

**DTIC-MdM Strategic Program: Data-Driven Knowledge Extraction** 

### Imaging biomarkers: Algorithms, open data and infrastructure for neurological disorders

Christian BARILLOT, CNRS Research Director Visages U746 INSERM/INRIA Unit, IRISA, Rennes









#### Plan

- Frontiers in ICT editorial activity
- Imaging biomarkers: application to Multiple Sclerosis
- E-infrastructure for data management and analytics

founded by **neuroscientists** at **EPFL**, Switzerland in 2007 **improve** the current practice of academic **publishing** expanded to a journal series in **31 fields** by 2014

to be launched in Sept 2014: Frontiers in ICT

www.frontiersin.org

www.frontiersin.org/ict





Frontiers is part of the Nature Publishing Group family

**Real-Time Peer Review** – interactive, transparent and thorough

World-Class Editors - disclosed on all accepted articles

Quality Open Access – Nature Publishing Group family

Fastest Publisher – on average, 3 months from submission

Post-Publication Evaluation – advanced analytics on real-time impact of your article

**Democratic Article Tiering** – distill the best published research

Indexing in Major Archives – PubMed Central, Google Scholar, and more Copyright Retention by Authors – freely print and distribute your articles International Reach – over 6 millions page views per month across 160 Countries





Frontiers is part of the Nature Publishing Group family

for scientists by scientists

### frontiers in

ICT

Digital Education Cliff Shaffer

Quantum Computing Alexandre Zagoskin

Human-Computer Interaction Anton Nijholt

Mobile and Ubiquitous Computing Paul Lukowicz **Computer Image Analysis** Christian Barillot and Patrick Bouthemy

Computer Network Security Sandro Etalle

**Cyber-Physical Systems** Boudewijn Haverkort

Formal Methods Joost-Pieter Katoen

#### **Virtual Environments**

in collaboration with Frontiers in Robotics and AI Mel Slater

#### **Digital Health**

in collaboration with Frontiers in Public Health Patty Kostkova

**Big Data** Nick Duffield

..and even more to come!

frontiers

Frontiers is part of the Nature Publishing Group family

for scientists by scientists



frontiers in ICT

#### Computer Image Analysis

Our aim is to publish research addressing the scientific and

engineering challenges of Image Processing, Image Analysis, Computer Vision, and Image Perception and their related application fields, with a specific attention to emerging interdisciplinary areas.

The challenge of Computer Image Analysis is to handle overflowing sets of images that convey imperfect, partial or noisy information on the observed scene, while aiming at contributing to recognition, decision or prediction purposes. To achieve these goals, image analysis must ensure quality and performance criteria capable of producing more robust and accurate perception of physical or human processes.

#### Applications

- Environmental Sciences
   Satellite Imaging
   Flow Visualization
- Life sciences Biological Imaging Medical Imaging
- Imaging Genetics

   Multimedia
  Photo and Video Editing
  3D Vision
  3D TV
  Mobile Multimedia
- Videosurveillance and Biometrics

#### **Basic research**

cope

- Analysis of Images and Shapes
- Change and Motion Detection
- Color and Texture
- Computational Photography
- Computed Imaging
- Datasets and Performance Evaluation
- Denoising, Restoration and Inpainting
- Early and Biologically-Inspired Vision
- Event and Action Recognition and Localization
- Face and Gesture Recognition
- Graphical Models and MRF for Image Analysis
- Image and Video Indexing and Databases
- Image and Video Processing
- Image Classification

- Object Detection and Recognition
- PDE and Variational Methods for Image Analysis
- Registration and Matching
- RGB-D Image and Video Analysis
- Scene Understanding
- Segmentation and Grouping
- Shape Representation and Analysis
- Spectral Methods
- Statistical Methods and Learning for Image Analysis
- Stereovision
- Tracking
- Vision for Graphics
- Vision for Robotics
- 2D and 3D Optical Flow

