

Automated Cell Division Detection and Classification in Early Mouse and Human Embryos

by
Jonas Malmsten

Submitted in partial fulfillment
of the requirements for the degree of
Doctor of Professional Studies
in Computing

at

School of Computer Science and Information Systems

Pace University

May 2019

ProQuest Number: 13896245

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 13896245

Published by ProQuest LLC (2019). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

We hereby certify that this dissertation, submitted by Jonas Malmsten, satisfies the dissertation requirements for the degree of *Doctor of Professional Studies in Computing* and has been approved.



05/15/2019

Date

Chairperson of Dissertation Committee

Dr. Juan Shan

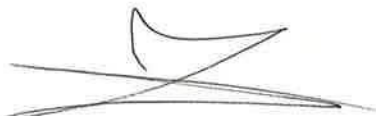


Dr. Charles Tappert

Dissertation Committee Member

5/15/2019

Date



Dr. Nikica Zaninovic

Dissertation Committee Member

5/15/2019

Date

Seidenberg School of Computer Science and Information Systems
Pace University

Abstract

Automated Cell Division Detection and Classification in Early Mouse and Human Embryos

by
Jonas Malmsten

Submitted in partial fulfillment
of the requirements for the degree of
Doctor of Professional Studies
in Computing

May 2019

Infertility affects millions of couples every year. The most effective treatment for infertility is in-vitro fertilization (IVF). Embryo selection, which has become a critical part of IVF treatments, is the process of selecting the most viable embryos to transfer and maximize the chance of pregnancy while minimizing the risks resulting from multiple births. Advancements in time-lapse microscopy have provided a stepping stone for better assessments of embryo quality but have also placed a greater burden on embryologists having to review a steadily growing number of images. Recent developments in artificial intelligence and convolutional neural networks (CNN) have shown great potential in computer-aided analysis of images. In this work, we explore ways to use CNNs to automate the embryo selection. In particular, we focus on predicting the timings of cell divisions in early human embryos, which have been shown to correlate well with pregnancy outcomes. In conclusion, we present a method to aid embryologists with annotating embryos and in a future version we think our method will be a part of a fully automated system for selecting the most viable embryos.

Acknowledgements

There are many people to whom I owe my thanks - advisors, professors, colleagues, and fellow students, but foremost they are also my friends. I would like to thank my advisor, Dr. Juan Shan, for taking on this project, always being supportive of my ideas, and tolerating my going back and forth while reviewing and publishing papers. The committee members, Dr. Nikica Zaninovic and Dr. Charles Tappert, for their advice, support, and help in making this happen. The Pace University CSIS faculty, Dr. Charles Tappert, Dr. Lixin Tao, and Dr. Ronald Frank for always being there and available as well as the students of DPS class of 2018 for many interesting and enjoyable discussions. My colleagues at Weill Cornell Medicine, in particular, Dr. Nikica Zaninovic, whom I hold in part responsible for starting this whole project, Dr. Qiansheng Zhan, Ty Chicketano, Volodymyr Kozin, and our director Dr. Zev Rosenwaks for their support. Last, but not least, I would like to thank my family for the support they have given me. My lovely wife, Delethia Malmsten, for understanding when I needed to stay home with my books, Sean Ridley, and my parents, Sten and Margareta Malmsten. Mom, I know you are proud.

Table of Contents

Abstract	iii
List of Tables	x
List of Figures	xii
Chapter 1 Introduction.....	1
1.1 In Vitro Fertilization	1
1.2 The IVF Treatment	2
1.3 Advances in ART.....	4
1.4 Embryo Development	6
1.5 Embryo Selection.....	8
1.6 Embryo Grading.....	9
1.7 Embryoscope Time-Lapse Microscopy	10
1.8 Problem Statement.....	11
1.9 Thesis Organization	11
Chapter 2 Background.....	12
2.1 The Embryoscope and Time-Lapse Microscopy	12
2.1.1 Hoffman Modulation Contrast.....	13
2.1.2 Dark-Field Illumination	13
2.2 Traditional Image Analysis.....	13
2.2.1 Image Format	13
2.2.2 Image Compression – JPEG and PNG.....	14
2.2.3 Histograms	15
2.2.4 Spatial Filters and Convolution	15
2.2.5 Edge Detectors	18
2.2.6 Sobel Filter.....	20
2.2.7 Median Filter.....	21

