

SUJET DE STAGE M2 : Etude par *deep learning* de composants tissulaires imagés en 3D *in vitro* par microscopie à fluorescence

Responsable : Thierry Delzescaux –PhD, Directeur de recherche CEA, HDR- (thierry.delzescaux@cea.fr)

Equipe d'accueil :

CEA-MIRCent, LMN

Laboratoire des Maladies Neurodégénératives

Equipe traitement de l'image

18, route du Panorama - BP N° 6

92265 - Fontenay aux Roses Cedex

<https://jacob.cea.fr/drif/francoisjacob/Pages/Departements/MIRCent/themes/alzheimer-vieillesse-cerebral-modelisation.aspx>

Durée de stage : 6 mois

Description du sujet de stage : L'objectif de ce stage ingénieur / Master 2 sera d'étudier les méthodes de segmentation de plusieurs types cellulaires (neurones, fibroblastes, cellules de Schwann) à partir de données 3D de microscopie optique de fluorescence dans le cadre de l'étude *in vitro* de tissus et de tumeurs (fibroblastomes, tumeurs NF1). L'étude des cellules et de leurs principales propriétés géométriques permettra de caractériser les modèles tumoraux ainsi générés et de mieux décrire les interactions entre les différents types cellulaires les constituant.

Une étude approfondie des méthodes de l'état de l'art de segmentation sémantique et d'instance cellulaire sera réalisée dans un premier temps en se concentrant sur les techniques de *Deep Learning* (*Cellpose* [1], architectures *U-Net* [2], etc.). Les données du projet seront mises à disposition par le partenaire de l'Institut Mondor de Recherche Biomédicale expert dans ce domaine. Une base de données 2D-3D de référence sera constituée et les solutions logicielles identifiées au cours de la recherche bibliographique seront testées afin d'évaluer leur capacité à mener à bien la première étape: 1) de segmentation sémantique (segmentation des pixels en un nombre limité de classes –neurones, cellules de Schwann, etc.-) et 2) de segmentation d'instances (individualisation des composantes cellulaires des tissus étudiés).

Le candidat.e aura accès à des ressources computationnelles adaptées au projet (station de travail, moyens de calcul intensif CPU/GPU voire cluster de calcul). Les verrous méthodologiques potentiels seront identifiés ainsi que les pistes envisagées pour les surmonter.

La seconde partie du stage aura pour objectif d'explorer les paramètres géométriques cellulaires qui pourraient être utilisés pour décrire au mieux les tissus produits *in vitro* et tout particulièrement pour détecter des modifications induites lors du développement de tumeurs (nombre de cellules, taille, distribution, distance inter-classes cellulaires, etc.). Le développement de tels outils ouvrirait de plus des perspectives d'application dans un cadre de développement et de test d'efficacité de nouvelles thérapies.

Compétences requises : Au cours de ce stage, le candidat sera amené à interagir principalement avec l'équipe de traitement de l'image de MIRCent (informaticiens, méthodologistes en traitement de l'image, etc.) et les neurobiologistes de la plateforme d'histologie. Il est demandé d'avoir une très bonne connaissance des environnements **Linux** et **Windows**, de maîtriser la programmation dans les **langages C, C++ et Python**, d'avoir des connaissances dans le domaine de la parallélisation et du calcul haute performance (CPU/GPU). Des connaissances générales des techniques de traitement de l'image (segmentation, *machine learning* – *deep learning*) et d'un outil de **gestion de sources (github)** serait un plus. Il est également nécessaire de savoir utiliser les suites bureautiques standards (Open Office, Office). Le stagiaire bénéficiera pour réaliser ces développements des savoirs faire, de l'encadrement des équipes du CEA et de l'infrastructure informatique existante (serveurs de calculs internes, plateforme logicielle). De bonnes capacités d'**adaptation** à des environnements multidisciplinaires, de **coordination** et de **travail en équipe** sont également requises pour ce projet.

Formation requise : Ecole ingénieur, Master de recherche dans les sciences de l'information, traitement de l'image et du signal (salaire calculé en fonction de la formation du/de la candidat.e).

Possibilité de poursuivre en CDD sur des thématiques intelligence artificielle / calcul intensif ou en thèse en fonction de l'obtention de demandes de financements qui sont en cours.

Bibliographie : [1] Stringer C. *et al.*, Nat Methods, 2021. [2] Falk T. *et al.*, Nat Methods. 2019