Artificial Intelligence in Medicine – AIM
Contributo Unità INFN di Pavia

Partecipanti - Unità di Pavia	FTE
Alessandro Lascialfari – PO - resp. Locale	0.4
Silvia Figini – PO – UNIPV & Mondino	0.2
Andrea Negri – PA – UNIPV	0.1
Anna Pichiecchio – PA – UNIPV & Mondino	0.2
Manuel Mariani – RU – UNIPV	0.3
Nicoletta Protti – RTDB – UNIPV	0.2
Marta Filibian – tecnico D1 – UNIPV	0.1
Ian Postuma – postdoc – INFN	0.1
Lisa Rinaldi – PhD – UNIPV	0.3
Davide Cicolari – PhD – UNIPV	0.3
Totale	2.2

Partecipanti esterni :

L. Preda, A. Filippi, UNIPV and IRCCS S. Matteo – Pavia
S. Bastianello, Istituto Neurologico Mondino - Pavia

Artificial Intelligence in Medicine

INFN - CSN5 proposal
2019-2021

Project framework (2019) :

National Responsible
Alessandra Retico (Pisa Division)

Participating INFN units
Bari, Bologna, Catania, Cagliari, Genova, Firenze, Pisa

External Research Centers and Hospitals involved in the proposal
IRCCS Azienda Ospedaliera Universitaria San Martino (Genova); IRCCS Fondazione Stella Maris (Pisa); Dip. Ricerca Traslazionale e delle Nuove Tecnologie in Medicina dell'Università di Pisa; Azienda Ospedaliera Universitaria Meyer (Firenze); Sezione Radioterapia Dip. Scienze Biomediche Sperimentali e Cliniche Mario Serio dell'Università di Firenze, IRCCS Arcispedale S. Maria Nuova Reggio Emilia; IRCCS Istituto Scientifico Romagnolo per lo studio e la cura dei tumori Meldola; IRCCS SDN (Napoli); IRCCS Giovanni Paolo II (Bari); Azienda Ospedaliera Universitaria Policlinico (Palermo).

Artificial Intelligence in Medicine

ABSTRACT

Analysis techniques based of Artificial Intelligence, including machine learning and, more recently, deep-learning approaches, are widely used in **medical diagnostics and therapy**. Starting from sensor data processing for image reconstruction, specific solutions include a variety of data mining, image segmentation, annotation and analysis applications, to end with intelligent systems for computed-aided diagnosis and image-guided therapy.

Big companies are quickly developing and placing on the market intelligent systems applications to assist clinicians in their daily tasks. Nevertheless, research institutions can provide relevant contributions in this still-open field of research. In particular, INFN can take advantage of its unique expertise in big data handling inherited for high-energy physics experiments and to the availability of extremely powerful computing centres mainly built to store and process those data. More importantly, a **network of fruitful interactions between INFN Physicists and Clinicians** of several Italian hospitals and clinical research centers has been built in the last two decades, thanks also to specific research initiatives funded by INFN-CSN5.

The **Artificial Intelligence in Medicine (AIM)** project aims to exploit the expertise of INFN and associated researchers on medical data processing and enhancement, and turn it in the development of advanced and effective analysis instruments to be eventually clinically validated and translated into products.

Keywords:
Interdisciplinary Physics, Medical Image and Data Analysis, Data Mining, Machine Learning and Deep Learning, Systems Medicine

INFN
Istituto Nazionale di Fisica Nucleare

Artificial Intelligence in Medicine

AIM

Objectives :

