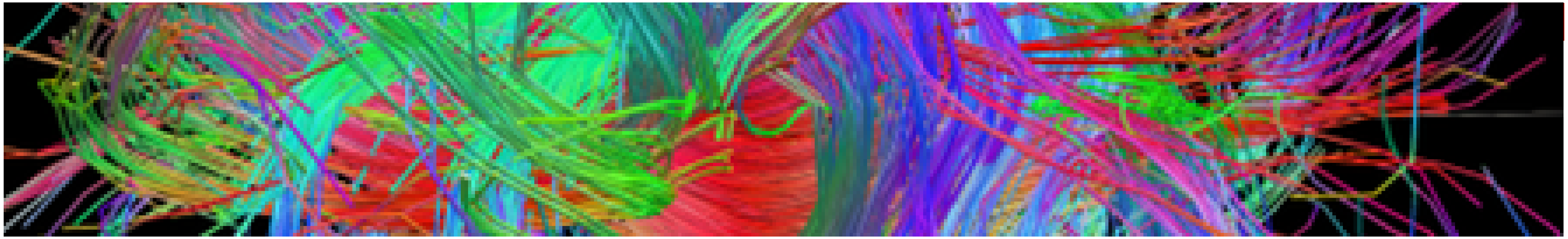


Medical Image Processing

Alessandra Retico

Istituto Nazionale di Fisica Nucleare - Sezione di Pisa

alessandra.retico@pi.infn.it

Workshop CCR: *Il Calcolo in Fisica Medica*

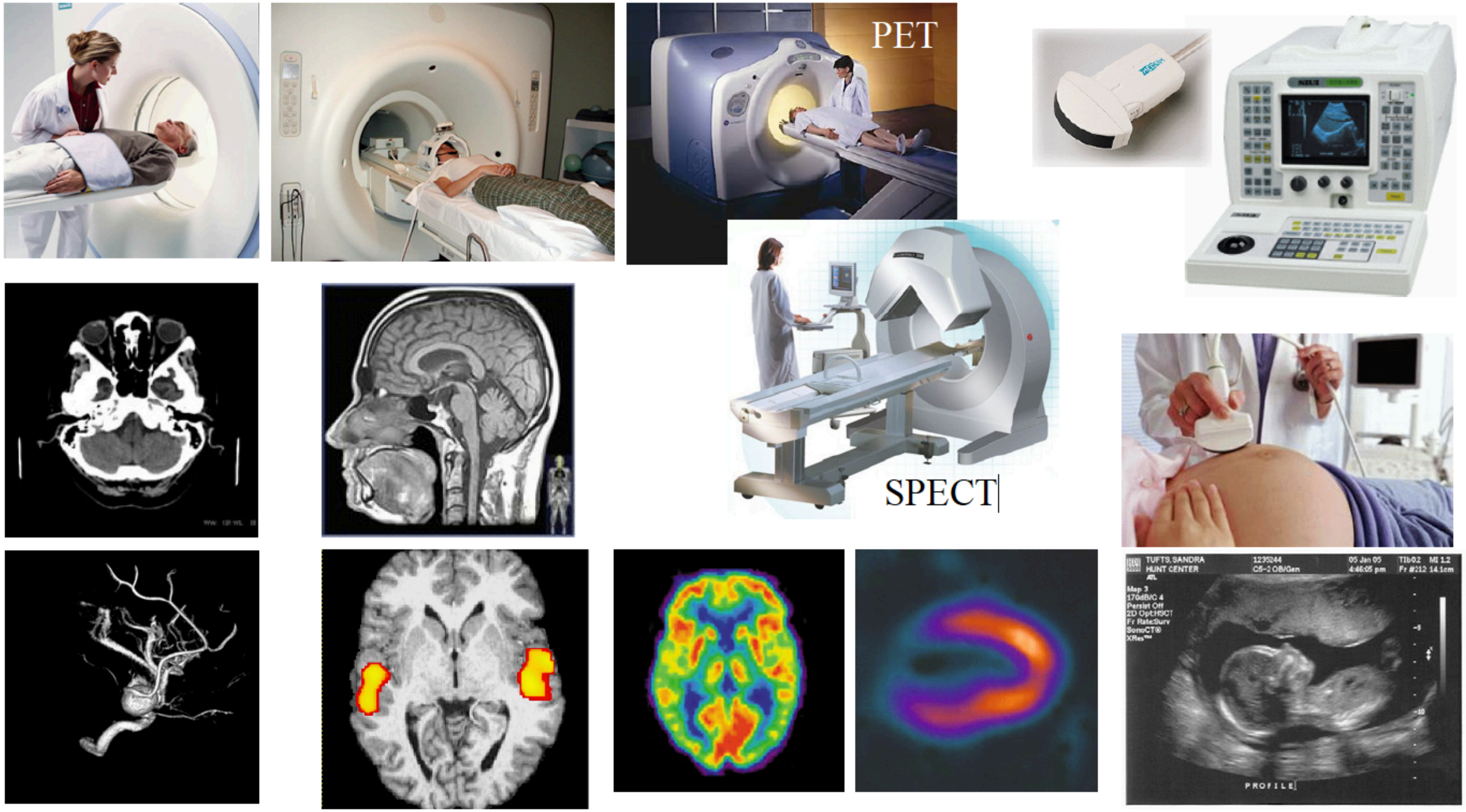
Trento, 17 Marzo 2016

Outline

- Diagnostic Imaging
- Medical Image Processing:
  - Visualization
  - Segmentation of anatomical/functional/pathological structures
  - Quantification
  - ➔Identification of Biomarkers
- Technological developments (more and more data!)
- Perspectives

Diagnostic imaging modalities

CT MRI / fMRI Nuclear Ultrasound



X-ray magnetic spin metabolic tracer X-ray emission sound waves

Historical perspective

from T. Beyer et al., Insights Imaging (2011)

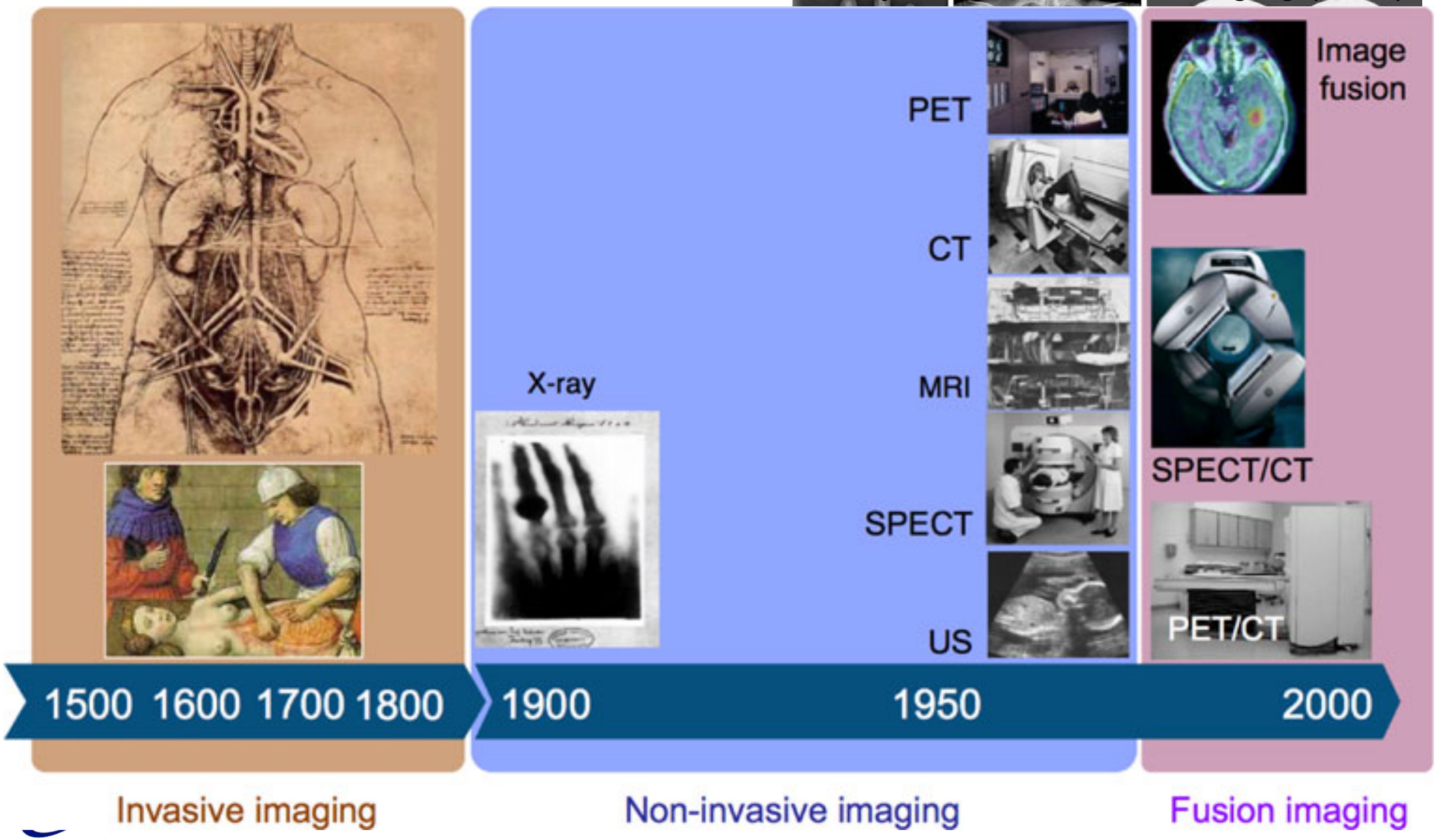


Image processing and machine learning

- Post-processing of raw data:
  - Visualization
- Quantification:
  - Image segmentation
  - Volumetric assessment of regions of interest (ROIs), lesions, etc.
- Follow up of pathological conditions:
  - Grow rate of lesions; assessment of treatment efficiency
- Fusing information:
  - Inter/intra-modality acquisitions (deformations, registration algorithms)

The aim is to assist clinicians in their tasks, not to replace them.

Societa Italiana di Neurologia (SIN)

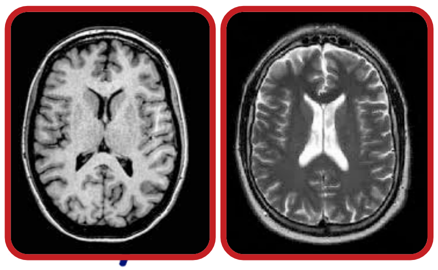
Diagnosi precoce di:

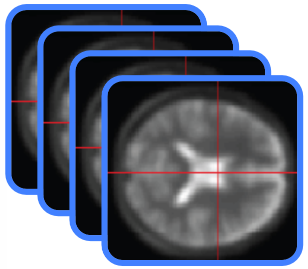
- Parkinson
- Alzheimer
- Sclerosi multipla