2.2.8	Min and Max Filter	21
2.2.9	Erode, Dilate, Close and Open.....	21
2.2.10	Object Recognition	22
2.2.11	Segmentation.....	22
2.2.12	Segment Representation.....	22
2.2.13	Thresholding	23
2.2.14	Fourier Smoothing	23
2.3	Artificial Neural Networks	24
2.3.1	Neural Network Architecture.....	24
2.3.2	Activation Function	26
2.3.3	Learning with Stochastic Gradient Descent.....	27
2.3.4	Backpropagation	29
2.3.5	Batches and Epochs	29
2.3.6	Training, Validation, and Testing.....	30
2.3.7	Overfitting and Underfitting.....	30
2.3.8	Sample Class Balance.....	31
2.3.9	Learning Rate, Decay, and Momentum	32
2.3.10	Optimizers.....	33
2.3.11	Vanishing and Exploding Gradients	33
2.3.12	Batch Normalization	34
2.3.13	Weight Initialization	35
2.3.14	Softmax Layer.....	35
2.3.15	Dropout Layer.....	36
2.4	Convolutional Neural Networks	37
2.4.1	Convolutional Neural Network Architecture.....	37
2.4.2	Convolutional Layer	38
2.4.3	Pooling Layer.....	39

2.4.4	Receptive Field	39
2.4.5	ImageNet Competition.....	40
2.4.6	Transfer Learning and Fine Tuning	40
2.4.7	Learning Challenges	41
2.5	Named CNN Models.....	42
2.5.1	LeNet.....	42
2.5.2	AlexNet.....	43
2.5.3	ZF Net	43
2.5.4	GoogLeNet and Inception.....	44
2.5.5	VGG Net	45
2.5.6	ResNet.....	45
2.5.7	Inception-ResNet	46
2.5.8	Inception V3.....	46
2.5.9	Fully Convolutional Networks.....	47
2.5.10	U-Net.....	48
2.5.11	Mask R-CNN	48
2.5.12	Autoencoders	48
2.6	Tools	49
2.6.1	Hardware.....	49
2.6.2	Tensorflow v1.1.0	49
2.6.3	Keras v2.1.3	50
2.6.4	Jupyter Notebook	50
Chapter 3	Previous Work	50
3.1	Stained Embryos	50
3.2	Non-Stained Embryos.....	51
3.3	Early Embryo Viability Assessment.....	51
3.4	Traditional Image Analysis for Embryo Study	52

3.5	Neural Network Models Used in Embryo Study	55
Chapter 4	Analysis of Mouse Embryos	57
4.1	Introduction.....	57
4.2	Dataset.....	58
4.3	Reproducing Previous Results	59
4.4	Image Pre-processing and Augmentation	62
4.5	Transfer Learning and Fine-tuning	63
4.6	Results and Discussion	65
Chapter 5	Analysis of Human Embryos	67
5.1	Introduction.....	67
5.2	Dataset.....	67
5.2.1	Image Export Program	68
5.2.2	Dataset for Segmentation and Classification	69
5.2.3	Dataset for Time-lapse Analysis.....	70
5.3	Segmentation.....	71
5.3.1	Segmentation Using Traditional Image Analysis Techniques.....	71
5.3.2	Segmentation Using Fully Convolutional Neural Networks	74
5.4	Training Inception V3 to Classify Human Embryos	78
5.4.1	Baseline Model	79
5.4.2	Adjusting the Image Size.....	80
5.4.3	Adjusting Batch Size.....	81
5.4.4	Multiple Focal Planes	82
5.4.5	Increasing the training data.....	87
5.4.6	Regression Model	88
5.5	Using Inception V3 to Analyze Human Time-Lapse Images.....	89
5.5.1	Time-lapse Optimization	90
5.5.2	Image Variance	96

