A Deep Learning Approach for Prediction of IVF Implantation Outcome from Day 3 and Day 5 Time-lapse Human Embryo Image Sequences

Mehryar Abbasi[1], Parvaneh Saeedi[1], Jason Au[2], and Jon Havelock[2]
[1]School of Engineering Science, Simon Fraser University, Canada
[2]Pacific Centre for Reproductive Medicine, Canada

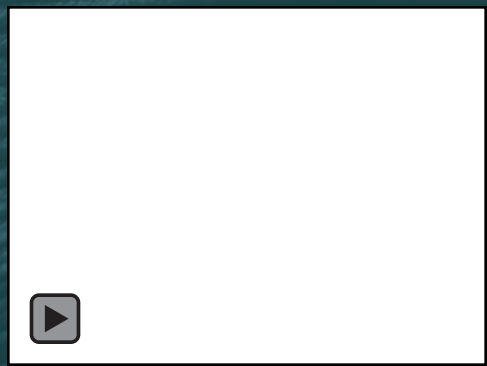
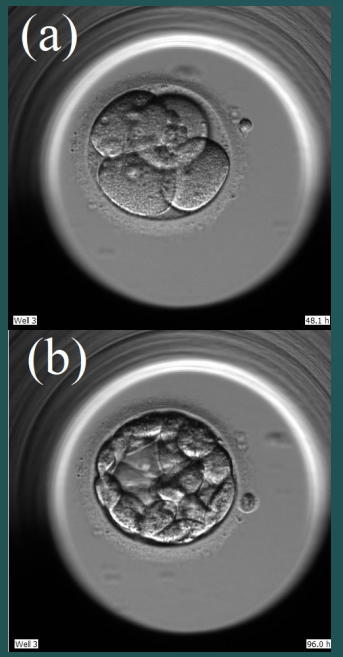


Table of Contents

In Vitro Fertilization

- One of the most common practices.
- OVAs are fertilized in a lab environment.
- Incubated for 3 or 5 days.
- Embryo grading :
  - Day 3 (a) [1].
    1. Number of cells.
    2. Quality of cells.
  - Day 5 (b) [2].
    1. ICM quality.
    2. TE quality.
    3. Blastocyst expansion.
- 30% Success rates [3]



Artificial Intelligent Assisted Embryo Selection

- Automatic embryo grading.
  - Compared against the accumulated decision of multiple embryologists.
  - Not completely indicative of the outcome.
- Implantation or Live-birth prediction.
  - Hard to gather data.
  - Mostly focused on single image analysis.
  - Few methods based on time-lapse analysis.
    - ! Frames are assumed to have the same deciding attributes.
    - ! Different time windows, different attributes.

Our Proposals

- Time-lapse analysis.
- Separated time window analysis.
  - Day 3.
  - Day 5.
- Combined final prediction.
- Data Length Scheduler (DLS)
  - Regulates the training process.
  - Suppresses the adverse effects of training on variable-length image sequences.

INTRODUCTION

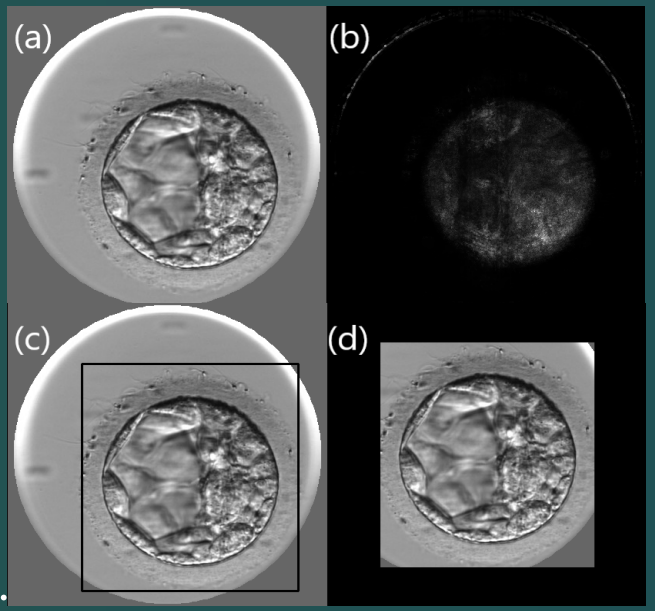
METHODOLOGY

RESULTS

CONCLUSIONS

Data

- Time-lapse image sequences of 130 transferred embryos with known outcome.
- 15-minute interval frame capture.
- 5-Fold cross-validation.
- Day 3:
  - Between 48-72nd Hour.
  - 96 Frames.
  - Training frames = 9984.
  - Test frames = 2496.
- Day 5:
  - Pass the 96th Hour.
  - Varied lengths (70 - 96).
  - Varied number of Train/test frames.
- Image preparation.
  - Cell crop & center:
    a) Input b) Optical flow median
    c) Detected ROI* d) Output
  - Resize from 500x500 to 224x224



