Iot4Health: Uncertainty Prediction for Personalized Health

Debora Gil, debora@cvc.uab.es,
www.iam.cvc.uab.es

Computer Vision Center

Imaging Knowledge

Only Center in Europe fully devoted to Computer Vision

23 Years

+130 Staff

€2,8 M€/year Income

40 Diffusion/year

+2000 Followers

Research and Innovation

+50 Publications/year

+6 Thesis/year

55 PhD Students

40 Researchers

Technological Transfer

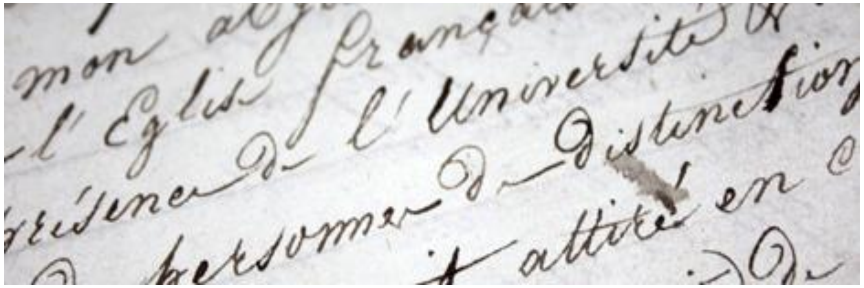
+15 Years Experience

11 Spin-offs

+40 New Clients/year

€1,0 M€/year Income

Areas of Excellence



Health and well-being

Computer assisted diagnosis, intervention and planning;

Well-being and ambient assisted living.

Mobility and transport

Advanced driving systems and autonomous driving;

Virtual worlds for ADAS;

Unmanned Aerial Vehicles.

Intelligent Content and Media

Cultural heritage (AR/VR)

Reading Systems – Document analysis

Surveillance

Industry 4.0

Quality control

AR/VR technologies for industry 4.0

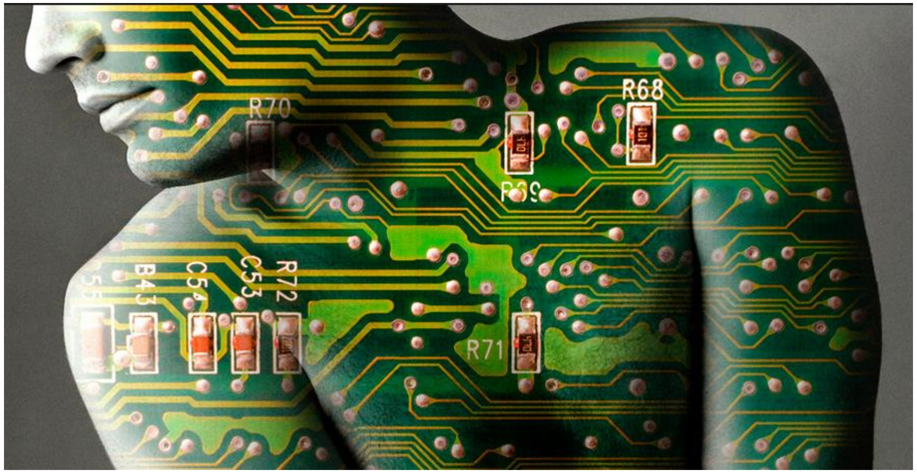
Robotic Vision

Courses

Data Science Engineering-UAB

Master in Computer Vision-UAB-UOC-UPF

Master in Internet of Things for eHealth-UAB



(on-line) Master in Big Data in Health-UAB-Tauli

Iot4Health: Uncertainty Prediction for Personalized Health

1. Introduction

2. State-of-Art AI Methods

3. Challenges and Hints

Introduction

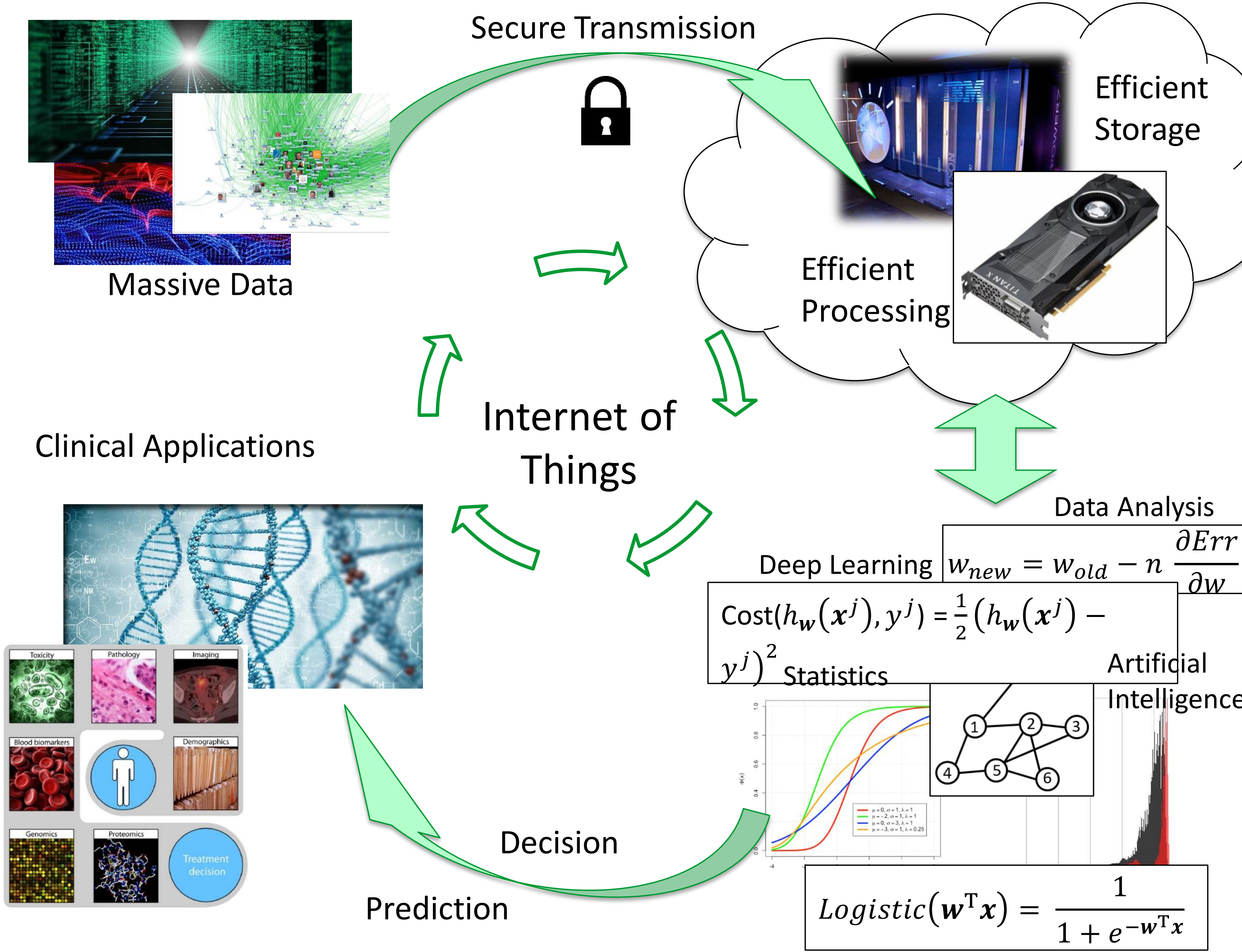
Clinical Issues

SUPPORT	GOAL	USE CASES
Diagnosis	Determine Lesion Pathology (degree of malignancy)	Cancer Diagnosis in 3D scans, in-vivo Diagnosis using endoscopy
Treatment	Predict Treatment Outcome/ Select Best Treatment	Personalized Cancer Treatments, Resynchronization Therapies
Intervenction	Planning / Guiding in Opertating Room without Altering Protocols	Biopsy using Endoscopy, Pace-maker to restore cardiac function

Standard Approach

Most clinical decisions are taken after human-based analysis of patient's data (scanners, blood analysis, ….) that requires highly specialized experts

- Inter and intra observer variability
- Analysis might be inconclusive
- Test repetition
- Patient anxiety


Secure Transmission
Efficient Storage
Efficient Processing
Massive Data
Internet of Things
Clinical Applications
Data Analysis
Deep Learning
$w_{new} = w_{old} - n \frac{\partial Err}{\partial w}$
$\mathrm{Cost}(h_{\boldsymbol{w}}(\boldsymbol{x}^j), y^j) = \frac{1}{2}(h_{\boldsymbol{w}}(\boldsymbol{x}^j) - y^j)^2$
Statistics
Artificial Intelligence
Toxicity
Pathology
Imaging
Blood biomarkers
Demographics
Genomics
Proteomics
Treatment decision
Decision
Prediction
$Logistic(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}) = \frac{1}{1 + e^{-\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}}}$

AI Support Systems

Computational systems that support in clinical decisions. Though final decision is taken by clinicians, AI systems can :

- Analyze data in systematic way
- Standardize criteria
- Reduce time to reach conclusion

REDUCE

- Variability across experts
- Non-experts training curve
- Inconclusive results

Computational Tasks in Clinical Support Systems

Systems manage and analyze data acquired from patients using medical devices: (3D) Radiological Data, Endoscopic Videos.

