

# Postdoc position: Deep learning for colon cancer histopathological images analysis

AiCOLO project - INSERM - Plan Cancer<sup>1,2,3</sup>

<sup>1</sup>ICube laboratory, Data science and Knowledge, Illkirch, France

<sup>2</sup>IRIMAS institute, Modeling and Data Science, Mulhouse, France

<sup>3</sup>Centre Georges Francois Leclerc, Cancer and adaptive immune response, Dijon, France

## ABSTRACT

Digital Pathology is currently regarded as one of the most promising avenues of diagnostic medicine. With the recent advent of Whole-Slide Imaging (WSI), the field of digital pathology produces daily a massive amount of images with related metadata (e.g. patient information, diagnosis, treatment). In the context of colon cancer, such images could be used for both diagnosis and to find some prognostic biomarkers. For example immune infiltrate are associated with better prognosis while high stromal content or tumor budding or poor differentiation status are associated with poorer outcome. Making these quantitative analysis is time consuming for pathologist and frequently lack of reproducibility. To assist experts, automatic analysis of whole-slide images (WSI) has been recently studied to predict survival outcomes or making tumor classification. The **AiCOLO project** aims to contribute to the development of new artificial intelligence techniques trained on a large cohort of clinical annotated colon cancer patients with a twofold objective. On the one hand, we will develop an innovative prognostic aid tool to automatically classify tissues constituting WSI pathological slides and to enumerate the various known prognostic markers like TILS immune infiltration, stromal content or eosinophil count in the different areas. The method will be trained to find image patterns in tumor tissue related to patients' outcome. On the other hand, we will also propose a resolutely new approach to predict RAS and BRAF genetic status from WSI. We aim to determine if artificial intelligence could detect patterns associated with such genetic features and could outperform clinical or immune infiltrate variables. The idea is to study the activation layers of a deep network trained to classify the patients in order to extract information to explain its decision.

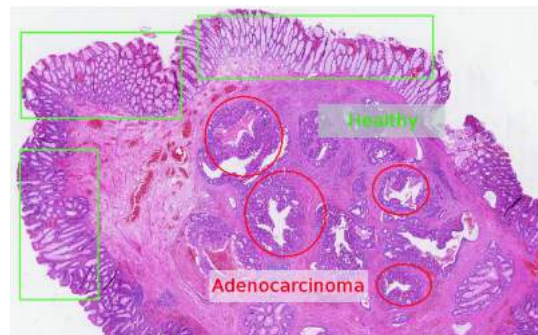
## Job description

The Engineering science, computer science and imaging laboratory (ICube, Strasbourg, France), associated with the Institute of Research in Computer Science, Mathematics, Automatics and Signal processing (IRIMAS, Mulhouse, France), opens a **postdoctoral position** for a computer scientist, in the field of artificial intelligence and histopathological whole slide images analysis, with a duration of **36 months (2019/12/01 - 2022/11/30)**.

In the context of the AiCOLO project described above, the appointee will work in close collaboration with the three partners of the project to develop the methodology for WSI analysis, spatial patterns extraction and the machine learning approach to classify automatically the genetic mutation from HES images.

More specifically, the objective is to develop a complete methodology enabling to assign a label to each image region. This problem will be tackled by two complementary approaches: a pixel-based method, in order to obtain a cartography of regions of interest for the studied pathologies (inflammatory zones or tumors area) and an object-based approach enabling to compute a list of biological objects with their contour, localization and attributes. The main workflow must be automatic in order to operate in an unsupervised manner, which constitutes a crucial and challenging aspect of this task. The WSI analysis system will rely both on previous work<sup>2,3</sup> and on novel techniques based on levellines decomposition of an image<sup>4</sup> and connected operators from mathematical morphology<sup>5</sup> based on hierarchical representations. These latter methods enable to analyze an image at the level of the connected components of its threshold sets (or other increasing transformation). These methods are relevant in this context since (i) they enable to process an image in a contrast invariant way; (ii) they prevent to alter the contour of objects; (iii) they permit to compute object based attributes.

The work will also consist on the identification of genetic prognostic/predictive markers on HES slide. BRAF and RAS



Extract of WSI of colon adenocarcinoma<sup>1</sup>

mutational status are mandatory required for the treatment of metastatic colon cancer. These markers are both prognostic and predictive of response to anti EGFR therapies. Recent data in lung cancer make the demonstration that gene mutations could affect the pattern of tumor cells on a lung cancer whole-slide image<sup>6</sup>. Training network using the presence or absence of mutated genes as a label revealed that there are certain genes whose mutational status can be predicted from image data alone: EGFR, STK11, FAT1, SETBP1, KRAS, and TP53 with good accuracy. The ability to quickly and inexpensively predict both the type of cancer and the gene mutations from histopathology images could be beneficial to the treatment of patients with cancer given the importance and impact of mutation in patient care. We propose to perform similar work on colon cancer and try to isolate feature detected by neural network to detect these mutation.

This research will be part of a collaboration between AI researchers of ICube and IRIMAS, and pathologists and biostatisticians of the Centre Georges Francois Leclerc, Cancer and adaptive immune response team (Dijon). Regular exchanges will take place either by videoconference or during stays in their structures, to benefit from their expertise. Within the ICube laboratory, the postdoctoral researcher will be integrated into the Data and Knowledge Science team under the supervision of Cédric Wemmert and Benoît Naegel, and co-supervised by Germain Forestier and Jonathan Weber (IRIMAS).

## Job summary

- Postdoc position in artificial intelligence and histopathological image analysis in the ICube laboratory (Strasbourg, France) and IRIMAS Institute (Mulhouse, France).
- The candidate should hold a PhD in computer science (preferably in computer vision or machine learning) and have excellent knowledge of the English language.
- Demonstrated experience in Python programming and Keras/Tensorflow libraries will be considered as an advantage.
- Salary: around €3000 (monthly gross) depending on the experience of the candidate.

## Contact

Please contact Pr. Cédric Wemmert by email ([wemmert@unistra.fr](mailto:wemmert@unistra.fr)) for any further question or application.

## References

1. Ponzio, F., Macii, E., Ficarra, E. & Di Cataldo, S. Going deeper into colorectal cancer histopathology. In *International Joint Conference on Biomedical Engineering Systems and Technologies*, 114–131 (Springer, 2018).
2. Apou, G., Naegel, B., Forestier, G., Feuerhake, F. & Wemmert, C. Efficient region-based classification for whole slide images. In *International Joint Conference on Computer Vision, Imaging and Computer Graphics*, 239–256 (Springer, 2014).
3. Krüger, J. M. *et al.* Combat or surveillance? evaluation of the heterogeneous inflammatory breast cancer microenvironment. *The J. pathology* **229**, 569–578 (2013).
4. Caselles, V. & Monasse, P. Geometric description of topographic maps and applications to image processing. *Lect. Notes Math* (1984).
5. Xu, Y., Géraud, T. & Najman, L. Connected filtering on tree-based shape-spaces. *IEEE transactions on pattern analysis machine intelligence* **38**, 1126–1140 (2015).
6. Coudray, N. *et al.* Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning. *Nat. medicine* **24**, 1559 (2018).

