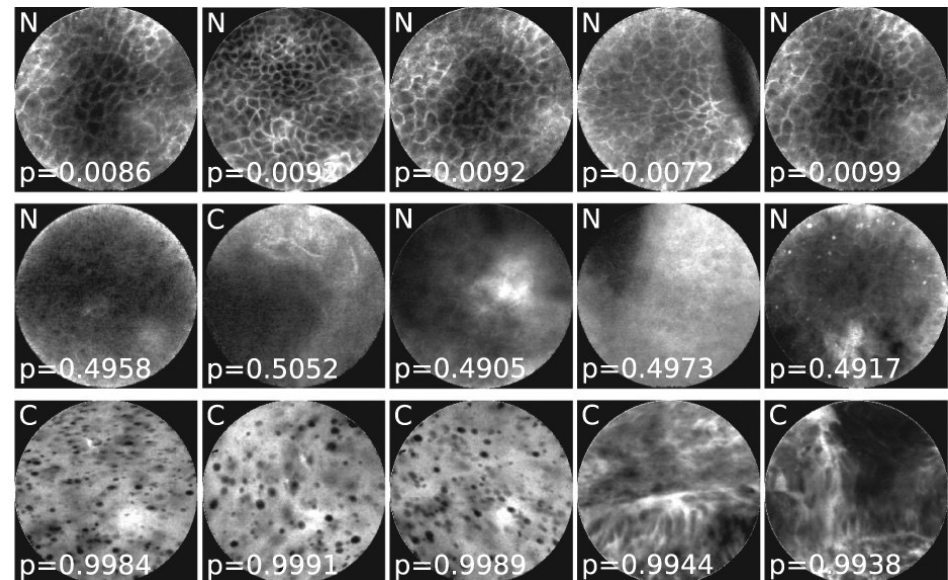


- Patches

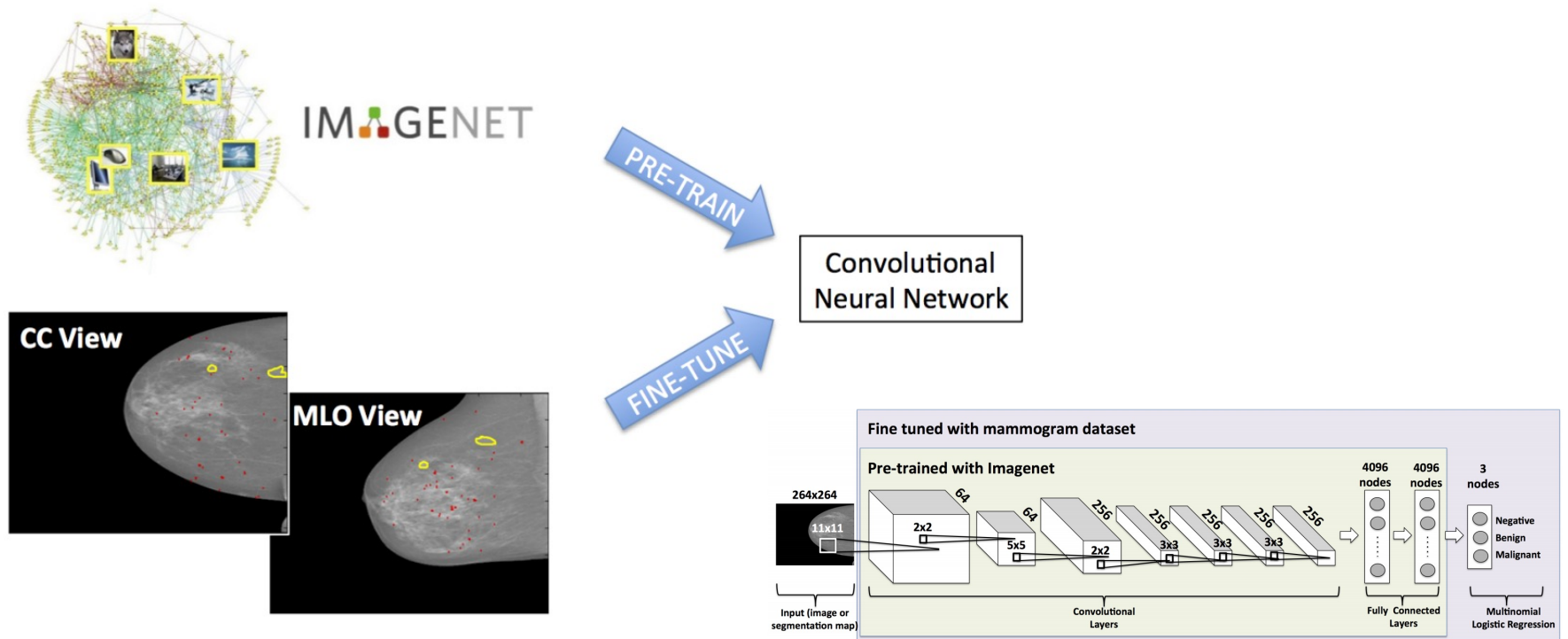


- N = normal, C = cancer

images



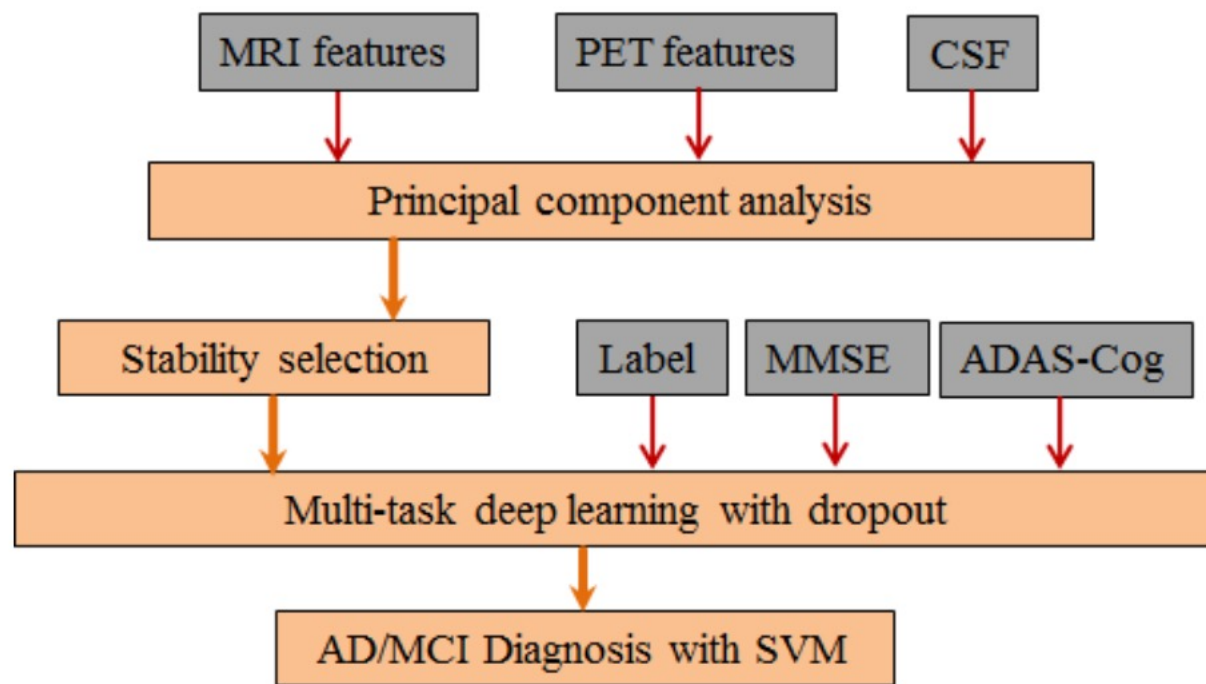
Classifying Mammogram Exams Containing Unregistered Multi-view Images and Segmentation Maps of Lesions



AD / MCI classification

- When and which MCI patients will develop AD ?
- Identify different progression stages of AD patients based on MRI and PET scans.
- MRI scans of 2146 subjects (803 for training and 1343 for validation) to predict MCI subjects' progression to AD dementia
- Databases :
 - ADNI (<http://adni.loni.usc.edu>) : MRI scans of 1711 subjects
 - AIBL (www.aibl.csiro.au) : 435 subjects.

A Robust Deep Model for Improved Classification of AD/MCI Patients



LEARNING :

- ADNI data set:
 - 51 AD patients,
 - 99 MCI patients
 - (43 MCI patients who converted to AD and 56 MCI patients who did not progress to AD)
- 52 healthy normal controls

Minimum Mental State Examination

Alzheimer's Disease Assessment Scale-Cognitive subscale

Fig. 1. Diagram of the proposed multi-task deep learning framework.