5.6	Prediction Analysis	99
5.7	Other Experiments	102
Chapter 6	Implementation	103
6.1	Architecture.....	104
6.2	Web Services Interface	105
6.3	Image Cache.....	106
6.4	Results Database	106
6.5	Model Compute Server	106
Chapter 7	Conclusion	107
7.1	Results.....	107
7.2	Discussion and Future Work.....	108
Appendix A	Human Embryoscope Time-Lapse Microscopy Images One embryo photographed in 5 focal planes at 20-minute intervals during division from 2-cell to 4-cell.	112
Appendix B	Mouse Time-Lapse Microscopy Images One mouse embryo photographed at 7-minute intervals during transition from 2-cell to 4-cell.....	113
Appendix C	Fully Convolutional Network.....	114
Appendix D	A random selection of images with miss-predicted labels from the human embryo dataset using Inception V3 trained on segmented augmented images.	115
Appendix E	Forward-backward dynamic programming algorithm for locating cell stage transitions in embryo TLM images.....	116
References	117

List of Tables

Table 1 National ART Statistics	2
Table 2 Overview of an IVF treatment	4
Table 3 Examples of one-hot encoding for embryos and animals.....	36
Table 4 The number of images available and used for training, validation, and testing. Oversampling is used to balance the classes in the used training set.	58
Table 5 Confusion matrices for training the LeNet inspired model with different pre- processing and data augmentation	63
Table 6 Confusion matrices for training InceptionV3 using transfer learning.	64
Table 7 All images from the validation set where classification was incorrect using the InceptionV3 model trained on cropped augmented images. Labels and predictions for neighboring frames are displayed above each image.	65
Table 8 Confusion matrix from evaluating the best model on the test set.....	66
Table 9 Distribution of the exported images at different cell stages and focal planes. Most images were captured in all 7 focal planes, but some were only captured in the central focal plane (focal 0) and some were only captured in 5 focal planes (± 45 were excluded). There are also a few discrepancies based on randomly missing images, but overall the classes are well balanced.	69
Table 10 Distribution of images after separation into training, validation, and test sets. The number of subjects is the number of unique embryos in each set.	70
Table 11 Validation loss (higher loss means less accurate segmentation) using different number of training examples.	77
Table 12 Model trained on the central focal plane using a subset of the human embryo dataset with cell stages 1 through 4.	80
Table 13 Model trained on the central focal plane using images from the human embryo dataset.	80
Table 14 Accuracies achieved when using different focal planes. Individual focal planes, as well as combinations of focal planes, were tested. Each model was trained for 100 epochs and the listed accuracy is from the epoch with the best result.....	87
Table 15 Confusion matrix for the validation data after training the regression model for 150 epochs.	89

Table 16. Five consecutive images and their probability for each cell stage as given by the model. The highest probability at each position is highlighted to trace a path from left to right.	92
Table 17. The sum of probabilities after the first pass of the forward-backward algorithm. The highlighted optimal path is found by tracing the highest sum of probabilities from the rightmost column moving backward.....	93
Table 18 Comparison of accuracies when predicting cell stages using the different methods to incorporate time-lapse information. The first accuracy is based on predicting individual images (no time-lapse information). Median is the median prediction from every three consecutive frames. Greedy is after enforcing the constraint that cell stages are always ascending. Global is using the forward-backward dynamic programming algorithm, and Global+ is using the enhanced version described in section 5.5.2. ..	94
Table 19 Accuracy of the model to predict cell stage transition within N number of frames from embryologists' annotations. The high lighted row indicates 93.4% accuracy to predict cell stage transition within 5 frames of the embryologists' annotation.	95
Table 20 Accuracy within N frames of embryologists annotation after applying a penalty for cell stage transitions based on image variance.....	98
Table 21 Comparison of the TLM optimization methods using accuracy within N frames of embryologists' annotation. The Global method uses the graph model described in 5.5.1 and the Global+ method includes the edge penalties from image variances introduced in 5.5.2.	98
Table 22 Model performance when used to analyze images on GPU compared to CPU.	104

