

# Forecasting Morphogenesis in Latent and Image Space

Alejandro Cobo  
Angel David Rodriguez  
Prof: Andres Morales-Navarrete, PhD

Yachay Tech University

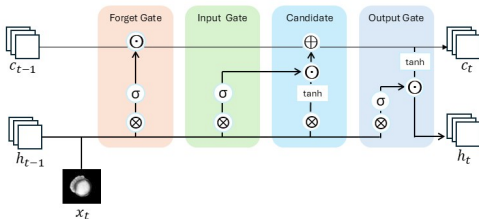
November 26, 2025

# 1.- Forecasting Morphogenesis: ConvLSTM Model & Data

## The Model:

- **Architecture:** ConvLSTM (Convolutional Long Short-Term Memory).
- **Complexity:** 70,000 parameters.

ConvLSTM Layer 1:



## The Data:

- **Dataset:** 6 Embryos Types.
- **Controls:** Control1, Control2.
- **Mutants:** Mutant1, Mutant2, Mutant3, Mutant4.
- **Temporal Dimension:** 450 Frames per embryo.
- **Range:**  $t = 120$  min to  $t = 960$  min (16 hours of development).

## 2.- Experimental Setup

### 1. Data Split (Generalization Focus)

- **Train/Validation (90/10):** Trained on Control1, Mutant1, Mutant3.
- **Test (100%):** Evaluated on Control2, Mutant2, Mutant4 (Completely unseen subjects).

### 2. Loss Function

- **SSIM (Structural Similarity Index):** Prioritizes structural shape.

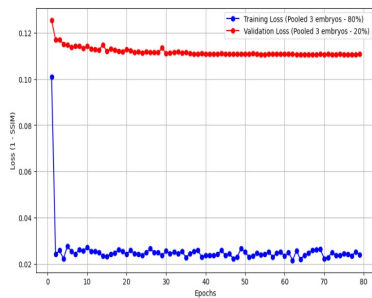
### 3. Robustness: Data Augmentation

Applied "On-the-fly" to the training set to teach spatial invariance.

- **SSIM (Structural Similarity Index):** Prioritizes structural shape.
- **Transformations (Lock-Step):** Flips: Horizontal & Vertical.  
Rotation ( $-180^\circ$  to  $180^\circ$ )  
Zoom ( $\pm 5\%$ )  
Shift ( $\pm 5\%$ ) (10 pixels approx)

# 3.- Model Performance on Unseen Data

## Learning Curves



## Quantitative Metrics:

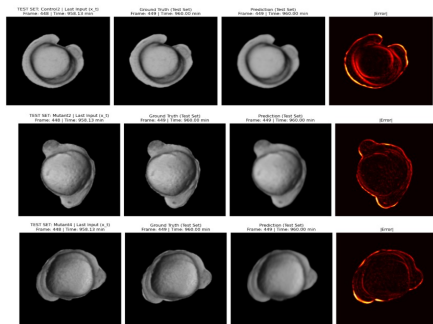
Avg. Test Loss for Control2: 0.07159310

Avg. Test Loss for Mutant2: 0.09640111

Avg. Test Loss for Mutant4: 0.11105474

Observation: The model successfully predicts future morphology on embryos it has never seen before.