Three main tasks:

1. Lesion Localization

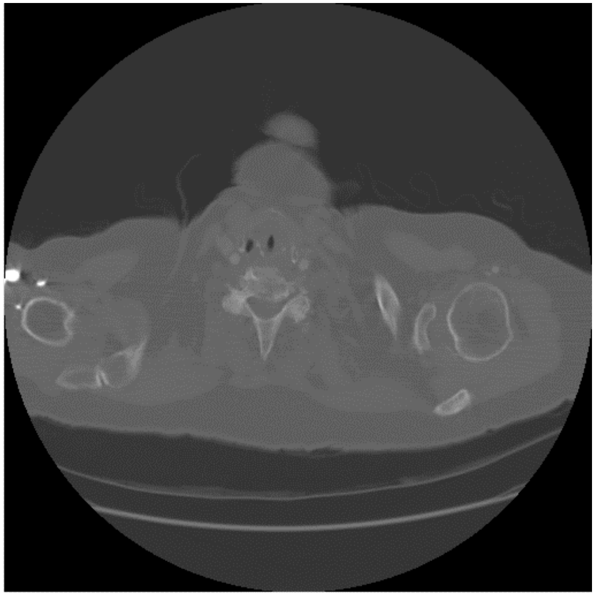
2. Lesion Segmentation

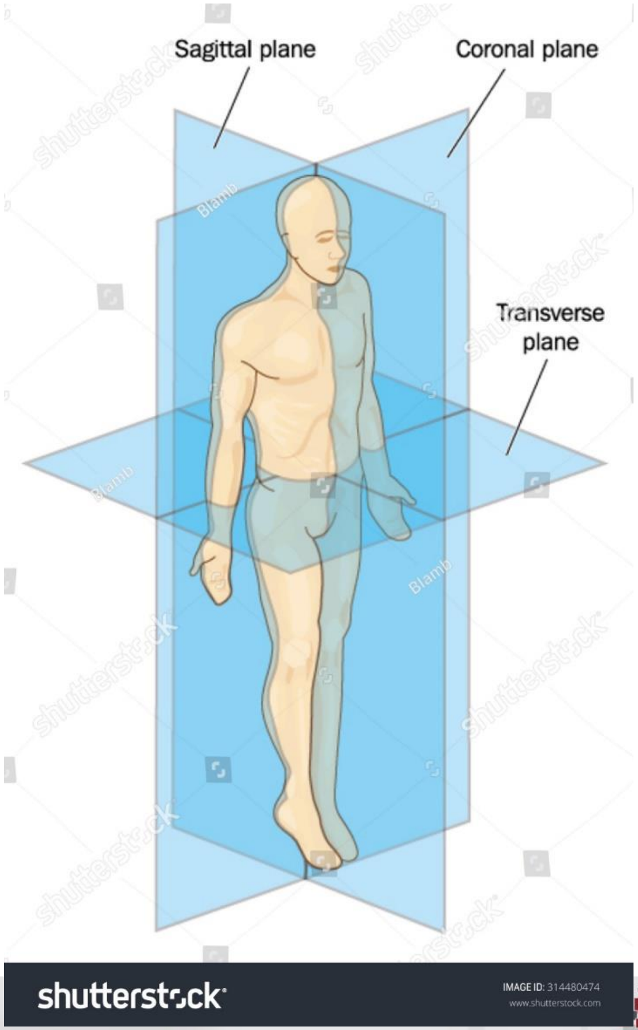
3. **Lesion Characterization**

Radiological Data

3D Volumes acquired from Magnetic Resonance (MR), Positron Emission Tomography (PET), Computerized Tomography (CT).

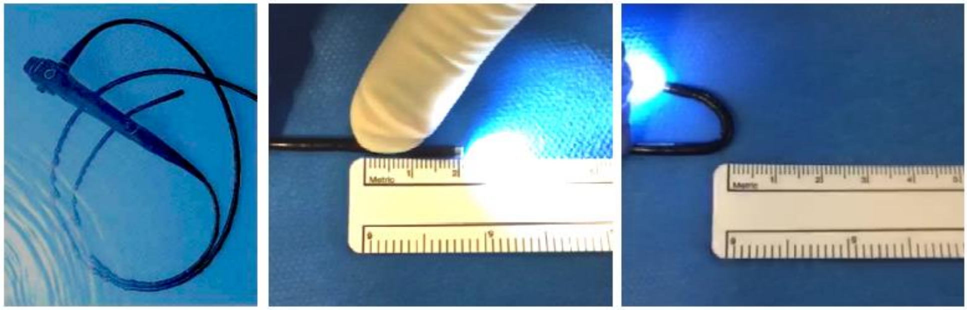
Thorax CT-scan Short Axis View

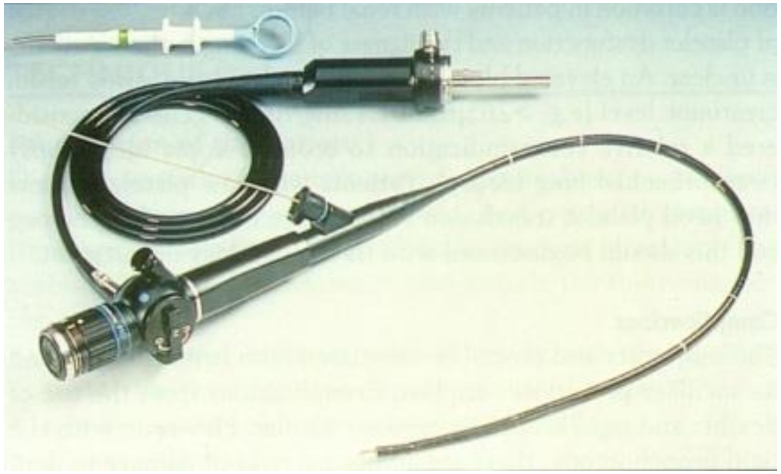




Endoscopic Data

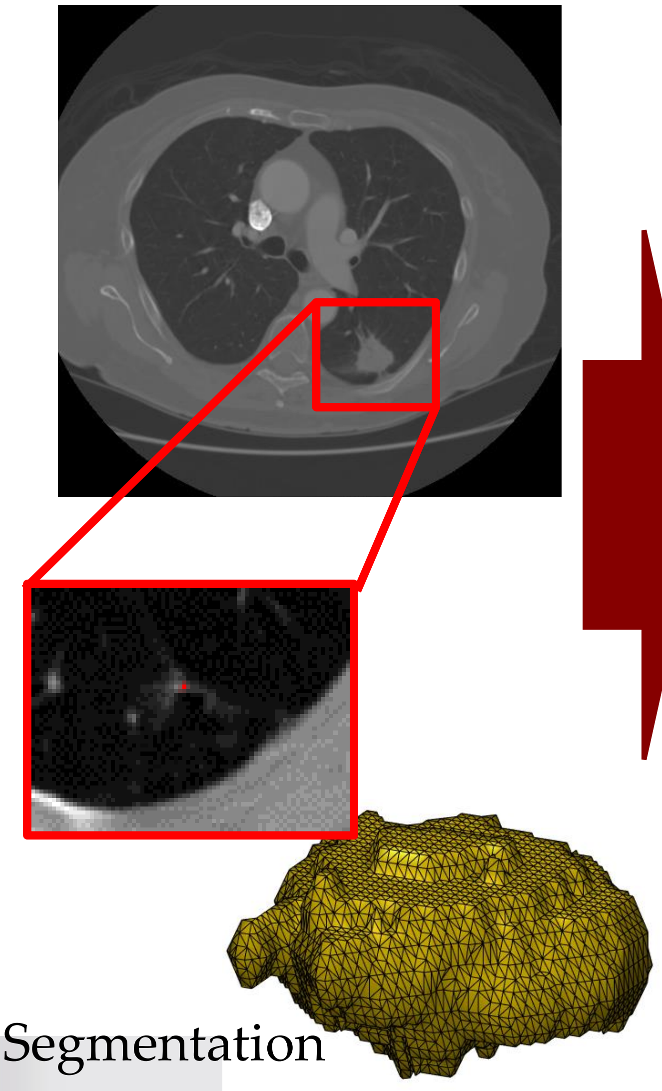
Videos of patient interior organs





Lesion Characterization

Detection



Segmentation

Radiomic Features

Volume
Area
Diameters
Regularity
Granularit
y
Brightness
Texture
Patterns

DATA ANALYSI S

Lesion Pathology

Treatment Outcome

Best Intervenction

AI Approach

Clinical issues are considered as a classification problem.

Lesions/patients are grouped into categories specific for each problem:

CLINICAL PROBLEM	GLOAL	IA APPROACH
Diagnosis	Lesion Pathology	Classify into bening, malign
Treatment	Treatment Outcome	Classify into responder, non-responder

AI Approach

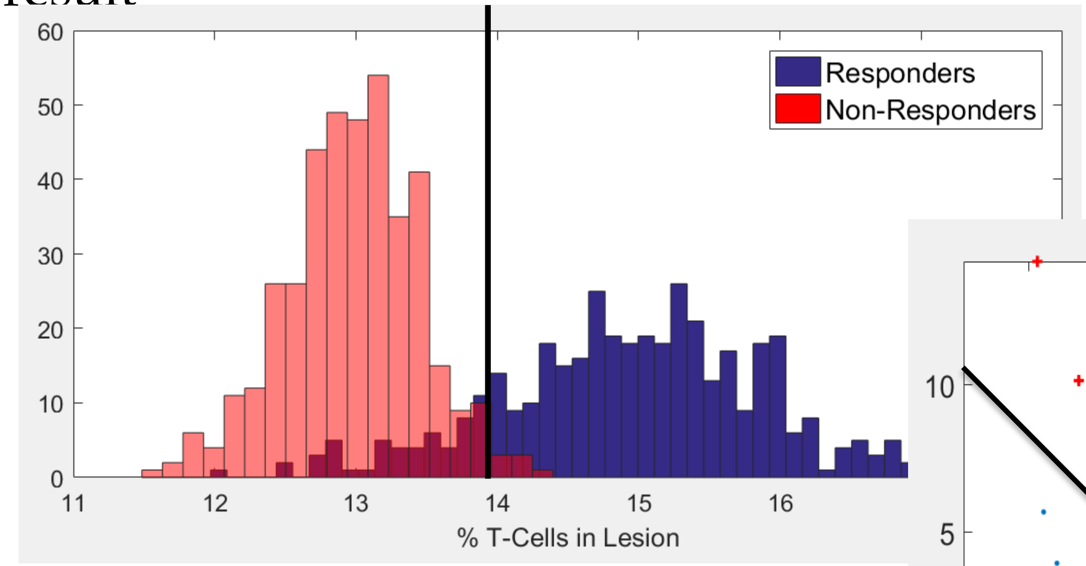
Cases (lesions, patients) are classified according to values of measures extracted from clinical data