#### SPECIALTY CHIEF EDITORS



ANALYSIS Christian Barillot IRISA CNRS, France



COMPUTER IMAGE ANALYSIS **Patrick Bouthemy** INRIA, France

#### ASSOCIATE EDITORS



**Vincent Lepetit** TU Graz, Austria



**Marleen De Bruijne** Erasmus MC, Netherlands



**Jean-Marc Odobez** Idiap Research Institute, Switzerland



Guillermo Sapiro Duke University, USA



**Peyman Milanfar** Google, USA



Florence Forbes INRIA France



**Marcello Pelillo** Ca' Foscary University, Italy



Shin'Ichi Satoh National Institute of Informatics, Japan



Kaleem Siddiqi McGill University, Canada



Dimitris N. Metaxas The State University of New Jersey, USA



Amar Mitiche INRS, Canada



Hayit Greenspan Tel-Aviv University, Israel



**Rita Cucchiara** Università degli Studi di Modena, Italy



#### Research Topics are key



#### MAPPING: MAnagement and Processing of Images for Population ImagING



### **Frontiers Spotlight Conference Award**

We commit to supporting the community

**100,000 USD** awarded to the Topic Editors of the most collaborative and outstanding Research Topic closed in 2016 to organize a conference on their Research Topic.



Read more in our blog: http://fron.tiers.in/go/3ZkS0w

#### Advantages of publishing in Frontiers



**Fast publication** Average 90 days from submission to publication



**Open access** Articles are free to read, for greatest visibility



**Copyright to authors** No limit to article distribution and re-use









Impact metrics Advanced metrics track your article's impact



#### Research network

Our network increases readership for your article

#### Global spread



Six million monthly page views worldwide



**Support** By our Swiss-based editorial team



#### Plan

- Frontiers in ICT editorial activity
- Imaging biomarkers: application to Multiple Sclerosis
- E-infrastructure for data management and analytics



### Imaging biomarkers: application to Multiple Sclerosis

What is a Biomarker?

Biomarkers are features that are objectively measured as indicators of normal biological processes, pathological changes, or pharmaceutical responses to a therapeutic intervention\*

\* Biomarkers definitions working group (2001) - http://www.biomarkersconsortium.org



## Imaging biomarkers: another way to use multimodal images

#### • Imaging biomarkers are used to:

- Detect pathologies
- Predict the level of risk
- Classified the extent of a disease
- Evaluate the therapeutic response
- Imaging biomarkers must be:
  - Quantitative
  - Accurate
  - Reproducible
  - Feasible over time



### «Brain Imaging Biomarkers»: From Bench to Bed





# **Imaging Biomarkers in Multiple Sclerosis**

- **Goal**: To guide the clinician (e.g. a neurologist) within the mass of information to integrate into the medical decision process
- Multiple Sclerosis (MS)
  - Chronic inflammatory-demyelinating CNS disease
  - Lead to acute handicap in young adults (high prevalence in Brittany)
  - Most frequent CNS disease in young adults
- Main Issues and Challenges
  - Early diagnostic and treatment of the pathology
  - Prevention of disease progression and future handicap
  - Better understanding of the pathology (new in-vivo classification of MS lesions)
  - Set-up and evaluate new therapeutic protocols (disease modifying drugs)



# Natural Evolution of Multiple Sclerosis (MS)





### **Imaging of Multiple Sclerosis**

#### State of the Art: White matter lesions as markers of evolution



D. Garcia. et al. IEEE-TMI 2011



# Segmentation of MS Lesions: a complex workflow



# **Imaging of Multiple Sclerosis: Estimation of Total Lesion Load**





Before



### **Imaging Biomarker of MS:** Quantification of temporal evolution

- Goal: Robust detection of evolving lesions
- Challenge: Intensity variations between two time points robust to outliers
  - T2, FLAIR, T1 acquisitions not quantitative



#### Intensity Normalization

[Karpate et al, 2014]: Longitudinal Intensity Normalization in Multiple Sclerosis Patients Robust Detection of Multiple Sclerosis Lesions from Intensity-Normalized Multi-Channel MRI Proceedings of MICCALCUP 2014



### Patient to Population Comparison for MS Lesions Detection

- Goal: Robust detection of lesions as deviation from local tissue intensity distributions
- Challenge: knowing at each voxel the multi-modal intensity distributions
  - Comparing multi-channel intensity of the patient to control intensities distributions
- Contribution:
  - Robust multi-modal intensity normalization
  - Atlas-based detection of intensity abnormalities in a patient (lesions)





