COMPUTER VISION FOR SEQUENTIAL NON-INVASIVE MICROSCOPY IMAGING CYTOMETRY WITH APPLICATIONS IN EMBRYOLOGY

A L MÖLDER

PhD   2016

COMPUTER VISION FOR SEQUENTIAL NON-INVASIVE MICROSCOPY IMAGING CYTOMETRY WITH APPLICATIONS IN EMBRYOLOGY

Anna Leida Mölder

A thesis submitted in partial fulfillment of the requirements
of the
Manchester Metropolitan University for the degree of
Doctor of Philosophy

Faculty of Science and Engineering
School of Computing, Mathematics and Digital Technology
Manchester Metropolitan University

15 November 2016

Declaration of Authorship

I, Anna Leida Mölder, declare that this Thesis, titled 'Computer Vision for Sequential Non-invasive Microscopy Imaging Cytometry with Applications in Embryology' and the work presented in it are my own. I confirm that:

- This work was done mainly while in candidature for the degree of Doctor of Philosophy at the School of Computing, Mathematics and Digital Technology (SCMT) at Manchester Metropolitan University.

- Where any part of this Thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.

- Where I have consulted the published work of others, this is always clearly attributed.

- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this Thesis is entirely my own work.

- I have acknowledged all main sources of help.

- Where the Thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

- Embryo images for evaluation were acquired as anonymized sequences donated to research with ethical approval from the Human Fertilisation and Embryology Authority (R0155).

Signed:

Date:

"I know it's a measurement error, because it's messing up my data"

Shawn Achor

Abstract

Many *in vitro* cytometric methods requires the sample to be destroyed in the process. Using image analysis of non-invasive microscopy techniques it is possible to monitor samples undisturbed in their natural environment, providing new insights into cell development, morphology and health. As the effect on the sample is minimized, imaging can be sustained for long un-interrupted periods of time, making it possible to study temporal events as well as individual cells over time. These methods are applicable in a number of fields, and are of particular importance in embryological studies, where no sample interference is acceptable.

Using long term image capture and digital image cytometry of growing embryos it is possible to perform morphokinetic screening, automated analysis and annotation using proper software tools. By literature reference, one such framework is suggested and the required methods are developed and evaluated. Results are shown in tracking embryos, embryo cell segmentation, analysis of internal cell structures and profiling of cell growth and activity. Two related extensions of the framework into three dimensional embryo analysis and adherent cell monitoring are described.

Acknowledgements

This Thesis would not have been completed without help, support and patience from the outside world. While wrapping it up, my thanks go in particular to:

My friend, support and most stubborn supervisor ever: Silvester Czanner, because you never falter and never stop smiling.

My assisting supervisors Nicholas Costen and Geraldine Hartshorne who did their best in enlightening me in the areas where my knowledge was waning.

Manchester Metropolitan University, for financing my PhD studies and making it all possible.

Johan, for holding me up with your enthusiasm and your ideas at every turn of the road. Will you believe me if I say I will have more time to play now?

Sarah for coming to me with interesting images.

Anette and Zahra for offering great working space, enthusiasm, tips-and-tricks, knowledge and interesting discussions.

Tomas, Mariana, Åsa and Christina, for putting up with my constant drift into science.

Dan, Adrian, Ezak, Gemma, Choon-Ching, Moi and all the other people at MMU for being a second family to me.

Karin, Emelie, Anders, Christopher, Maria, Emil, Raffe, Johan, Morgan, Efva, P-O, Håkan, Andreas and to all the rest of the people at AF Tech Åmål and VCRS for being a third family to me. You all have a special place in my heart.

Kersti, Sven-Göran, Karin, Jan, Göran and all the rest of you who kept my interest in science alive over the years.

My parents for always supporting me whether I deserved it or not.

Frida, for being there when I needed it.

Viva, Zinner and Lyra for all the joy and trouble you bring me.

Jeanette, Henrik and Christina for your help in relieving the trouble.

My dear C, for being with me as long as you could.

And to everyone else I may have forgotten . . .