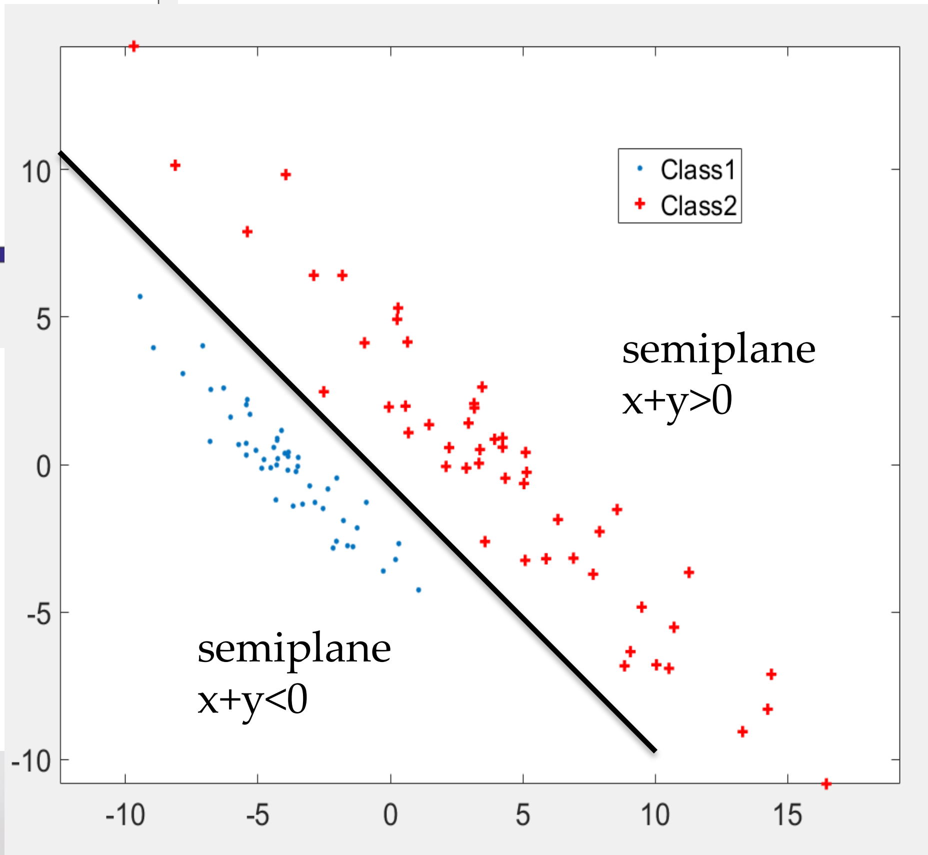
Compute (learn) the ranges (extreme values) of each measure that best differentiate (discriminate) each category:

CLINICAL PROBLEM	MEASURE	IA APPROACH
Diabetes Diagnosis	% Glucose in Blood	50<%<75 → Grade A Diabetes 75<% → Grade B Diabetes
Cancer Treatment	% Inmune Cells in Lesion	%<50 → Non-Responder

AI Approach

Separation in n-dim space with best compromise across classification errors using using probabilistic distribution of training population with known result



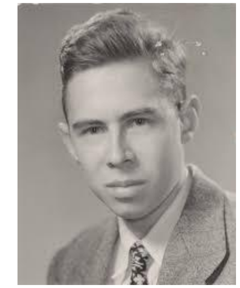


New case classified according to semi-plane it belongs to

State-of-Art AI Methods

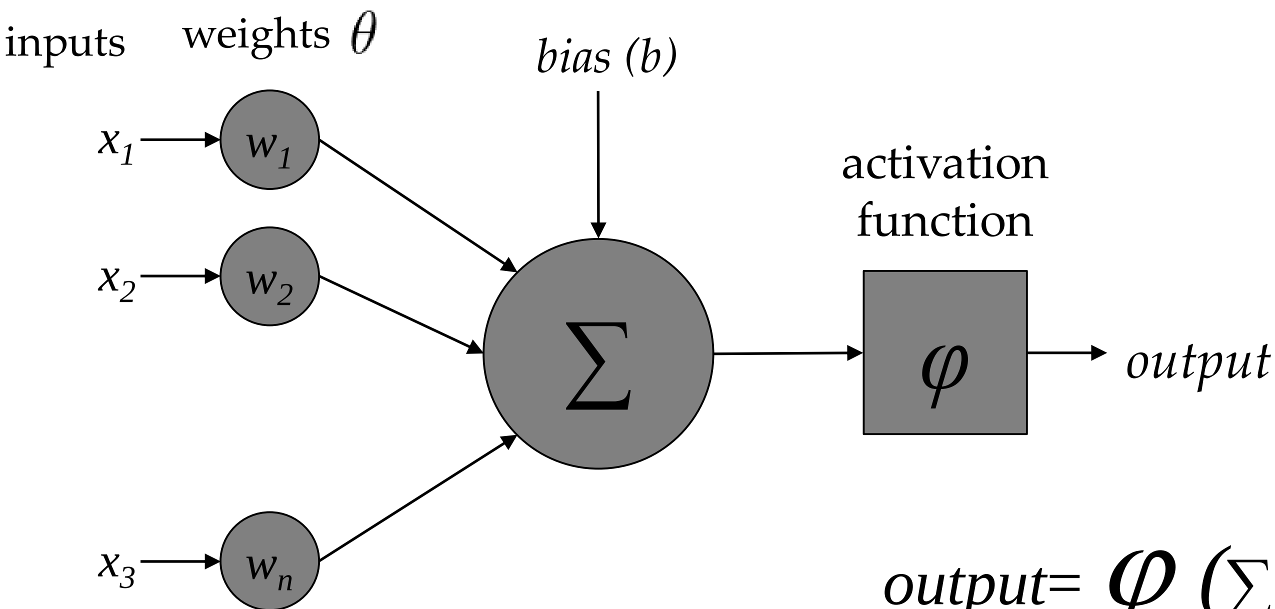
Neural Networks

Frank Rosenblatt (1957)