Measuring structure and function of the brain

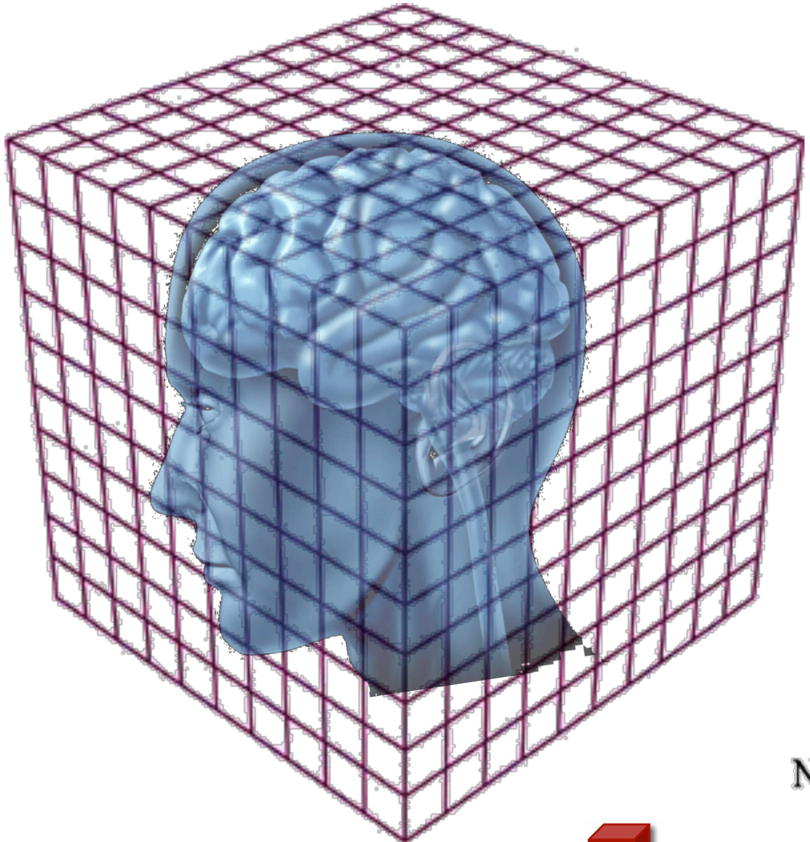
Non-invasive methods

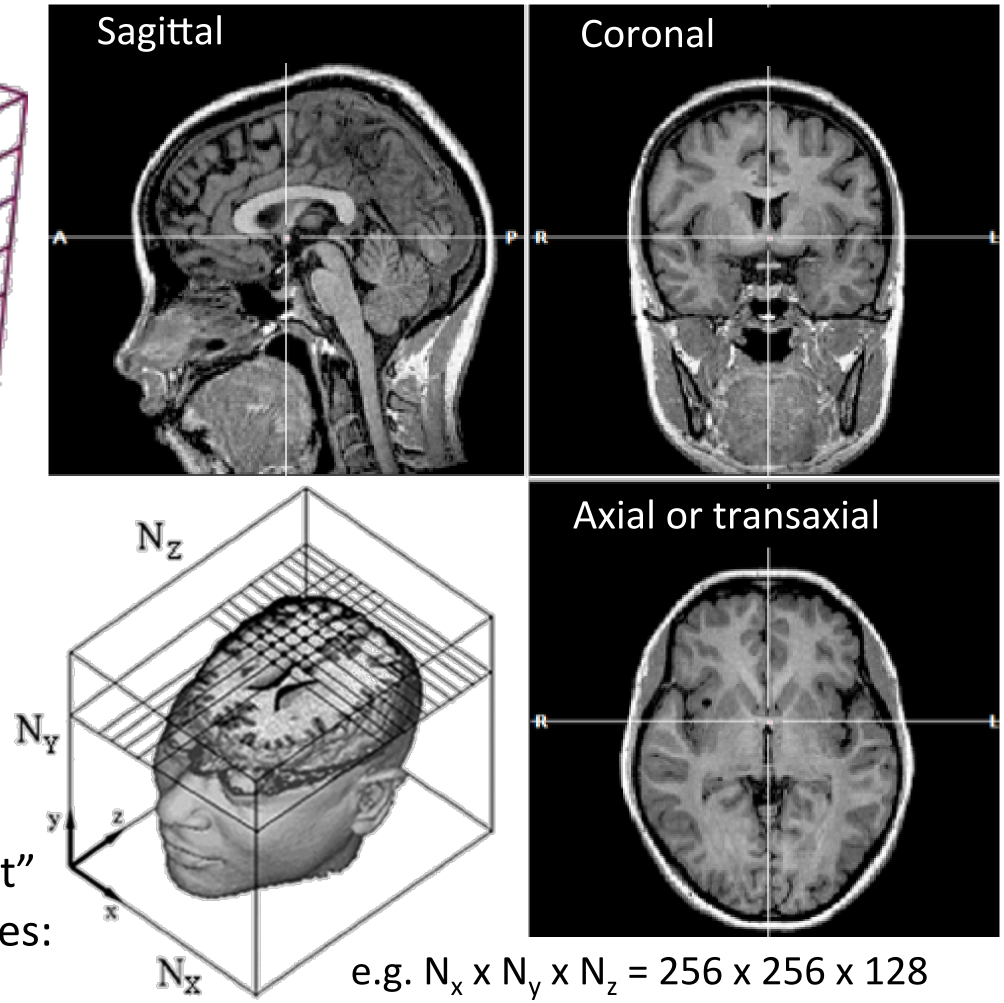
- **Structure**
  - Structural Magnetic Resonance imaging — MRI
  - Diffusion Tensor Imaging — DTI
  - Computed Tomography — CT
- **Function**
  - Functional Magnetic Resonance Imaging — fMRI
  - Positron Emission Tomography — PET





Structural MRI (sMRI) T_1-w images





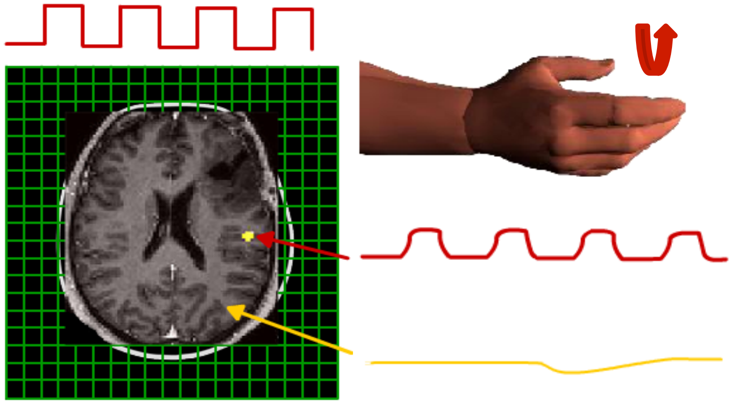
3D array of data:

- The voxel is the “volume element”
- “High resolution” MRI T_1-w images: $v_x \times v_y \times v_z \approx 1 \times 1 \times 1$ mm^3

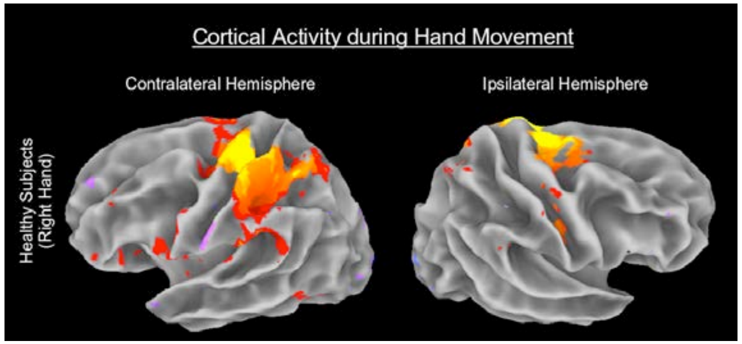
e.g. N_x x N_y x N_z = 256 x 256 x 128
16 bit **16 MB**

Functional MRI (fMRI)

- BOLD Response: blood-oxygen-level-dependent (BOLD) contrast
- 1 volume per second ($3x3x3mm^3$), for 4-5 min:~ 320 MB



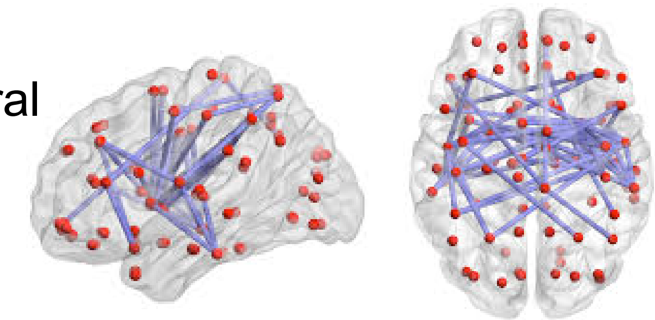
- Stimuli (visual, auditory, tactile, …) are presented during the scan

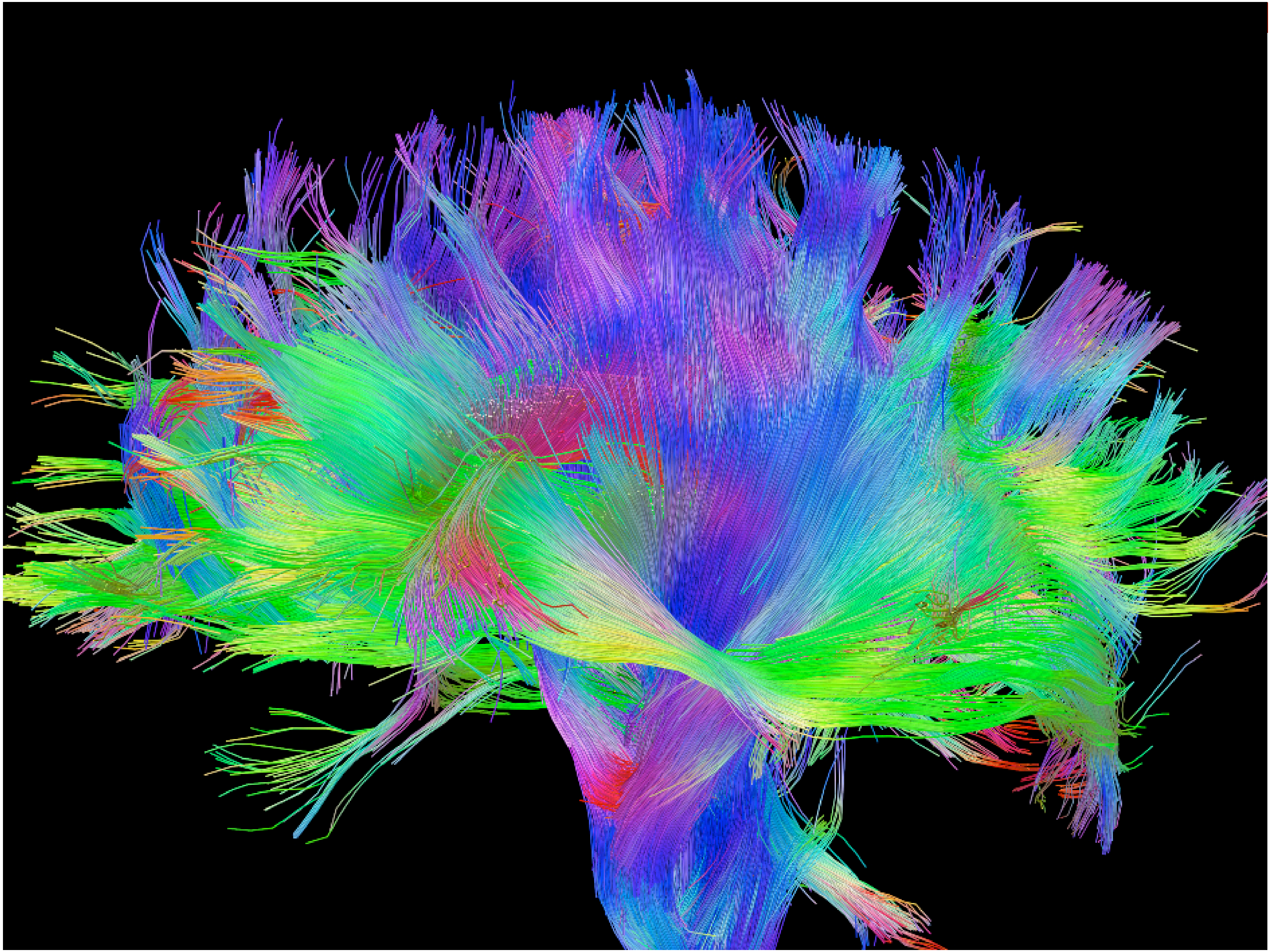


- Analysis of data time series to look for up-and-down signals that match the stimulus time series

Functional connectivity

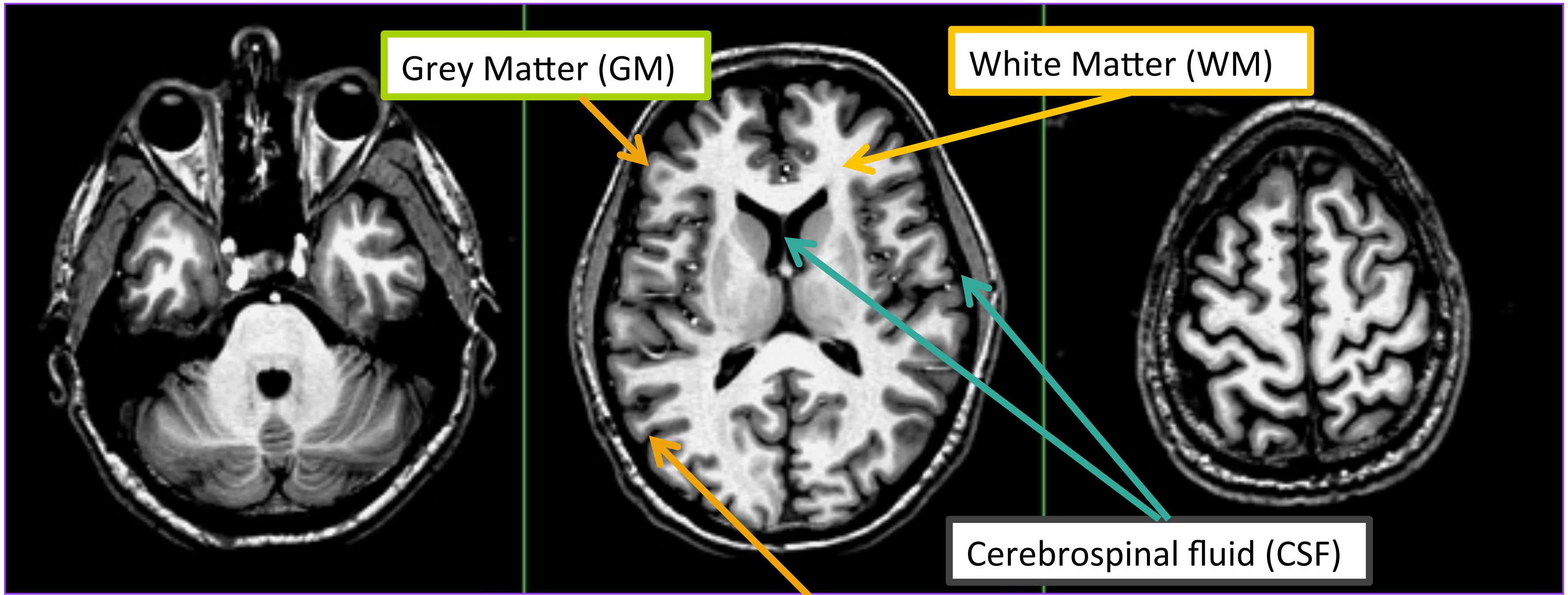
Resting state rs-fMRI: study of temporal correlations between spatially remote neurophysiological events





T_1-weighted brain images

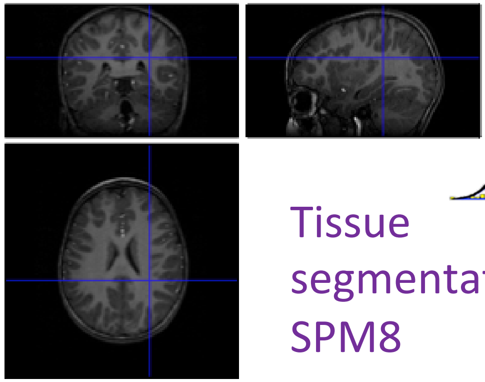
Axial slices of a human head with spatial resolution of 1 mm^3

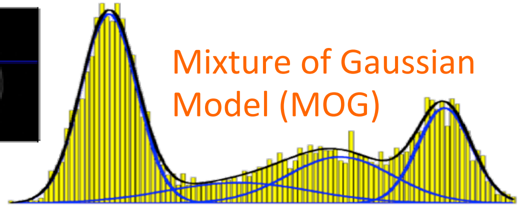


Grey Matter (GM) cortex can by followed and cortical thickness can be evaluated to investigate GM involvement in pathological conditions.