Publications

I. A Mölder, S Czanner and N Costen. Focal Plane Selection in Microscopic Embryo Images. *Conference on Computer Graphics & Visual Computing*, 2014. [1]

II. A Mölder, S Czanner, N Costen and G Hartshorne. Automatic detection of circular structures in human embryo imaging using trigonometric rotation of the Hough Transform. *22nd International Conference on Pattern Recognition*, 22:3239–3244, 2014. [2]

III. A L Mölder, N Costen, G M Hartshorne and S Czanner. Method for segmentation of single blastomeres in three dimensional microscopy image stacks. *In print.* [3]

IV. A Mölder, S Czanner and N Costen. Multidimensional visualisation to improve the understanding of biological data sets. *Proceedings of Central European Seminar on Computer Graphics*, 17:121–128, 2013. [4]

V. A Mölder, S Drury, N Costen, G M Hartshorne and S Czanner. Semiautomated analysis of embryoscope images: Using localized variance of image intensity to detect embryo developmental stages. *Cytometry Part A*, 87(2):119–128, 2015. [5]

VI. A L Mölder, S Drury, N Costen, G M Hartshorne and S Czanner. Efficient multiclass Support Vector Machine for the classification of embryo time-lapse imaging. *Submitted for publication*, 2016. [6]

VII. A L Mölder, J Persson, Z El-Schich, S Czanner and A Gjörloff-Wingren. Machine learning for the study of etoposide-treated in vitro adherent cells based on non-invasive imaging growth characteristics. *Submitted for publication*, 2016. [7]

VIII. A L Mölder, S Drury, N Costen, G M Hartshorne and S Czanner. Three Dimensional Visualisation of Microscope Imaging to Improve Understanding of Human Embryo Development. *Visualization in Medicine and Life Sciences III*, Springer International Publishing, Cham, 249–264, 2016. [8]

Contents

List of Figures

List of Tables

Abbreviations

CAD	**C**omputer **A**ided **D**iagnosis
DAG	**D**irectional **A**cyclic **G**raph
DAPI	4',6-**D**i**A**midino-2-**P**henyl**I**ndole
DF	**D**ark **F**ield (microscopy)
DH	**D**igital **H**olographic (microscopy)
DNA	**D**eoxyribo**N**ucleic **A**cid
GLCM	**G**rey **L**evel **C**o-occurence **M**atrix
HMC	**H**offman **M**odulation **C**ontrast (microscopy)
HT	**H**ough **T**ransform
IOU	**I**ntersection **O**ver **U**nion
IVF	**I**n **V**itro **F**ertilization
LC	**L**ow **C**oncentration
MAP	Maximum **A** **P**osteriori
MET	Minimum **E**rror **T**hresholding
MTS	3-(4,5-di**M**ethyl**T**hiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-**S**ulfophenyl)-2H-tetrazolium
NPV	**N**egative **P**redictive **V**alue
OT	**O**tsu **T**hresholding
PPD	**P**osterior **P**robability **D**istribution
PPV	**P**ositive **P**redictive **V**alue
RF	**R**andom **F**orest
ROC	**R**eceiver **O**perator **C**haracteristc
ROI	**R**egion **O**f **I**interest
SS	**S**uccesive **S**licing (SVM)
SVM	**S**upport **V**ector **M**achine
TLI	**T**ime **L**apse **I**maging

Physical Constants

pi	π	=	3.14159
micro	μ	=	$1 \cdot 10^{-6}$

Symbols

∞	infinity	
$\int$	integral	
$\Re$	real space	
sig	logistic sigmoid function	
Λ	activation state mapping function	
Σ	sum	
Ω	gated sigmoid	
α	training set index	
β	observation index	
δ	infinitesimal unit	
θ	angle	rad
λ	result classification group	
μ	mean	
ρ	number of parameters	
σ	standard deviation	
ϕ	activation parameters	
φ	angle	rad
ψ	learned model parameters	
ω	activation function	