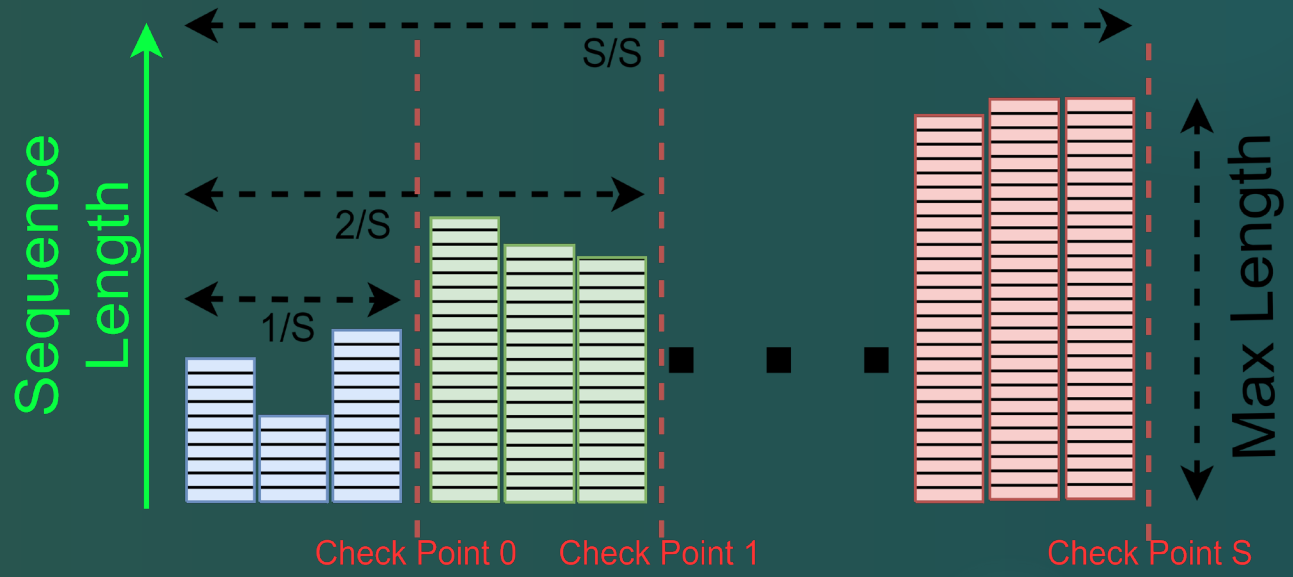
*Region Of Interest

Data Length Scheduler

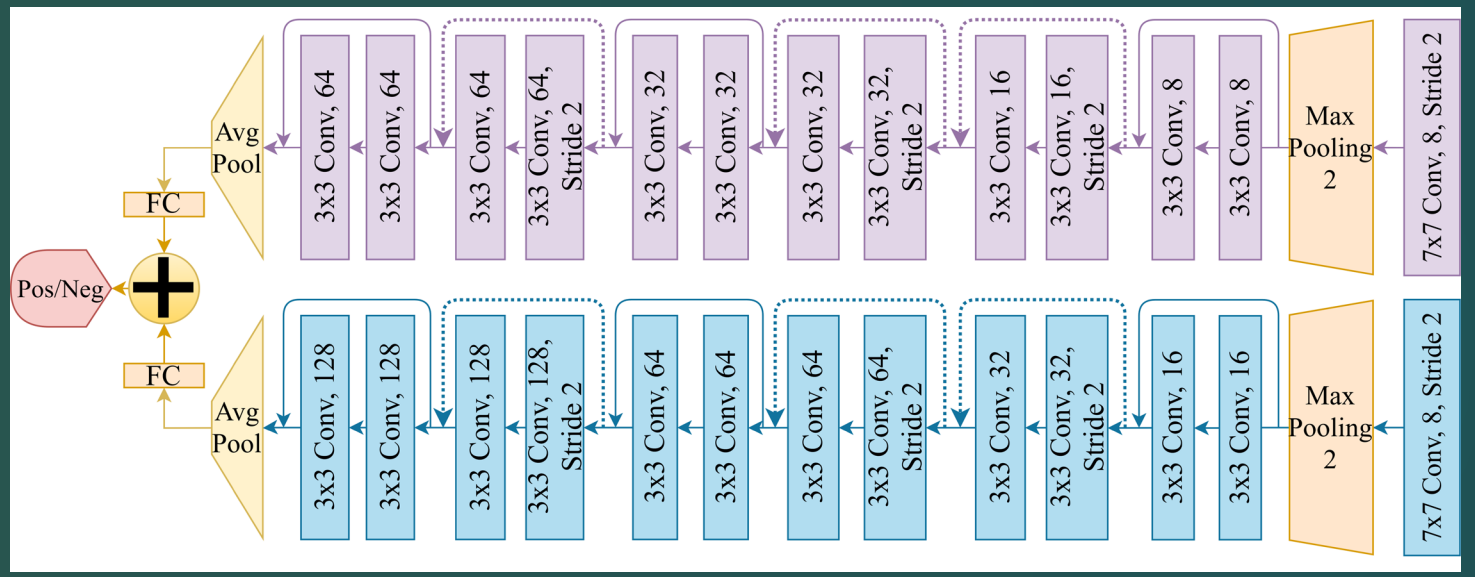
- Different development speeds, different sequence lengths.
- Slower developing embryos:
  - Similar frames ~ repeated samples.
  - Unbalanced training.
- A data regularization method.
- Regulating the training data based on sample's lengths.