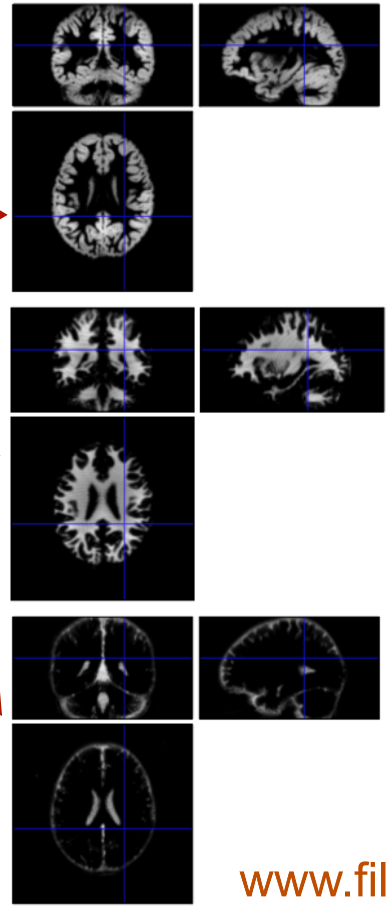
Segmentation of brain components

Grey matter
White matter
Cerebrospinal fluid

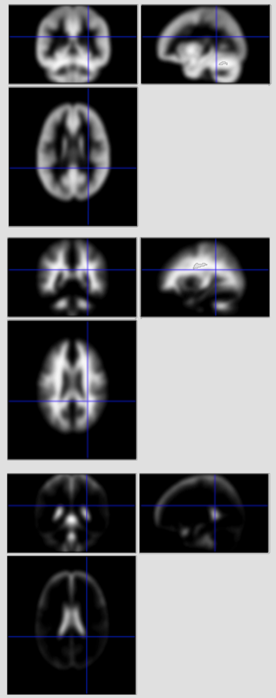




Tissue segmentation in SPM8



Tissue probability maps (TPMs) in SPM8



- In a simple MOG, the probability of obtaining a voxel with intensity y_i given that it belongs to the k^{th} cluster, i.e. Gaussian ($c_i = k$), and that the k^{th} Gaussian is parameterized by μ_k and σ_k^2 is:

$$P(y_i|c_i = k, \mu_k, \sigma_k) = \frac{1}{\sqrt{2\pi\sigma_k^2}} exp\left(-\frac{(y_i-\mu_k)^2}{2\sigma_k^2}\right)$$

- The prior probability of any voxel, irrespective of its intensity, belonging to the k^{th} is:

$$P(c_i = k|\gamma_k) = \gamma_k \qquad \sum_{k=1}^{K}\gamma_k = 1$$

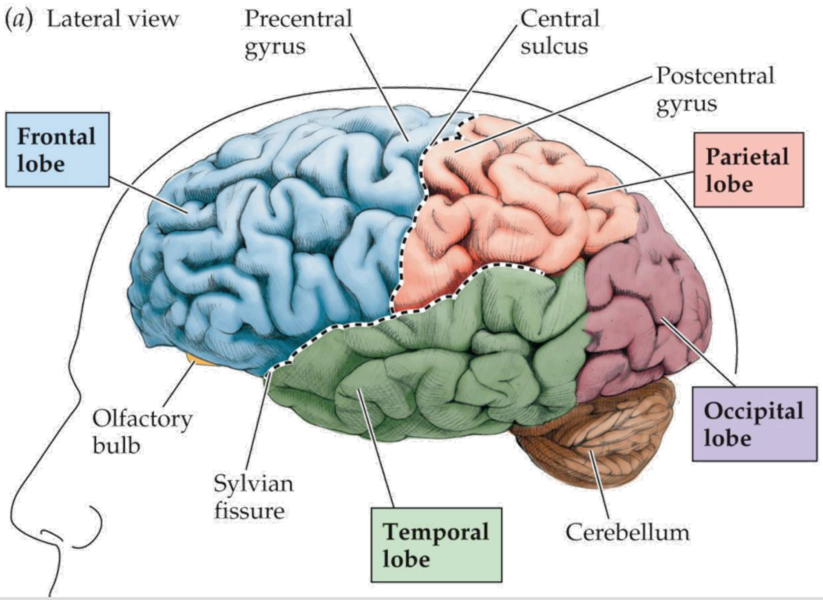
- Using Bayes rule, the joint probability of cluster k and intensity y_i is:

$$P(y_i, c_i = k|\mu_k, \sigma_k, \gamma_k) = P(y_i|c_i = k, \mu_k, \sigma_k)P(c_i = k|\gamma_k)$$

⇒ Sum over K clusters, accounting for each voxel and maximize with respect to the unknowns (μ, σ)

www.fil.ion.ucl.ac.uk/spm/

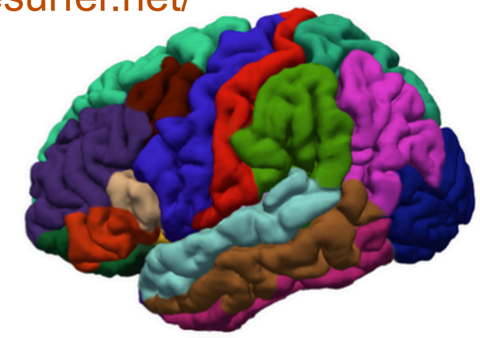
Brain parcellation



FreeSurfer

Wiki Documentation | Download | Support | Courses | People | Publications

http://freesurfer.net/



FreeSurfer Software Suite

An open source software suite for processing and analyzing (human) brain MRI images.

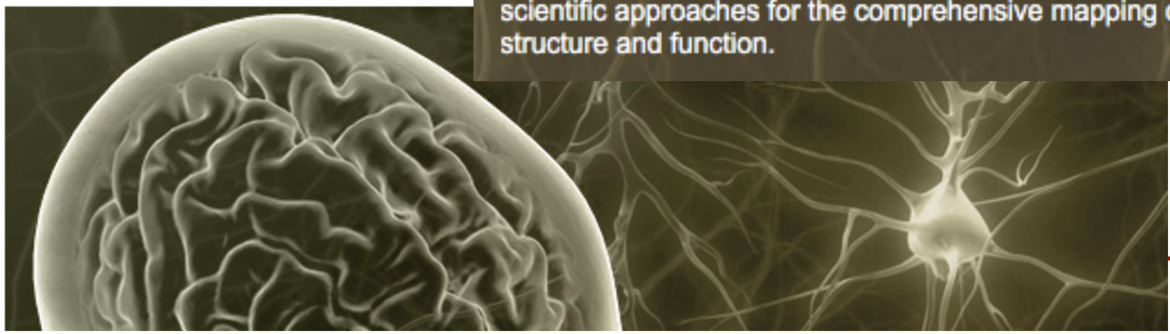
- Skullstripping
- Image Registration
- Subcortical Segmentation
- Cortical Surface Reconstruction
- Cortical Segmentation
- Cortical Thickness Estimation
- Longitudinal Processing
- fMRI Analysis
- Tractography
- FreeView Visualization GUI
- and much more ...

LONI ® Laboratory of Neuro Imaging

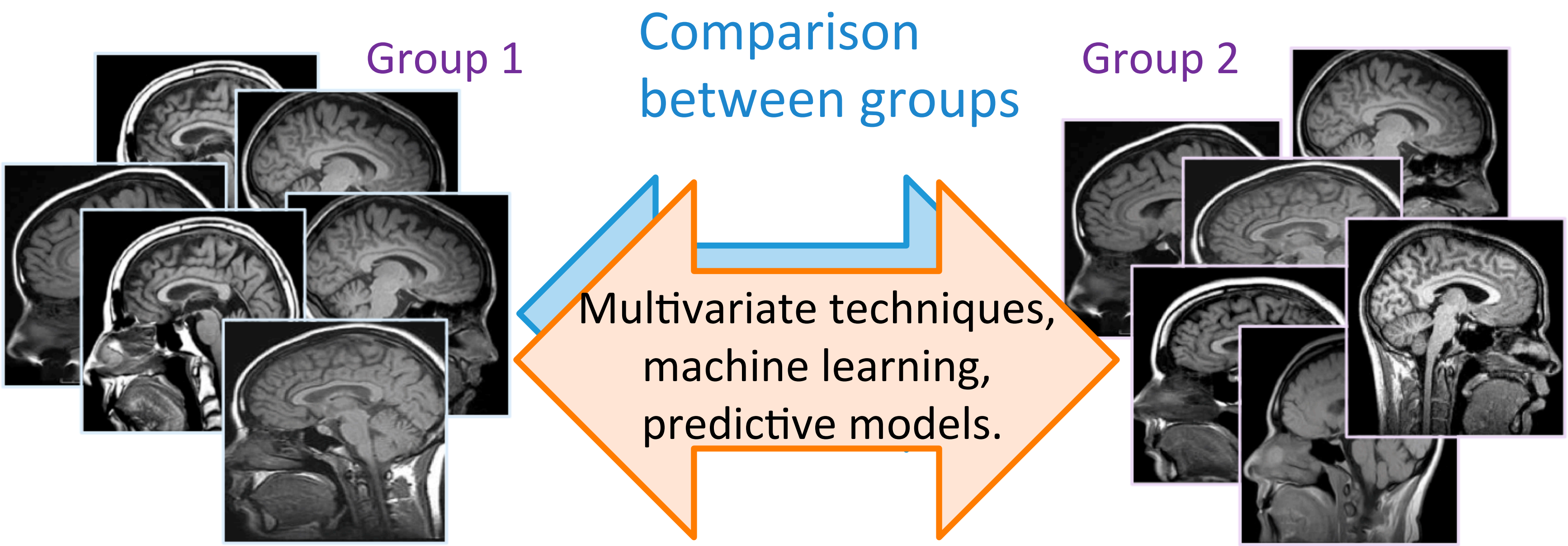
http://www.loni.usc.edu/

Home About LONI

LONI seeks to improve understanding of the brain in health and disease. The laboratory is dedicated to the development of scientific approaches for the comprehensive mapping of brain structure and function.



Group comparison (voxel/regional level)

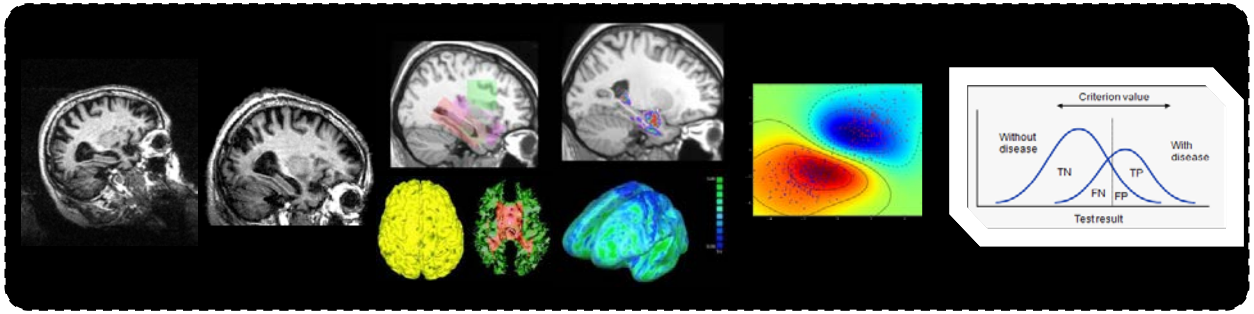


- Patients (group 1) vs. Controls (group 2):
  - Pathology specific brain alterations
- Longitudinal studies (same group after a time delay)
  - Effect of aging
  - Effect of treatments

Biomarkers

from A. Chincarini's slides: nextMR experiment (INFN-CNS5)

Images ain't markers!
(at least not immediately)



Signal contribution	Physics, scanner, protocols	Image modality	Clinical input	Information theory	Data set
Processing	Quality assessment	Normaliza tion	Characteriza tion	Feature extraction	Classifier
Noises	Acquisition noises	Physiological	Processing, Gold standard	Data processing	Gold standard

Raw data 10^7 DoF / image

Abstraction layer $10^2 - 10^3$ DoF

Marker 1-2 DoF

Raw data → Synthesis

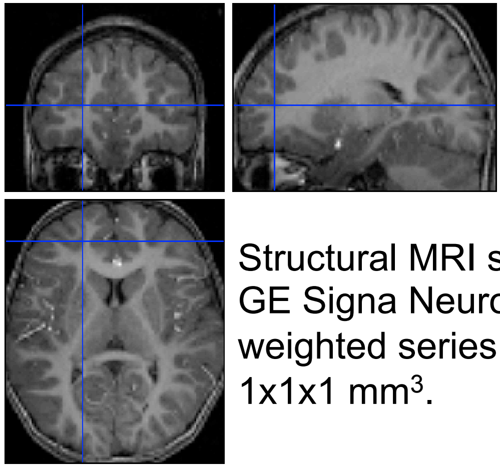
Abstraction layer

MARKER

Analysis of sMRI in Autism Spectrum Disorder

GR-2317873

SUPPORTING AN EARLY AUTISM SPECTRUM DISORDER (ASD) DIAGNOSIS THROUGH THE SUPPORT VECTOR MACHINES (SVM)



Structural MRI scans acquired with 1.5 T GE Signa Neuro-optimized System, T1-weighted series (FSPGR) with voxel size 1x1x1 mm^3.