#### Machine learning: Probabilistic One Class SVM for Automatic Detection of MS Lesions

- Goal: Propose an automatic framework for MSL Detection based on multichannel MRI patch based information
- State-of-the-art machine learning algorithms:
  - SVM [Vapnik et al.1995], Logistic Regression[Zhang et al.2002], Neural Network...
  - Works well in practice when training examples in classes are balanced
- If not ?
  - Class Imbalance  $\Rightarrow$  under-/over-fitting of the Classifier [Chawala 2005]
  - Class imbalance between Normal Brain Tissues and MS lesions
  - Solution : A higher misclassification penalty on the minority class (MS lesion)



Toy example of SVM for balanced and unbalanced classes, Courtesy : www.scikit-learn.org.



#### Machine learning: Probabilistic One Class SVM for Automatic Detection of MS Lesions

#### Methodology



[Karpate et al, 2015]: Probabilistic One Class Learning for Automatic Detection of MS Lesions. Proceedings of ISBI 2015



### Probabilistic One Class Learning for Automatic Detection of MS Lesions

- Goal: Robust detection of lesions as deviation from normal appearing tissues
- Challenge: overcome learning approaches problems with MS lesions:
  - Two-class imbalance problem (much more normal samples than lesions)
- Contribution:
  - Robust spatio-temporal multi-modal intensity normalization for T1-Gd and longitudinal MS lesion detection
  - One class learning for lesion detection from multidimensional MRI
    - Dimensionality reduction of the feature space
    - Lesions modeled as the complementary of the normal class
    - Testing by comparing patient patches characteristics to the *pdf* of the normal class
  - Able to robustly detect lesion locations in patients



15 [Karpate et al, 2015]: Probabilistic One Class Learning for Automatic Detection of MS Lesions. Proceedings of ISBI 2015



# Detection of MS lesions via competitive Dictionary Learning

- Goal: New sparse representation and dictionary learning method for classification
- Challenge: competitive dictionary learning
- Contribution:
  - Adaptation of dictionary size to each class complexity: improved over standard DL or discriminative methods
  - Detection of Multiple Sclerosis Lesions by classification of multimodal MRI images





#### Imaging Biomarkers in Multiple Sclerosis: Current Limits

Multiple sclerosis: auto-immune demyelinating disease



- Clinical radiological paradox [1]
  - Lesion study on conventional MRI (T2, FLAIR) not enough
  - Why? Missing information (lesion severity, position...)



# Imaging Biomarkers in Multiple Sclerosis: A Paradigm Shift

- Current bio-markers are not enough: The clinical-radiological paradox
- Measuring what the human eye cannot see:
  - Studying the early deposit of macrophages/microglia
  - Studying the diseases by characterization of axonal degeneration
- Imaging the cells and the microstructure:
  - Nano carriers of iron oxide can tag the macrophages activity (USPIO) in MRI
  - New ligand can tag microglia activity in PET
  - Quantitative MRI can characterize the damaged neuronal microstructure





# Imaging neuro-inflammation in MS: spatio-temporal analysis

- Longitudinal Analysis of inflammatory lesions in MRI (USPIO + Gd)
- Discovery of lesion classes to prospectively stratify MS population
- Analysis of the first 2 time points (before any treatment)
- 1. Selection of spatio-temporal patterns
- 2. Use machine learning framework to classify patients

#### Results

- Able to predict 2-years evolution of individual patients from the onset

#### Conclusion

Early lesion patterns characterization allows population stratification and suggests that belonging to a specific group can have an incidence on the future evolution of the disease.

$$f = [\lambda_{x0}, \lambda_{y0}, \lambda_{z0}, \lambda_{x1}, \lambda_{y1}, \lambda_{z1}, \lambda_{x2}, \lambda_{y2}, \lambda_{z2}, \lambda_{y2}, \lambda_{z2}, \lambda_{y2}, \lambda_{z2}, \lambda_{y2}, \lambda_{z2}, \lambda_{z2}, \lambda_{z2}, \lambda_{z2}, \lambda_{z2}, \lambda_{z2}, \lambda_{z3}, \lambda_{z3},$$