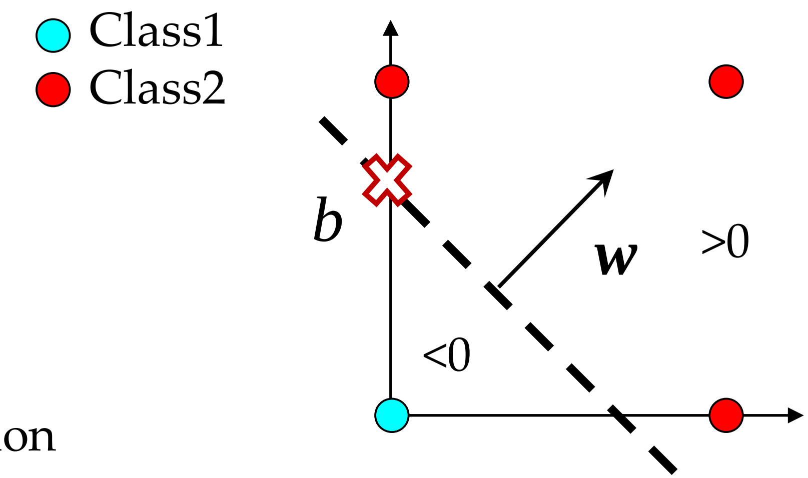
Perceptron

- Structure of an artificial neuron

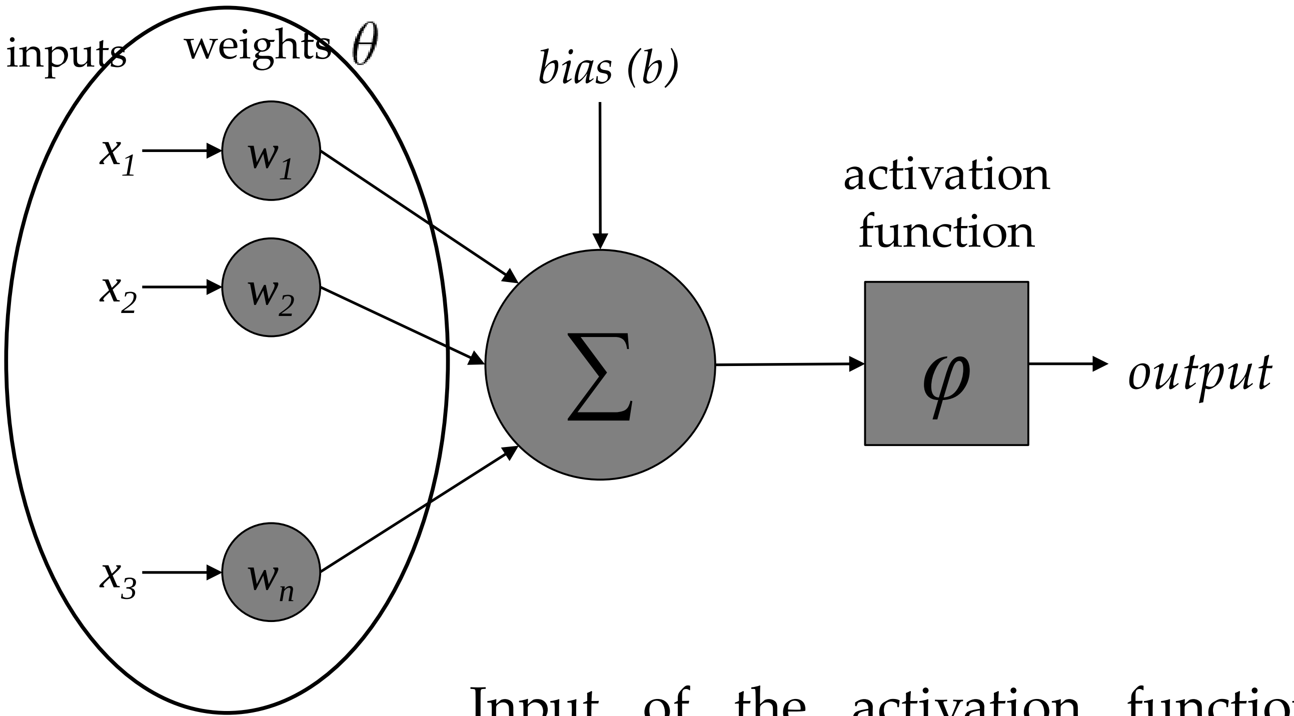


$$output = \varphi \left(\sum w_i x_i + b \right)$$

Neural Networks



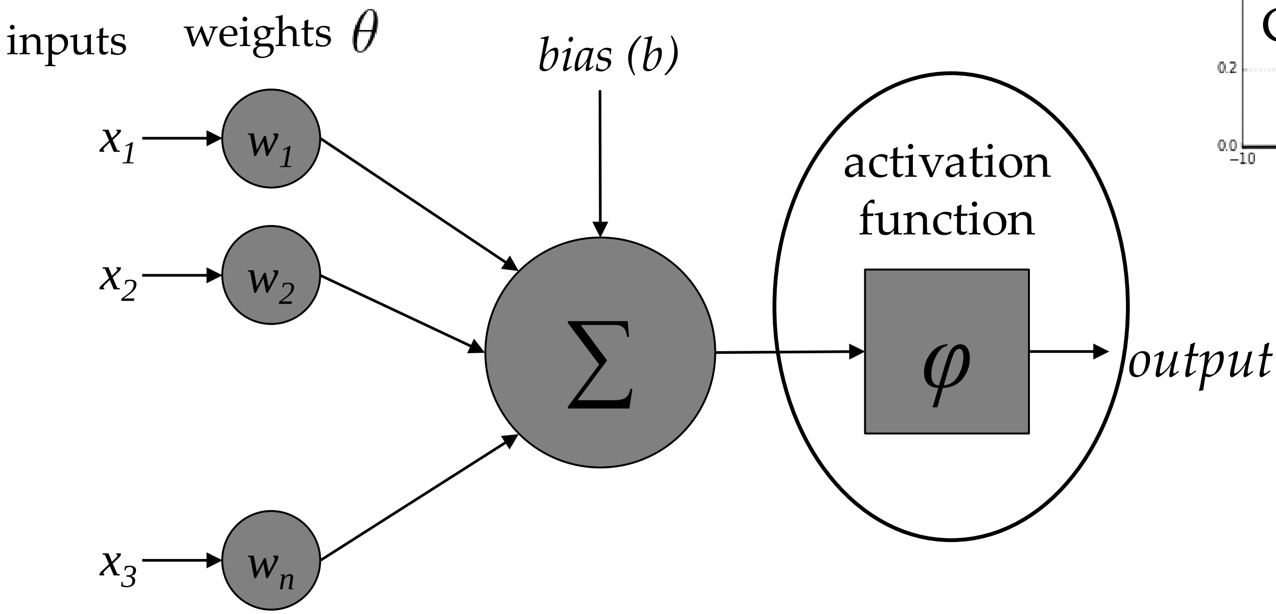
$$\mathbf{wx} + b = \sum_i w_i x_i + b$$

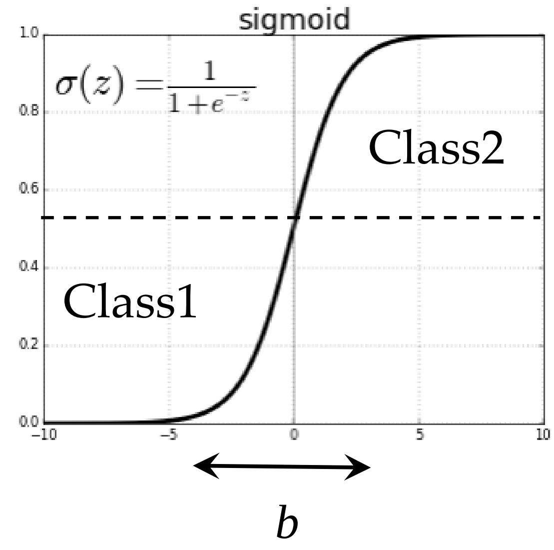


Input of the activation function is a linear function of neuron input data:

$$\mathbf{wx} + b = \sum_i w_i x_i + b$$

Neural Networks



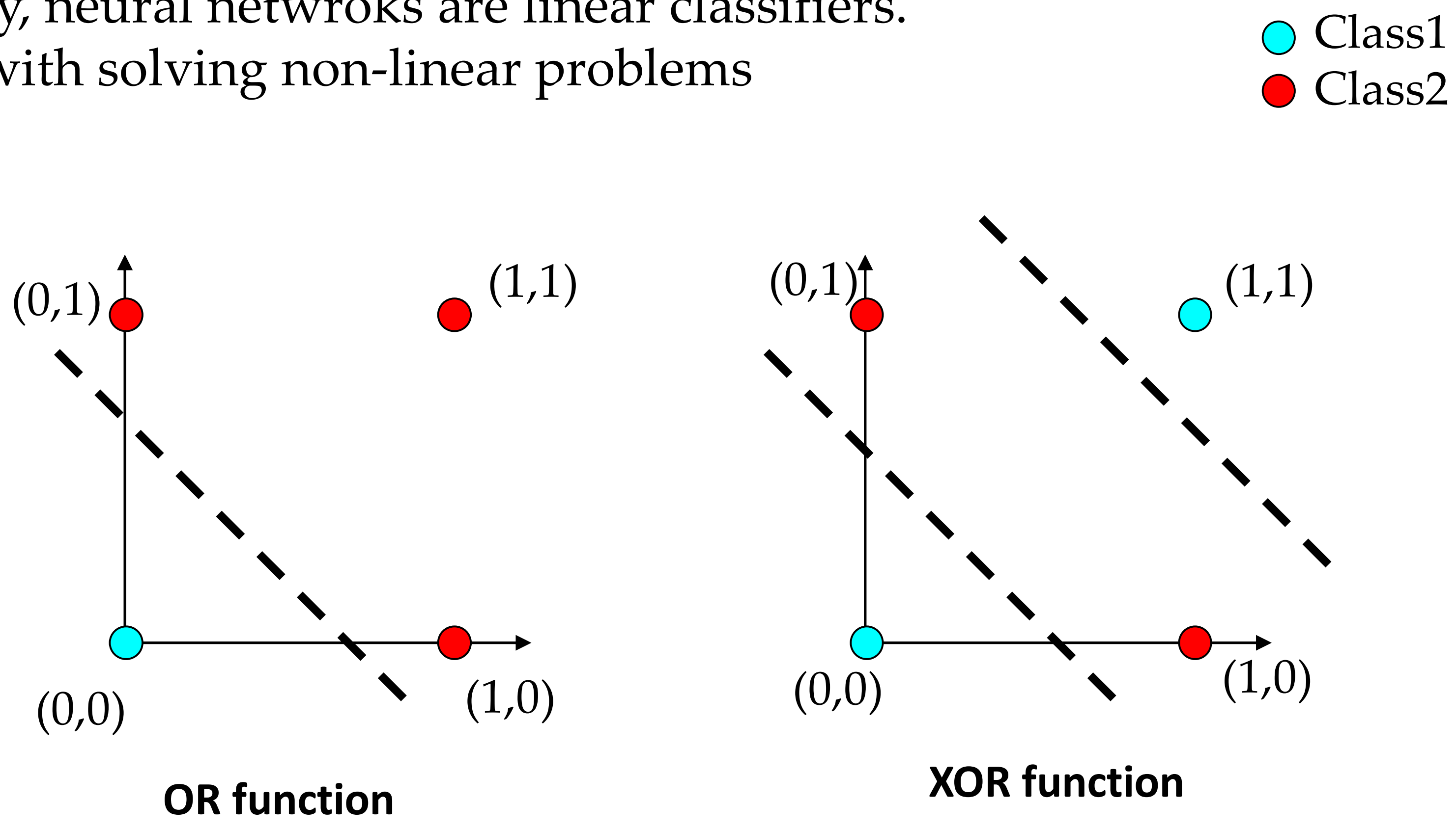


Linear classifier is modulated by (non-linear) activation function: sigmoid, tanh, ReLU, etc.