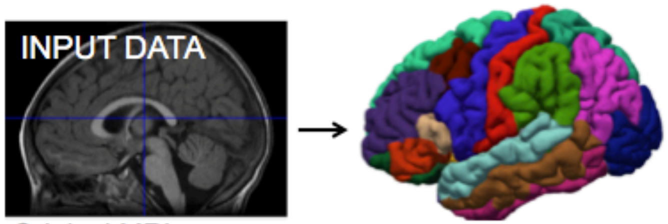
Young children (preschoolers)

Variable	Subject group, mean ± std [range]			
	ASD (n=41)		Controls (n=40)	
Age (months)	49 ± 12 [28-70]		49 ± 14 [22-72]	
NVIQ	73 ± 22 [34-113]		73 ± 23 [31-123]	
	Males (n=21)	Females (n=20)	Males (n=20)	Females (n=20)
Age (months)	50 ± 10 [34-70]	48 ± 13 [28-69]	48 ± 13 [24-70]	50 ± 16 [22-72]
NVIQ	75 ± 22 [40-113]	70 ± 23 [34-113]	73 ± 23 [32-123]	71 ± 24 [31-106]

Legend: autism spectrum disorder (ASD); non-verbal intelligence quotient (NVIQ).

Freesurfer preprocessing

Parcellation of the cerebral cortex into 62 structures.



Original MRI scan

Computation of:

- 5 Region-of-Interest (ROI)-based features for each structure:
(1) area; (2) volume; (3) average thickness; (4) thickness standard deviation; (5) mean curvature.

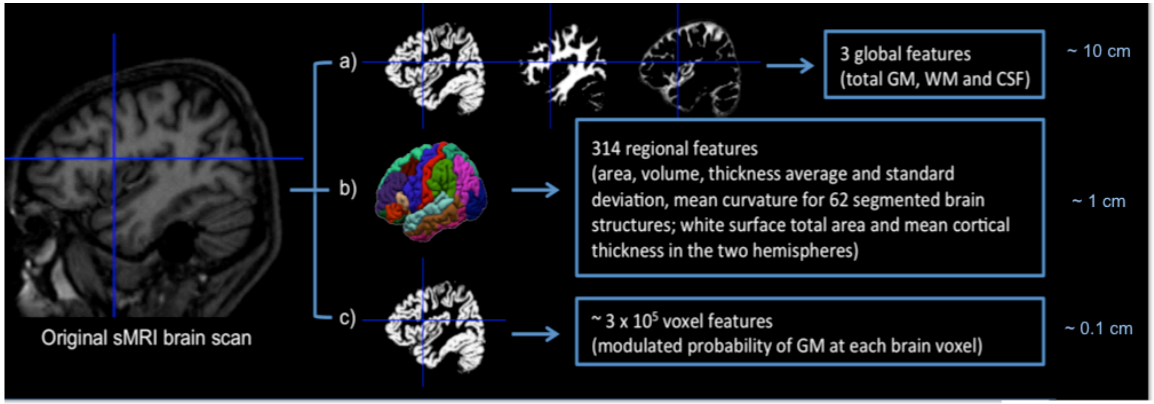
- White surface total area and mean thickness of cerebral cortex for both emispheres.

314 ROI-based features
for each subject

http://surfer.nmr.mgh.harvard.edu/

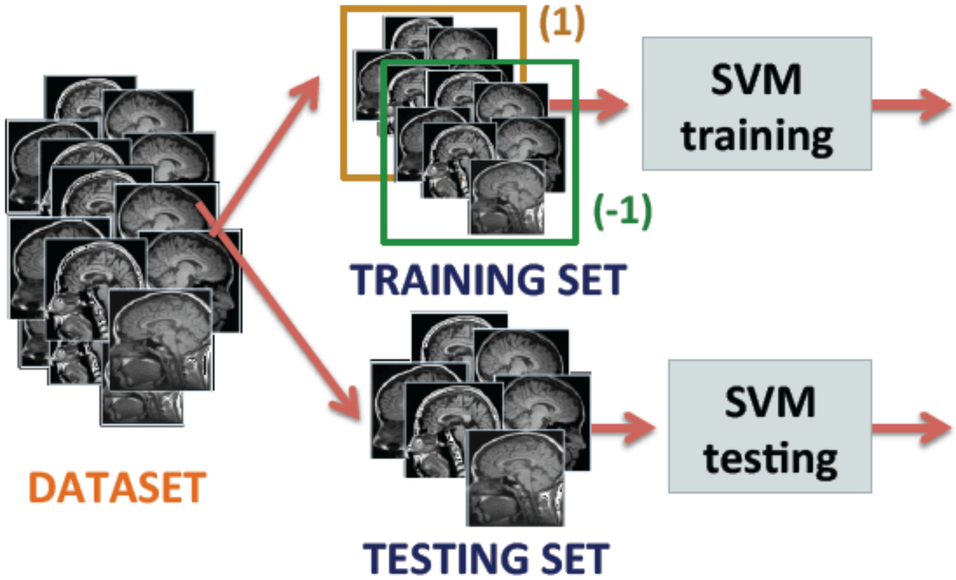
sMRI data analysis

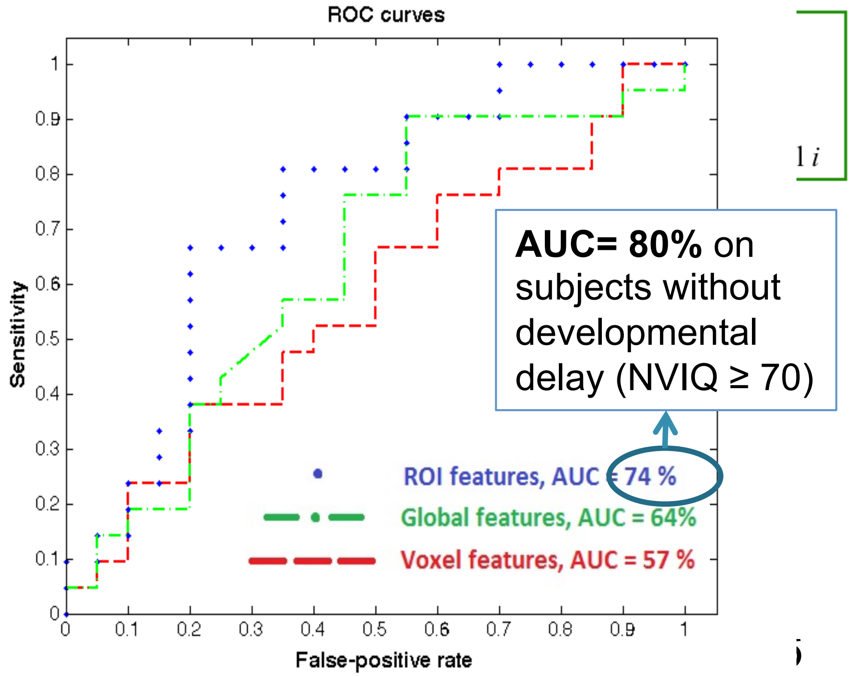
- Two-class classifiers with Support Vector Machines (SVM)



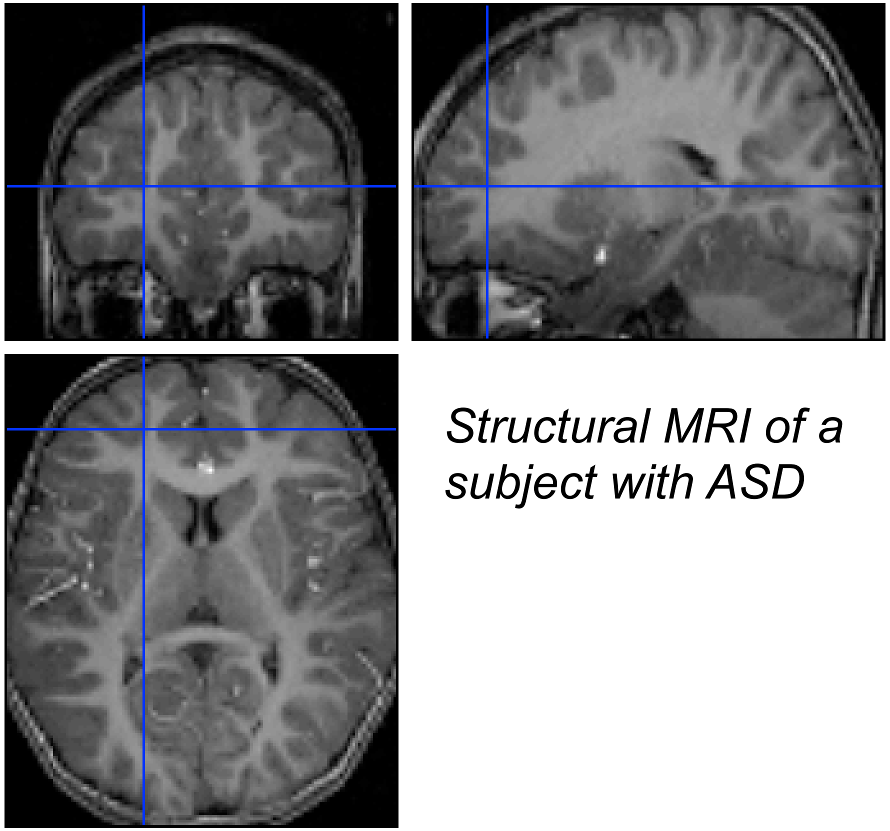
Comparison of different MRI analysis pipelines:

- different preprocessing
  - SPM, http://www.fil.ion.ucl.ac.uk/spm/
  - Freesurfer, http://surfer.nmr.mgh.harvard.edu/
- different features: Global, ROI, voxel
- different scales of analysis

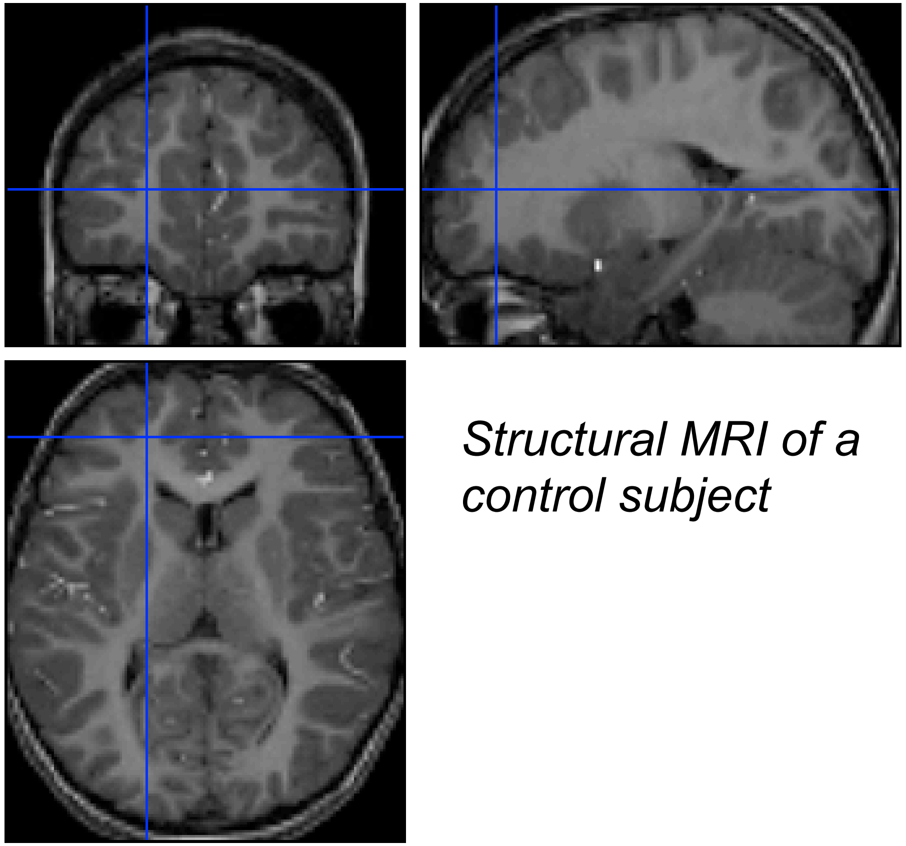




ASD vs. control *visual* comparison



Structural MRI of a subject with ASD



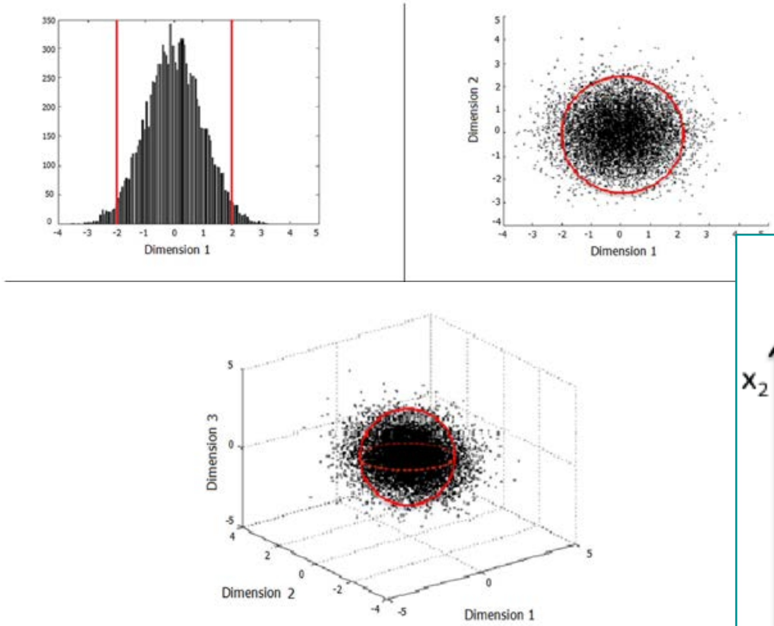
Structural MRI of a control subject

Multivariate techniques can catch subtle correlations hidden within large amount of data (whole-brain voxels/features)