		chroC: Low Risk of evolution		
Patient	lesion clusters (and cardinality)	hypointense at m24 (cm3)	(cm3)	Group
6	C2(x1), C3(x1), C1(x38)	13,4	18,9	А
11	C1(x1)	1,8	9,7	В
9	C2(x1)	1,42	6,72	А
10	C2(x2), C1(x8)	1,37	4,82	А
16	C3(x1)	1,46	4,7	А
24	no active lesions at m0 and 3	3,35	4,63	С
	<u>C2(x1), C1(x2)</u>	1,77	4,20	^
13	C1(x2)	0,12	3,54	В
21	C1(x2)	0,74	3,5	В
18	C1(x4)	0,75	3,4	В
25	C1(x2)	0,96	3,32	В
7	C1(x9)	0,82	2,1	В
12	C1(x6)	0,34	2,1	В
2	C1(x1)	0,46	1,9	В
19	C1(x1)	0,27	1,73	В
5	C1(x2)	0,14	1,7	В
22	C1(x4)	0,1	1,27	В
17	C1(x1)	0.52	1,18	R
1	no active lesions at m0 and m3	0,12	1,18	С
8	C1(x1)	0,31	1,14	В
3	no active lesions at m0 and m3	0,28	0,98	С
15	no active lesions at m0 and m3	0,13	0,68	С
20	no active lesions at m0 and m3	0,18	0,54	С
23	no active lesions at m0 and m3	0,06	0,49	С
1.4	no potivo lociono at m0 and m2	0	0.20	C

**High Risk of evolution** 

B: Medium Risk of

A. Crimi. et al. PLOS-One 2014

T1

T0



# Imaging Biomarkers in MS: Brain microstructure can help

- Brain degeneration is also occurring at a diffuse level
  - Multiple sclerosis  $\Rightarrow$  axonal degeneration
- Challenging problem: overcome limitations of current methods:
  - Capability to be used in clinical routine (standard sequence) and on retrospective studies
  - Capability to exhibit crossing, diverging or kissing bundles and to track these bundles
  - Capability to characterize properties of neural circuits (instead of properties of voxels)





# Multi compartment model of MRI relaxometry for myelin quantification

- Brain tissues have different relaxation properties
- Challenging problem
  - Very long sequences, subject to artefacts
  - Pathological events leading to Neuro disease are first reflected as quantitative changes at the microstructure level.





otro



# Imaging Biomarkers in MS: Pooling of data for population studies





#### Plan

- Frontiers in ICT editorial activity
- Imaging biomarkers: application to Multiple Sclerosis
- E-infrastructure for data management and analytics



# Medical Imaging: a Big Data application

- Medical image databases
  - 20-40% increase of storage per year
  - A regular hospital produces 100-300 Tb/year of images

Needs for adapt computational solutions for the management of medical images

- Involve Data analytics: Image processing and machine learning
- Involve Data protection: complex legal constraints
- Very different to the e-commerce domain : less instances, more data / people



### Big Data and Medical Imaging: Where are we?

- What to do with the data produced?
  - Explosion of production and exchange solutions for imaging data
  - But ... "Information" does not mean "Knowledge"
- Known Issues:
  - How to exploit this mass of information easily?
  - How to deal with the mass of images?
- For the moment, approaches rather based on:
  - a descriptive analysis than on statistical one,
  - the search of correlations
  - the idea that the *mass* compensates the *quality*



# Big Data and Medical Imaging: Where do we go?

#### • Generalization of digital infrastructures on the Internet

- Towards "PACS 3.0"
- Local Storages are overpassed
- Dissociation acquisition / storage
- Remote viewing and analytics

#### • Emergence of dedicated digital infrastructures

- What operators?
- What costs and cost models?
- Emergence of virtual communities of users
- Emergence of new e-services on top of the image
- Emergence of new usages
  - New ways of working → image is shared
  - Emergence of virtual care networks
  - Evolution of the concept of territorial coverage (the image goes closest to the expert) → towards specialized imaging centers?
  - Image sharing and processing → Standardization of imaging protocols
  - Potential economies of scale