AIM+: Networking and continuous training

AIM1: Data harmonization

AIM2: Quantification

AIM3: Predictive models

… AIMn: Future aims to be agreed within the collaboration

Methods

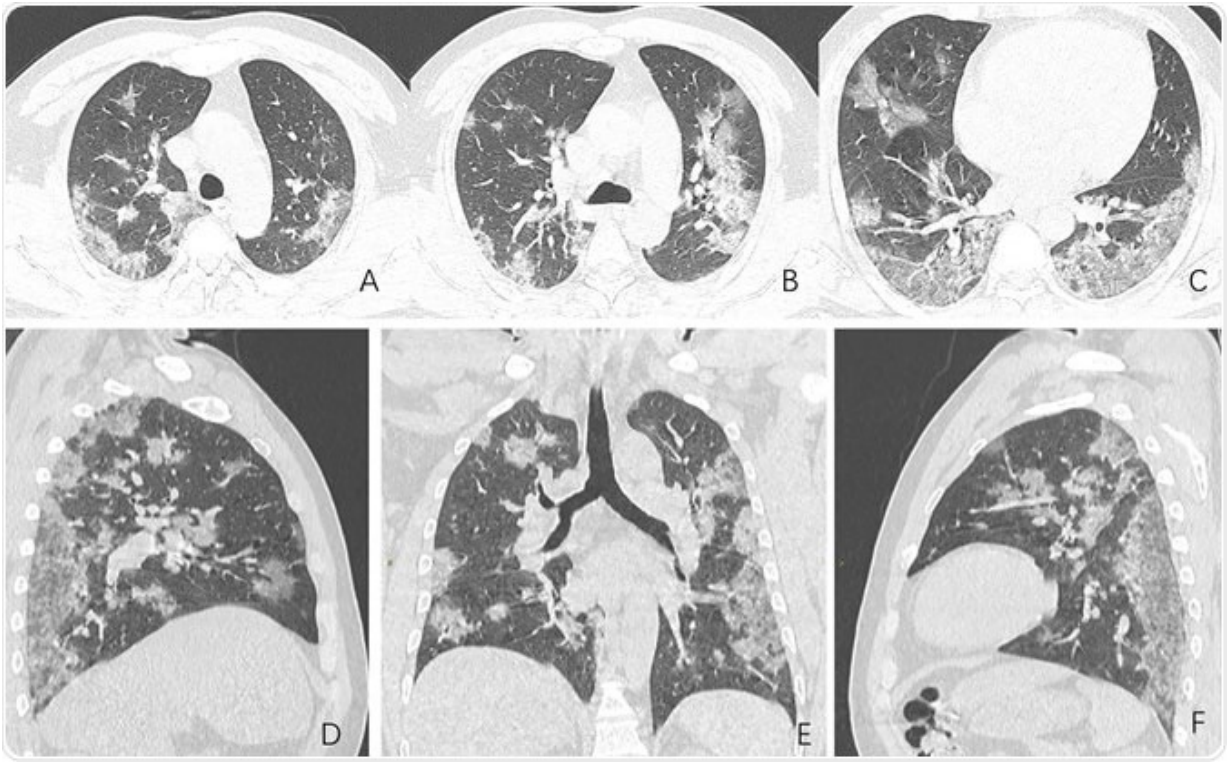
Three main ingredients are fundamental to the success of the AIM proposal: original strategies to mine the data, relevant data samples and suitable computing resources.
The subsections below first report the already-available analysis pipelines/data samples/computing facilities, which are the baseline of the project, upon which AIM researchers will grow a data mining knowledge in the medical field which will serve a large variety of future clinical applications.

Pavia Unit contribution (I)

COVID-19 data (in collaboration with IRCCS S. Matteo)

Radiomics/Machine Learning analysis :

- Segmentation (crucial to «shape» the lung)
- Individuation of shape and n^{th}-order features characterizing the data
- Percent of lung damage (fraction of ground glass) and its follow-up



Esempio

Immagini in un uomo di 44 anni che si è presentato con febbre, e sospettato di aver contratto la polmonite COVID-19.
A-C : Immagini assiali di CT (fetta 1-mm) : si vede l'opacità «a vetro smerigliato», irregolare e multipla (peribronchiale e subpleurale). Linfoadenopatia assente.
D-F : ricostruzione multiplanare; si nota la distribuzione diffusa delle lesioni.

