PhD position 2018-2021

Machine learning for cancer mapping based on multimodality medical imaging

Scientific context

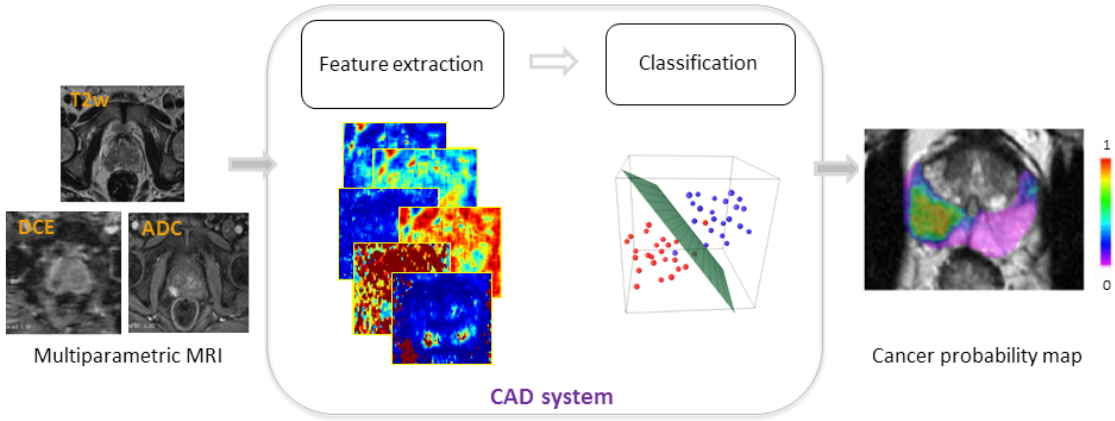

Figure 1: Automatic cancer mapping based on multiprametric MRI and machine learning. A first CAD prototype was developed based on a clinical collaboration with Lyon University Hospital (HCL) and the financial support provided by two research grants from the INCa (Institut National de Recherche sur le Cancer)

Machine Learning (ML) has emerged as a field of artificial intelligence (AI) to provide tools and methods that allow learning to make prediction on simple tasks based on learning samples. Application of ML to healthcare is among the most challenging ones with the potential to exploit information provided by exponentially growing mass of heterogeneous data (image, semantic information, biological parameters etc..).

In recent years, machine learning has received a lot of attention to explore and structure multidimensional and multi-modality medical imaging data for the purpose of image segmentation, registration or automated detection and characterization of pathology.

CREATIS has developed strong skills in developing clinical decision systems (CAD) for cancer [1, 2, 3, 4] and brain imaging [5, 6, 7] based on the most advanced machine learning techniques. Such systems are designed to assist clinicians in their diagnosis by highlighting abnormal regions in an image. One active project in the 'Images and Models' team concerns the **prototyping of a computer-aided diagnosis system for prostate cancer screening based on multiparametric magnetic resonance imaging (MRI)**. As illustrated on Figure 1, this system allows performing a probabilistic prostate cancer mapping based on a standard machine learning methodology where the decision model is learned at the voxel level on a series of feature vectors selected from normal and pathological locations on patient scans [8, 9]. Despite an important improvement brought by such systems for the problem of cancer mapping, they still suffer from limitations that restrict them from being used at a larger scale.

Project aim

The purpose of the PhD project is to go one step further and address **two challenges**

- **to predict not only a presence/absence of cancer but also the degree of its aggressiveness**. The binary prediction does indeed not suffice for active surveillance of patients with low aggressiveness cancers or patients that could benefit from focal therapy. The main challenge is that the different classes, corresponding here to the different levels of lesion aggressiveness, are highly correlated and interdependent which challenges standard multi-class classification algorithms [8].

- **to develop a system whose performance generalize well with data potentially coming from different populations**, as can be encountered with imaging data pooled over different clinical centers (ie acquired on different MR and/or with different sequences and parameters). In our context, we have shown that a CAD system exhibiting good performance when trained and tested on the same data distribution can drop significantly when tested on data from a different distribution [9, 10].

From the methodological point of view, we plan to **explore new machine learning algorithms** that tackle the problem induced by the presence of **highly correlated and interdependent outcomes in multi-class classification** as well as **heterogeneous data**.

One research axis that will be investigated is to explore the potential of **deep learning** to address both questions of multi-class learning with heterogeneous and correlated data. Deep learning is one approach of machine learning that is currently focusing increasing attention in the medical scientific community, since it has been recently shown to outperform traditional approaches in most of the applications cited above. Our objective will be to investigate **novel deep architectures** that will efficiently fit our needs, including the derivation of novel objective functions for **multi-class and transfer learning**. We will particularly focus on **semi-supervised networks** allowing to operate on **partially labeled data**, which is a major characteristic of medical data.