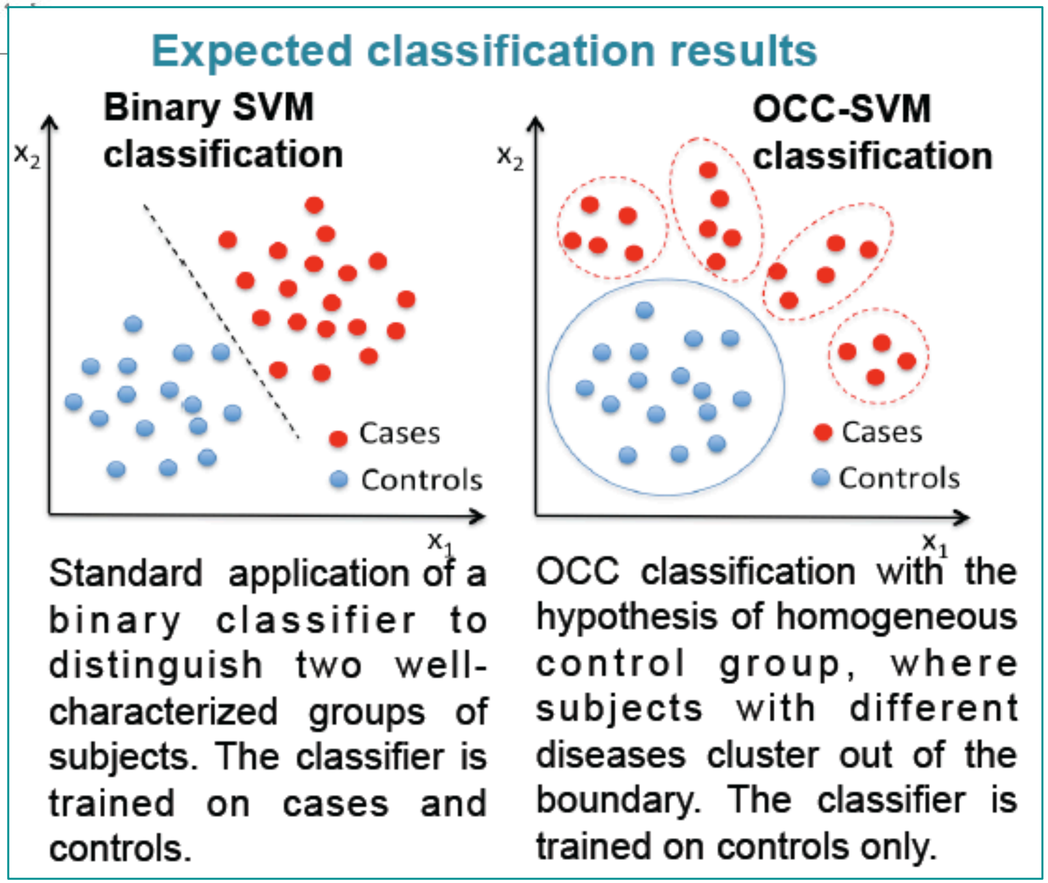
How to handle the ASD heterogeneity ?

sMRI data analysis

- One-class classification with SVM: data description method



Schölkopf *et al.*, Neural Comput 13 (2001) 1443 – 71
Tax & Duin, Pattern Recog Letter 20 (1999) 1191-1199
Mourañ-Miranda *et al.*, NeuroImage 58 (2011) 793 – 804
Sato *et al.*, Frontiers in Neuroscience 6 (2012) 178



Patient classification as an outlier detection problem

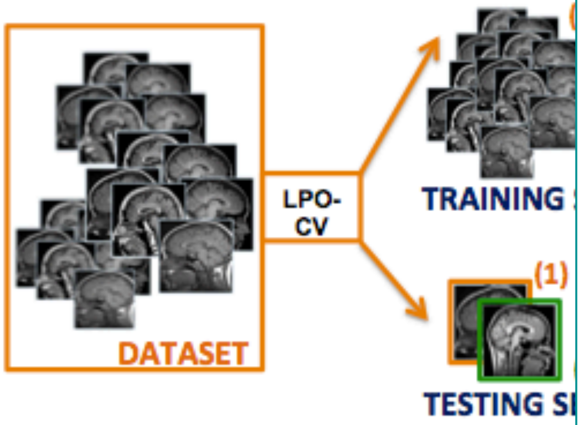
Unsupervised learning

Non-linear kernel SVM (RBF)

Finding the smallest hypersphere enclosing data

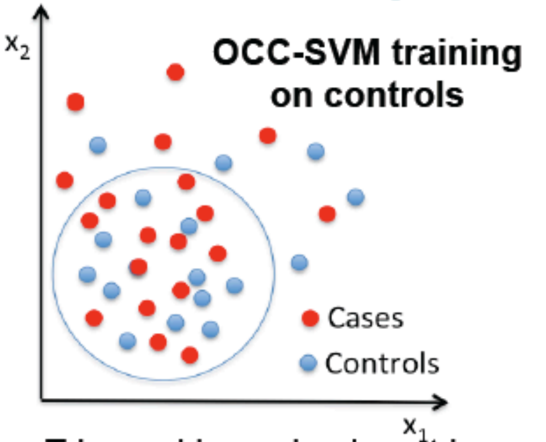
One-Class Classification (OCC) of brain features

We applied OCC to the vector of 314 characteristics extracted for each subject of our datasets, separately on the **male subset**, on the **female subset** and on the **entire dataset**.

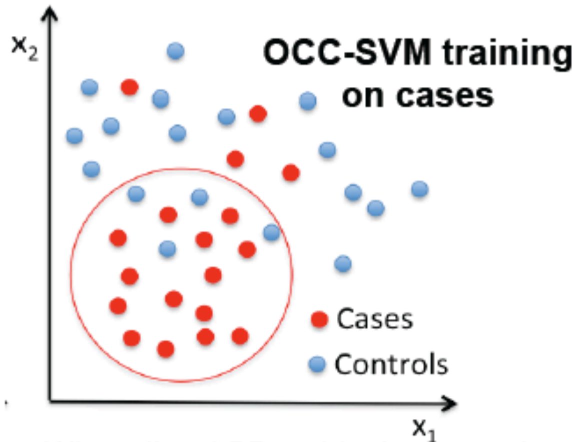


- Radial Basis Function (**RBF**)
- Optimization of the **paramete** **boundary) and ν** (it controls
- **Leave-pair-out cross-valida**

OCC performances



The discrimination performaces obtained using Controls as target class were not above the chance level, i.e. the controls do not have characteristics of homogeneity allowing recognizing ASD examples as outliers.



When the ASD subjects are set as the training class, the ASD vs. Controls discrimination performance achieved in terms of the area under the receiver operating characteristic curve (AUC) is:

AUC=0.74 for the **male subset**

AUC=0.68 for the **female subset**

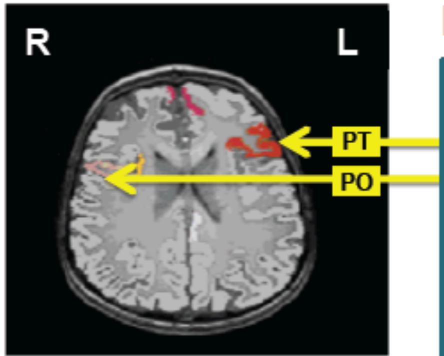
AUC=0.64 for the **entire dataset**

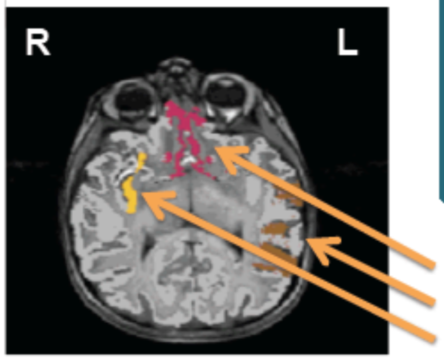
There is a common structure among the ASD subjects that the OCC-SVM could capture.

RESULTS AND CONCLUSIONS

Maps of most discriminant brain regions with permutation test ($p<0.05$)

When using SVM with non-linear kernel (e.g. RBF), the separating hyperplane is obtained in the feature space. We used the approach proposed by [6] to approximate the *preimages* for OCC-SVM with RBF. To identify the features and the neuroanatomical regions which drive the SVM boundary definition, we tailored the permutation testing [3] to the OCC-SVM.





Male population

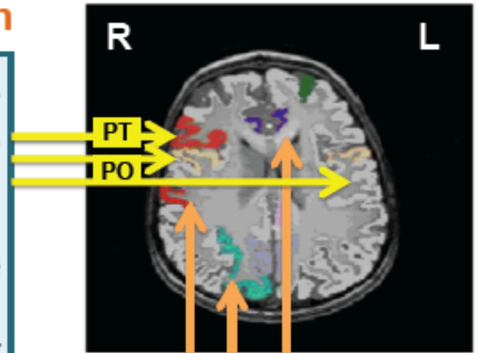
Female population

Language regions and **regions from the default mode network** (posterior cingulate cortex, **pars opercularis (PO)**, **pars triangularis (PT)**, and transverse temporal gyrus) contributed most to distinguishing individuals with ASD from controls, arguing for the crucial role of these areas in the ASD pathophysiology.

Pars opercularis and **pars triangularis** together comprise **Broca's area**, which is primarily involved in higher-order abilities such as action imitation, language, attribution of mental states and empathy, which are critically altered in ASD.

Other involved regions in the male population:

- L and R medial orbitofrontal cortices;
- L middle temporal cortex;
- R insula.



Other involved regions in the female population:

- L and R rostral anterior cingulate;
- R superior parietal cortex;
- R postcentral gyrus.

Best poster award at Bioimaging 2016, Rome, Febr 21-23

Accepted abstract at Human Brain Mapping 2016 Geneve, June 26-30

Technological developments

- New instrumentation for data acquisition (finer spatial and temporal resolution)
  - ➔ More and more data to store and process
- Large Consortia for data sharing
  - ➔ "Big data" processing

Suitable algorithms to extract information from data and assist the clinicians.

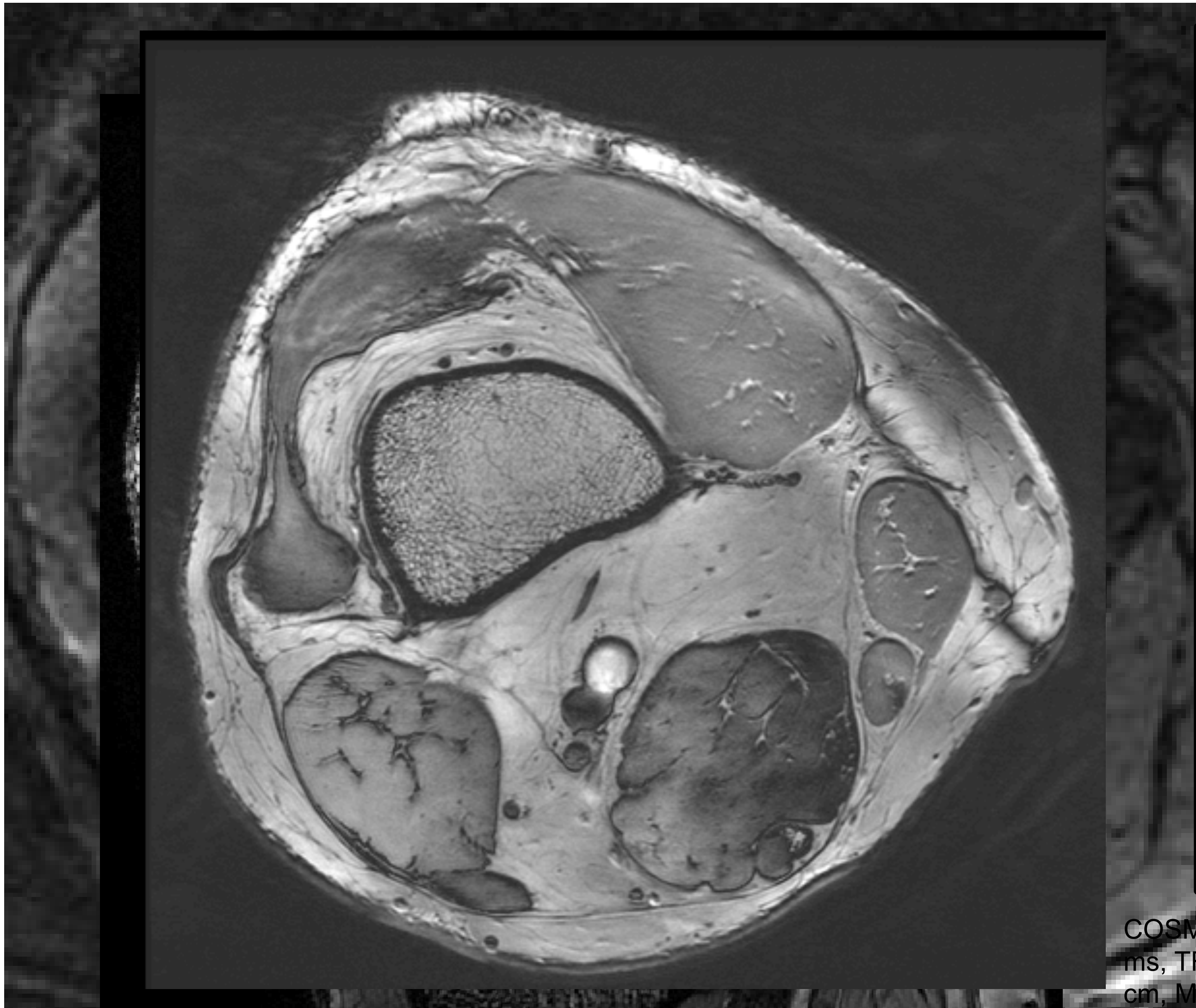
Accessible and user-friendly environment to store and analyze data.