List of Figures

Figure 1 A normal developing embryo at different cell stages with mean hours since fertilization (based on our dataset).....	8
Figure 2 Left: 8-cell embryo with 5 clearly visible blastomeres, ZP, and some fragmentation highlighted. Right: Blastocyst with Expansion, ICM, and TE highlighted.	10
Figure 3 Matrices of intensity values between 0 (black) and 255 (white) are used to represent digital images. From left to right, a section of the original image is zoomed in to reveal some of the intensity values. Notice how higher numbers represent brighter shades of gray.....	14
Figure 4 Left: A dark image and its histogram. Middle: The same image after stretching the histogram. Right: The same image after equalizing the histogram.....	15
Figure 5 A convolution kernel is slid across the top row of the image from Figure 3 to produce a new image (bottom). Repeat the sliding motion for each row (2 more times) to complete the convolution.....	16
Figure 6 Result after applying an average filter to image in Figure 5. The convolution example showed how to calculate the highlighted partial 3x3 matrix on the right. Because of the high resolution, the visible difference in the left image is very small, but the difference is clear in the middle and right images.	18
Figure 7 Discrete function, $f(x_i)$, using values from the middle row in Figure 6.	19
Figure 8 Normalized image after applying an edge detector to Figure 3.	20
Figure 9 Feedforward neural network with 3 layers: one input, one hidden, and one output. The network has 5 neurons and 6 trainable weights.....	25
Figure 10 Sigmoid (left) and ReLu (right) activation functions shown in blue with their derivatives in orange.....	27
Figure 11 Common CNN architecture for classification, the first and last layers follow the same pattern.	38
Figure 12 2-dimensional illustration of the receptive field at different depths of a 4-layer CNN: 5x5 convolution (blue), 2x2 pooling (blue+green), 5x5 (blue+green+orange) convolution, and 2x2 pooling (blue+green+orange+yellow). The blue bell-shaped curve in the background represents the effective receptive field. To reduce clutter, not all paths to the first layer are drawn.....	40

Figure 13 Inception module as described by Szegedy et al. [44] (image recreated and modified from the original). The colors represent each of the four paths that the data can travel through the module.....	44
Figure 14 Inception V3 architecture based on the implementation in Keras.....	47
Figure 15 CNN architecture inspired by LeNet.....	59
Figure 16 Median filter subtraction. Left: Original image. Middle: Subtle smoothing of the original image by a small median filter. Right: The normalized difference between the original and the smoothed image.	72
Figure 17 The resulting image after applying median filters is dilated and eroded to locate the outer boundary of the embryo.....	73
Figure 18 Left: Threshold applied to Figure 17C. Right: Final segmentation laid over the original image. The white line follows the border of the initial rough segmentation, and the green line is the segmentation after Fourier smoothing.	73
Figure 19 A fully convolutional network for segmenting embryos. The input to the model is a 224x224 pixel embryo image. The first layer uses a 3x3 convolution with 16 filters followed by 2x2 max pooling. The next four layers repeat this process while doubling the number of filters until reaching the middle section with 512 filters and 7x7 pixel patches. The process is then reversed using 3x3 convolutional layers and 2x2 up-sampling until the original 224x224 image size is restored. A precise definition of the model is available as Python code in Appendix C.....	75
Figure 20 Examples with differences between the two segmentation methods. The first column is the original image. The second and third columns are the segmentation results from the first and second method, respectively. The last column shows the original image overlaid with the outline from both segmentation methods. Green is used for the first method and red is used for the second method, resulting in yellow where both methods overlap.	76
Figure 21 ZP segmentation result using our FCN model. The first image is from the validation dataset. The second and third images are the manual and model segmentations, respectively. The last image shows the original image overlaid with the outline from both segmentation methods.....	78
Figure 22 Percent accuracy after training the model at various image resolutions.	81
Figure 23 Three examples of RGB images constructed using three focal planes; -30, 0, and 30 for red, green, and blue, respectively.....	83
Figure 24 A simplified illustration of our CNN model using three focal planes to predict up to eight cell stages. The number of focal planes can be adjusted dynamically. A TLM image of an embryo in multiple focal planes is provided as an input to the model. The model breaks the image down into features and uses the convolution operator to generate feature maps. This process is repeated several times, creating a hierarchy of	