Scientific coordination and collaborations

The candidate will be supervised by Carole Lartizien from CREATIS who has an acknowledged experience with multi-modal medical image processing and have developed strong skills in machine learning for medical imaging over the past few years (creatis.insa-lyon.fr/∼lartizien).

This PhD project is part of a 8 Meuros research project recently funded by the French National Research Agency (ANR) to develop diagnostic and therapeutic tools for prostate cancer. The successful PhD candidate will benefit from multiparametric MR imaging database including more than 300 prostate cancer patients, build by Pr Rouvière from Hospices Civils de Lyon (HCL). The PhD student will also benefit from ongoing collaborations with external experts in the machine learning domain, as well as the scientific emulating environment of the 'Images and Models' team whose research interest are focus on medical image processing with machine and deep learning.

Scientific fields

Machine learning, deep learning, multi-class learning, transfer learning.
Biomedical imaging, MRI, Prostate cancer mapping.

Scientific and work environment

The postdoctoral position will take place at the CREATIS laboratory (www.creatis.insa-lyon.fr). The succesful candidate will join the 'Image and Models' team (https://www.creatis.insa-lyon.fr/site7/en/Images_Models)

Profile of the applicant

The candidate is expected to have strong knowledge either in **machine learning** or **image processing** and a good experience in both fields. Some prior experience with medical image processing would be appreciated but is not required. **Good programming skills** are also required. The available code is written in Matlab and Python but other languages can be used. We are looking for an enthusiastic and autonomous student with strong motivation and interest in multidisciplinary research (image processing and machine learning in a medical context).

Salary

Around 1700 euros net per month (+ teaching)

Applications

Interested applicants are required to send a cover letter, CV and any other relevant documents (reference letter, recent transcripts of marks,...) to:
carole.lartizien@creatis.insa-lyon.fr

References

[1] C. Lartizien, M. Rogez, E. Niaf, and F. Ricard. Computer aided staging of lymphoma patients with fdg pet/ct imaging based on textural information. *IEEE Journal of Biomedical and Health informatics*, 18(3):946–955, 2014.

[2] E. Niaf, O. Rouviere, F. Mege-Lechevallier, F. Bratan, and C. Lartizien. Computer-aided diagnosis of prostate cancer in the peripheral zone using multiparametric mri. *Physics in Medicine and Biology*, 57(12):3833–3851, 2012.

[3] E. Niaf, R. Flamary, O. Rouviere, C. Lartizien, and S. Canu. Kernel-based learning from both qualitative and quantitative labels: Application to prostate cancer diagnosis based on multiparametric mr imaging. *IEEE Transactions on Image Processing*, 23(3):979–991, 2014.

[4] E. Niaf, C. Lartizien, F. Bratan, L. Roche, M. Rabilloud, F. Mège-Lechevallier, and O. Rouvière. Prostate focal peripheral zone lesions: Characterization at multiparametric mr imaging—influence of a computer-aided diagnosis system. *Radiology*, 271(3):761–69, 2014.

[5] Meriem El Azami, Alexander Hammers, Julien Jung, Nicolas Costes, Romain Bouet, and Carole Lartizien. Detection of lesions underlying intractable epilepsy on t1-weighted mri as an outlier detection problem. *PloS one*, 11(9):e0161498, 2016.

[6] M. El Azami, C. Lartizien, and S. Canu. Converting svdd scores into probability estimates: Application to outlier detection. *Neurocomputing*, 268:64–75, 2017.

[7] Z. Alaverdyan and C. Lartizien. Feature extraction with regularized siamese networks for outlier detection: application to epilepsy lesion detection. In *Conférence sur l'apprentissage automatique (CAp 2017)*, 2017.

[8] J. Lehaire, R. Flamary, O. Rouviere, and C. Lartizien. Computer aided diagnostic for prostate cancer detection and characterization combining learned dictionaries and supervised classification. In *IEEE international conference on image processing (ICIP)*, pages 2251 – 2255.

[9] Rahaf Aljundi, Jérôme Lehaire, Fabrice Prost-Boucle, Olivier Rouvière, and Carole Lartizien. Transfer learning for prostate cancer mapping based on multicentric MR imaging databases. In *MLMMI@ICML*, pages 74–82, 2015.

[10] L. Gautheron, I. Redko, and C. Lartizien. Adaptation de domaine pour la détection automatique du cancer de la prostate en imagerie irm multiparamétrique. In *Colloque GRETSI*, 2017.