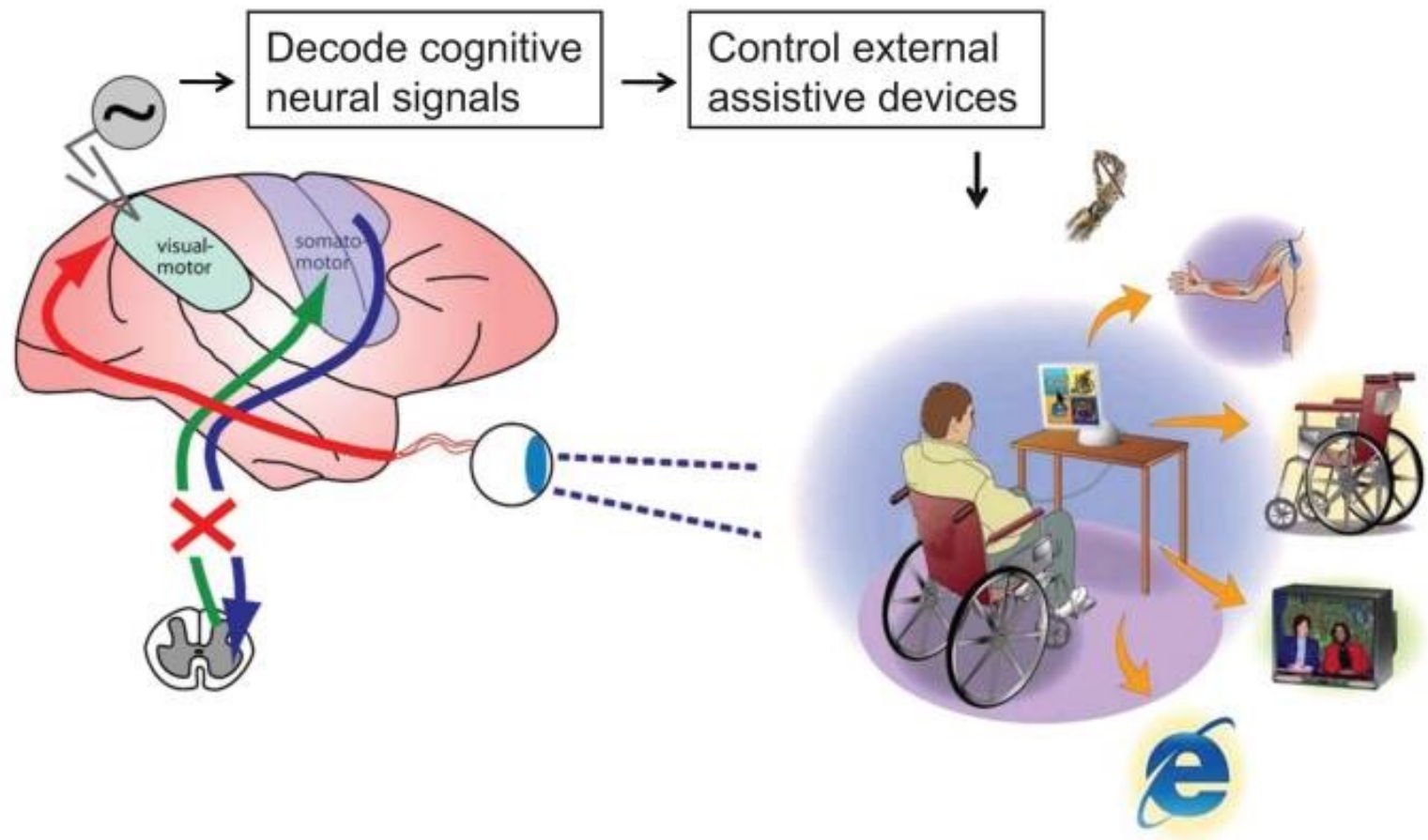
Results : accuracy of classification

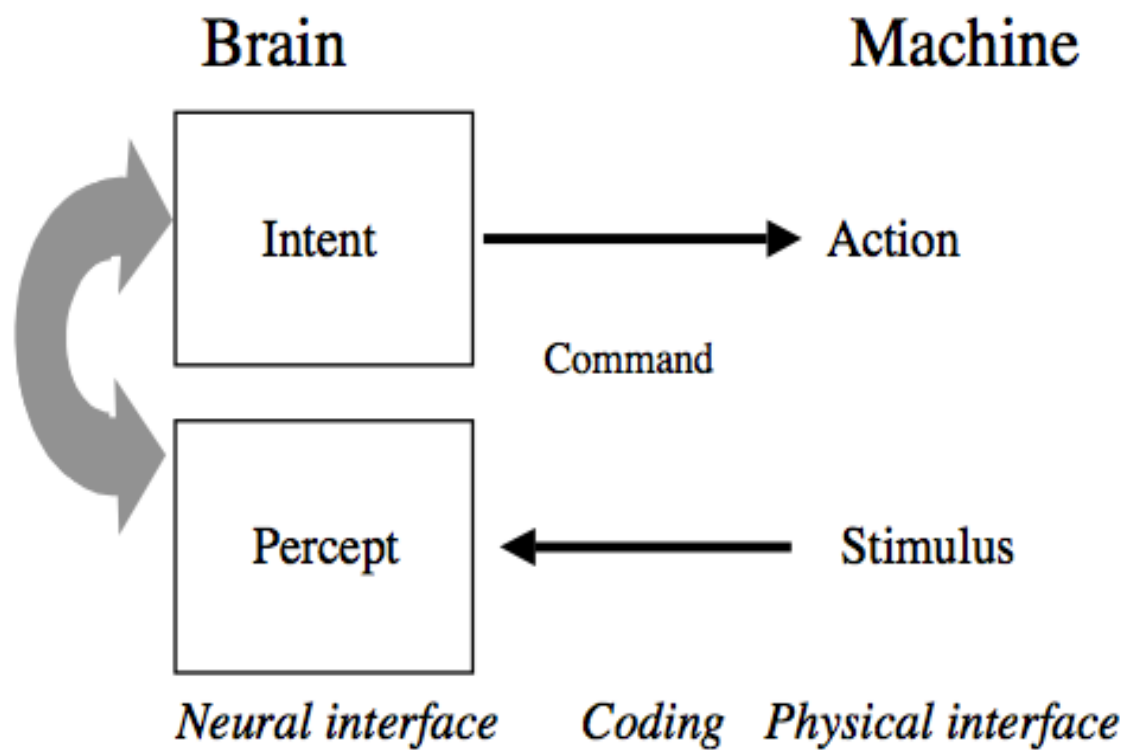
Tasks		Proposed
AD vs HC		91.4(1.8)
MCI vs HC		77.4(1.7)
AD vs MCI		70.1(2.3)
MCI.C vs MCI.NC		57.4(3.6)
Average		74.1

AD patients vs Healthy Control subjects (AD vs HC),
MCI patients vs HC (MCI vs HC),
AD patients vs MCI patients (AD vs MCI)
MCI-converted vs MCI-non converted (MCI.C vs MCI.NC).

2 – DL and biosignals

- Brain Computer Interface





BrainGate : Kathy Hutchinson

Table 2
 For each medical area this table lists the DL tasks, the obtained contributions and the main used DL models.

Area	Task	Contribution	Model
ECG	Arrhythmia	Classification and detection of arrhythmia levels	RNN [198], LSTM [200], CNN [218], CNN+RNN [219]
	Glucose	Evaluation of low glucose levels	CNN [204]
	Ventricular	Data relative to different races	CNN [201]
	Atrial	Choice of low-dimensional datasets	CNN [199]
		Use of ECG with different lengths	CNN [220]
	Foetal	Restriction of data only one ECG channel	AE [221]
	Apnoea	Classification of obstructive sleep apnoea and hypopnoea	CNN [203]
EEG	Epilepsy	Multimodal approach	CNN+LSTM [222]
		Increase in the SNR	CNN [223]
	Emotions	Improvement in performance in heterogeneous samples	DL [222].
PCG	Heart abnormalities	Optimal identification of heart errors	CNN [208]
PPG	Blood analysis	Use of signals correlations	LSTM [211], CNN [212,213]
EMG	Muscle conditions	Improvements in signals accuracy	CNN [214], AE [215]
NNS, SS	Biosignals analysis	Accurate analysis of biosignals	CNN [216,217]

3 – Electronic health records and DL

- CBR , **Content-Based Retrieval** : technique for knowledge discovery in massive databases (« Big Data »)
- Generating reports from images
- Finding similar cases in previous records
- COVID + connected watches

- EHR are growing
- 10 million patients over a decade
- A single hospitalisation = 150000 pieces of data
- Then : understand questions in natural language such as :
what is this patient's problem list ?
- Help to transcript patient visits (automatic speech recognition)

Table 4

This table lists the medical tasks, the contributions provided and the DL models most frequently used in EHR.