Multifactorial challenge:

- Growing size of data
- Sociological and logistic sharing issues (data security)
- Multi-site, multi-datatype archiving
- Exploring and mining data (multimodal neuroimaging, biological and clinical data, environmental effects)


In-vivo
in-plane res
195x195 μm^2

IMAGO7 & INFN Collaboration



COSMIC 3D (slice thickness 600 um, TE 1. ms, TR 6 ms. Field of view (FOV) 14 cm x 14 cm, Matrix 512 x 512, interpolated to 1024 ... 137μm

3D Merge (multi-echo recalled gradient echo) TR=30 ms, TE=12.2 ms, FOV= 16 cm x 16 cm, Matrix 320x320, slice thickness 2.2 mm. In-plane resolution:0.5x0.5 mm²

Data storage

~ 1 GB

- 1.5 Tesla

Acquisition protocol

- Structural MRI (e.g. T1-w, T2-w): 1x1x1 mm^3
- 20 MB per volume

- fMRI: 3x3x3 mm^3=27 mm^3 (5 min)
- …

More than 25000 MRI scanners are in use worldwide and more than 60 million examinations are conducted annually

~10 GB

- 7 Tesla

Acquisition protocol

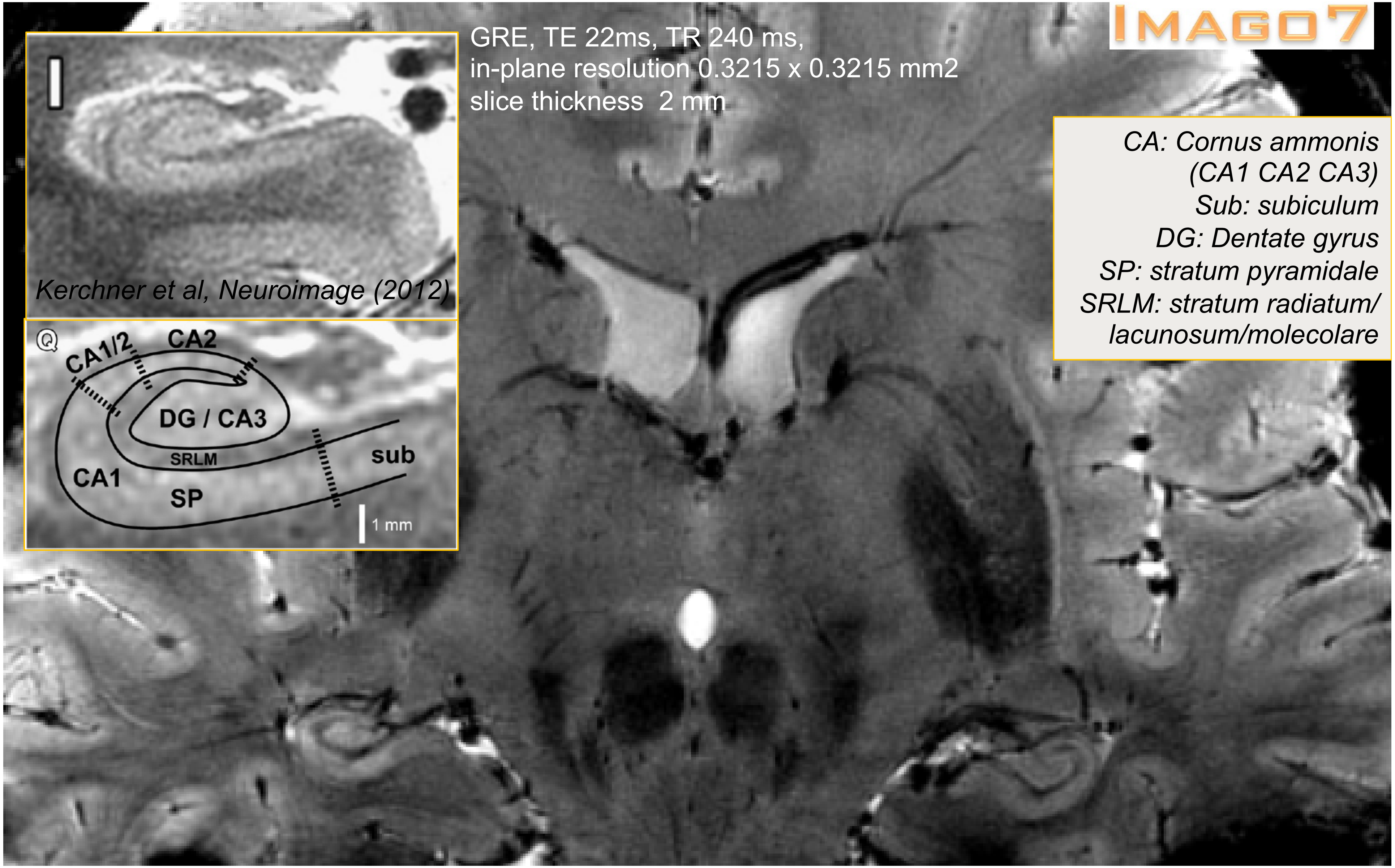
- Structural MRI (e.g. T1-w, T2-w): 0.6x0.6x0.6 mm^3
- 100 MB per volume

- fMRI: 1x1x2 mm^3=2 mm^3 (3 min) – only a brain slab

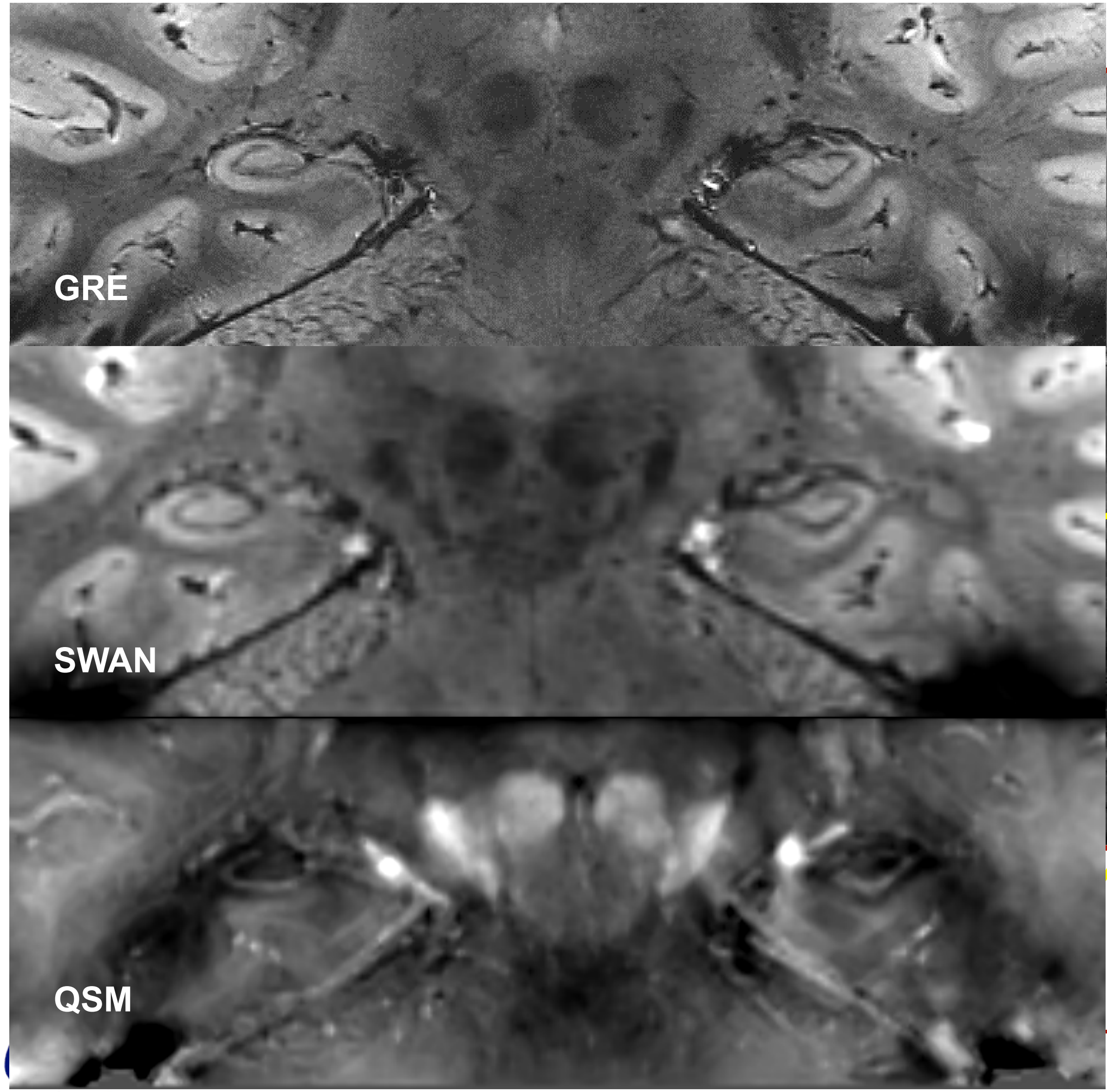
- + SWI with different phase-weightings, etc.
- …

~ 60 PB/year, comparable with LHC data

Human hippocampus at 7T



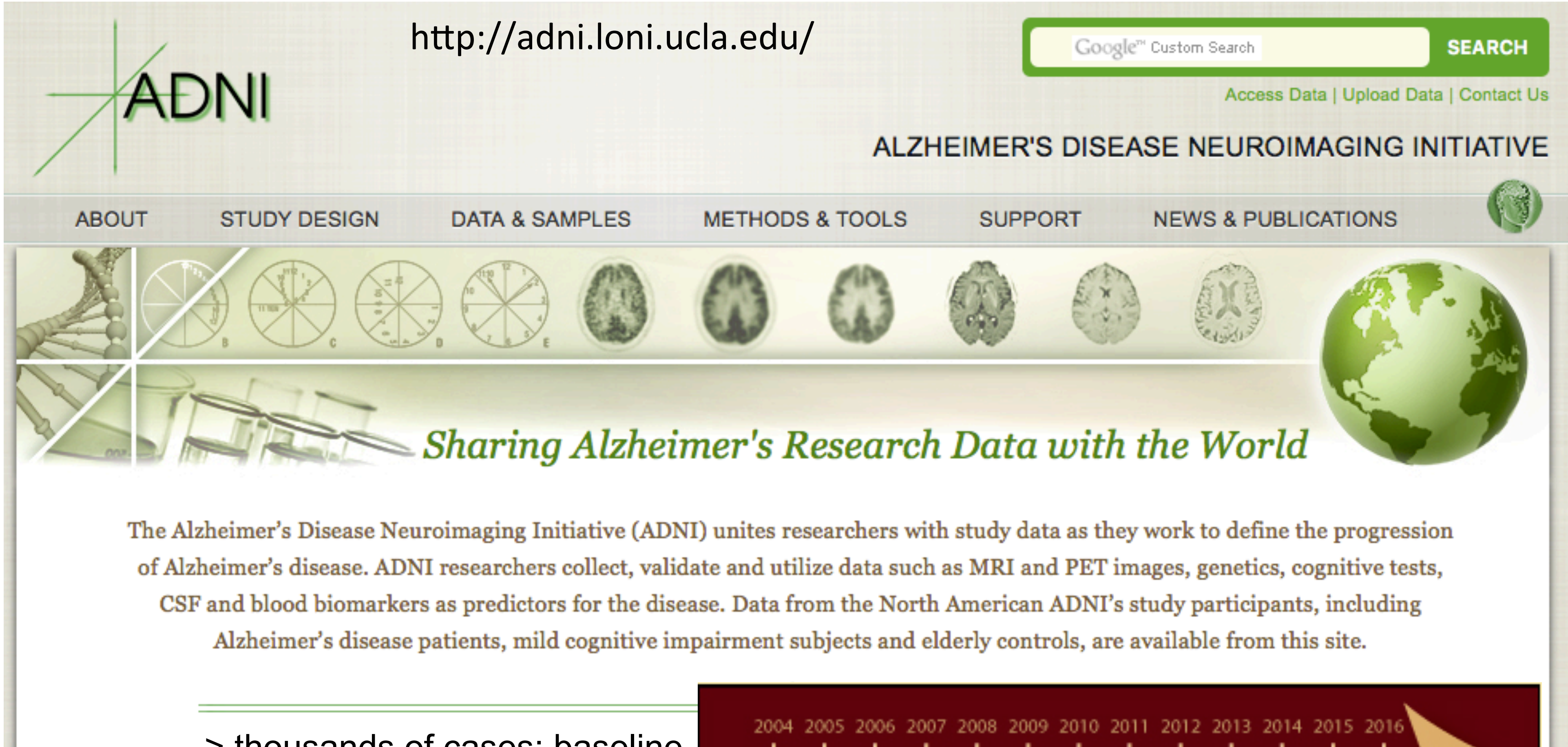
Estimating the size of hippocampus sub-structures at 7T



Analysis of 10 MCI patients: the SRLM thickness significantly correlated with the Mini-Mental State Examination (MMSE) score, and with the Free and Cued Selective Reminding Test (in particular free recall FCSRT-FR).

