Project 05 — Forecasting Morphogenesis

From Brightfield Time-Lapse Images to Predictive Models of Zebrafish Embryogenesis

Team 5: Alejandro C. Angel R.

Proposal

1 Scientific Context

Embryonic development is a process where a single fertilized cell gradually transforms into a complete organism. In vertebrates, this involves the self-organization of thousands of cells that move, divide, and change their shape to form tissues and organs. The zebrafish has become one of the most important model organisms to study these processes because its embryos are transparent. This transparency allows scientists to directly observe how the embryo changes over time using simple brightfield microscopy.

During early zebrafish development, the embryo passes through characteristic stages: blastula, gastrula, and segmentation. Each can be identified by clear changes in the shape and structure of the embryo [1]. At these stages, several fundamental biological processes occur, such as:

- **Epiboly:** This is one of the first large-scale movements in the zebrafish embryo. During epiboly, the outer cell layers gradually spread over and around the yolk. It marks the beginning of visible tissue organization.
- **Convergence:** After epiboly, many cells start migrating toward the central midline of the embryo. This movement, called convergence, brings cells closer together along the sides, narrowing the embryo's shape.
- **Extension:** While convergence narrows the body, extension lengthens it. Cells move and intercalate along the embryo's head–to–tail axis, stretching the developing body. This process transforms the embryo from a roughly circular shape into an elongated form, establishing the basic layout of the future fish body plan.

These movements are controlled by molecular signaling pathways, including Nodal and BMP, which coordinate how cells communicate and organize themselves. When these signals are disrupted, embryos show visible differences in their morphology, such as changes in body shape or incomplete tissue formation.

Recent advances in computer vision and deep learning have made it possible to automatically recognize such developmental patterns from images. For example, the EmbryoNet project demonstrated that neural networks can classify zebrafish embryos according to which signaling pathway is disrupted [2]. However, most of these studies focus on identifying what has already happened. In contrast, our goal is to look one step ahead: to predict how the embryo's shape will evolve in the near future. This shift from recognition to forecasting opens the possibility of quantitatively studying the dynamics of morphogenesis instead of only static snapshots.

2 Problem Statement and Hypotheses

The central idea of this project is to learn how to predict future stages of zebrafish development from earlier observations.

Task. We start from brightfield microscopy videos that show the embryo from early to mid-development. Each video contains 150 frames, and for each embryo, we also have a mask showing its outline and some features. The task is to use this information to predict the next stage of the embryo, either as an image or through the features.

Inputs. The input data can be considered as short sequences of (a) brightfield frames, (b) segmentation masks, and (c) features of the embryo.

Outputs. The model should produce either (1) an estimate of how the next image will look, or (2) a prediction of the next values of the features.

Hypotheses. We expect that temporal machine learning models will capture how embryos move and deform over time. In particular:

- **(H1)** We expect that machine learning models designed to learn from sequences of images (such as ConvLSTMs or Transformers) will be able to better understand how the embryo's shape changes over time.
- **(H2)** We also expect that the way the model predicts future shapes will depend on the type of embryo. Embryos with mutations in the Nodal or BMP signaling pathways develop differently from normal embryos, showing changes in how their cells move and organize. Therefore, the prediction errors and learned patterns may reveal these biological differences.

3 Work Plan and Approach

To achieve our goal of forecasting embryo morphology, we will work in three simple stages during the first phase of the project. The goal is not yet to obtain perfect predictions, but to build understanding and prepare a clean, reproducible workflow that can later be improved.

Stage 1: Understanding the data. We will begin by exploring the dataset of brightfield time-lapse videos. Each embryo will be visualized as both images and numerical features such as area, body length, and movement of its center. By plotting how these quantities change over time, we hope to recognize basic developmental patterns across normal, Nodal-mutant, and BMP-mutant embryos.

Stage 2: Designing a first temporal model. After establishing the baseline, we will build a first prototype of a temporal deep learning model that can learn motion and deformation over time. We plan to start with a small ConvLSTM model, using short image sequences to predict the next frame or the next set of numerical features. The model will be trained on a subset of embryos and tested on others to ensure that it can generalize beyond individual examples.

Evaluation and interpretation. To check if the model learns meaningful patterns, we will compare predicted and real images using simple metrics such as Mean Absolute Error (MAE) and Structural Similarity Index (SSIM). We will also visualize how well it reproduces biological features like area or axis length over time. Finally, we will compare results between genotypes to see whether the model detects the characteristic differences of Nodal and BMP mutants.

Expected outcome. By the end of this first phase, we expect to have a small but working framework that can load the data, explore it, and produce the first short-term forecasts. This foundation will serve as the starting point for more advanced modeling and interpretation in the following weeks.

3.1 First Steps

- Crate a data loader.
- Plot time evolution.
- Overview of ConvLSTM model.
- Implementation of a basic ConvLSTM model on data sets.

References

[1] Charles B. Kimmel, William W. Ballard, Seth R. Kimmel, Bonnie Ullmann, and Thomas F. Schilling. Stages of embryonic development of the zebrafish. *Developmental Dynamics*, 203(3): 253–310, 1995. doi: 10.1002/aja.1002030302.

[2] Embryonet: using deep learning to link embryonic phenotypes to signaling pathways. *Nature Methods*, 20(8):1031–1043, 2023. doi: 10.1038/s41592-023-01873-4. Authors as listed in the published article.