features (using several layers). The features from the first layer (illustrated above) represent simple shapes that form parts of an embryo. Features at deeper layers consist of patterns from the resulting feature maps and usually do not make visual sense to us. After the final layer, there is a SoftMax function converting the final feature maps to a normalized probability distribution for each cell stage (the sum of probabilities for all cell stages equals one). The convolutional layer in the figure can be substituted for the Inception V3 model in Figure 14, section 2.5.8.	84
Figure 25 Training and validation accuracy at each epoch when training the model for different focal planes. The training accuracies are represented by the smooth curves at the top while the validation accuracies are represented by the more irregular curves below. The same color is used for each pair of training and validation curves. For example, the red curve representing focal -15 had the highest validation accuracy, but not the highest training accuracy. The highest training accuracy was achieved by focal 30 (yellow curve), which also had the lowest validation accuracy.....	85
Figure 26 Validation accuracy when training the model using randomly selected smaller subsets of the available training data. Because the expectation is a logarithmic curve, we also show the logarithmic trendline fitted to our data using least squares (thin dotted line).	88
Figure 27 The data from Table 16 illustrated as a directed graph where the constraint, that cell stages may not decrease from one frame to the next, is enforced by the directed edges. A path following the edges from left to right in the graph will never reduce the number of cells.....	92
Figure 28 Variance calculated between every two subsequent frames over 476 TLM images from one embryo. The orange line indicates the mean of the variance and cell divisions are annotated below. Cell divisions always happen at or near when the variance peaks above the mean, but not all peaks indicate cell divisions.	96
Figure 29 The normalized negative logarithm of the image variance curve from Figure 28 is used to add a penalty to all cell stage transitions. Cell stage transitions correlated with low image variance are given a higher penalty.....	97
Figure 30 Model prediction confidence levels for correctly versus incorrectly predicted images. Img True and Img False represent the model's confidence levels for correctly and incorrectly predicted images when analyzing images one at a time. Opt True and Opt False represent the model's confidence levels after using the global optimization to adjust predictions.	100
Figure 31 Receiver operating characteristic (ROC) curve plotted for each cell stage predicted by the CNN model. It is clear that the 3-cell stage is less accurate than the other cell stages.....	101
Figure 32 Architectural overview of how the model can be implemented in a laboratory setting.....	104

Chapter 1

Introduction

1.1 In Vitro Fertilization

Robert Edwards and Patrick Steptoe invented and performed the first successful in-vitro fertilization (IVF) procedure in 1978 [1]. In vitro means in glass, or in general outside of the living organism. Oocytes (unfertilized eggs) are surgically retrieved from a patient's ovaries, fertilized in vitro (union of female and male gametes), and transferred back to the uterus to achieve pregnancy. Infertility is often caused by problems with the uterus or fallopian tubes, preventing sperm from reaching the oocyte. The IVF procedure directly circumvents these and, as we shall see, many other problems as well. Robert Edwards received the Nobel Prize for his achievements in 2010.