A	area	m^2
C	connectivity	
D	image space dimensionality	
E, E'	real, predicted value	
F	image function	
G	GLCM	
H	number of gray levels in GLCM	
I	maximum gray level	
K	maximum number of classification groups	
L	number of α training pairs	
$\widehat{L}$	Laplace filter	
P	posterior probability distribution	
$\widehat{P}$	Prewitt filter	
Q	inference	
R	radius	m
S	property of GLCM	
T	threshold	
X, Y, Z	image size in x, y, z dimension	pxl
a	ellipse major axis	
b	ellipse minor axis	
c	concentration	mol·m^{-3}
d	distance	m
g	filter	
i, j	gray scale value	
k	class index	
m	arbitrary constant	
n	arbitrary counter	
p	part (ratio)	
r	radius	m
v	observed image state	
w	world state	
x, y, z	dimension variables in 3D space	pxl

For the Kids

Chapter 1

Aims & Objectives

1.1 Introduction

The human body is a gigantic symbiotic collaboration between individual cells. The study of cellular behavior is of great importance to our understanding of human body functions, development, response to pathogens and the effect of treatments. Each cell functions as an individual entity with its own resource consumption, energy production, waste disposal and reproductive cycle. At the same time, several cells cooperate to form tissues and organs and a cell is far from an isolated organism. T.W. Deacon has described a multi-cellular organism as *"critically incomplete"* [9]. Cells interact with their neighbors and with the rest of the body through chemical, electrical or tactile signaling, and it is this ability for communication and self-organization which makes it possible for us to function as one organism. Understanding one cell's function is far from understanding the body. It is by studying individual cells in their living context, that we can best understand cellular interaction.

Computer vision is the technique of extracting semantics from images. With biological staining in combination with image-based analysis it is possible to extract genetic, proteomic, metabolic, molecular and morphological information of cells. In some cases, staining with fluorescence or dyes is not suitable, due to the sensitive nature of the sample or particular requirements of the experiment. When stains cannot be used, the computer analysis task is challenged by the transparent nature of the biological samples, low image contrast and high noise levels. Overcoming these obstacles and obtaining a robust marker-free image-based automatic analysis opens up new possibilities in cytometry; long-term and continuous monitoring of individual cells under various conditions, comparative studies of the behavior of different cell types in the same image field, automatic monitoring of stem cell and tissue culture for implantation and organ growth,

operator-free in-the-background gathering and analysis of large amounts of single cell data and a better analysis of especially sensitive tissue material or material restricted by ethical considerations. One such case is the clinical study of human embryo growth.

Embryos as well as the individual embryonic cells display highly individual growth characteristics and are an excellent model for three-dimensional inter-cellular interaction and for developmental biology. Human embryos are particularly difficult to study because of ethical concerns. At the same time the study of human tissue has the greatest relevance to the understanding of human health and development. Many studies can be performed on animal tissue but differences exist, and in order to extrapolate correctly from animal models, it is necessary to understand the differences between the human and the animal model. Also, for long-term monitoring of human embryos for fertility treatment, no alternatives to non-invasive analysis exist. The procedure for embryo evaluation is today to a large extent performed as manual assessment, due to a lack of properly evaluated software tools. This is a time-consuming and subjective process requiring expert knowledge. In recent years advances in computation and hardware have aided to overcome many of the technological issues in continuous cell and embryo monitoring, and biology is slowly changing from a qualitative to a quantitative field. With an increase in data gathering and availability of data storage, the problem has shifted from limitations in data generation to limitations in data analysis and understanding.

1.2   Definition of Aim

The overall aim of this Thesis is:

The development of a technological base for a unified framework for image analysis of human embryos in vitro.

1.3 Background

1.3.1 The Cell

The eukaryote cell is a complex organism of compartments, membranes and vesicles, controlled by its genetic code. Each cell in our body carries the same genetic code and many cells function in collaboration using an intricate system of signaling pathways. All cells originate from the same single cell, the *zygote*. During development, the zygote divides into several cells forming an embryo. In its early stages it consists of a small body of identical daughter cells. At this stage every cell has the capacity to form any cell type in the body. As an embryo grows into a fetus, the cells successively differentiate and are eventually only capable of carrying out the specialized tasks of their own tissue. Several tissue types are then organized into organs, which function together to carry out our bodily functions.

