PhD in BIOINGEGNERIA / BIOENGINEERING - 38th cycle

PNRR_351_DOTT_RICERCA Research Field: EXPLAINABLE ARTIFICIAL INTELLIGENCE THROUGH HIGH ENERGY PHYSICS DATA FOR MEDICAL IMAGING IN ONCOLOGY

Monthly net income of PhDscholarship (max 36 months)
€ 1250.0 In case of a change of the welfare rates during the three-year period, the amount could be modified.

Context of the research activity	
Motivation and objectives of the research in this field	Over the last decade, medical imaging in oncology evolved towards the personalisation of patient treatment thanks to the use of advanced computational techniques, as Machine Learning (ML) and Deep Neural Networks (DNNs), coupled with radiomics and biomarkers identification to attain high diagnostic and therapeutic precision. However, despite their superior performance, ML/DNNs are often untrusted because of their complex structures, lack of interpretability, and black-box nature, thus demanding proper explanation and validation procedures before being used as supportive tools by clinical end-users. The main objective of the project is therefore to optimally describe the behaviour of ML models for tumour characterization and treatment outcome prediction, exploiting the support of novel methodologies based on explainable artificial intelligence (XAI), designed to explain and interpret the causal relationship between input features and outcome of a ML model. Specifically, different pathologies will be considered (with a main focus on brain and lung tumours) to allow for reproduction of the XAI models on different use cases. XAI techniques will be implemented and validated using the High Energy Physics (HEP) domain as a testbed, as it offers a vast library of datasets with solid and well-understood underlying causal theory, available also in the format of simulated events conformal to the data, ideal for a quantitative and measurable validation of XAI models; it

	is nowadays at the forefront of sophisticated big data analysis at the CERN Large Hadron Collider (LHC). The validated XAI models will then be applied to the oncology domain for an exquisite in-vivo patient-specific tumour description, by exploiting for the first time in an informed manner ML/DNNs tools to explain quantitative imaging radiomics or deep features selected by the model and their associations with clinical and biological information.
Methods and techniques that will be developed and used to carry out the research	Medical imaging data will be firstly collected from the publicly available dataset of the Cancer Image Archive (https://www.cancerimagingarchive.net/), exploiting datasets containing multi-parametric imaging data along with clinical/biological information and mainly focusing on head and lung cancers (two of the main widespread cancers) to allow for reproduction of the XAI models on different use cases. XAI methods will be then implemented and investigated by relying on literature-based approaches (e.g. SHAP, Grad-CAM, LIME) and/or implementing novel XAI models either incorporating inductive learning or adopting post-hoc analysis. Most of these approaches are global or local model-agnostic interpretation methods and allow for describing the contribution of individual features in the prediction. Identifying which features have the largest impact in the model output will allow one to trim useless ones and, focusing on the relevant inputs. The models will be implemented and validated on HEP dataset composed of Monte-Carlo ground truth simulations and collected data samples from the LHC experiments, through a key collaboration with CERN, INFN and University of Milano Bicocca. The evaluation will rely on the agreement between the ground truth HEP data and the estimated classification/prediction performed by the model, whereas the explainability of the models will be evaluated relying on quantitative metrics derived, for example, from saliency maps. XAI models developed and validated on HEP data, will be then adopted in tumour characterization and treatment outcome prediction in oncology, starting from information extractable from medical imaging. The models will be trained and tested on medical imaging datasets publicly

	available on the Cancer Image Archive. Quantitative imaging biomarkers will be derived from multi-parametric imaging, such as CT and MRI, relying on ad-hoc signatures, radiomics or deep features quantified on contours delineated on the tumour region to obtain in-vivo patient-specific tumour description. XAI models will be adapted and applied to characterise tumour and predict treatment outcome by means of classification and regression ML/DNNs models relying on the selected imaging biomarkers. Features robustness will be tested and model explainability and feature interpretability will be made possible by comparing model estimates and saliency maps with available surrogates of ground truth information based on clinical/biological endpoints, i.e., evaluation with respect to clinical feedback or tumour grading/pathology. This will allow for improved description of the ML model and identification of relevant and meaningful in-vivo imaging features for informed clinical decision-making thanks to validate XAI models.
Educational objectives	Educational objectives include that: - The PhD student will be involved in educational courses provided by the PhD school of Bioengineering at Politecnico di Milano; - By working in a collaborative environment involving different high-standing research units, the PhD students will be able to interact with professionals coming from different sectors to make the most of XAI advantages available in both HEP and medicine. - The PhD student will be able to understand the main concepts of machine learning and XAI applied in medical imaging in oncology - The PhD student will learn how to implement and validate XAI models relying on HEP as a test bed. - The PhD Student will participate in national and international conferences and schools, as those provided by CERN. He/she will be also involved in producing manuscripts to be submitted in top-ranked peer-reviewed indexed journals in Open Access modality.