Assisted reproductive technology (ART, including IVF) is a very controversial topic with ethical concerns, from both political and religious views, and whether we should tamper with the creation of life. Despite the controversy, the market for ART has grown continuously since inception. It is no longer unusual for people to seek professional help with infertility and ART has become more socially accepted. IVF became available to patients in the US in 1981 in Norfolk, VA. Society for Assisted Reproductive Technology (SART) was founded in 1985 to provide guidelines, represent, and promote the ART community, in which they have been very successful. Since 1992 all ART procedures should be reported to the Center for Disease Control and Prevention (CDC), who publish

yearly national statistics. In 2015, the US fertility market was estimated to be between \$3 and \$4 billion [2], and there were 231,936 ART procedures performed using IVF resulting in 60,778 live births and 72,913 children [3]. With 3,978,497 births reported in the same period, 1.8% of all children born in the US were from IVF. There is no indication that the trend of continuous growth is about to change, Table 1.

Table 1 National ART Statistics

Year	Total Children Born in the US	ART Cycles	Births from ART	Children from ART	%Twin Births from ART	%Children from ART
2015	3,978,497	231,936	60,778	72,913	20.0	1.83
2014	3,988,076	208,786	57,332	70,352	22.7	1.76
2013	3,932,181	190,773	54,323	67,996	25.2	1.73
2012	3,952,841	176,247	51,267	65,160	27.1	1.65
2011	3,953,590	151,923	47,818	61,610	28.8	1.56
2010	3,999,386	147,260	47,090	61,564	30.7	1.54
2009	4,130,665	146,244	45,870	60,190	31.2	1.46
2008	4,247,694	148,055	46,326	61,426	32.6	1.45
2007	4,316,233	142,435	43,412	57,569	32.6	1.33
2006	4,265,555	138,198	41,343	54,656	32.2	1.28
2005	4,138,349	134,260	38,910	52,041	33.7	1.26

1.2 The IVF Treatment

There are many variations of how to perform an IVF treatment. Table 2 provides an overview of the most common steps. For each treatment, there is a time-line starting with day 1 of the patient's menstrual cycle. The average length of the cycle is 28 days. Embryo selection, which is the main topic of this study, is done during the 3-5 day incubation period starting usually on day 16. The days indicated are average numbers and may vary slightly from one treatment to another. For example, the length of the follicular phase is determined based on the results of sonograms (ultrasound) and blood tests. If HCG trigger (ovulation

trigger) is injected on day 12 instead of 14, retrieval will be on day 14 and transfer will be on day 17 or 19.

The first pregnancy test is performed on, or after, day 28. The first sign of a successful pregnancy is a positive β HCG test; an elevated level of beta-human chorionic gonadotropin (β HCG) in the bloodstream. A clinical pregnancy is established by observing sacs or fetal hearts during a sonogram, and may not be visible as early as day 28. A visible sac in the uterus is proof of implantation; that the embryo has implanted in the uterus wall. A heartbeat should be visible soon after, at which point the patient is usually done at the IVF clinic and referred to a regular OB/GYN clinic to follow-up on their pregnancy. When measuring the number of successful pregnancies, the implantation rate is sometimes used instead of actual live births. In part because live birth data is not always available, but also because there are many non-embryo related events that can cause a pregnancy to fail after implantation.

Table 2 Overview of an IVF treatment

Day	Short Name	Description
1		The first day of bleeding.
2-3	Start Stims	Start hormone injections to stimulate follicles to produce more oocytes. Injections continue daily for about 12 days.
2-14	Follicular Phase	Sonogram and blood tests are used several times to monitor the development of follicles in the ovaries. Each follicle contains one oocyte and grows in size as the oocyte matures.
14	HCG Trigger	Human chorionic gonadotrophin (HCG) is injected to trigger ovulation; the release of oocytes from the ovaries into the fallopian tubes where natural fertilization would occur. Day 14 is an average; the previous monitoring determines the exact day. HCG will induce ovulation;
16	Retrieval	36 hours after HCG trigger a surgical procedure is used to retrieve the oocytes from the follicles just before ovulation.
16	Fertilization	Retrieved oocytes are fertilized with sperm and placed in an incubator. They remain in the incubator for 3-5 days.
19 or 21	Transfer	The embryos (fertilized oocytes) are transferred back to the uterus to implant and achieve pregnancy.
28		The first pregnancy test.