Data Length Scheduler

- *S* percentile groups of sequences based on the length.
- Training starts with the first group.
- Group replacing checkpoints.
- Checkpoint modes:
  1. Passage of *n* epochs.
  2. No validation loss decrease after *p* epochs



Model's Structure



- Top path = Day 3 model.
- Bottom path = Day 5 model.

Model's Structure

- Training:
  - Independent path training.
  - Automatic extraction of Day 3 and Day 5 sequences.
  - Sequences are divided into frames.
  - Batch construction: only one frame per sequence.
- Testing:
  - Automatic extraction of Day 3 and Day 5 sequences.
  - Each sequences passes through its respective path.
  - Score are averaged over the temporal dimension.
  - Day 3 and Day 5 averaged together.
- Data Length Schedular (DLS) used in Day 5 training:
  - Mode 1.
  - $S = 4$.
  - $n = 10$.

INTRODUCTION

METHODOLOGY

RESULTS

CONCLUSIONS

Results

- DLS improves Day 5 model accuracy by 1.6%.
- Applying DLS to Day 3+ Day 5 model = 4.6% accuracy increase.
- Comparison against state of the art:
  - 6% accuracy increase in implantation prediction.
  - 2.6% accuracy improvement against live-birth predictor model.

Table 1: Performance comparison on Embryo outcome prediction

Row No	Model	Label format	Precision	Recall	Jaccard-Index	Accuracy
1	Day 3 model	Implantation	63.9	67.4	50.6	68.5
2	Day 5 model	Implantation	70.6	69.0	52.6	69.2
3	Day 5 model + DLS	Implantation	72.6	70.4	54.2	70.8
4	Combined Day 3 and Day 5	Implantation	72.6	72.3	56.7	72.3
5	Combined Day 3 and Day 5 + DLS	Implantation	**79.6**	**76.4**	**61.8**	**76.9**
6	Image CNN classifier [4]	Implantation	63.6	63.6	46.7	62.8
7	Image + Segmentation CNN classifier [4]	Implantation	71.1	72.7	56	70.9
8	Handmade feature classifier [5]	Live-birth	61.5	60.5	44.0	62.0
9	Image + Morphological factors CNN [6]	Live-birth	70.2	71.4	55.3	74.3

INTRODUCTION

METHODOLOGY

RESULTS

CONCLUSIONS

Conclusions & Future Work

- Our approach:
  - A deep-learning based system.
  - Capable of processing time-lapse embryo image sequences.
  - Predict embryo implantation outcome.
  - Individual Day 3 and Day 5 analysis.
  - More accurate than using only one of the stages or only single images.
- DLS algorithm is a way to suppress the adverse effects of training on length variant image sequences.
- Future works:
  - Time window range analysis.
  - AI-based time series analysis of embryo sequences.

Thank You.

References

[1] Allison E Baxter Bendus, et al., “Interobserver and intraobserver variation in day 3 embryo grading,” Fertility and sterility, vol. 86, no. 6, pp. 1608–1615, 2006.

[2] David K Gardner, et al., “Blastocyst score affects implantation and pregnancy outcome: towards a single blastocyst transfer,” Fertility and sterility, vol. 73, no. 6, pp. 1155–1158, 2000.

[3] Silke Dyer, et al., “International committee for monitoring assisted reproductive technologies world report: assisted reproductive technology 2008, 2009 and 2010,” Human reproduction, vol. 31, no. 7, pp. 1588–1609, 2016.

[4] Reza Moradi Rad, et al., “Predicting human embryos’ implantation outcome from a single blastocyst image,” in 2019 41st Annu. Int. Conf. of the IEEE Eng. in Med. and Biol. Society. IEEE, 2019, pp. 920–924.

[5] Alejandro Chavez-Badiola, et al., “predicting pregnancy test results after embryo transfer by image feature extraction and analysis using machine learning,” Scientific Reports, vol. 10, no. 1, pp. 1–6, 2020.

[6] Yasunari Miyagi, et al., “Feasibility of predicting live birth by combining conventional embryo evaluation with artificial intelligence applied to a blastocyst image in patients classified by age,” Reproductive Medicine and Biology, vol. 18, no. 4, pp. 344–356, 2019.