



Développement d'approches d'intelligence artificielle pour la segmentation et l'analyse comparative d'embryons en microscopie 3D+temps

Contexte

L'intelligence artificielle (IA) et plus particulièrement l'**apprentissage profond** transforment l'analyse d'images, notamment biomédicales. Mais les méthodes actuelles peinent encore à traiter efficacement des données volumineuses et complexes comme celles issues des suivis temporels et en 3D de processus biologiques par microscopie (3D+temps) qui se sont considérablement améliorés ces dernières années.

Les microscopes de nouvelle génération sont notamment capables de capturer en temps réel la dynamique du développement d'embryons entiers en 3D. Ces données ont une résolution suffisante pour suivre le développement embryonnaire cellule par cellule, avec une précision sans précédent. Toutefois, la richesse de ces ensembles de données pose un défi majeur : l'imagerie d'un seul embryon, contenant des milliers de cellules en interaction dynamique, peut générer plusieurs téraoctets de données.

L'apprentissage profond devient fondamental pour analyser ces données et notamment pour :

- **Segmenter automatiquement les cellules** (distinguer et délimiter chaque cellule dans le volume),
- **Suivre leur trajectoire au cours du temps** ainsi que celle de leurs cellules filles après division.
- **Comparer objectivement le développement de différents embryons** pour comprendre sa dynamique et la structure de la variabilité interindividuelle du développement, à différentes échelles (globale, tissus, cellules).

C'est à cette frontière entre intelligence artificielle et biologie du développement que se place ce projet de thèse en informatique. Néanmoins des connaissances avancées en biologie du développement ne sont pas requises.

Objectifs

La doctorante ou le doctorant développera de nouvelles méthodes d'IA permettant :

1. **Segmentation et suivi cellulaire en 3D+temps** : Développer des architectures end-to-end intégrant à la fois la segmentation et le suivi cellulaire, afin d'obtenir des reconstructions capables de capturer avec précision la variabilité individuelle
2. **Représentation morphologique non supervisée** : à partir des données d'imagerie et des reconstructions spatio-temporelles de la morphologie embryonnaire, développer un espace latent compact permettant de comparer objectivement les embryons, indépendamment de leur taille ou de leur orientation, et de construire ainsi des trajectoires développementales typiques ou atypiques.
3. **Intégration dans un cadre unifié** : produire une méthodologie complète permettant l'analyse comparative et quantitative de données embryonnaires complexes.
4. **Diffusion des résultats** : proposer un outil logiciel open-source accessible à la communauté scientifique internationale.

Méthodologie

La thèse explorera deux axes principaux :

- **Axe 1 : Segmentation et suivi spatio-temporel** : Développement de réseaux convolutionnels épars (*sparse CNN*) combinés à des mécanismes d'attention et à des architectures hybrides CNN-Transformer. Cette approche permettra de modéliser explicitement les dynamiques cellulaires (divisions, mouvements collectifs) et de renforcer la cohérence temporelle des reconstructions.
- **Axe 2 : Apprentissage non supervisé et représentations latentes** : Utilisation d'autoencodeurs 3D et de représentations implicites (Neural Radiance Fields, DeepSDF) pour modéliser les formes embryonnaires. L'objectif est de filtrer les variations non pertinentes (bruit, orientation, artefacts) pour ne conserver que les structures morphologiques essentielles.

Les modèles seront entraînés et validés sur une **base unique d'images d'embryons d'ascidie déjà acquises**. Cet organisme modèle de groupe frère des vertébrés a été choisi car son embryogenèse est rapide, très stéréotypée.

Résultats attendus

- Une **méthodologie robuste** et objective pour comparer des morphologies embryonnaires complexes et leur dynamique.
- Un **réseau neuronal innovant** dédié à la segmentation et au suivi en 3D+temps.
- Un **logiciel open-source** pour l'analyse d'embryons, mis à disposition de la communauté.
- Des **publications internationales** en IA, imagerie biomédicale et biologie du développement.

Ce projet ouvrira un changement de paradigme pour l'étude automatisée et quantitative de l'embryogenèse, avec des retombées dans d'autres systèmes biologiques.

Encadrement et environnement

La thèse sera co-encadrée par :

- **Emmanuel Faure (LIRMM, CNRS Montpellier)** – Informaticien, spécialiste de l'IA et de l'analyse d'images 3D+temps. emmanuel.faure@lirmm.fr
- **Patrick Lemaire (CRBM, CNRS Montpellier)** – Biologiste, expert en embryologie et morphogénèse. patrick.lemaire@crbm.cnrs.fr

Candidature:

https://adum.fr/as/ed/voirproposition.pl?site=adumR&matricule_prop=67285

Le doctorant bénéficiera :

- d'un poste de travail avec GPU, d'un NAS sécurisé pour le stockage et d'un accès aux ressources nationales de calcul haute performance (supercalculateur Jean-Zay, GENCI),
- d'un environnement pluridisciplinaire dynamique au croisement de la biologie et de l'IA,
- d'une mission internationale par an (conférence, école thématique ou workshop).