Pavia Unit contribution (II)

Neuromuscular disorders

Facio-Scapulo-Humeral Dystrophy (FSHD) data (in collaboration with Istituto Neurologico Mondino)

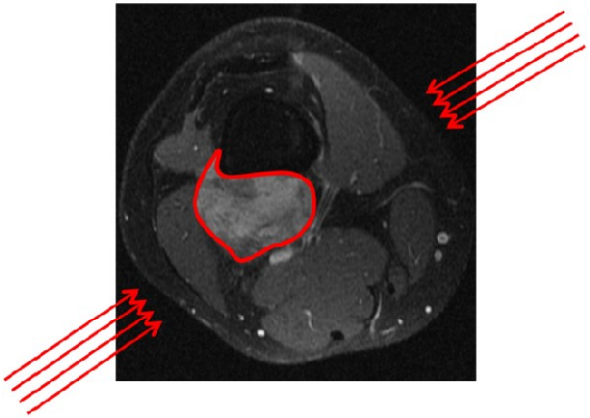
Data set : about 30 patients (till 24 months of follow-up)
Quantitative data (single muscle, inferior limbs) : fat fraction (FF), water-T2 (w-T2)
Images : MRI on thigh and leg (disease is diffused, FF is proportional to gravity and follows the inflammation, with unknown times). Different sequences, like e.g. STIR.

Radiomics analysis :
(i) shape, volume, energy, entropy features ;
(ii) FF and w-T2 ;
(iii) Connections among features and clinical scores, also “active phase” and “worsening”
(iv) Extraction of new features predictive of muscle deterioration and clinical changes over time

Pavia Unit contribution (III)

improve ongoing research with ML

Tumor contouring for TPS



- Recover clinically relevant CT and MRI scans
- accumulate a training set of tumor contours by physicians
- Standardise selection

Track counting for ^{10}B uptake measurements

- track selection algorithms depend on light and irradiation conditions
- Increase the set of training images
- find the appropriate algorithm

Unità INFN - Pavia - budget

- **Missioni (Due meeting annuali gruppo AIM (3 persone). Tre/quattro dottorandi/postdoc per una settimana presso altri gruppi) 3.5 k €**
- **Hard disks per storage dati. GPU. Licenze software (piattaforme Google, Amazon) 1.5 k €**

TOTALE 5 k €

External funds for Imaging projects : Progetto COVID FISR (INFN-PIGEPV, UNIPI, UNIPV), Progetto Ricerca Corrente S. Matteo, Istituto Mondino (personale)

The end

AIMs

AIM+: Networking and continuous training

The basic tool for expertise sharing consists in two annual workshops: a collaboration-wide event, dedicated to knowledge exchange and training of young researchers; and an application workshop, open to the INFN community and to the clinical world.
We also envision an internal PhD exchange program, so that our young researchers can be exposed to a greater variety of scientific problems.
In addition, the AIM collaboration shall look for collaboration with the INFN-sponsored software schools. The intent is to contact the scientific and organization boards in order to add a dedicated training session on applied data analysis.
The AIM collaboration will be oriented to applications for external funds. These can span both regional, nation-wide and EU calls. We expect to reshape the collaboration into ad-hoc research units, functional to the need of the calls, and leveraging the strong link with our clinical partners.

AIM1: Data harmonization

AIM1 is focused of data harmonization problems, and includes the development, optimization and implementation of algorithms (including ML- and DL-based ones) devoted to the **management of the systematic errors** introduced by different acquisition machine (e.g. MR scanners, RX tubes, etc.) **in multicenter studies**. Case studies include a large data samples of multicenter structural and functional MRI data, and a multicenter data sample of breast mammographic (screening and clinical) and Digital Breast Tomosynthesis (DBT) images. Also the integration of high-throughput data from different platforms (e.g. microarray and Next-Gen Sequencing) will be relevant for this aim.

AIMs

AIM2: Quantification

AIM2 deals with methodological developments in extracting quantitative information that can directly relate-to / impact-on a clinical outcome. Quantification does not usually involve ML or other data mining tools that need validation on an external dataset. Rather, it seeks direct data transformations either onto a small set of indexes (typically within dimensionality reduction techniques) that have direct and unequivocal clinical meaning, or directly into physically measurable quantities. Effective developments rely on the incorporation of external knowledge (i.e. clinical expertise/ex vivo/phantom measurements) rather than on a data-driven approach (typical of ML). Quantification methods need to be both robust (with respect to provenance and data acquisition protocols) and applicable to small and large dataset alike.