1.3 Advances in ART

The technology used to perform IVF has evolved significantly over the past decades [4], resulting in both increased success rates and the ability to treat different problems. Patients chose IVF for various reasons. The most common reason is that they are unable to achieve pregnancy on their own within a certain period of trying. Advances in the surgical procedure to retrieve oocytes and the use of stimulated cycles have had a huge impact on IVF success rates. During a regular menstrual cycle, the ovaries release on average one oocyte each month. The number of oocytes retrieved in the original natural (non-stimulated) cycles from 1978 was only about 0.7 with each attempt [4]. In a modern stimulated cycle, the patient injects daily hormones for about two weeks before retrieval increasing the number of oocytes tenfold; retrieval of 10-20 oocytes is common. After

fertilizing all oocytes, creating embryos, only a few of them are selected for transfer back to the patient. The process of selecting the most viable embryos for transfer is referred to as embryo selection and has become a critical part of IVF that will be explained further in the next section.

Cryopreservation is a technology that has opened up for new reasons to perform IVF, as well as improved safety and efficiency of regular IVF. Oocytes, sperm, and embryos can be cryopreserved (frozen) and kept indefinitely before they are thawed and used. Fertility preservation (FP) is a process when patients chose IVF just for cryopreservation without intent to become pregnant. Some chose it for social reasons, like delay having children. Others have medical reasons, for example, cancer patients who know they will become infertile after cancer treatment.

Elective single embryo transfer (eSET) is a way to reduce cost and risk when doing IVF with the intent to become pregnant. eSET was made available by advancements in embryo cultures, cryopreservation and the general increase in IVF success rates. During a traditional IVF cycle, it is common to transfer two or more embryos to increase the chance of success on the first attempt. However, this also increases the chance of multiple pregnancies (twins and triplets). Multiple pregnancies have an increased risk of complications for both mother and children. There is an increased rate of early births resulting in longer hospital stays and higher costs. The idea behind eSET is to transfer only one embryo from a fresh IVF cycle and cryopreserve remaining embryos. If the IVF cycle fails, another attempt is made using a frozen embryo transfer (FET). FET is cheaper and less invasive because there is no need for hormone stimulants or oocyte retrieval. The

benefits of not having twins have been shown to outweigh the cost of performing the extra procedure [5]. Embryo selection becomes even more important with eSET.

Advances in preimplantation genetic testing (PGT) can be used to de-select unwanted abnormal embryos. A small sample, or biopsy, is taken from each developing embryo and sent to a genetics laboratory. The laboratory can detect genetic mutations or chromosomal abnormalities, like aneuploidy, which often results in miscarriages. DNA can be sequenced and analyzed for inheritable diseases. PGT can also be used to select gender and other traits. Deselecting embryos to eliminate well-documented diseases is more common, and more socially and ethically acceptable, than selecting embryos for particular traits [6].

Male infertility is about as common as female infertility [7]. Female infertility often results from problems with ovaries, fallopian tubes, or hormonal imbalance, which are inherently solved by the IVF procedure. Male infertility may result from low sperm count or poor sperm motility, which can be treated using intracytoplasmic sperm injection (ICSI) [8]. ICSI means using a fine needle to inject one spermatozoon (5 μ m) into each oocyte. Other procedures for male infertility include testicular sperm extraction (TESE).

Donation programs have expanded the market for ART further. Patients who are medically unable to produce oocytes or sperm can use a donation program. Donation programs also allow for singles or same-sex couples to become parents.