Job opportunities	After the PhD, different job opportunities will be available as Postdoc or Research Scientist in national or international institutions. Positions as clinical bioengineer will be also possible within clinical institutions making use of AI tools to improve patient care and clinical workflows. Careers related to machine learning applied to medical imaging are recently evolving in many enterprise organisations and recently founded spin-off companies, demonstrating the excellence and actuality of the proposed project. Jobs as data scientists, big data engineers and machine learning engineers will be additional opportunities as the implemented and validated XAI models can be also replicated in other socially-relevant fields (e.g. security and finance).
Composition of the research group	1 Full Professors 1 Associated Professors 2 Assistant Professors 1 PhD Students
Name of the research directors	Prof. Chiara Paganelli

Contacts
mail: chiara.paganelli@polimi.it phone:+390223999021 web-page: www.cartcas.polimi.it

Additional support - Financial aid per PhD student per year (gross amount)	
Housing - Foreign Students	--
Housing - Out-of-town residents (more than 80Km out of Milano)	--

Scholarship Increase for a period abroad	
Amount monthly	625.0 €
By number of months	6

National Operational Program for Research and Innovation	
Company where the candidate will attend the stage (name and brief description)	-
By number of months at the company	0
Institution or company where the candidate will spend the period abroad (name and brief description)	CERN; home.cern
By number of months abroad	6

Additional information: educational activity, teaching assistantship, computer availability, desk availability, any other information
Attinenza alla tematica prescelta del bando ex D.M. 351, artt.6-9, comma 1 The project is relevant to the PNRR as it aims at bringing a significant development of knowledge in biomedical scientific research, specifically related to medicine in oncology. The main objective of the project is indeed to implement and validate Explainable Artificial Intelligence (XAI) models to be used in medical imaging in oncology, to unveil which and why specific quantitative imaging features derived with radiomics or deep learning based approaches are selected by a model. The XAI algorithms will be developed in collaboration with High-Energy Physics (HEP) experts, relying on the huge samples of well-formatted real and simulated data in that field to be used as a test bed. In agreement with the PNRR requirements, the project will allow for a more conscious and trustworthy use of machine learning tools in medical imaging in oncology, pushing towards an informed decision-making in oncology in patient stratification, treatment optimization and personalization to improve first of all patient care, clinical workflows, and thus also to reduce costs. Such an ambitious aim will be reached only through a multi-disciplinary approach, involving the different scientific and technological sectors of medical imaging in oncology and HEP. All developed ML models and algorithms will be made available to the scientific community and to the public, as they will run on datasets from the CERN Open Data Project (http://opendata.cern.ch) and on clinical open data present in Cancer Image Archive (https://www.cancerimagingarchive.net/), to allow for universal reproduction of the results. All new data samples will be generated according to the FAIR Data Principles. **Ente, università, azienda, centro di ricerca presso cui si svolgerà il periodo di studio e ricerca all'estero** The secondment period (up to six months) will be conducted at CERN (https://home.cern/), the European Organization for Nuclear Research, one of the world's largest and long-lived centres for scientific research in particle physics and its spinoffs. In particular, the development of technologies initially designed for particle physics lead to major contributions to the development of instrumentation for biomedical research, diagnosis and therapy, including particle beam technology and medical imaging. A collaboration with INFN and University of Milano Bicocca will be also present to conduct the project, therefore additional exchange periods are also foreseen with these institutions. A shared desk and a PC will be given to the student for the time needed to carry out research. A limited budget will be available for travelling and purchases, too.