PhD project:
Development of methods for the creation of efficient and robust AI models in medical imaging
Contact: irene.buvat@universite-paris-saclay.fr

Description

Context

The development of Artificial Intelligence (AI) methods to facilitate the interpretation of medical images and to better exploit their content is growing fast. However, these developments face recurrent obstacles. For example, the quantity of medical images available for the learning databases is often limited, especially in comparison to the volumes of "natural" images that are available to train algorithms for recognizing objects, animals, plants, etc. This limited volume of images available for training is even more frequent in the context of precision medicine where diseases are stratified by an increasing number of criteria, which makes it difficult to collect homogeneous image bases with respect to all these criteria. Beyond the quantity of images composing the learning bases, most AI methods developed in medical imaging are currently based on supervised learning. They require a tedious work of annotation of the images by experts. The annotations can be of variable quality, noisy and biased. These limitations are pervasive and call for new solutions.

The hypothesis of this thesis is that AI methods developed for applications other than medical imaging could be revisited and adapted to the specific context of medical images, with the aim of exceeding the performances currently achieved by algorithms developed from databases limited in quantity and quality.

Objectives

The objective of the thesis is to exploit recent advances in the field of AI, notably concerning weakly supervised and unsupervised learning or transfer learning, to solve the obstacles systematically encountered in the development of AI applications in medical imaging, namely the small volume of imperfectly annotated data. It will also demonstrate the relevance of the methods developed on use cases in PET/CT and MRI imaging.

Methods

Several databases of annotated PET/CT images (lymphoma, lung cancer and breast cancer) and MRI images (breast cancer, brain tumors) are available in the laboratory, to which questions of medical interest are associated. These databases, each containing the data of 100 to 500 patients, have been exploited to build AI models by supervised learning whose performances have been precisely characterized (1-4). The thesis work will build on this existing work to provide solutions to the following questions:

- Can we artificially enrich these learning bases to build better and more robust models? For this, original methods of data augmentation or synthetic image generation can be developed.
- Can we characterize the quality of annotations and better exploit imperfect annotations to improve the performance of the models? To this end, some databases have been partially annotated by several experts, which makes it possible to evaluate the variability of the annotation and to study its impact.
- Can we pre-select the images to be annotated in a preferential way to ensure sufficient variability in the annotated data to build robust models? To facilitate this study, we have several independent databases for which the same question is asked.
- How can we better leverage additional data (new tests acquired in the same patient, additional patients) to improve the performance and robustness of an existing model? For several databases, we have longitudinal follow-up images of patients (images taken before and during treatment).
- Can we predict whether a model can be successfully applied on a new patient cohort? We do have several datasets for challenging the model ability to generalize to different samples.

Expected results
The thesis work should lead to new methods to facilitate the design of efficient and robust AI models in medical imaging, and to demonstrate of the relevance of these methods on clinically-relevant use cases, responding to highly topical problems in the context of precision medicine.

References
1. Saint Martin MJ, Frouin F, Malhaire C, Orlhac F. Decrypting the information captured by MRI-radiomic features in predicting the response to neoadjuvant chemotherapy in breast cancer. 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 3227-3230, 2022. DOI: 10.1109/EMBC48229.2022.9871844.
2. Rahimpour M, Saint Martin MJ, Frouin F, Akl P, Orlhac F, Koole M, Malhaire C. Visual ensemble selection of deep convolutional neural networks for 3D segmentation of breast tumors on dynamic contrast enhanced MRI. Eur Radiol, in press, 2022. DOI: 10.1007/s00330-022-09113-7
3. Girum KB, Rebaud L, Cottereau AS, Meignan M, Clerc J, Vercellino L, Casasnovas O, Morschhauser F, Thieblemont C, Buvat I. 18F-FDG PET maximum intensity projections and artificial intelligence: a win-win combination to easily measure prognostic biomarkers in DLBCL patients. J Nucl Med. 63:1925-1932, 2022. DOI: 10.2967/jnumed.121.263501
4. Wallis D, Soussan M, Lacroix M, Akl P, Duboucher C, Buvat I. An FDG-PET/CT deep learning method for fully automated detection of pathological mediastinal lymph nodes in lung cancer patients. Eur J Nucl Med Mol Imaging. 49:881-888, 2022. DOI: 10.1007/s00259-021-05513-x
5. Bradshaw TJ, Boellaard R, Dutta J, Jha AK, Jacobs P, Li Q, Liu C, Sitek A, Saboury B, Scott PJH, Slomka PJ, Sunderland JJ, Wahl RL, Yousefirizi F, Zuehlsdorff S, Rahmim A, Buvat I. Nuclear Medicine and Artificial Intelligence: Best Practices for Algorithm Development. J Nucl Med. 63:500-510, 2022. DOI: https://doi.org/10.2967/jnumed.121.262567

Supervision, collaborations, dissemination

Supervision
The PhD thesis will be supervised by Irène Buvat (DR CNRS) and Frédérique Frouin (CRHC Inserm). The project will take place in the context of the AI.DReAM research program, led by General Electric and funded by BPI France. The PhD student will participate in weekly team meetings to follow the progress of the work carried out in the unit and will be expected to regularly present his/her results on this occasion. He/she will participate in bi-monthly meetings with General Electric. He will also be followed by a tutor, a permanent member of the laboratory but not involved in the doctoral project, to prevent any difficulties that may arise during the thesis.

Collaborations
For 10 years, the unit has been involved in the field of artificial intelligence for medical image analysis and collaborates with other teams, at the Institut Curie, Ecole des Mines, and Neurospin, experts in bioinformatics and artificial intelligence. It also has close collaborations with teams from the Medical University of Vienna (Austria), the University of Amsterdam (Netherlands) and international labs involved in the Society of Nuclear Medicine and Molecular Imaging AI Task Force (US).

Dissemination
The doctoral student will present his work via communications in national and international congresses in the field of imaging and oncology. He/she will also have the opportunity to publish his/her results in international journals.

Funding

The PhD is funded through the AI.DReAM contract, supported by BPI France (2021-2026), and managed by the Institut Curie Research Center. The PhD student will be recruited by the Institut Curie Research Center.

The doctoral student will have access to the unit's computer resources. He/she will also have access to anonymized imaging and clinical data respecting the RGPD regulation thanks to the collaboration established between the LITO and Institut Curie. In addition to the training courses offered by the Doctoral School, the doctoral student will participate in conferences (one per year) subject to acceptance of a communication as first author. These participations will be supported by the AI.DReAM contract. The PhD student will comply with the internal rules of the unit. He/she will not be exposed to specific risks (chemical, biological or radiological).

The PhD will take place within the framework of the Doctoral School EOBE (Electrical, Optical, Bio-physics and Engineering) of Paris Saclay University. The subject of the thesis is listed available at: https://www.adum.fr/as/ed/voirproposition.pl?site=PSaclay&matricule_prop=45977

Location of the work

The thesis will take place in the Laboratory of Translational Imaging in Oncology (lito-web.fr) currently located in the Institut Curie Research Center, Building 101B, on the Paris Saclay University Campus, 3 min walk from the RER B, Orsay-Ville station.
The laboratory will move in 2025 to the Institut Curie Research Center in Saint-Cloud, 2 minutes walk from the Saint Cloud train station (Transiliens L and U), Tramway T2, Metro line 10 Boulogne - Porte de Saint Cloud.
Teleworking is authorized for a maximum of 2 days per week.