b acts like a threshold on the activation function

Neural Networks

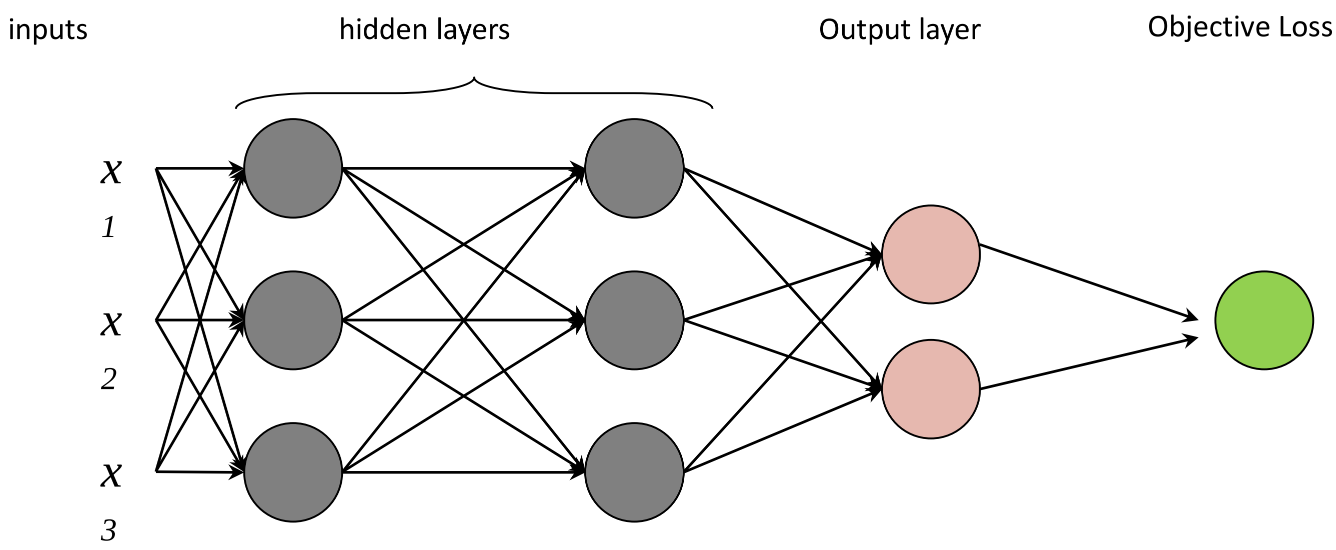
Intrinsically, neural netwroks are linear classifiers.
Problems with solving non-linear problems



Neural Networks

Multi Layer Perceptron (MLP)

- Connect neurons in multiple layers to model non-linear functions

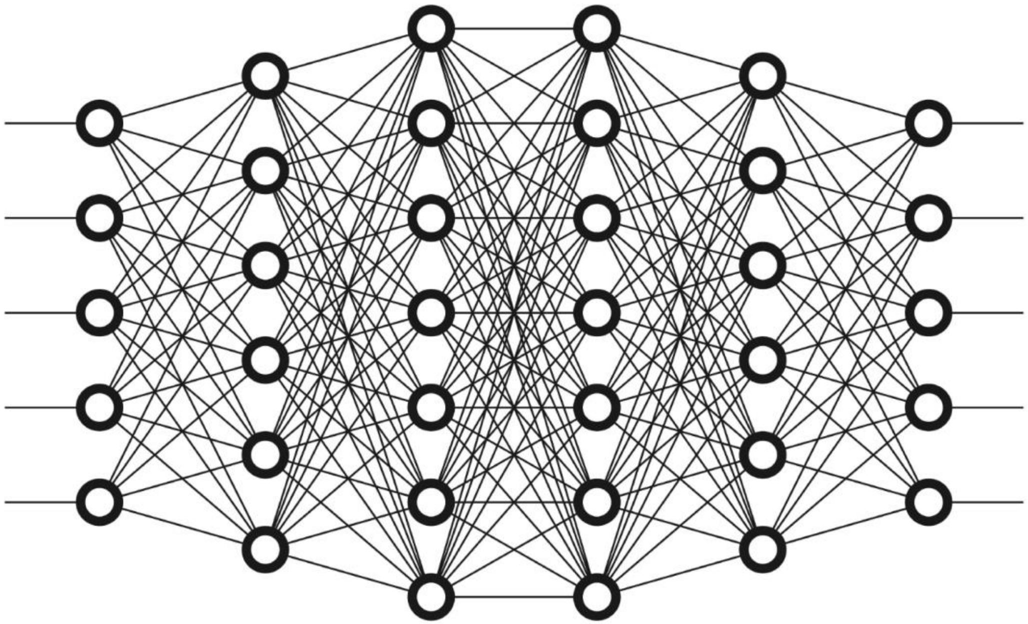


Each hidden layer models
an hyperplane

Neural Networks

Deep Neural Network

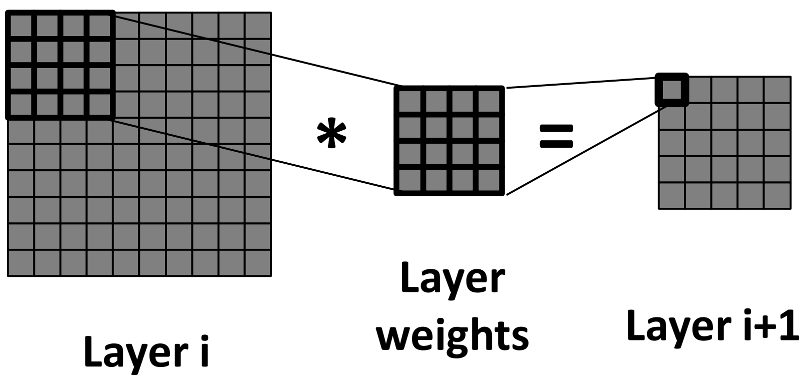
- Hierarchy of multiple layers of artificial neurons that processes information using non-linear transformations.



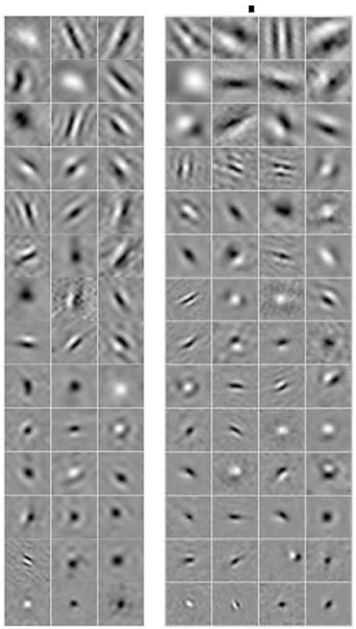
Neural Networks

Convolutional Neural Network (CNN)

- Instead of modelling the whole signal, neurons model signal in a region (act as convolution filters)



LeCun (1990s)



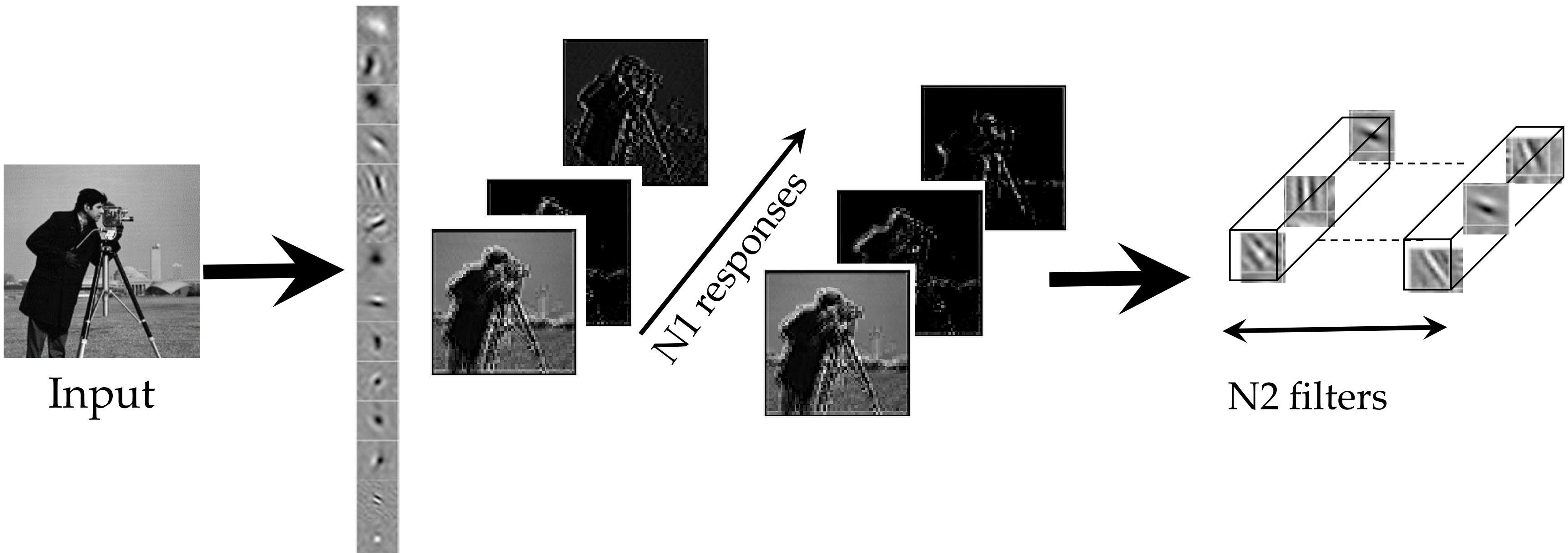
(c) Joel Zylberberg et al.

