

MASTER 2 – ENGINEERING INTERNSHIP

Thematics : Medical AI ; Computer Vision ; Imagomics

Object : Weakly supervised classification of primary liver malignancies using neural networks based on 2D-3D multi-scale data sources

Duration : 6 months, 35 hours/week

Period : Feb - Aug 2023

Compensation : SMIC, 1603,12 €/month

Place : Inria OPIS team, CVN - CentraleSupélec, Gif sur Yvette & AP-HP Beaujon Hospital, Clichy

CONTEXT AND TEAMS

MOSAIC is an innovative digital health project at the interface of multi-scale medical imaging, oncology and artificial intelligence (AI). We are leading this project as an interdisciplinary hospital-university research federation (FHU) including AP-HP (Paris Hospitals) and his Health Data Warehouse (EDS), Inserm, Université Paris Cité, and industrial partners.

MOSAIC aims to identify prognostic and theranostic signatures of digestive tumours (with a specific focus on primary liver cancers and digestive neuro-endocrine tumors) through the development of an innovative multiparametric imaging approach: imagomics (combination of radiology, pathology, and nuclear medicine). Using several artificial intelligence approaches, this project will improve tumour characterization and disease monitoring. It will allow us to reveal correlations between imaging markers and tumour aggressiveness/response to therapy that remain hidden to the human eye until now. We hope to overcome the current need for invasive diagnostic and prognostic modalities by using an AI-understood, image-driven virtual biopsy. MOSAIC is a proof of concept with the ambition to transfer this approach to other cancers and diseases.

The mission will be to be part of our initiative by participating in the development and set-up of these AI algorithms. Our multidisciplinary MOSAIC team includes clinicians, researchers in biology and physics, AI and bioinformaticians offering the fellow a comprehensive training environment.

This research project is more particularly carried out in partnership with the OPIS Inria team. This team is hosted by the CVN (Center for Visual Computing), a joint CentraleSupélec research unit. Its main topic of research is the design of advanced optimization methods for the analysis of complex massive datasets, in the context of inverse problems and machine learning problems, by focusing on biomedical data. These activities are organized around 3 targeted themes: (i) large-scale optimization algorithms, (ii) graph mining, (iii) neural networks.

STUDY

Primary liver cancers are the fourth leading cause of cancer death worldwide, with an increasing incidence in most Western countries. PLC define a heterogeneous group of tumours associated with distinct risk factors, clinical findings, and imaging, histologic and molecular characteristics. Among these tumours, hepatocellular carcinoma (HCC) and intrahepatic cholangiocarcinoma (iCCA) are by far the most common and represent the two ends of the tumor spectrum. In between, combined hepato-cholangiocarcinoma (cHCC-CCA) is recognized as a rare tumor sharing features of HCC and iCCA, the prognosis of which is still unclear.

While imaging and tumor biopsy separately achieve a good performance for the accurate diagnosis of HCC and iCCA, the diagnostic performance of cHCC-CCA based on a single modality is low. Histological

analysis represents the gold standard for the diagnosis of cHCC-CCA, but even from a histological point of view, this diagnosis can be very challenging, in particular in biopsy samples. Recently, we have shown that the combination of imaging and tumour biopsy could improve the diagnosis of cHCC-CCA, highlighting the interest in using multiscale imaging data. We previously developed a weakly supervised classifier based on biopsy samples to distinguish different tumor samples based on their histological composition.

OBJECTIVES AND STUDY QUESTIONS

The goal of this study is to apply various weakly supervised learning methods based on deep convolutional neural networks (CNN) to classify primary liver cancers using multiphasic CT scans.

1. Can we devise an efficient weakly supervised classification approach for learning from multiphasic CT abdomen scans samples?
2. What is the optimal neural network architecture to obtain the « best » classification performance?
3. Which interpretation can be given to the classes extracted by the proposed method?

AI TECHNOLOGY AND DATA ANALYSIS

The methodological approach will be grounded on recent developments in AI concerning weakly supervised learning from features extracted by suitably trained deep neural networks.

FAISABILITY

Beaujon hospital through the clinical, radiology and pathology departments is a referral centers for all materials required for the current project (CT) are available for analysis on a dedicated AP-HP digital online work environment. 120 cases have been selected and regulatory approval already obtained.

PROFIL AND SKILLS :

Hard Skills :

- Languages : Python, PyTorch
- Machine learning
- Image processing and computer vision
- Statistical analysis
- Good practices in ethics and computer security

Soft Skills :

- Identify, analyse, prioritize, and synthesize information
- Written and oral communication...
- Team spirit, interpersonal skills
- Autonomy, rigor, innovation
- Curiosity, ability to adapt

Sensitivity :

- Interest in the field of medical research and data sciences
- Adherence to the values and missions of public service

CONTACT :

- Mondet Kevin, FHU Mosaic project manager, kevin.mondet@aphp.fr
- Gregory Jules, FHU Mosaic clinical assistant, jules.gregory@aphp.fr
- P^r Paradis Valérie, FHU Mosaic project leader, valerie.paradis@aphp.fr
- P^r Pesquet Jean-Christophe, Inria OPIS leader, jean-christophe.pesquet@centralesupelec.fr
- Dr Nora Ouzir, Assistant Professor, nora.ouzir@centralesupelec.fr