A cell multiplies through the process of cell division, or *mitosis*. In mitosis, the cell grows in size and duplicates its DNA in the replication process. Finally the DNA, organized in chromosomes, is divided between the two daughter cells. The sequence from one mitotic division to the next is referred to as the *cell cycle*, a process carefully regulated by several signal substances. *Gametes* (sperm or eggs) are produced by specialized cells in the process of *meiosis*, in which the amount of DNA and chromosomes in each daughter cell is halved.

Given the complicate nature of the cell division and the sheer amount of DNA to be correctly duplicated, it is perhaps not surprising that faults often occur. The cell cycle contains a large number of check points and regulatory mechanisms correcting for replication and division errors. In some cases - when correction is not possible - cell death is induced by means of *apoptosis*. When replication errors occur in an embryo of only a few cells in total, it can lead to the demise of the entire embryo. On occasion the embryo can continue developing normally despite an early division error, leading to syndromes such as Down's. When replication errors occur at later fetal stages or in the full-grown adult body, it can lead to the development of tumors and sometimes cancer.

1.3.2 Visualizing the Cell

The main advantage of manual image analysis is the human's superior ability to recognize properties of interest in images of widely varying appearance. The main drawback is the tedious process of manually marking and counting structures of interest, but also a subjective bias, leading to differences in results obtained from different observers. Using

computing, the throughput can be increased and the subjectivity reduced. On an abstract level, *computer vision* is the use of image data to infer something about the world. In medical imaging, a common task is to identify structures in an image and then infer something about the health of the patient, the effect of a drug, or identifying positions of anatomic or cellular structures. Image cytometry is the computerized measurement of cell properties using images. Properties measured can be anything from counting cells or cellular structures, to performing statistical analysis of pixel values and extracting shape information from relevant areas. Measurements can be performed manually or with the aid of a computer, in which case we call it *digital image cytometry.*

In non-visualizing imaging techniques and tomography, such as magnetic resonance imaging, techniques for visualizing data have existed for some time. In microscopy the gathered data is directly visualized in two dimensions and few other ways to present the data have evolved. Presentation tools are usually limited to a variety of two-dimensional image displays and interactivity is often lacking. Many methods exist for enhancing the visual appearance of cells. Some of these make use of attaching fluorophores, dyes or other contrast-enhancing agents to visualize structures of interest within a cell. Dyes and flurophores are added to cells in a complex chemical process, often involving several steps of incubation and washing, which takes a lot of time and makes the analysis costly and cumbersome. In some cases, the use of dyes is not possible at all because of the nature of the sample or its intended use after analysis. Examples include living tissue intended for transplantation, *in vivo* imaging, or imaging of tissues where the dye or the process of attaching it can be expected to interfere with measurements. Moreover, the process of adding the dye is a step in the analysis itself and introduces further sources of error. All sorts of interference with the sample must be assumed to influence the analysis in some way, and we can never be completely sure how representative the image of a dye is of the prevalence of the underlying structure.

Over time, numerous methods have evolved to non-invasively enhance the contrast of biological specimen. A few of the most common non-invasive microscopy techniques are described in Chapter 2.2. Non-invasive microscopy is appealing because of the inherent ability to study samples completely without interference, and the possibility to study cells in their natural environment. Label-free analysis also has the potential of significantly reducing laboratory preparation time and lowering cost, especially if combined with high throughput computer aided analysis.

1.3.3   Time Lapse Imaging

Time lapse imaging (TLI) is the capture of images in sequence to register a temporal process within the sample. Time lapse imaging is desirable not only for the study of dynamic processes, but is in fact necessary when doing automatic image capture, in order to make sure all events of interest are recorded. The philosophy being: if we do not know when an event occurs, the best approach is to capture the whole process. Some technical challenges exist when performing automatic time lapse imaging of biological samples. One is the interference between the imaging setup and the incubation chamber used to keep the sample under favorable growth conditions. Either the optical set-up must incorporate a climate chamber to accommodate the living cellular material or - if the microscope is instead meant to sit inside an incubator or other external chamber - the optics must be shielded to protect it from the humidity, temperature and pH of the chamber. When imaging over prolonged periods of time, technical limitations such as drifting focus must also be considered. Another reason for caution is potential sample damage due to repeated exposure to the light source. There is a trade-off between information gathered and potentially harmful sample exposure, and the frequency of image capture must be carefully chosen depending on the study end-point and the expected frequency of the dynamics under study. In the case of simultaneous monitoring of multiple samples, two solutions exist. In *scanning*, either the imaging hardware or sample is moved and re-positioned at each image capture. In this case there is a trade-off (limited by the moving mechanics) between samples imaged and images captured per sample. In *full-field* the image captured includes all samples simultaneously. In this case, there is instead a trade-off between the number of samples imaged and the image resolution available to each sample.