Biologically inspired (Receptive Fields of macaque V1 neurons)

Neural Networks

Convolutional Neural Network (CNN)

- Combine several convolution-activation layers
- After some convolution-activation layers, signal is downsampled



Conv1:
Filter bank with N1 filters of size n1xn1

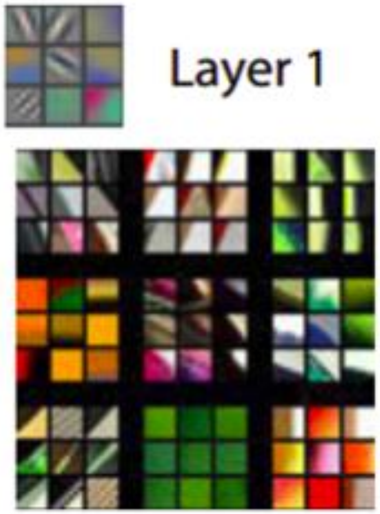
Activation Layer: threshold on Conv1 responses

Conv2: Filter bank with N2 filters of size n2xn2xN1.
They blend the N1 responses to Conv1 filters

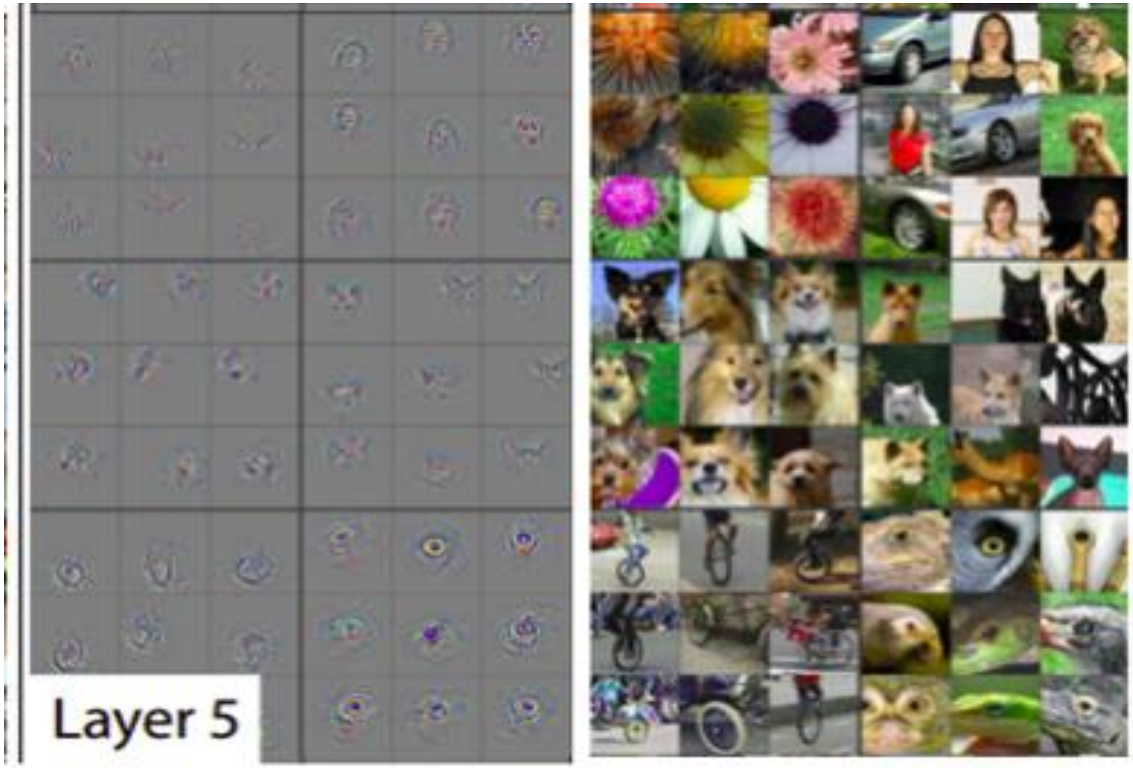
Neural Networks

Convolutional Neural Network (CNN)

- Hierarchy of neuron layers that mimic the brain and provide a multiescale decomposition ("wavelet"-like) of input data



First layers provide low-level local descriptors

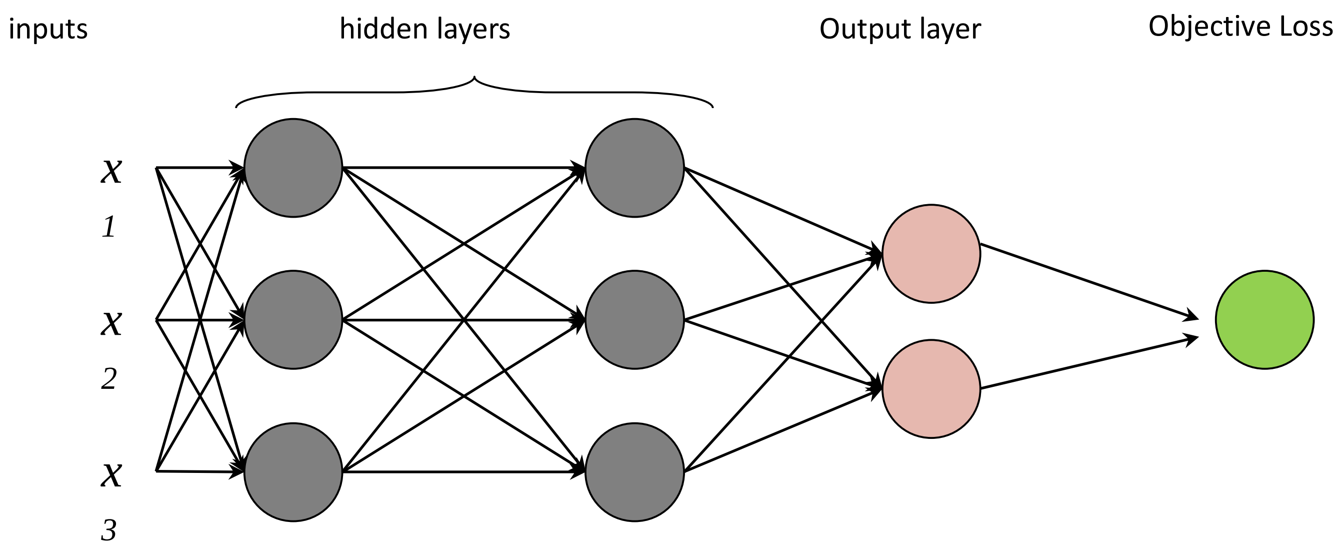


Deeper layers provide high-level global descriptors

Neural Networks

Training/Learning

- Network weights are learned (adjusted) to optimize loss: cross entropy (classification), square difference (representation spaces)

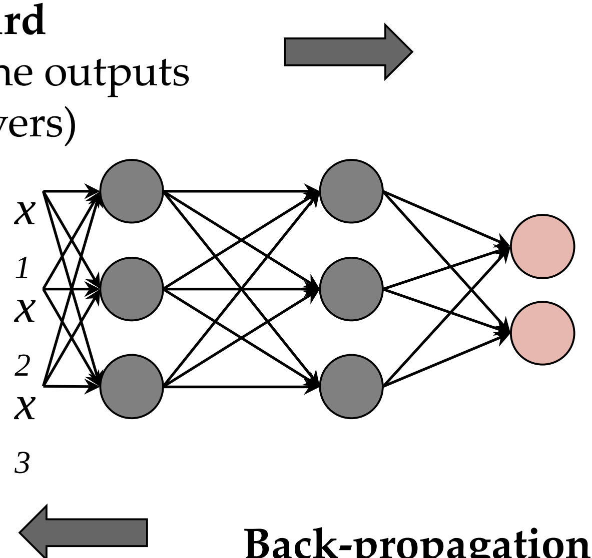


Neural Networks

Back Propagation

- Optimization by (Stochastic) Gradient Descent Scheme
- Convergence depends on:
  - initial weights
  - convexity of the loss function
  - population sampling (batch) for stochastic gradient

Feed forward
(Forward the outputs through layers)

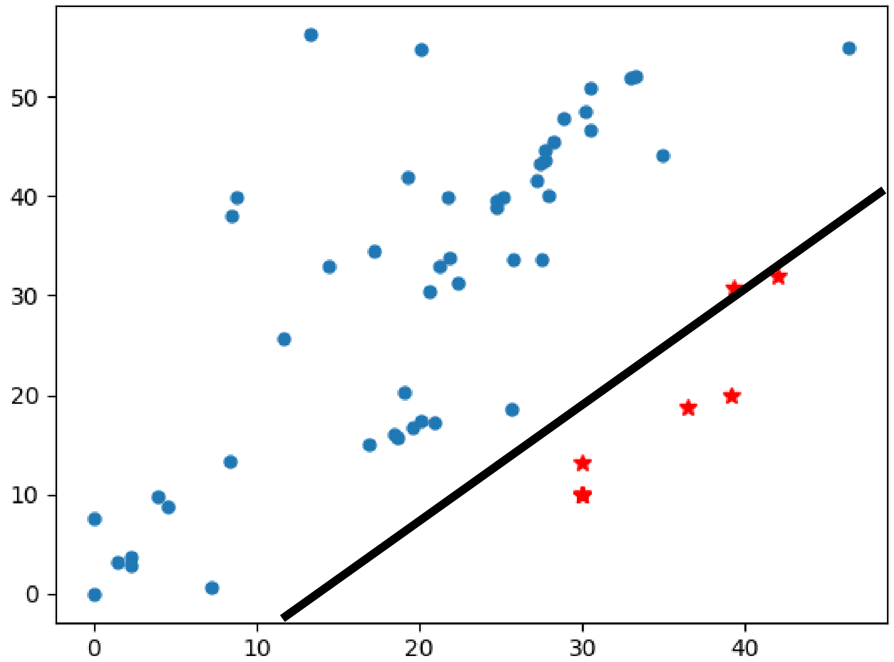


Back-propagation
(Propagate back the gradients to update weights)

Challenges & Hints

Challenges in Health Problems

Small sample size (SSS) unbalanced problems with several sources of uncertainty (variability) in data like acquisition parameters or intra-observer variability in annotations



Extreme values become highly influential

Model has low generalization (reproducibility) power (over-fiting)

Approaches to Challenges

Techniques to avoid them

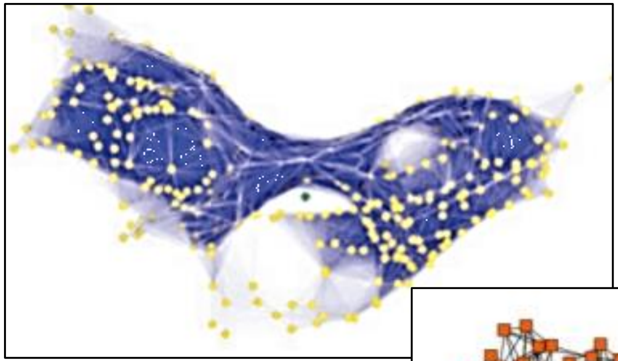
- **Sampling Strategies:**
    - Data Filtering
    - Data Augmentation