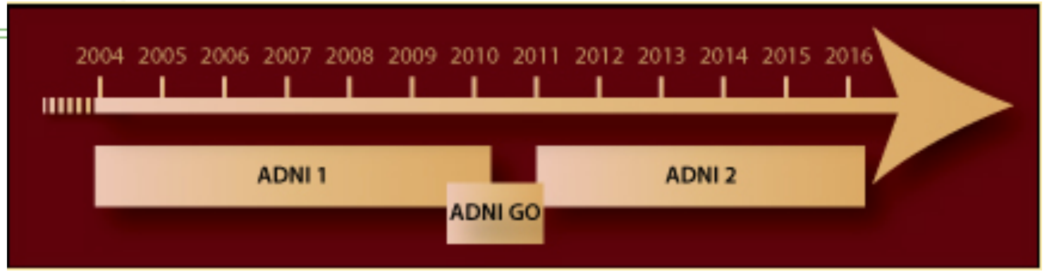
Presented at European Congress of Radiology (ECR) 2016

The Alzheimer's Disease Neuroimaging Initiative (ADNI)



> thousands of cases: baseline, follow up; AD, MCI, Controls; MRI, PET

~ 800 cases with genetic data



The Autism Brain Imaging Data Exchange (ABIDE)

ABIDE

Largest public dataset of subjects with Autism Spectrum Disorder (ASD).

- storing and sharing rs-fMRI and sMRI datasets collected in 17 international sites
- includes 1112 exams and phenotypic information common across all sites.
- age: 7 to 64 years

This data sharing initiative has already provided important contribution to the ASD research by unraveling inconsistent results about the strength of connectivity in ASD brains.

http://fcon_1000.projects.nitrc.org/indi/abide

Original Article

Molecular Psychiatry

Molecular Psychiatry **19**, 659-667 (June 2014) | doi:10.1038/mp.2013.78

The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism

A Di Martino, C-G Yan, Q Li, E Denio, F X Castellanos, K Alaerts, J S Anderson, M Assaf, S Y Bookheimer, M Dapretto, B Deen, S Delmonte, I Dinstein, B Ertl-Wagner, D A Fair, L Gallagher, D P Kennedy, C L Keown, C Keysers, J E Lainhart, C Lord, B Luna, V Menon, N J Minshew, C S Monk, S Mueller, R-A Müller, M B Nebel, J T Nigg, K O'Hearn, K A Pelphrey, S J Peltier, J D Rudie, S Sunaert, M Thioux, J M Tyszka, L Q Uddin, J S

Whole-brain analyses reconciled seemingly disparate themes of both hypo- and hyperconnectivity in the ASD literature; both were detected, although hypoconnectivity dominated, particularly for corticocortical and interhemispheric functional connectivity.

Human connectome project (HCP)

- Quarter 1 (Q1) data released (68 subjects):
  - 10 GB/subject (raw data)
  - 16 GB/subjects (proprocessed)
- Total: ~ 2 TB (68 subjects)
- Expected: 400 subjects/year

HCP Data Sizes (per Subject)

Session	Format	.zip File Size
Structural	Unprocessed	70.99 MB
	Preprocessed	1.19 GB
Resting State fMRI (each of 2 sessions)	Unprocessed	2 GB
	Preprocessed	3.24 GB
Task fMRI (avg per Task)	Unprocessed	490 MB
	Preprocessed	771 MB
(all 7 Tasks)	Unprocessed	3.43 GB
	Preprocessed	5.4 GB
Diffusion	Unprocessed	2.18 GB
	Preprocessed	2.81 GB
Group-Average on Unrelated 20	Additionally Processed	289 MB
Total (per Subject)	Unprocessed	9.81 GB
	Preprocessed	15.77 GB
	Both	25.58 GB
Total (5 Subjects)	Unprocessed	62.16 GB
	Preprocessed	78.83 GB
	Both	141 GB
Total (20 Subjects)	Unprocessed	247.34 GB
	Preprocessed	315.05 GB
	Both	562.39 GB
Total (68 Subjects)	Unprocessed	815.4 GB
	Preprocessed	1.058 TB
	Both	1.873 TB

http://enigma.ini.usc.edu/

ABOUT ENIGMA | NEWS | PUBLICATIONS | WORKING GI
ENIGMAVIS | PROTOCOLS | TRAINING | JOIN ENIGMA

ENIGMA

Enhancing NeuroImaging Genetics through Meta-Analysis

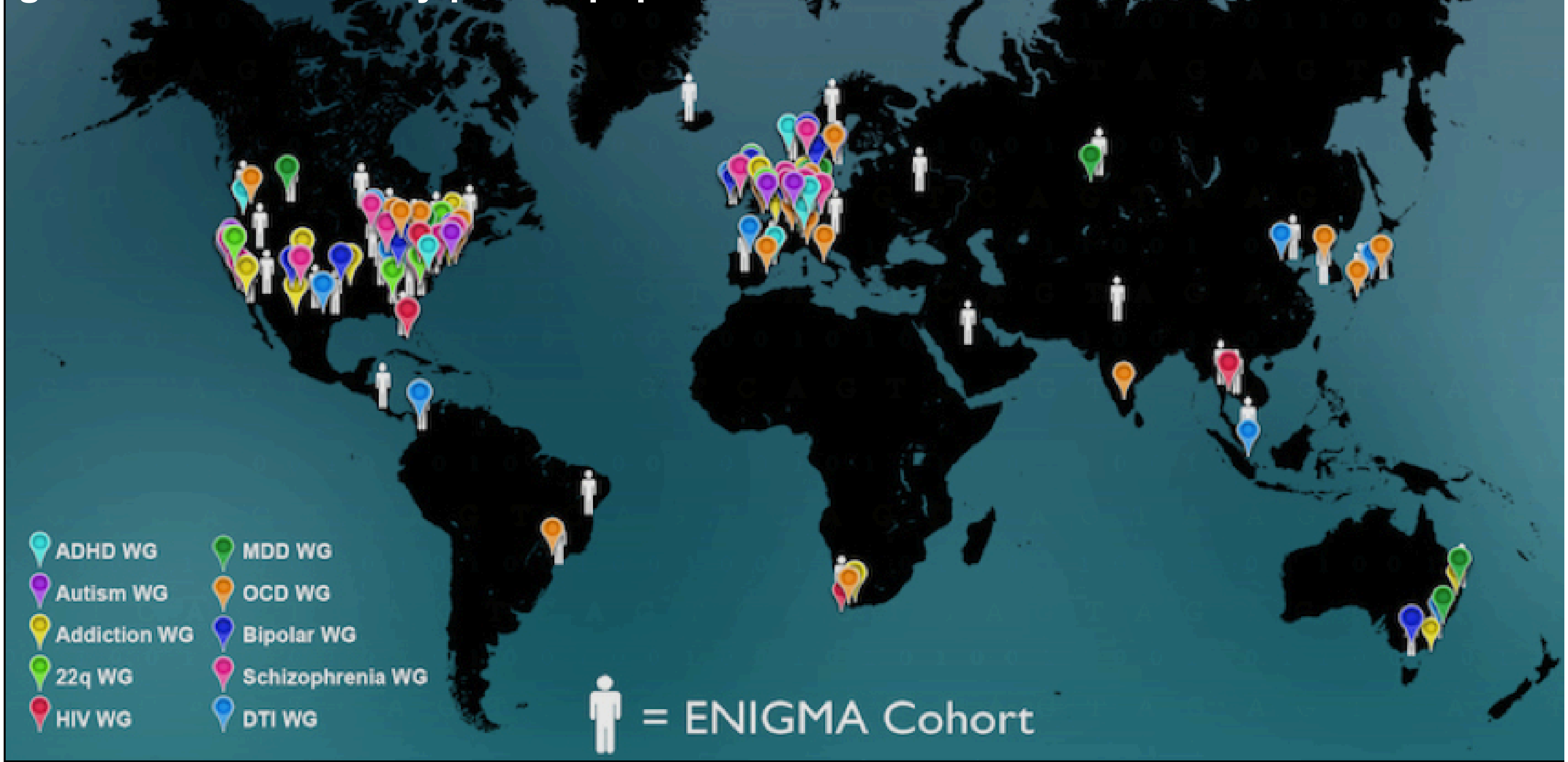
ENIGMA ASD

The ENIGMA – ASD working group will analyze data relevant to autism from people of all ages including case-control comparisons. Going beyond this, the group will also examine the influence of gender and covariates of disorders.
The ENIGMA meta-analytic approach will be used to aggregate data from case-control and developmental cohorts to examine the relative contribution of various genetic and brain correlates.

Please contact Daan van Rooij and Jan Buitelaar for information on joining ENIGMA-ASD.

ABOUT ENIGMA

The ENIGMA Consortium is an international effort by leaders worldwide. The Network brings together researchers in imaging genomics, neurology and psychiatry, to understand brain structure and function, based on MRI, DTI, fMRI, genetic data and many patient populations.



enigma.ini.usc.edu/wp-content/uploads/2015/04/enigma_map.jpg

NEW

ARIANNA

Ambiente di **R**icerca **I**nterdisciplinare per l'**A**nalisi di **N**euroimmagini **N**ell'**A**utismo

IRCCS Fondazione Stella Maris (Pisa)

Istituto Nazionale di Fisica Nucleare (Pisa)

Istituto di Teoria e Tecniche dell'Informazione Giuridica, CNR (Firenze)

I+ S.r.l (Firenze)

Net7 S.r.l (Pisa)

Regione Toscana

FAS
Fondo Aree Sottoutilizzate
2007-2013

REPUBBLICA ITALIANA

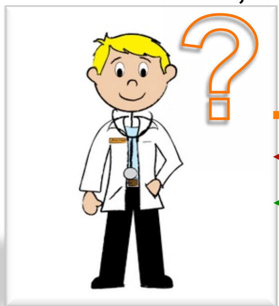
INFN

Interdisciplinary research in ARIANNA

Team of medical experts

Sharing research objectives, not only data;
data re-purposing and re-analysis.

User-friendly data browsing and preview; safely archiving; integration of additional data; queries on meta-data; statistics on data variables.



Team of data analysts

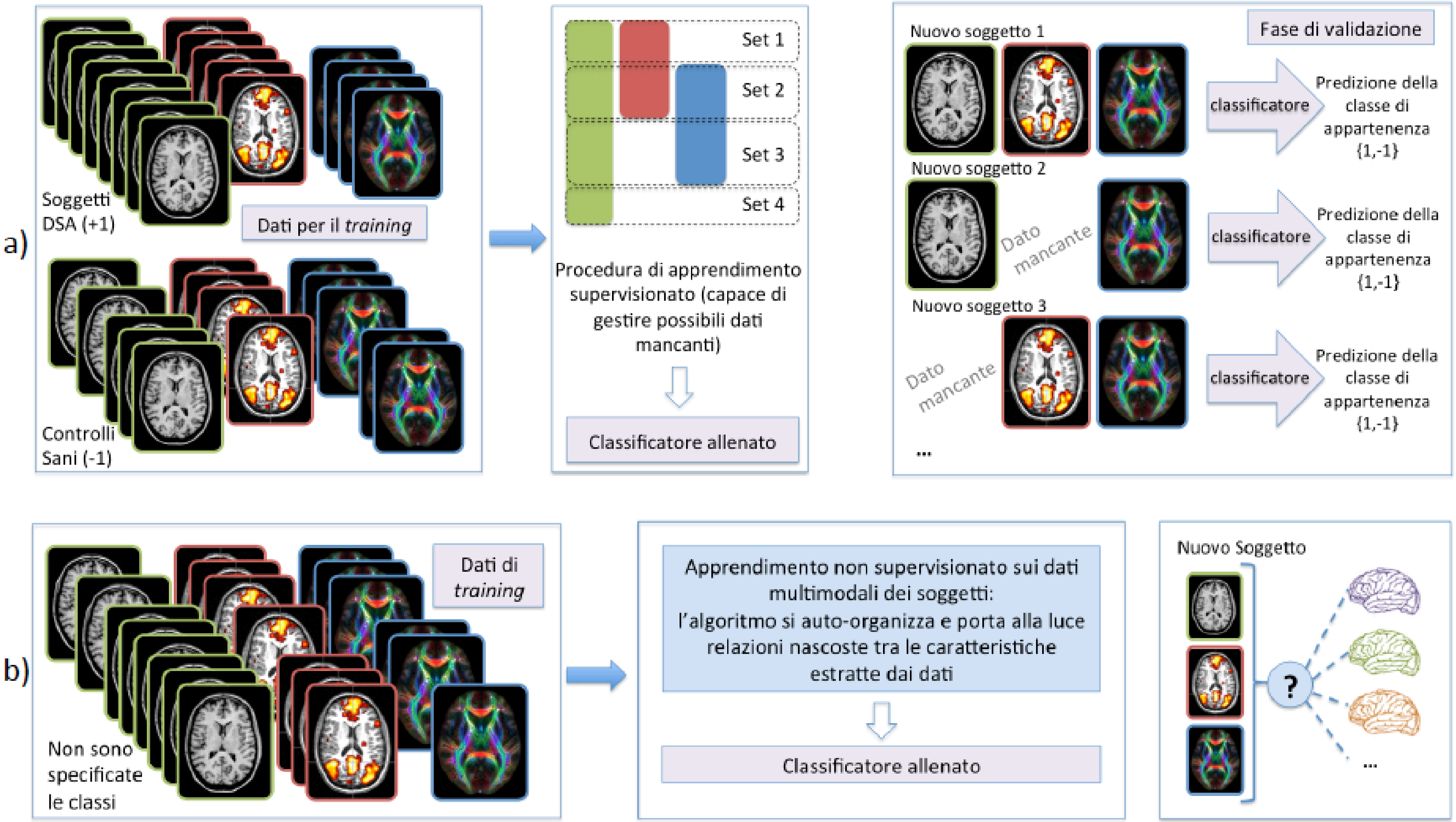
Developing appropriate data mining techniques.

