Istituto di Calcolo
e Reti ad Alte Prestazioni

CDS Computational Data Science Group

Computational Data Science Approaches for Biomedical Images and Biological Data

L. Antonelli, E. Falbo, M. Giordano, I. Granata, M.R. Guarracino, L. Maddalena

Aims

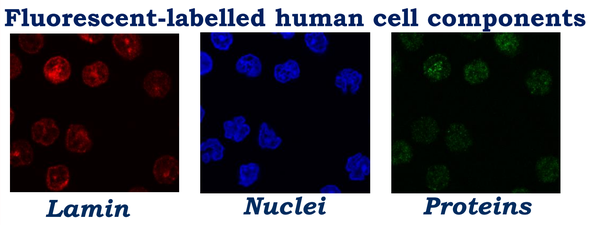
Solving real-life problems in different application fields by developing models, algorithms, and software tools to discover, understand, and gain insight into scientific phenomena through analyzing data produced in experiments and simulations, especially biomedical images and biological data, characterized by complexity, heterogeneity, and massive size.

Biomedical images

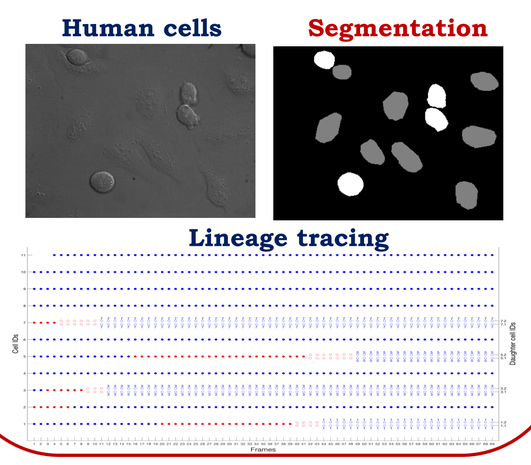
Biomedical images visualize human organs with different scales down to the cell resolution. They are essential data for studying diseases, discovering new therapies, and improving human health care. The most used imaging tools rely on X-rays (**CT scans**), magnetism (**MRI**), sound (**ultrasound**), radiopharmaceuticals (**SPECT**, **PET**), or light (**endoscopy**, **OCT**, **optical microscopy**). Beyond their specific aim in research studies and diagnostic practices, **computational data science approaches** are needed to process and analyze huge amount of imaging data for solving problems such as denoising and deblurring, classification, detection, segmentation, lineage tracing, and tracking in different application fields.

Detection and analysis of intranuclear protein patterns from fluorescence microscopy image stacks [1]

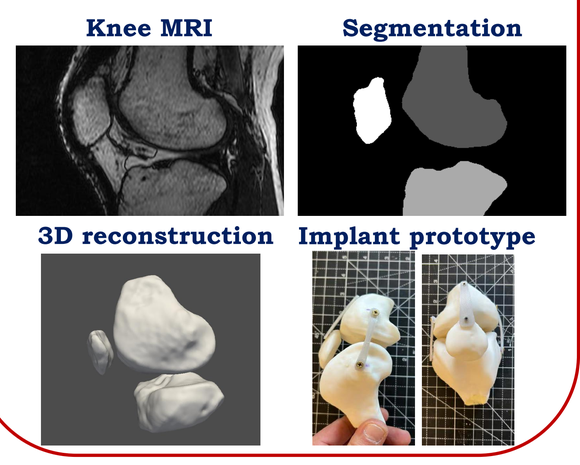
Fluorescent-labelled human cell components



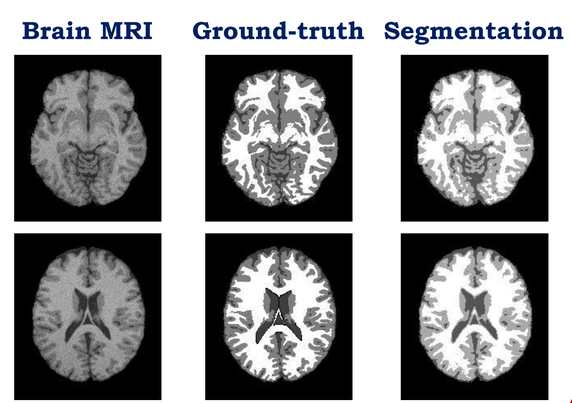
Cell phenotype classification and lineage tracing from time-lapse phase-contrast image sequences [2]



3D reconstruction of patient-specific models from MRI data for knee replacement surgery [3]



Segmentation of specific tissue regions from brain MRI data corrupted by noise and artifacts [4]