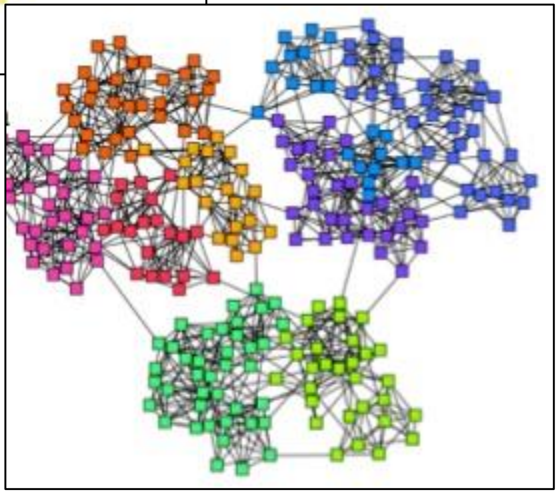
- **Uncertainty Modelling:**
    - MultiTask Learning
    - MC-Dropout
    - Feature Uncertainty Measures

Sampling Strategies

Data Filtering

- Usual approaches detect outliers using probabilistic global descriptions of population sample → Bad pose in SSS problems





Hint *:

Use algebraic topology and analysis of communities in social networks to provide a local description of feature space

* This is on-going research under project ToPiomics funded by the ATTRACT project under EC Grant Agreement 777222

Sampling Strategies

Data Augmentation

- Alter training images using known transformations

Original

Affine transformations (scale, rotate, translate)

Color shifts

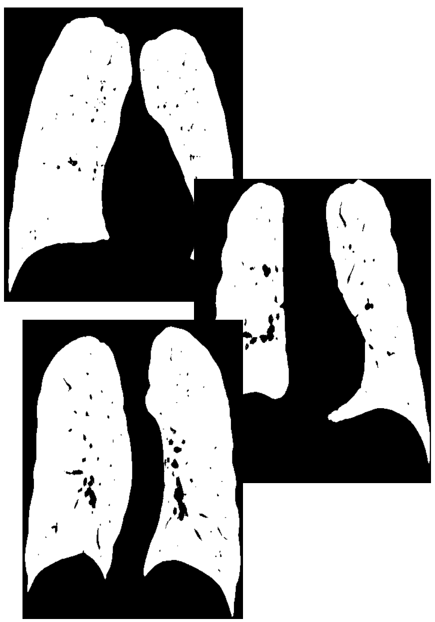
Flip

- Network will be invariant to the set of transformations
- Augmented data is highly correlated

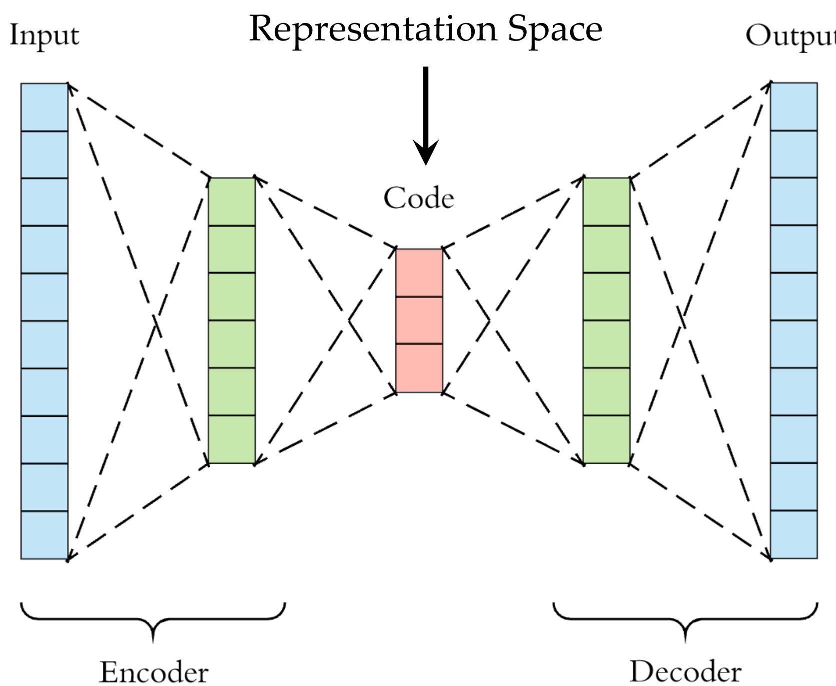
Sampling Strategies

Data Augmentation

Hint *: Generate new data from existing data sets using Auto-Encoder network to define a low dimensional representation space and statistical model (PCA)



Existing Samples



Generated Samples

* This is on-going research under project Up4Health funded by the Spanish Government under RETOS coordinated project RTI2018-095209-B-C21

Uncertainty Modelling

MultiTask Learning

- Problem: neurons adapt too much to inputs
- Solution: Learn different tasks simulateously

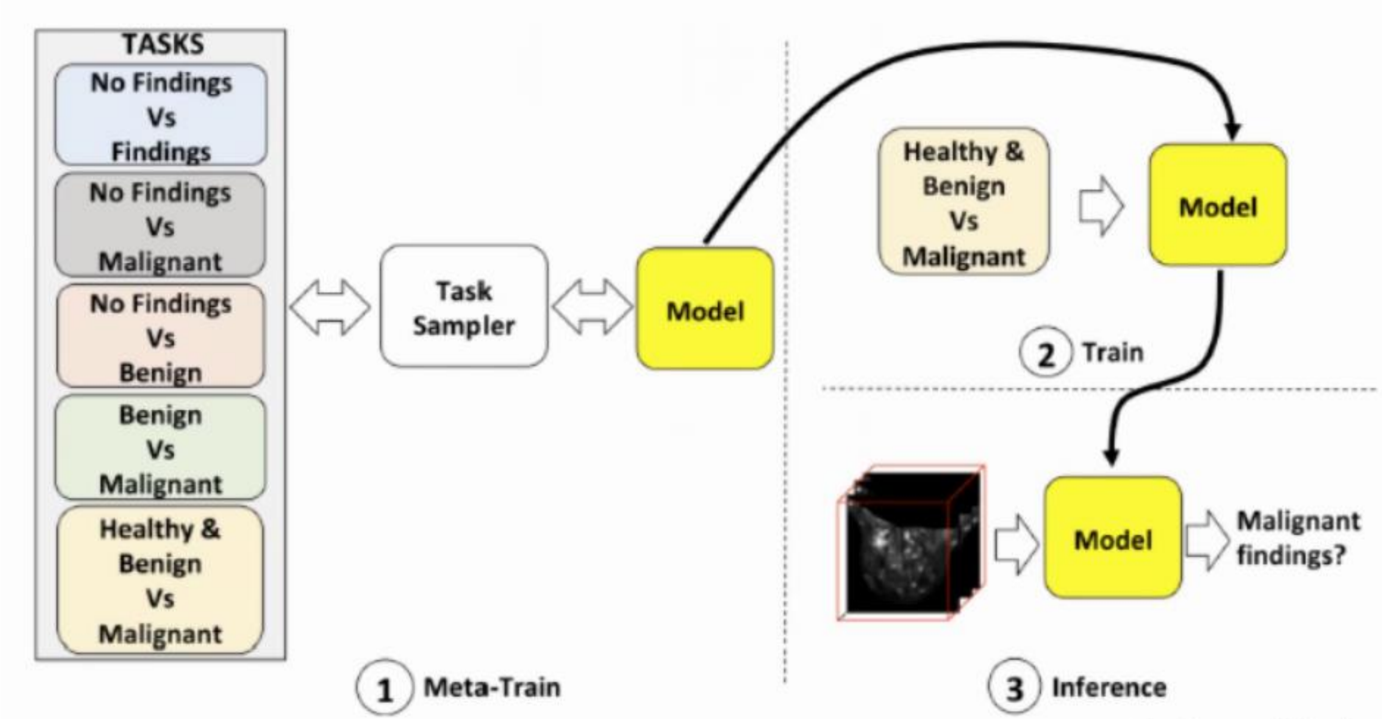
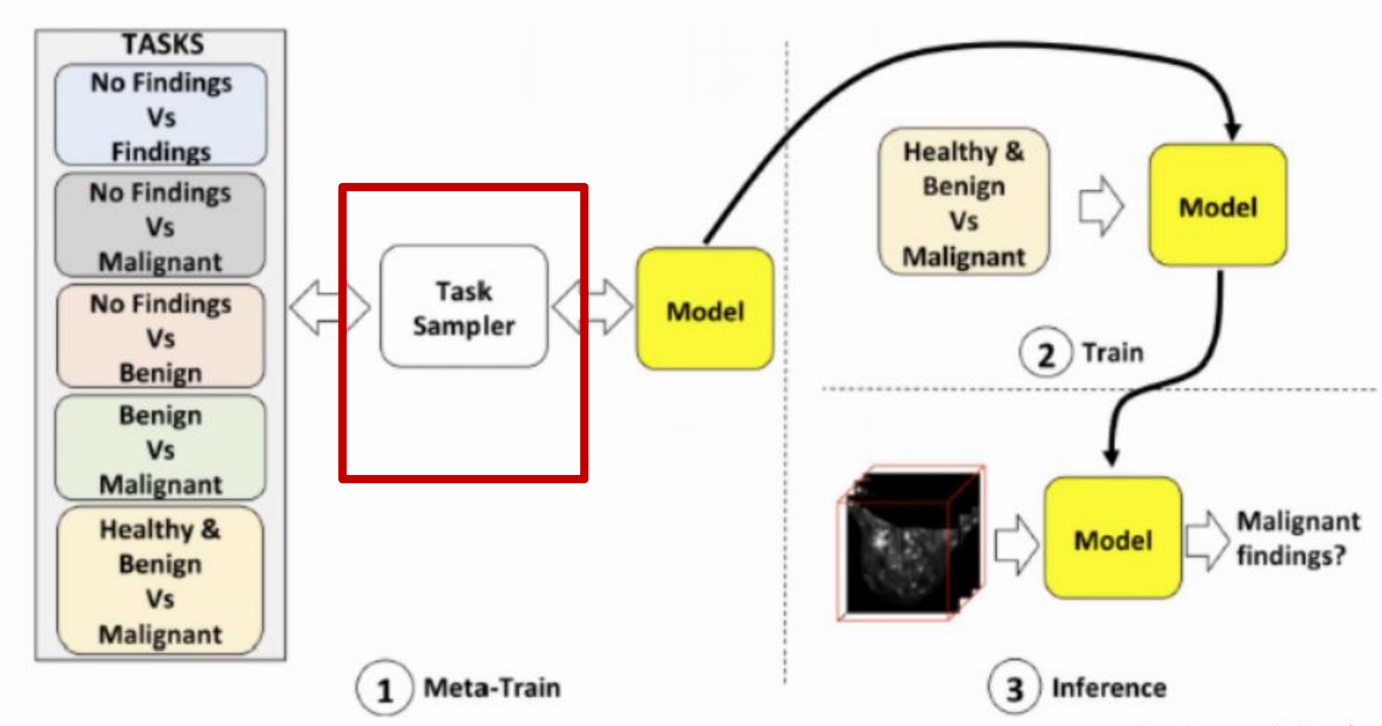