Area	Task	Contribution	Model
EHR	Disease prediction	Integration of different medical data in cancer analysis	CNN [263]
		High accuracy in dyslipidemia prediction	LSTM [261]
		Improvement in the imbalance problem in heart failure	CNN [264]
		Impact of external databases in sleep staging evaluation	CNN [265]
		Mortality risk estimation by using patients' historical information	LSTM [262]
		Optimal heart failure prediction	LSTM [51,259]
		Integration of structured and unstructured data for the prediction of acute kidney injury subtypes	LSTM [260]
	Risk analysis	Organization of the data on the basis of semantic spheres	GCNN [266]
		Feature extraction by reducing the data dimensions	AE [268]
	Treatment plans	Construction of a treatment plans by using small population sets	LSTM [267]
		Estimation of a treatment plan by exploiting data correlations	CNN [6]

4 – Other applications

- Predict 3D structure of proteins from aminoacids :
AlphaFold
 - understanding diseases
 - designing drugs
 - developing personalized medicine.



Deep learning: new computational modelling techniques for genomics

- As a data-driven science, genomics can use machine learning to capture dependencies in data and derive novel hypotheses.
- The ability to extract new insights from the exponentially growing volume of data requires more expressive machine learning models.
- DL : used for, for example, predicting the impact of genetic variation on gene regulatory mechanisms such as DNA accessibility and splicing.

Table 3
For each medical area this table lists the medical tasks, the contributions obtained and the DL model most frequently used.

Area	Task	Contribution	Model
Genomics	DNA structure	Prediction of DNA missing values from dependences	RNN [230]
		Identification of DNA regions by exploiting the spatial configuration	GCNN [230]
		Simplification of DNA expressions by reducing noisiness	AE+RNN [232], AE [231]
	Disease prediction	High accuracy in the identification of sub-kinds of tumour; Personalized Treatments	Deep Triage [44]
		Improvement in cancer prediction in very sparse molecules	SAE [231]
		Parallel extraction of features from pure DNA expressions	CNN+RNN [234]
		Data integration	DAE [233], AE [232]
Transcriptomics	RNA structure	Reduction of data dimensionality and sparsity	AE [235,237]
		Exploitation of the spatial configuration of RNA molecules	CNN [235]
		Accurate classification of the RNA components	CNN [249], RNN [236]
	Disease prediction	Classification of tumour types	CNN [238]
		RNA variation analysis	CNN [239]
		Heterogeneous data integration	SAE [240]
	Drug discovery	New drug–target interaction identification	DL [250]
Proteomics	Protein structure	Molecular region identification	CNN [251], DeepGSH [252]
		Protein identification	CNN+LSTM [253], AE [254]
	Drug discovery	drug–target interactions	GAN [242]
		Scoring function construction	CNN [241]
Metabolomics	Diseases prediction	Improvement of prediction models	CNN [244]
	Drug discovery	Determination of optimal targets; decrease in drug toxicity	AE [246,255]
		Optimal molecular interactions	CNN [247]

Table 4
This table lists the medical tasks, the contributions provided and the DL models most frequently used in EHR.

Area	Task	Contribution	Model
EHR	Disease prediction	Integration of different medical data in cancer analysis	CNN [263]
		High accuracy in dyslipidemia prediction	LSTM [261]
		Improvement in the imbalance problem in heart failure	CNN [264]
		Impact of external databases in sleep staging evaluation	CNN [265]
		Mortality risk estimation by using patients’ historical information	LSTM [262]
		Optimal heart failure prediction	LSTM [51,259]
		Integration of structured and unstructured data for the prediction of acute kidney injury subtypes	LSTM [260]
	Risk analysis	Organization of the data on the basis of semantic spheres	GCNN [266]
		Feature extraction by reducing the data dimensions	AE [268]
	Treatment plans	Construction of a treatment plans by using small population sets	LSTM [267]
		Estimation of a treatment plan by exploiting data correlations	CNN [6]

Some crucial points

- quality and size of the dataset :
 - unavailability of dataset
 - Annotations take time and agreement between experts
 - rare diseases are underrepresented in the data sets.
- Privacy and Legal Issue
 - share the medical data ?
 - Anonymisation, but stil...
- Data Interoperability and Data Standards
 - nature of data differ from hardware to hardware
 - combine several dif-ferent datasets for better algorithms learning and accuracy. I
 - Health data should be standardized and shared between providers
- Black Box and Deep Learning



Original Investigation | Health Informatics

Large Language Model Influence on Diagnostic Reasoning A Randomized Clinical Trial

Ethan Goh, MBBS, MS; Robert Gallo, MD; Jason Hom, MD; Eric Strong, MD; Yingjie Weng, MHS; Hannah Kerman, MD; Joséphine A. Cool, MD; Zahir Kanjee, MD, MPH; Andrew S. Parsons, MD, MPH; Neera Ahuja, MD; Eric Horvitz, MD, PhD; Daniel Yang, MD; Arnold Milstein, MD; Andrew P. J. Olson, MD; Adam Rodman, MD, MPH; Jonathan H. Chen, MD, PhD

Abstract

IMPORTANCE Large language models (LLMs) have shown promise in their performance on both multiple-choice and open-ended medical reasoning examinations, but it remains unknown whether the use of such tools improves physician diagnostic reasoning.