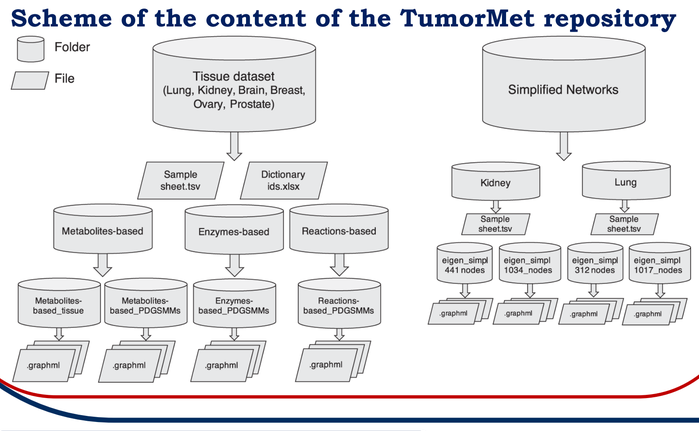


Biological data

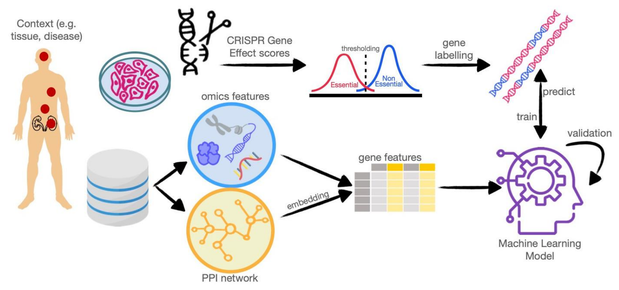
High throughput experiments produce massive and deeply informative amount of **biological data**, highly different in complexity, scale and format. The extraction of knowledge from these data in a comprehensive and holistic manner requires to adopt strategies aimed at integrating multi-modal and multi-source data. **Omics data** (genomics, proteomics, metabolomics, and transcriptomics) are considered by big data sciences, and their integration with multimodal imaging data has improved diagnosis and treatments in complex diseases, such as Alzheimer's and Parkinson's diseases and cancer, leading toward precision medicine. According to system biology, **biological data** can be organized in structures able to describe the role of each biological factor as part of a complex and highly interconnected system (organism, tissue, cell, disease). For the goal of **precision medicine**, these structures, alias networks, can refer even to a single patient or context. The principles and methods of **graph theory**, coupled with the **machine and deep learning approaches**, allow to read and interpret the richness of information contained in networks, as well as to capture the distances between networks describing different contexts. In the view of an open and successful science, sharing data and methods with the scientific community strongly contributes to the progression of knowledge and methodologies.

TumorMet: a repository of tumor metabolic networks extracted from context-specific genome-scale metabolic models [5]

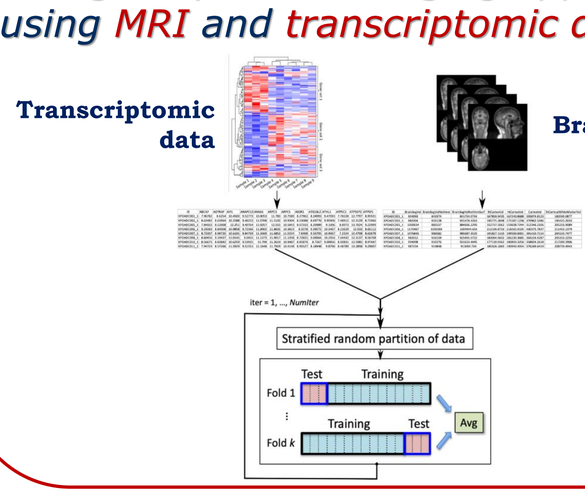
Scheme of the content of the TumorMet repository



Identification and prediction of human Essential Genes through a multi-omics and network-based approach. [6]



Classification of Alzheimer's Disease stages by omics imaging approach using MRI and transcriptomic data [7]



Acknowledgments

- ✓ FSC Project N. PON03PE_00060_5 MEDIA
- ✓ GNCS-INdAM group
- ✓ the ICAR-CNR INdAM Research Unit
- ✓ Institute for Cancer Research at Candiolo

References

[1] Gregoretti, F., Lucini, F., Cesarini, E., Oliva, G., Lanzuolo, C., Antonelli, L. In: MET in MOL BIOL (2023)
[2] Antonelli, L., […], Maddalena, L., and Guarracino, M.R., SCI DATA (*to appear)*.
[3] Antonelli, L., Gregoretti, F., Maddalena, L., Romano D., ICAR-CNR TECH REP (*to appear)*
[4] Antonelli, L., De Simone, V., Viola, M. In: COMPUTER VISION (2023)
[5] Granata, I., Manipur, I., Giordano, M. et al. SCI DATA **9**, 607 (2022)
[6] Manzo, M., Giordano, M., Maddalena, L., Guarracino, MR., Granata, I. In: STUDIES IN COMPUT INT (2023)
[7] Maddalena L., Granata, I., Giordano, M. et al. SN COMPUT. SCI. **4**, 249 (2023)

XIX WORKSHOP CNR-ICAR

6-8 SETTEMBRE 2023 RESORT TUI MAGIC LIFE CALABRIA