

Master 2 internship: Machine learning based classification for identifying metastatic tissue in histopathologic images

Required skills

Training level: Master 2 or Engineer fifth year
Good knowledge on machine learning techniques and image processing
Strong capability of coding using Python or C/C++ or Matlab is appreciated
Good knowledge on web development is a plus

Localization ISEN-Lille, 41 boulevard Vauban 59800 Lille, France/ IEMN CNRS laboratory/ digital systems and life sciences team.

Duration 5 to 6 months from 1st March 2019. With possibility to pursue a Ph.D thesis.

Application send your CV and your cover letter before February 25, 2019 to halim.benhables@yncrea.fr and feryal.windal@yncrea.fr

Context and goals

Developing histopathologic image analysis algorithms represents a real scientific challenge. This is mainly due to the lack of representation structure in these images. More precisely, these images are registered in the form of a pixel matrix in which no information is provided on the nature of the tissue and its microenvironment. Additionally, the variation of environmental conditions during the acquisition process of these images using microscopes will generate a noise that may affect the analysis results.

One of the promising directions to face the previously mentioned issues is the **integration of artificial intelligence in the developed algorithms**. This can be done using learning techniques to describe and characterize the collected data. In the histopathologic image analysis field, exploiting this type of techniques has become an obvious choice for boosting the performance of analysis algorithms [1].

More generally, in the medical image analysis field, deep learning techniques which are mainly based on a convolutional neural network (CNN) architecture have shown high performance in multiple difficult tasks including segmentation, classification and retrieval [2,3].

In this context, the main goal of the internship is to develop two machine learning based classification algorithms (a handcrafted method and a deep learning method) to identify metastatic cancer in a large histopathological image dataset. The dataset is provided within the frame of a Kaggle competition¹ for the machine learning community. The results obtained by the two algorithms will be submitted to the competition.

The funding of this internship is covered by a European project (Interreg 2 Seas) for which **we also have obtained a funding for a Ph.D thesis** that will start after the internship. The successful candidate will have the opportunity to pursue, if he/she wants, a Ph.D thesis. He/she will be asked during the internship to develop a website for communication on the project.

References

- [1] D. Komura and S. Ishikawa, "Machine Learning Methods for Histopathological Image Analysis," *Computational and Structural Biotechnology Journal*. 2018.
- [2] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. W. M. van der Laak, B. van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Medical Image Analysis*. 2017.
- [3] H. Benhables, K. Hammoudi, F. Windal, M. Melkemi, and A. Cabani, "A Transfer Learning Exploited for Indexing Protein Structures from 3D Point Clouds," in *Processing and Analysis of Biomedical Information (in conjunction with MICCAI 2018)*, 2018.

¹ <https://www.kaggle.com/c/histopathologic-cancer-detection/data>