Access to information on previous achievement with data samples (including negative results!).

INFN-PI computing center

Secure data handling;
Fast data processing.

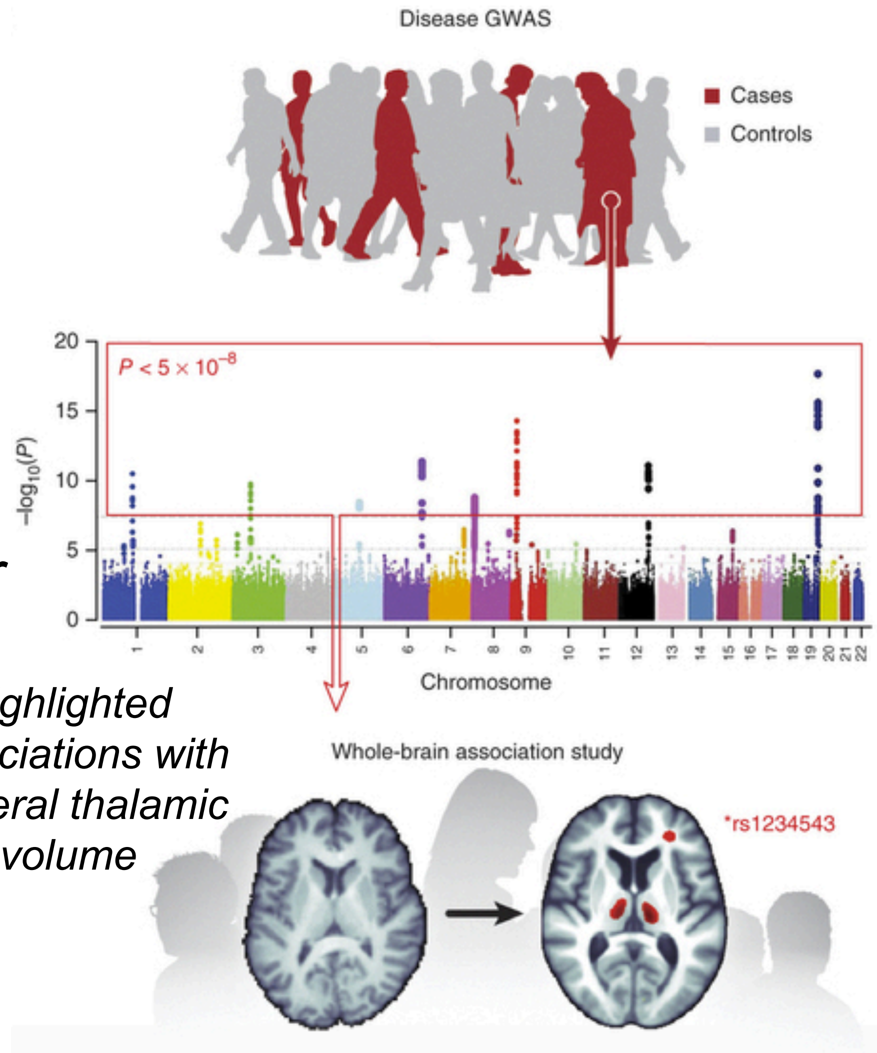
New analysis techniques



a) supervised learning; b) unsupervised learning.

Neuroimaging genetics

- Linking the influence of genes on the brain
- Genome-wide association scans (GWAS)
- Next Generation Sequencing (NGS) methods (several GBs per genome)
  - need for dedicated data compression algorithms
- Joint analysis of gene and imaging data: new tools needed



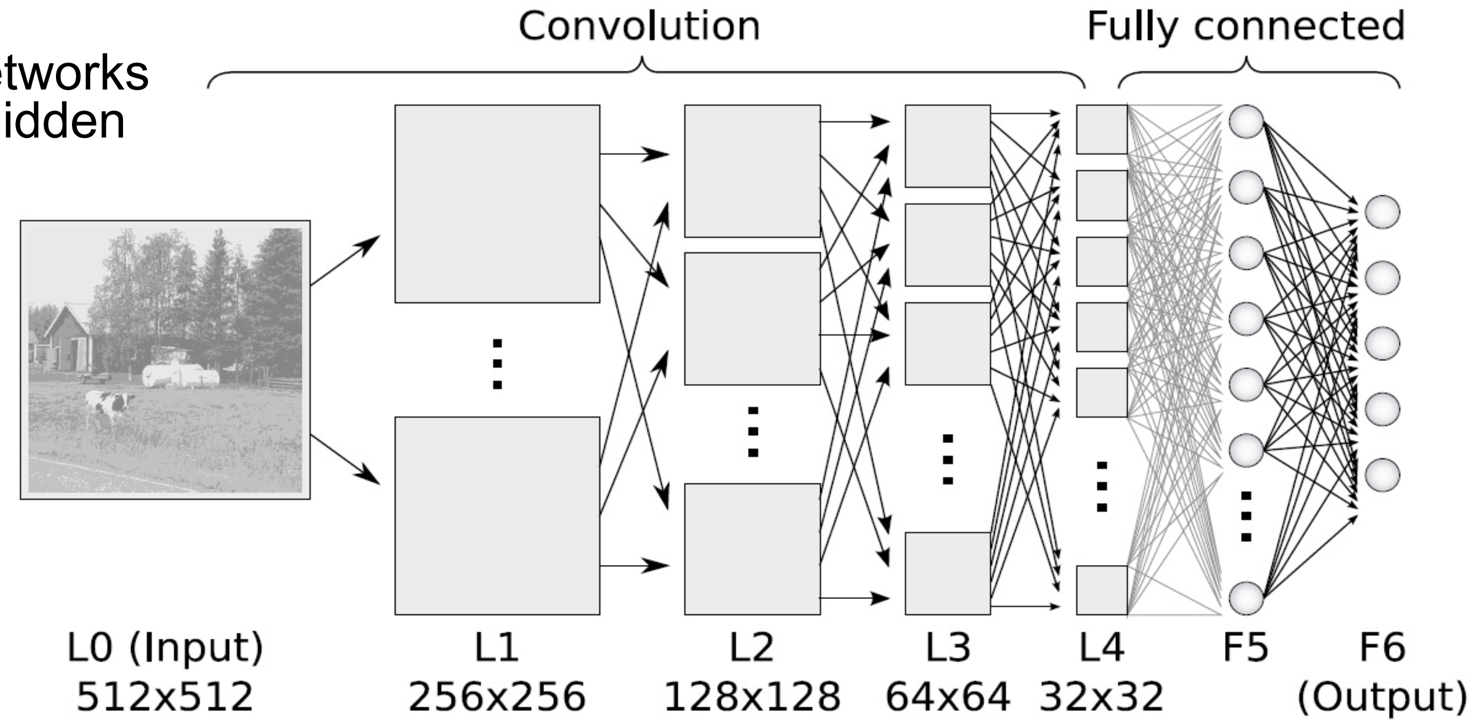
highlighted associations with bilateral thalamic volume

Medland et al., Nature Neuroscience, 2014

Deep learning

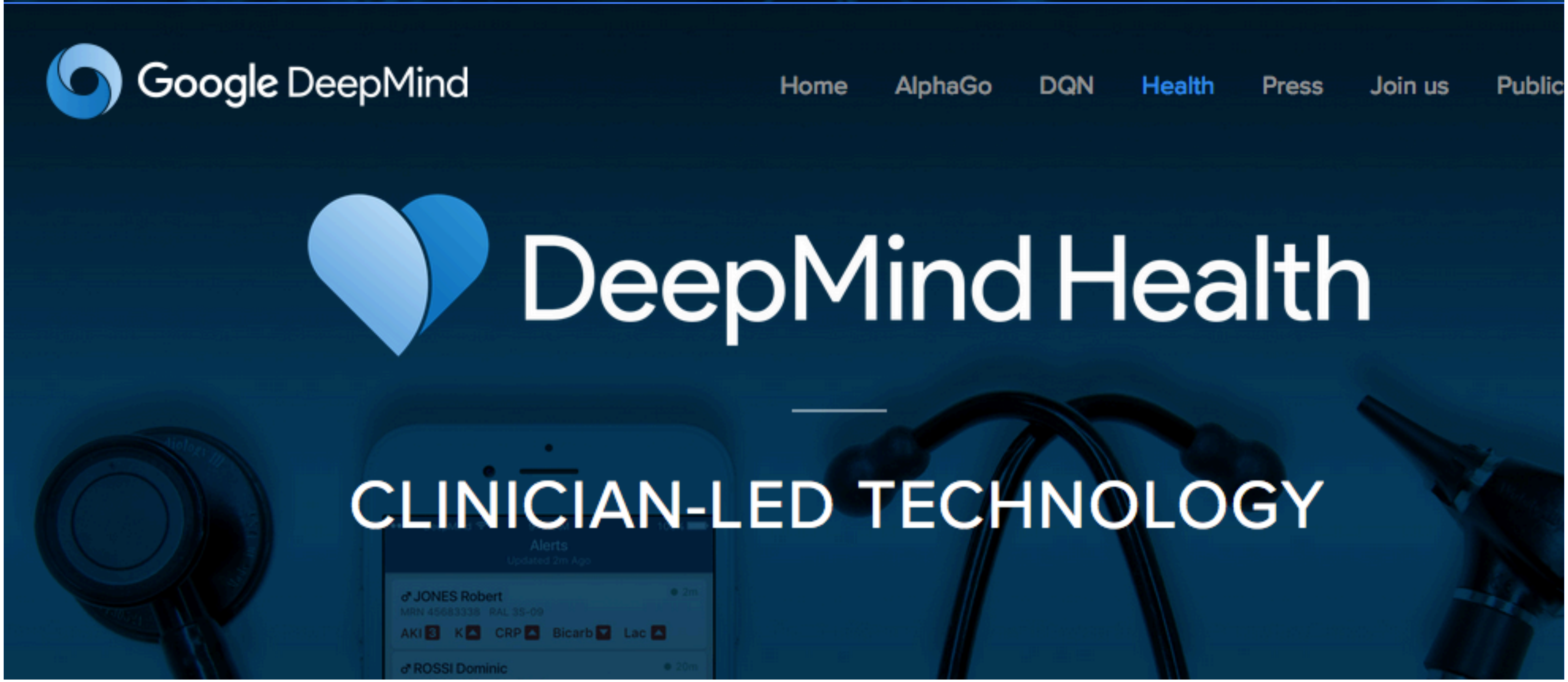
- An algorithm to make ANN learn in multiple levels of representation, corresponding to different levels of abstraction.

Neural networks with 6-9 hidden layers



- Convolutional neural networks (CNNs) represent mid-level and high-level abstractions obtained from raw data.

https://deepmind.com/index.html



Public funds and private initiatives

Big funded projects releasing data

Human Brain Project

ADNI

HUMAN Connectome PROJECT

- Commercial solution for data processing are being offered by Google, Amazon and Microsoft

→ Running on remote CPU clusters dedicated to brain research is essential to achieve real findings.

Conclusions

- Medical diagnostic/research imaging daily leads to an incredible amount of digital information
  - not fully exploited neither for diagnosis nor for research!
- Clinicians should be supported in the:
  - Diagnosis and monitoring of a wide range of disease conditions
    ➔ personalized medicine.
- New instrumentation and large-scale research consortia will be a valuable tool for research (e.g. neuroscience studies).
- New computational approaches are needed to handle and to mine multimodal data (imaging, genetics, clinical, demographic, etc.)

BIG data ➔ BIG science

Thank you for your kind attention!

I acknowledge all collaborators of the Stella Maris and IMAGO7 Foundations (Pisa) and all members of the SEVEN, TESLA and nextMR INFN-CNS5 experiments.

FONDAZIONE STELLA MARIS
ISTITUTO DI RICOVERO E CURA A CARATTERE SCIENTIFICO

IMAGO7

INFN
Istituto Nazionale di Fisica Nucleare

alessandra.retico@pi.infn.it