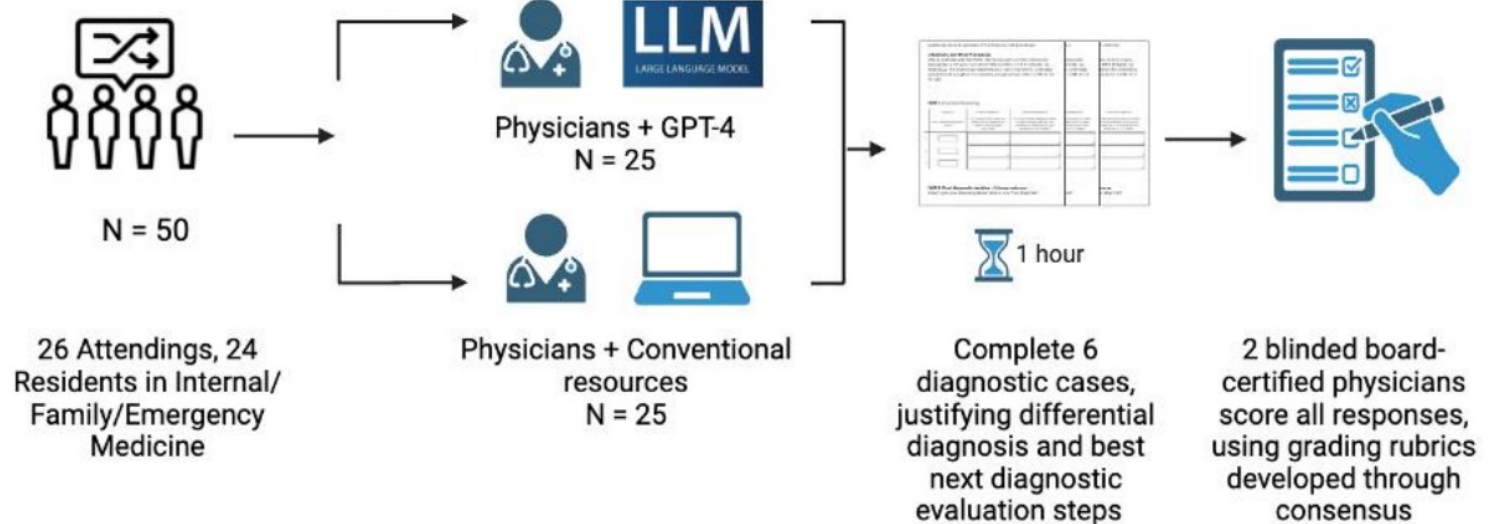
OBJECTIVE To assess the effect of an LLM on physicians' diagnostic reasoning compared with conventional resources.

Key Points

Question Does the use of a large language model (LLM) improve diagnostic reasoning performance among physicians in family medicine, internal medicine, or emergency medicine compared with conventional resources?

Findings In a randomized clinical trial including 50 physicians, the use of an LLM did not significantly enhance diagnostic reasoning performance compared with the availability of only conventional resources.

Meaning In this study, the use of an LLM did not necessarily enhance diagnostic reasoning of physicians beyond conventional resources; further development is needed to effectively integrate LLMs into clinical practice.



- **Type d'étude :** Essai clinique randomisé en simple aveugle
- **Participants :** 50 médecins (26 seniors et 24 résidents) en médecine interne, médecine générale et urgences.
- **Intervention :**
 - **Groupe LLM :** Accès à GPT-4 + ressources conventionnelles
 - **Groupe contrôle :** Accès uniquement aux ressources conventionnelles
- **Tâche :** Résolution de **6 cas cliniques** en 60 minutes
- **Critères d'évaluation :**
 - Score de raisonnement diagnostique (exactitude du diagnostic différentiel, arguments pour et contre, étapes suivantes)
 - Temps passé par cas
 - Exactitude du diagnostic final

Diagnostic Vignette

History Of Present Illness

A 76M comes to his PCP complaining of pain in his back and thighs for 2 weeks. He has no pain sitting or lying, but walking causes severe pain in his low back, buttocks and calves. He feels febrile and tired. He was told by the referring cardiologist that his recent tests results since the pain started showed a new anemia and azotemia. A few days before the onset of the pain he had undergone coronary angioplasty. Heparin was administered for 48 hours.

Past Medical History

Ischemic heart disease had first been diagnosed ten years earlier, at which time a coronary artery bypass procedure was done.