AIM3: Predictive models

AIM3 has the objective of putting in place **predictive medicine** solutions, and it includes at present: the search of biomarkers of pathology in a variety of disease conditions investigated with the most informative and appropriate diagnostic modalities (e.g. RX, MRI, PET); the identification of drug targets by in silico modelling personalized with patient profile; the identification of predictive features of treatment response in trans-cranial-MR-guided Focused Ultrasound Surgery; the prediction of treatment outcome in oncologic patients, relying on CT, MRI, and clinical dosimetric data of patients.

… AIMn: Future aims to be agreed within the collaboration

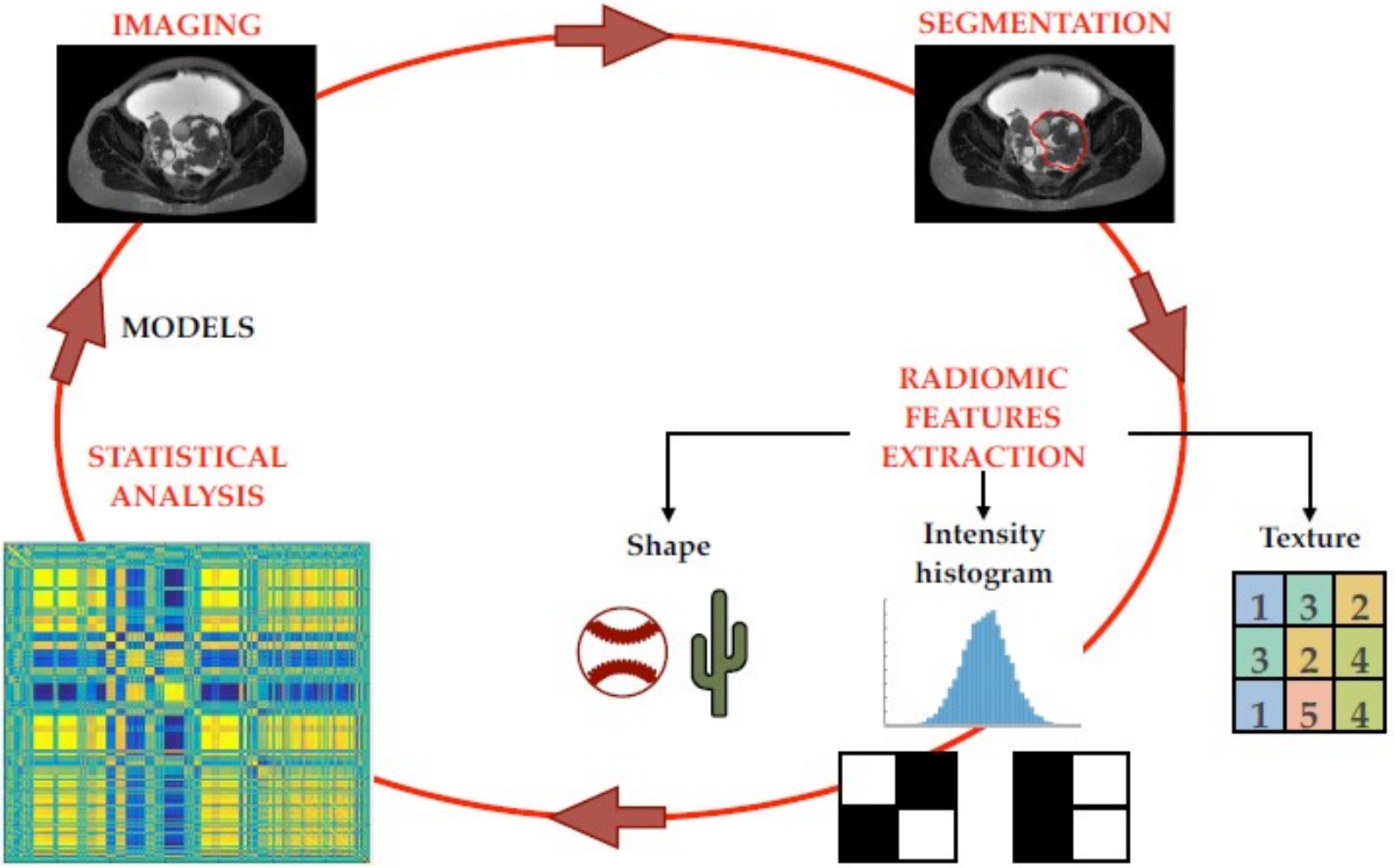
During the project execution, should a new clinically-driven general AIM arise, and the necessary test data samples made available, more AIMs would be added to the project workflow. Their inclusion in the workflow will be mainly conditioned to the available human resources, which strongly rely on undergraduate and PhD students, as dedicated fellowships for young researchers are not foreseen in this request for funding.

