PhD position

Integrating multimodal data (PET and CT imaging + genomics) for the prognosis of lung cancer

Institutions: Laboratoire de Traitement de l'Information Médicale (LaTIM), INSERM UMR 1101, Brest; SOPHiA GENETICS, Bordeaux

Superviors: D. Visvikis (LaTIM), T. Colin (SOPHiA GENETICS)

Duration: 3 years, starting September/October 2018

The significant advances these last few years in oncology has paved the way for a new therapeutic approach with the development of targeted therapies. These treatments allow to target an identified molecular anomaly at the level of the tumor. The recent evolution of these personalized patient management has revolutionized how certain types of cancers, such as metastasized lung cancer, are treated. This revolution was made possible by the higher availability and generalization of genomics analysis of tumors. At the same time, imaging protocols have become systematic and standardized for this pathology. A given patient now undergoes successively CT and PET imaging sessions. The generalization of PET/CT multimodal scanners now allows for the acquisition of morphological and functional data in the same exam. However, as these fields progress, it becomes more and more evident that the therapeutic decision is harder and harder to make for a clinician. The number of potential targets and their different prognostic impact combined with the previous therapy (received treatments, efficacy and side effects) render very difficult the optimal therapy choice for each patient at a given time. The evaluation of therapy response must also be adapted to each treatment type and the usual RECIST (radiology) or PERCIST (nuclear medicine) criteria may not be the most adapted for the targeted therapies, and finer metrics must be developed. The recent rise of immunotherapy increases the complexity of the field even further. The clinicians therefore face a decision scheme that must take into account i) quantitative data regarding dozens of genes, ii) a potentially complex clinical history and iii) multiple morphological and functional images that need interpretation. It is obvious the clinicians will need help to exploit efficiently this amount of data. The evolving nature of these data (they will be constantly updated during the patient follow up with new imaging exams, new biopsies and in the future liquid biopsies) implies that the usual data analysis techniques may not be sufficient to address their complexity. The richness of the available data for each patient as well as the knowledge acquired by analyzing cohorts of patients can only be exploited through advanced artificial intelligence and data mining methods in order to develop predictive models (for survival, response to therapy, etc.).

The objective of this PhD is to develop an artificial intelligence framework able to combine efficiently genomics, clinical and imaging data. This development will be carried out in two steps. The first is a radiomics-type analysis relying on automatic image segmentation and tumor characterization combined with already existing models for prediction of tumor growth. The development will especially focus on robustness and reproducibility of the radiomics workflow. The second step will consist in combining radiomics with genomics as they are provided by separate analyses. This coupling will rely on statistical learning methods exploiting the available clinical, genomics and radiomics data. The developed methods will be validated using lung cancer data available through partnerships with several clinical centers. More specifically, the clinical goals during the PhD are:

1) To provide a personalized prognosis. Exploiting the available data at diagnosis, the tools will be able to provide an accurate prognosis that can help guide and decide the treatment strategy. This

prognosis will be further estimated thanks to the early response to therapy assessment. It could regard disease-free survival, overall survival, toxicity prediction, or quantifying specific risks.

2) To provide a personalized follow-up: exploiting the clinical, genomics and radiomics data available during the course of treatment, the tools will be able to anticipate treatment failures such as disease progression (metastases) in order to improve their control. In particular, this includes response to targeted therapies with an evaluation combining imaging and liquid biopsies.

Environment for the PhD: this position will be funded by the society SOPHiA GENETICS in collaboration with the LaTIM (University of Brest, INSERM UMR 1101) and co-supervised by Dimitris Visvikis and Thierry Colin.

The LaTIM is the laboraty of medical information processing in Brest. The objective of the ACTION team within the LaTIM is to integrate multimodal data including imaging for therapeutic oncology. One of the major research themes of the team is radiomics, from image pre-processing and segmentation to the building and validation of multiparametrics predictive models. SOPHiA GENETICS is a worldwide leading company on data-based medicine that operates in 60 countries and has already performed genomics analyses of more than 200 000 patients. The Nenuphar teams in SOPHiA GENETICS is based in Bordeaux and dedicated to the radio-genomics studies.

We are looking for candidates with an Engineer or Master degree in computer sciences or applied mathematics with a strong expertise in software development and numerical experiments.

Requisites: C++ programming, basics of image processing, knowledge in database, OpenGL, QT and/or expertise in statistical learning.

Contact: Send a CV, letter of motivation and recommendation letters along with diplomas to Dimitris Visvikis (dimitris@univ-brest.fr) and Thierry Colin (TColin@sophiagenetics.com)