### Big Data and Medical Imaging: Towards the Open Data?

- Need to access more and more datasets, for:
  - Building atlases, learning models
  - Data mining, search for similar cases
  - Continuing education / training of health professionals
  - Evaluation / validation of image analytics
  - Certification of digital solutions around the image
  - Encouragements from funding agencies for the Open Data model
  - Emergence of population imaging
- Major Issues
  - What type of operators? (Public, Private / National, Global)
  - Which economic model (who bears the cost?)
  - What Standards?
  - Data quality control (*mass does not compensate quality*)
  - Emergence of new players (network operators, data centers, startup, GAFAM, ...)
  - Evolution/adaptation of the regulations
  - Ethical issues



# Medical Imaging as a Service

- Medical imaging applications have specific requirements for cloud computing:
  - Data:
    - Are heterogeneous
    - Are multistage
    - Have a strong semantic
    - Need acquisition protocols normalization
    - Are distributed over different sites (medical and/or academic)
    - Are confidential (security issues)
    - Need long term sustainability
  - Data Analytics / Image Processing:
    - Are often correlated (workflows)
    - Need automation for large cohorts (robustness, scalability, ..)
    - Need quality assessment: analytics transform images from qualitative to quantitative information and provide reference values (*imaging biomarkers*)
    - Computation time can be high on population cohorts
    - Computation time can also be sometimes critical (e.g. real time simulation, intervention, emergency, ...)



### Medical Imaging as a Service: Where do we go?

An example of an cloud solution for image data management : SHANOIR

#### **Two examples of e-infrastructures :**

National Cohort OFSEP

5

- MUSIC : a regional virtual care network
- Nation-wide Infrastructure France-Life-Imaging



#### : a Software as a Service Environment to Manage Population Imaging

- Neuro-imaging: large quantity of data still growing
  - Modalities: MRI, PET, Scanner, ...
  - Examples: MRI standard clinical examination: ~ 1400 images (> 100Mo)
     MRI clinical research examination: ~ 4500 images (> 2Go)
  - Clinical research: multi-centers studies
    - → Need for data storage and archive
    - → Need for data structuration
    - → Need for data sharing and accessing
- Existing solutions
  - CDs: simple solution but
    - Limited lifetime (~ 5 years), Data integrity not guaranteed
    - Physical storage
  - PACS: Picture Archiving and Communication System
    - Local (hospitals), regional or regional network  $\rightarrow$  limited access
- SHANOIR: secure server accessible from <u>any web browser</u>
  - Shanoir is currently operated and used in more than 50 research and clinical centers. W han 200 users connected, it hosts more than 80 multicenter research studies, with around 3500 patients patients and more than 200k data sets











#### SHAring NeurOImaging Resources An open source web platform for population imaging



**Shanoir**: this is more than 50 centers / 200 users connected / 80+ studies / 3500+ subjects connected and more than 200k data sets



# in the OFSEP Cohort

platform



	MRI Manufacturer - Model		
8	Siemens - Aera 1,5T		
7	Philips - Achieva 3T		
	General Electric - DISCOVERY MR750w		
2	3T		
2	Philips - Ingenia 1,5T		
2	Siemens - Avanto 1,5T		
2	Siemens - Skyra 3T		
1	General Electric - Signa HDxt 3T		
1	Philips - Achieva 1,5T		
1	Philips - Ingenia 3T		
1	Siemens - Espree 1,5T		
1	Siemens - Spectra 3T		
1	Siemens - Symphony Tim 1,5T		
1	Siemens - Trio 3T		
1	Siemens - Verio 3T		
31			

PACS Manufacturer	Model
GE Healthcare	Centricity PACS
Maincare Solutions	McKesson Radiology
Agfa	IMPAX
Telemis	Telemis
Carestream	Carestream Vue PACS
Global Imaging	GXD5 Pacs
	PACS Manufacturer GE Healthcare Maincare Solutions Agfa Telemis Carestream Global Imaging

nationwide, **clinical**, OFSEP is a cohort, representing about half of the MS patients population living in France, for a longitudinal follow-up (clinical, biological and neuroimaging data). Shanoir has been chosen to be the OFSEP neuroimaging data management



