

Ingénieur·e junior en apprentissage auto-supervisé pour l'analyse d'images biomédicales

Contrat : CDD 12 à 24 mois

Lieu : Toulouse – entre le B612 (ANITI) et le laboratoire RESTORE

Prise de poste : Septembre / octobre 2025

Niveau : Bac +5 (école d'ingénieur ou master en IA, informatique, traitement d'image ou biomédical)

Langue : travail en français, production en anglais

Environnement : recherche académique et valorisation (INSERM, ANITI)

Objectif du poste

Vous participerez à un projet de recherche appliquée visant à concevoir une pipeline d'apprentissage auto-supervisé (self-supervised learning) pour reconnaître automatiquement le type de noyaux cellulaires à partir d'images de microscopie. Ce travail s'inscrit dans un projet interdisciplinaire à l'interface entre l'intelligence artificielle, la biologie et le traitement d'image biomédicale, dans le cadre d'une collaboration entre la chaire ANITI CALM et l'équipe STROMAGICS du laboratoire RESTORE.

Missions principales

- Construction de la base de données
 - Rechercher, évaluer et agréger des bases de données existantes pertinentes (images de noyaux cellulaires, conditions expérimentales, etc.)
 - Étudier les caractéristiques techniques et les licences associées
 - Définir les modalités de publication open source (choix de licence, structuration, métadonnées)
 - Mettre en œuvre des outils de segmentation automatique (e.g. Cellpose, SAM) pour extraire les noyaux
 - En collaboration avec l'équipe de stromagic, créer une base de données de validation pour la tâche et pour valider de nouvelles techniques de multi-marquage
 - Génération de nouvelles données à l'aide d'algorithmes de diffusions
 - Créer une bibliothèque pytorch d'accès à ces données
- Développement de la pipeline d'apprentissage
 - Implémenter une pipeline d'apprentissage self-supervised avec PyTorch en collaboration avec un doctorant de la chaire

- Intégrer et tester des architectures modernes (ViT, SimCLR, etc.)
- Gérer l'apprentissage sur des données de grande taille
- Participer à la mise en place d'un pipeline reproductible (code, logs, reproductibilité)
- Recherche, valorisation et rédaction
 - Participer aux réunions de recherche (chaire ANITI, RESTORE)
 - Contribuer à la rédaction d'articles scientifiques
 - Participer à la valorisation du projet auprès d'acteurs comme Inserm Transfert ou ANITI

Compétences requises / à renforcer

- Connaissances en deep learning, notamment en vision par ordinateur
- Expérience en Python, PyTorch, NumPy
- Notions ou intérêt pour le traitement d'images biomédicales
- Goût pour la recherche appliquée, le travail reproductible et la valorisation open science
- Bonne capacité d'adaptation dans un environnement interdisciplinaire (biologie + IA)
- Rigueur, autonomie, et capacité à documenter son travail

Ce que le poste vous apportera

- Une expérience complète dans un projet de recherche ambitieux mêlant IA, biomédical et open science
- Des compétences recherchées dans l'industrie : deep learning, traitement d'image, gestion de données à grande échelle
- Une implication concrète dans un projet open source (base de données, code, publication)
- Une excellente base pour candidater à un doctorat ou un poste en R&D IA / data science