1.4 Embryo Development

After retrieval, the oocytes are fertilized with sperm and placed in an incubator providing a suitable environment for the embryos to develop. Temperature, pH, and gas levels are

regulated while minimizing exposure to light. The embryos remain in the incubator for 3-5 days before they are ready to be transferred back to the patient. During this time they are monitored closely. It is common to observe embryos through a microscope to assess embryo quality at regular intervals. However, with the development of the time-lapse microscopy (TLM) embryos can be monitored continuously without disturbance. Figure 1 illustrates some of these TLM images captured at different stages of embryo development. The first image is a zygote, that is an oocyte immediately after fertilization. There is no visual difference between an oocyte and a zygote until several hours later. A zygote consists of one cell surrounded by a thin shell called the zona pellucida (ZP). One of the early indications of successful fertilization is the appearance of 2 pronuclei (2PN, male-sperm and female-oocyte origin) as shown in the second image. The pronuclei are the small circular regions slightly above the center. These are the male and female chromosomes about to combine. Once combined, the cell starts to divide to form 2, 4, and 8 -cell embryos (cells called blastomeres), referred to as cell stages. Cell stages 3, 5, 6, and 7 are also possible, but usually for a much shorter duration. Blastomeres become too small to identify after 8-cell stage. As depicted in the last image, after about five days, the embryo forms a blastocyst. A blastocyst consists of three main parts: Trophectoderm (TE), inner cell mass (ICM), and the blastocoele. The TE is the surrounding shell of the embryo that will form placenta, the ICM is the lump of cells seen towards the left of the center of the image that will form fetus, and the blastocoele is the fluid-filled cavity between the ICM and the TE.

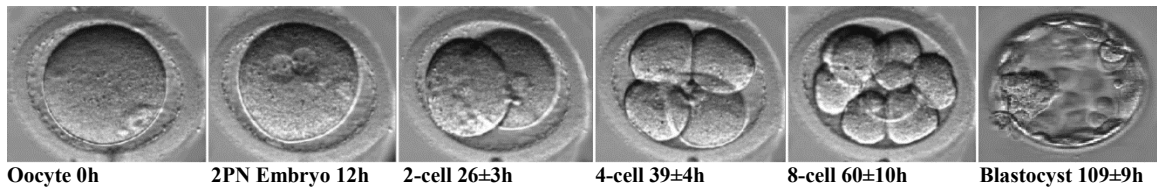


Figure 1 A normal developing embryo at different cell stages with mean hours since fertilization (based on our dataset).

1.5 Embryo Selection

Embryo selection is the process of selecting which embryos to transfer back to the patient, and its impact on IVF success rates is well documented [9-11]. Embryos are sometimes transferred back to the patient after three days of incubation, at around the 8-cell stage. Other times they are left in the incubator for about five days and transferred back to the patient as blastocysts. When the embryos are transferred is usually decided by the embryologist and physician and is determined by embryo number and quality. Which embryos to transfer is usually determined by an embryologist observing several indicators of embryo quality at various cell stages. The observations are combined into an embryo grade, explained in the next section. Many miscarriages and failed pregnancies are due to genetic defects. There is a correlation between certain genetic defects and the timing of blastomere divisions seen as early as the 8-cell stage [12]. The objective of this research is to develop a method for automatic detection of blastomere division events in a series of digital images of developing embryos.

1.6 Embryo Grading

The goal of embryo grading is to assign a score to each embryo, such that the scores are highly correlated with positive pregnancy outcomes [13]. A positive pregnancy outcome can be live birth or that the embryo implanted. Many things can go wrong between implantation and live birth, and these problems are often not related to embryo quality. Grading is done just before transfer, after three or five days of incubation. The grading on day three is quite different from that of a day five. After three days, the embryo is expected to have around eight blastomeres. A score between 1 (good) and 5 (poor) is assigned by an embryologist based on visual inspection of the embryo. The embryologist typically looks for features such as the number of blastomeres, their sizes and symmetry, fragmentation (cytoplasmic blebs), blastomere lineage and timing of cell division events, ZP thickness [14], zygote and pronuclear morphology [15]. A day five embryo is expected to have formed a blastocyst. Many clinics follow the Gardner blastocyst grading system [16], where individual scores are assigned for expansion, ICM, and TE. The expansion grade is assigned a number between 1 and 6, ICM and TE are both graded with letters A, B, or C, where 3, 4, or 5AA is the best possible grade. Figure 2 illustrates some of the observed features.