## Blind Test Predictions



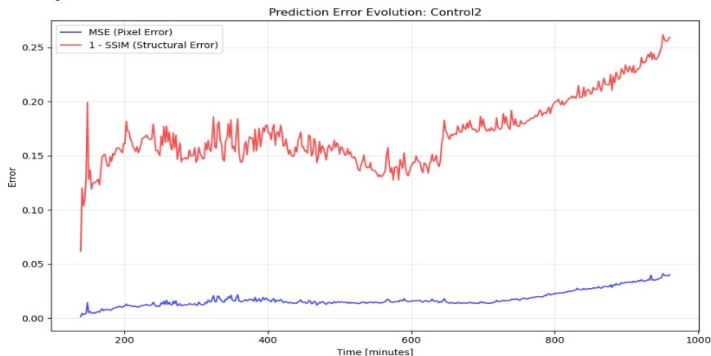
## 4.- Long-term Recursive Forecasting

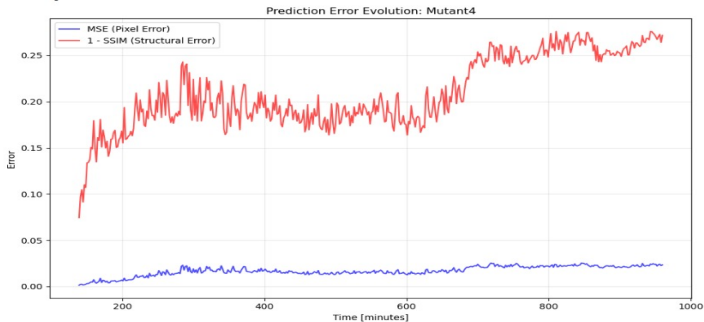
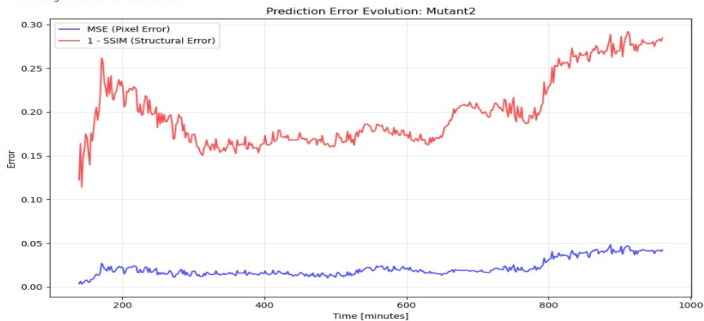
### The Challenge

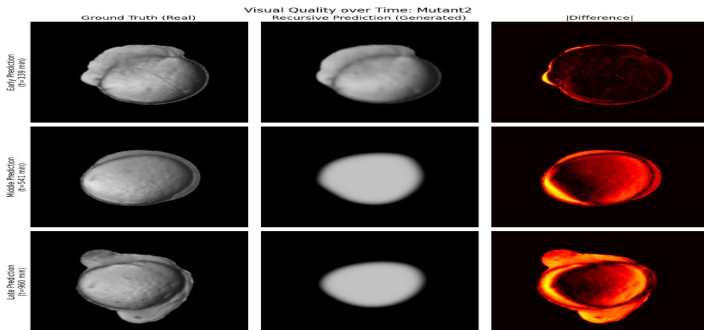
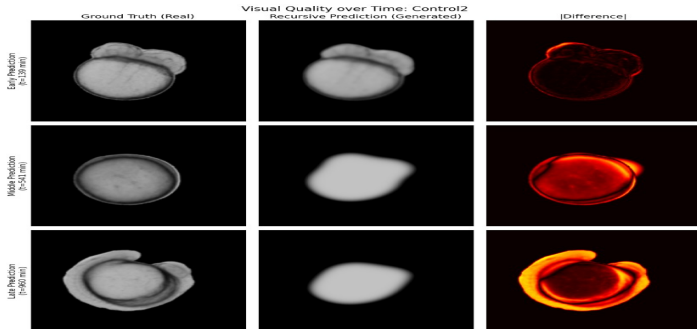
**One-Step Prediction:** Uses real ground truth to predict  $t + 1$  (Easy).

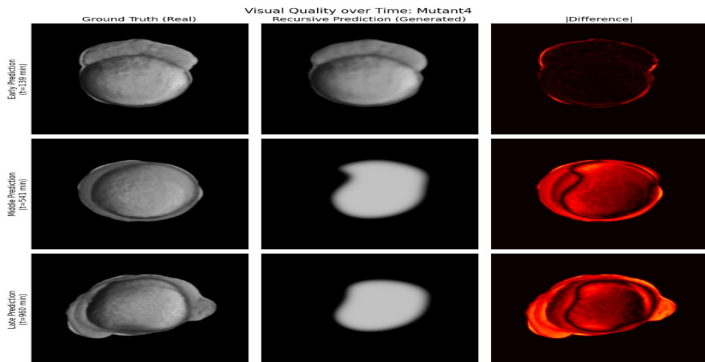
**Recursive Prediction:** Uses its own previous predictions to forecast 400+ frames into the future (Hard).

**Results:**









## Insights

Decay in SSIM reflects the "blurring" effect common in long-term recursive loops.

- Event Detection: Spikes in the error curve correlate with complex biological events.



## 5.- Future Work & Next Steps

### 1. Scaling the Dataset

- **Objective:** Incorporate more biological samples into the training set.
- **Hypothesis:** Increasing the number of training subjects will generalize better to new geometries and significantly improve sharpness in long-term predictions.

### 2. Scheduled Sampling

- **Objective:** Train the model using its own predictions.
- **Methodology:** Instead of always feeding the Ground Truth, we will gradually introduce the model's own previous outputs as inputs during training.