Fig. 1: The model is first meta-trained using several tasks containing relatively small training sets. The meta-trained model is then used to initialize the usual training process for breast screening (i.e., healthy and benign versus malignant). The probability of malignancy is estimated from a forward pass during the inference process.

Uncertainty Modelling

MultiTask Learning

- Criteria for task selection during training

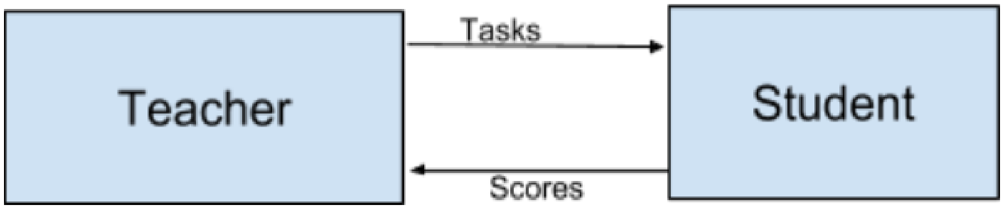


Uncertainty Modelling

MultiTask Learning

- Criteria for task selection during training
  - Random
  - All
  - **Teacher-Student Curriculum Learning**

Sample tasks that can achieve a higher improvement on their performance



Training Data

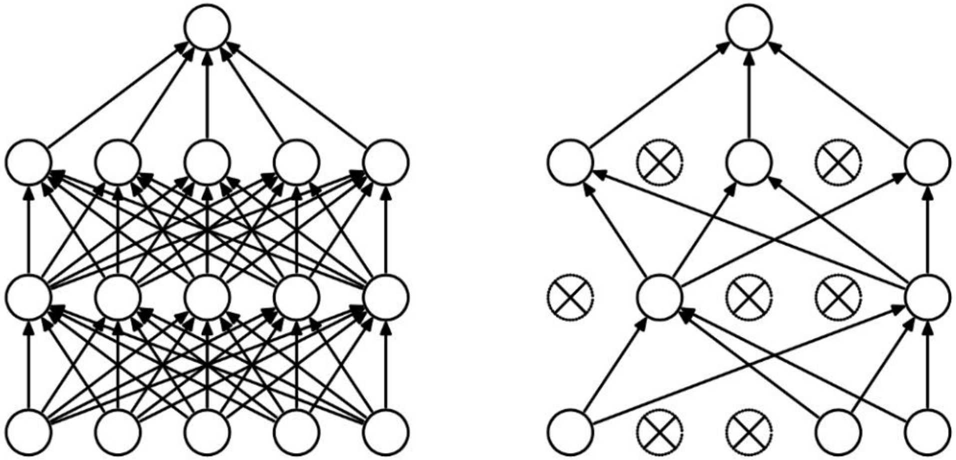
Network

Uncertainty Modelling

Hinton (et al.), 2012

Dropout

- Problem: neurons adapt too much to inputs
- Solution: drop out



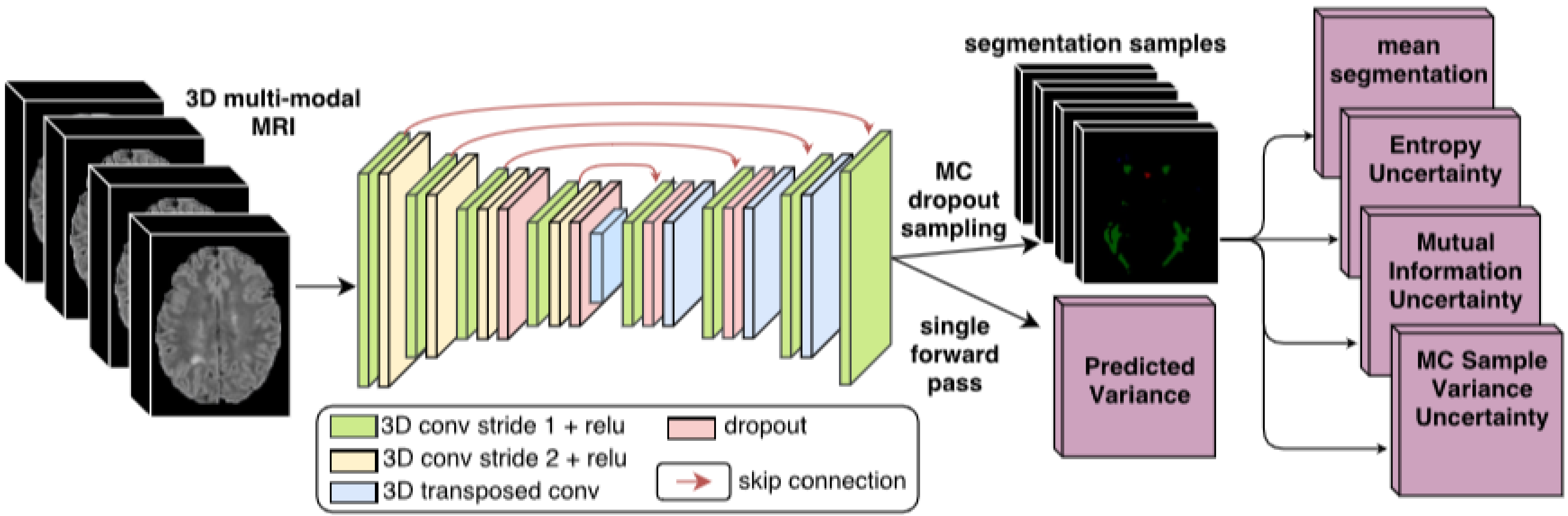
(a) Standard Neural Net (b) After applying dropout.

Figure 1: Dropout Neural Net Model. **Left**: A standard neural net with 2 hidden layers. **Right**: An example of a thinned net produced by applying dropout to the network on the left. Crossed units have been dropped.

Uncertainty Modelling

Dropout

- Allows estimating neuron variability (Bayesian networks) and define measures of output uncertainty



© Nair, MICCAI 2018

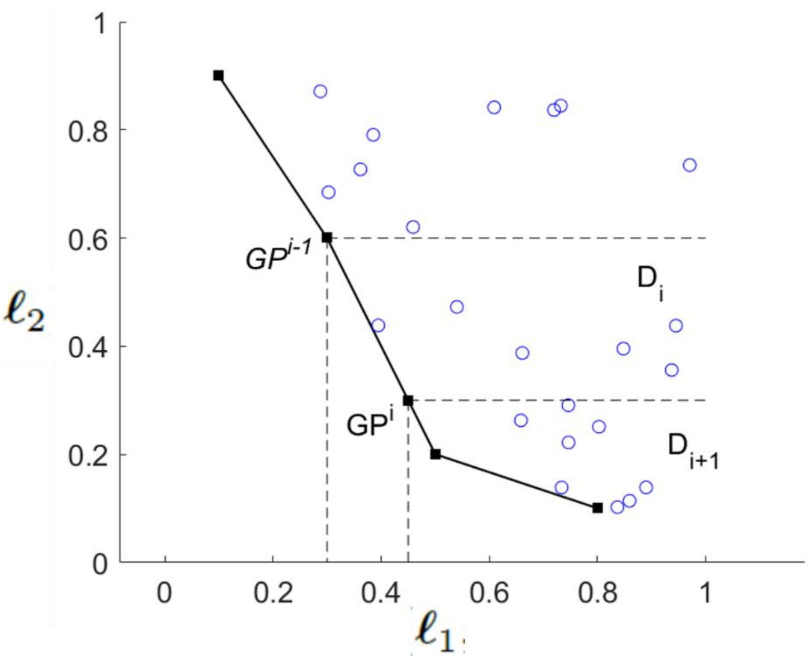
Uncertainty Modelling

Feature Uncertainty

- Estimated uncertainty is used as post-processing filter to either select reproducible features or define classifier cut-off threshold

Hint *: Incorporate uncertainty measures during training to obtain reproducible netwroks

Use a Pareto-like multi-task strategy for a multi-objective approach



* This is on-going research under project Up4Health funded by the Spanish Government under RETOS coordinated project RTI2018-095209-B-C21

Thanks 4 your time!!!!

Debora Gil, debora@cvc.uab.es,
www.iam.cvc.uab.es