Physical Examination

VITALS: 99.6° F.; pulse was 94/min and regular; BP was 110/88 mmHg.

GEN: Well appearing

CARDS: There is a grade III/VI apical systolic murmur.

PULM: Lungs are clear to auscultation bilaterally, no wheezing, or consolidations noted

ABD: Soft, non-tender to palpation

MSK: He does not have tenderness of his spine or pelvis. Spinal mobility is normal, as is the mobility of his hips. Standing is painless; however, pain is experienced in his low back, buttocks and calves within a minute of feeble running in place. The pain disappears shortly after exercise is discontinued.

Laboratory

WBC of 11.5 x 103 cells /μL; differential of 64% segs, 20% lymphocytes, 3% monocytes, 12% eosinophils and 1% basophil. The hematocrit was 28% and the platelet count was 315 x 103 /μL. The erythrocyte sedimentation rate was 99 mm/h. Urinalysis was normal except for 2+ proteinuria. Serum creatinine was 4.0 mg/dL; sodium was 145 mEq/L; potassium 4.0 mEq/L; chloride 105 mEq/L. SGOT was 27 U/L; GGT was 90 U/L; alkaline phosphatase was 153 U/L.

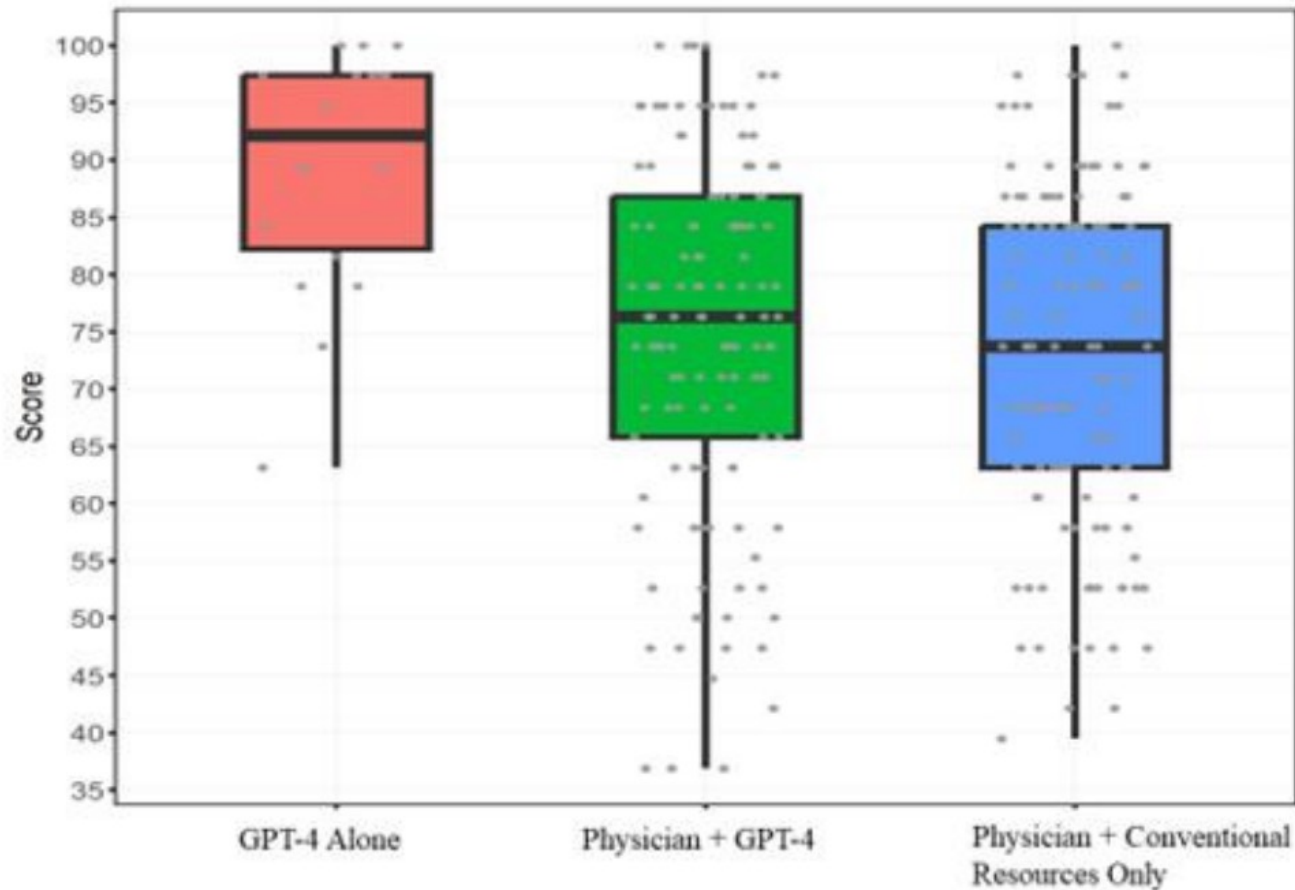
High Scoring Response

Part 1 – Structured Reasoning			
Questions		High Scoring Example	Scores
Question 1	Diagnosis - List 3 Possible	Cholesterol embolism	1/1
	Diagnosis	Acute interstitial nephritis	1/1
		Peripheral arterial disease	1/1
Question 2	Support Diagnosis - For each possible diagnosis listed, provide findings/risk factors supporting this hypothesis	Cholesterol embolism - Recent PCI, multi organ involvement (peripheral vasculature, renal, skin), rash, renal failure, eosinophilia, known CAD	2/2
		Acute interstitial nephritis - Renal failure, rash, +/- anemia, proteinuria, all after intervention with many medications (including ASA)	2/2
		Peripheral arterial disease - Leg pain associated with ambulation, improves with rest	2/2

Low Scoring Response

Part 1 – Structured Reasoning			
Questions		High Scoring Example	Scores
Question 1	Diagnosis - List 3 Possible	Interstitial nephritis	1/1
	Diagnosis	Contrast-induced nephropathy	0/1
		Pyelonephritis	0/1
Question 2	Support Diagnosis - For each possible diagnosis listed, provide findings/risk factors supporting this hypothesis	renal failure, eosinophilia, fever, rash	2/2
		contrast exposure, acute renal failure	2/2
		fever, aki	1/2
Question 3	Opposing Diagnosis - For each possible diagnosis listed, provide	Elevated GGT	0/2

eFigure 1. Distribution of Diagnostic Performance Scores of Physician + GPT-4 vs. Physician + Conventional Resources Only



Caption: Participants were randomized to access GPT-4 in addition to conventional diagnostic resources or to just conventional resources. GPT-4 alone scored 15.5 percentage points (95% CI 1.5 to 29, $p=0.03$) higher than the conventional resources group.