Candidate topics to be discussed for eventual future inclusion in AIM comprise:

- Data-driven image reconstruction,
- Multimodal data integration,
- Adaptive hadrontherapy.

Diagnostics and radiomics

"Radiomics cycle"



Models for radiomics

Repeatability and robustness of radiomic features extracted from Magnetic Resonance images of pelvic district: a phantom study

L. Bianchini[1], F. Botta[2], D. Origgi[2], M. Cremonesi[2], P. Arosio[1], A. Lascialfari[1]

[1]Dipartimento di Fisica and INSTM, Università degli Studi di Milano
[2]IEO, Istituto Europeo di Oncologia IRCCS, Milano

GIDRM XLVII NATIONAL CONGRESS ON MAGNETIC RESONANCE
Torino, 19th-21st September 2018

From scratch
First : phantom development for MRI

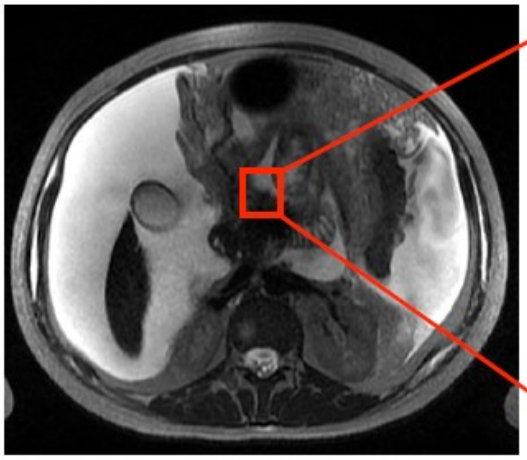
1. Need for phantom studies

2. Data from standard phantom

3. Development of “ad hoc 3D phantom”

What is radiomics?

512 x 512 matrix



313	289	305	293	266	312	407
302	279	293	271	228	270	376
285	265	274	252	205	236	340
271	255	264	250	209	227	318
267	257	268	264	231	239	315
273	264	278	281	248	248	313
285	267	280	288	255	244	297
295	265	266	283	260	239	271
301	261	245	275	276	250	254
300	259	232	269	296	276	256
293	262	232	265	302	292	266
282	265	241	261	287	283	269
262	251	242	252	261	260	266
231	219	228	241	242	248	269
195	183	211	236	244	260	283
169	165	207	243	260	279	289
169	176	225	261	275	281	274
194	213	258	281	277	265	246
223	252	287	291	270	247	230
239	276	299	291	264	243	233
242	284	295	281	261	246	240

Slides from
Linda Bianchini,
PhD, UNIMI

IEO members
F. Botta, D. Origgi,
M. Cremonesi

UNIMI
P. Arosio
F. Orsini

Data analysis

Software:

IBEX: An open infrastructure software platform to facilitate collaborative work in radiomics

Lifei Zhang
Department of Radiation Physics, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030

David V. Fried, Xenia J. Fave, and Luke A. Hunter
Department of Radiation Physics, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030 and The University of Texas Graduate School of Biomedical Sciences at Houston, Houston, Texas 77030

Jinzhong Yang
Department of Radiation Physics, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030

Laurence E. Court[a)]
Department of Radiation Physics, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030 and The University of Texas Graduate School of Biomedical Sciences at Houston, Houston, Texas 77030

(Received 27 October 2014; revised 15 January 2015; accepted for publication 2 February 2015; published 25 February 2015)

Pre-processing: "3σ normalization"

7 categories of features → shape, intensity histogram, texture

- Software
- Pre-processing
- Features selection

Machine learning
Deep learning
Artificial intelligence

Fundamentals :
Choice of phantom : 3D printing !
Data choice and collection
Data analysis