	<b>Brain MR imaging protocol</b> (At least one every 3 years)	Spinal cord MR imaging protocol (At least one MRI every 6 years)				
Recommended						
■ ■ ₽ ■	Sagittal enhanced 3D T1 Axial DWI with ADC map Axial 2D TSE T2/DP <u>or</u> 3D T2 Gadolinium injection (0.1 mmol/kg) 3D FLAIR ( <u>or</u> 2D FLAIR if not available) Contrast-enhanced 3D T1	<ul> <li>Sagittal T2</li> <li>In case of lesion occurrence</li> <li>Axial T2 GRE</li> <li>Axial T1 TSE with gadolinium injection</li> </ul>				
Recommended for first and differential diagnosis						
•	Axial 2D T2 GRE	<ul> <li>Sagittal T1 with gadolinium injection</li> </ul>				
Advanced sequences						
	DTI $\ge$ 15 directions (replace the DWI)	<ul> <li>Sagittal STIR</li> </ul>				



# e-Infrastructure

#### FLI + OFSEP = 2016 MICCAI Challenge of MS

FLĭ

in the

#### https://portal.fli-iam.irisa.fr/msseg-challenge/





France Life Imaging - IAM node

38





# Archiving and Sharing of images: How to change the usages?

- The gathering of data remains an issue
  - Reconciliation of image sources, reconciliation of images with clinical context, image fusion, quality control, harmonization of protocols
- Image analysis remains an emerging field
  - Robustness of tools, computation time, reproducibility of results, responsibility for use
- Who wants to share their data?
  - Conservatism of the community
  - Regulatory constraints
- Who can share their data?
  - Going beyond the "club" of insiders
  - Offer customized solutions
  - Certification of provided solutions
  - Accept the fair cost for these new uses.
- How to ethically manage data sharing and open data?
  - Anticipate before the problems arise: collectively (legally) and individually (e.g. consent)
- Security is and will always remain a challenge
  - Should not be underestimated
  - Do not use excuses for not taking care
  - Properly dimension the response to the risk so as not to kill the use
  - Multidisciplinarity: IT security is not just a medical issue
- Invent new jobs and develop them
  - Nothing can be done without a strong integration between Engineers, researchers, lawyers and doctors







#### THANK YOU FOR YOUR ATTENTION

- Elise BANNIER, PhD [U. Rennes I / CHU]
- Isabelle BONAN [Univ Rennes I / CHU]
- Fang CAO [INRIA]
- Laurence CATANESE [Inria]
- Emmanuel CARUYER CNRS]
- Sudhanya CHATTERJEE [Univ. Rennes I]
- Benoit COMBES [INRIA]
- Alessandro CRIMI [INRIA]
- Olivier COMMOWICK [INRIA]
- Isabelle COROUGE, PhD [Univ. Rennes I]
- Gilles EDAN [Univ Rennes I / CHU]
- Rene-Paul DEBROIZE [Univ. Rennes I]
- Hrishikesh DEHSPENDE [Inria]
- Quentin DUCHE [Inria]
- Ines FAKHFAKH [Inria]

- Jean-Christophe FERRE [Univ Rennes I / CHU]
- Daniel GARCIA-LORENZO [INRIA]
- Justine GUILLAUMONT [Inria]
- Renaud HEDOUIN [Inserm]
- Lea ITMI [SIEMENS]
- Angelique JARNOUX [Inria]
- Michael KAIN [Inria]
- Yogesh KARPATE [INSERM]
- Anne KERBRAT [Univ Rennes I / CHU]
- Jean-Yves GAUVRIT [Univ Rennes I / CHU]
- Juan Francisco GARAMANDI[Inria]
- Pierre-Yves JONIN [CHU Rennes]
- Jeremy LECOEUR [INRIA]
- Florent LERAY [Inria]
- Olivier LUONG [INRIA]
- Meena MANI [INRIA]

- Camille MAUMET [Univ. Rennes I]
- Pierre MAUREL [Univ. Rennes I]
- Marsel MANO [Inria]
- Cedric MEUREE [CEA]
- Thomas OBERLIN [Inria]
- Guillaume PASQUIER [Inria]
- Lorraine PERRONNET [Inria]
- Jan PETR [INRIA]
- Sylvain PRIMA [INRIA]
- Maia PROISY [CHU Rennes]
- Helene RAOULT [CHU Rennes]
- Guillaume RENARD [Inria]
- Aymeric STAMM [U. Rennes I]
- Yao YAO [Inria]
- Lei YU [Inria]