eTable 4: GPT Prompt and Responses for Diagnostic Case 1

GPT-4 Prompt	GPT-4 Responses
<p>Part 1 – Structured Reasoning</p> <p>You are an expert internal medicine physician solving a complex medical case for a test. You are going to receive a case vignette. After reading the case, I want you to give three parts of information.</p> <p>Here is the first part:</p> <p>First, you will list three possible diagnoses. Please number these 1., 2., 3.</p> <p>Next, for each of the three diagnoses listed, you will provide the findings or risk factors that support this hypothesis.</p> <p>Next, for each of the three diagnoses listed, you will provide findings opposing this hypothesis, or findings that were expected but not present.</p> <p>Part 2 – Final diagnostic decision</p> <p>Here is the second part: Please answer this question: Based upon your reasoning above, what is your final diagnosis?</p> <p>Part 3 – Additional Steps</p> <p>Here is the third part: Name up to 3 additional steps you would take in your diagnostic process.</p> <p>Diagnostic Case</p> <p>Here is the case:</p>	<p>Based on the provided medical case vignette, here are the three possible diagnoses along with supportive and opposing findings:</p> <p>Cholesterol Embolization Syndrome (CES)</p> <ul style="list-style-type: none"> • Supportive Findings: <ul style="list-style-type: none"> • Recent history of coronary angioplasty with heparin administration. • New onset anemia and azotemia post-procedure. • Purple, red, lacy rash over low back and buttocks (livedo reticularis). • Elevated erythrocyte sedimentation rate (ESR) and proteinuria. • Opposing Findings: <ul style="list-style-type: none"> • Lack of findings typically associated with CES, such as renal failure, gastrointestinal symptoms, or neurological deficits. <p>Peripheral Artery Disease (PAD)</p> <ul style="list-style-type: none"> • Supportive Findings: <ul style="list-style-type: none"> • Pain in the lower back, buttocks, and calves that worsens with exercise (claudication). • History of ischemic heart disease, suggesting atherosclerotic disease elsewhere. • Symmetrically reduced peripheral

Conclusions

The availability of an LLM as a diagnostic aid did not improve physician performance compared with conventional resources in a diagnostic reasoning randomized clinical trial. The LLM alone outperformed physicians even when the LLM was available to them, indicating that further development in human-computer interactions is needed to realize the potential of AI in clinical decision support systems.