Profil recherché

- Master 2 ou diplôme d'ingénieur en informatique, mathématiques appliquées, bio-informatique, imagerie biomédicale ou équivalent.
- Compétences en deep learning (PyTorch, TensorFlow), programmation (Python requis) et analyse d'images.
- Intérêt marqué pour l'interdisciplinarité et bonne maîtrise de l'anglais scientifique.

Ce projet s'adresse à des étudiant·e·s passionné·e·s par les **méthodes avancées d'intelligence artificielle** et curieux·ses de contribuer à une meilleure compréhension des mécanismes fondamentaux du **développement embryonnaire**.

Development of AI approaches for segmentation and comparative analysis of embryos in 3D+time microscopy

Context

Artificial intelligence (AI), and in particular deep learning, is transforming image analysis, especially in the biomedical field. However, current methods still struggle to efficiently process large and complex datasets such as 3D+time microscopy recordings of biological processes, which have dramatically improved in recent years.

Next-generation microscopes can now capture in real time the full 3D dynamics of whole embryos. These datasets have sufficient resolution to follow embryonic development cell by cell with unprecedented precision. Yet their richness also poses a major challenge: imaging a single embryo, containing thousands of dynamically interacting cells, can generate several terabytes of data.

Deep learning is becoming fundamental to analyzing these datasets, particularly for:

- **Automatically segmenting cells**, i.e. distinguishing and delineating each individual cell within the volume.
- **Tracking cell trajectories over time**, including daughter cells after division.
- **Objectively comparing embryonic development across individuals**, in order to understand developmental dynamics and the structure of inter-individual variability, at multiple scales (global, tissue, cellular).

This PhD project is positioned precisely at the frontier between artificial intelligence and developmental biology. Advanced knowledge in developmental biology is not required.

Objectives

The PhD candidate will develop new AI methods to:

1. **Segmentation and cell tracking in 3D+time**: Design end-to-end architectures integrating both segmentation and tracking, in order to achieve reconstructions that accurately capture individual variability.
2. **Unsupervised morphological representation**: From imaging data and spatio-temporal reconstructions of embryonic morphology, develop a compact latent space that enables objective comparison of embryos, independent of their size or orientation, and detection of typical or atypical developmental trajectories.
3. **Integration into a unified framework**: Produce a comprehensive methodology for the comparative and quantitative analysis of complex embryonic data.

4. **Dissemination of results:** Deliver an open-source software tool accessible to the international scientific community.

Methodology

The thesis will explore two complementary axes:

- **Axis 1: Spatio-temporal segmentation and tracking:** Development of sparse convolutional neural networks (sparse CNNs) combined with attention mechanisms and hybrid CNN–Transformer architectures. This approach will explicitly model cellular dynamics (divisions, collective movements) and reinforce the temporal coherence of reconstructions.
- **Axis 2: Unsupervised learning and latent representations:** Use of 3D autoencoders and implicit representations (Neural Radiance Fields, DeepSDF) to model embryonic forms. The goal is to filter out irrelevant variations (noise, orientation, artifacts) while preserving the essential morphological structures.

The models will be trained and validated on a unique dataset of ascidian embryos already acquired. Ascidians, a sister group of vertebrates, were chosen as a model organism because their embryogenesis is rapid and highly stereotyped.

Expected results

- A robust and objective methodology for comparing complex embryonic morphologies and their dynamics.
- An innovative neural network dedicated to segmentation and tracking in 3D+time.
- An open-source software tool for embryo analysis, made available to the community.
- International publications in AI, biomedical imaging, and developmental biology.

This project will represent a paradigm shift in the automated, quantitative study of embryogenesis, with potential applications to other biological systems.

Supervision and environment

The PhD will be co-supervised by:

- **Emmanuel Faure (LIRMM, CNRS Montpellier)** – Computer scientist, expert in AI and 3D+time image analysis. emmanuel.faure@lirmm.fr
- **Patrick Lemaire (CRBM, CNRS Montpellier)** – Biologist, expert in embryology and morphogenesis. patrick.lemaire@crbm.cnrs.fr

Application:

https://adum.fr/as/ed/voirproposition.pl?site=adumR&matricule_prop=67285

The PhD candidate will benefit from:

- a workstation with GPU, secure NAS storage, and access to national high-performance computing resources (Jean-Zay supercomputer, GENCI),
- a dynamic interdisciplinary environment bridging biology and AI,
- at least one **international mission per year** (conference, summer school, or workshop).

Candidate profile

- Master's degree (MSc) or engineering diploma in computer science, applied mathematics, bioinformatics, biomedical imaging, or equivalent.
- Skills in **deep learning** (PyTorch, TensorFlow), programming (Python required), and image analysis.
- Strong interest in interdisciplinary research and good command of scientific English.

This project is aimed at students passionate about advanced AI methods and eager to contribute to a better understanding of the fundamental mechanisms of embryonic development.