1.3.4   The Embryo: A Three-dimensional Tissue Model

A developing embryo is an example of a three-dimensional model for cell-cell interaction. Embryos as well as the individual embryonic cells display highly individual growth characteristics with interactions in three dimensions. An *embryo* is defined as the phase in a baby's development from the time of the first mitotic splitting of the zygote into two cells until the end of the eighth week of gestation after which it is called a *fetus*. The developmental stages of the early human embryo can be generally described (Figure 1.1), but the growth pattern of every embryo is unique and little is know about the details of the mechanism behind it. Throughout embryo growth and until the moment of hatching, the embryo is surrounded by a protective enclosure, the *zona pellucida*. The first stage - the zygote - has two pronuclei; one originating from each parent, which
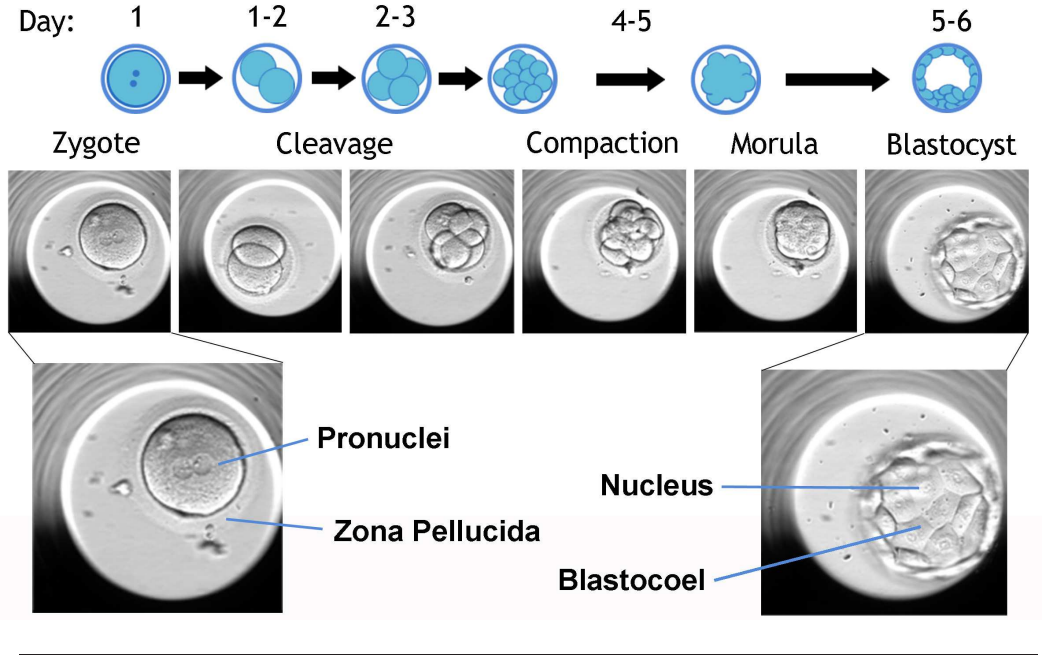
FIGURE 1.1: Overview of embryo growth [5].

will both disappear in preparation for the first mitotic cell division, a process known as *syngami*. The embryo now enters the cleavage stage, where a number of successive cell divisions increase the cell count. A cell within the embryo is called a *blastomere*. After a few days, the cells start to interact and form close membrane-to-membrane connections with one another. At this stage it is called a *morula*, and individual blastomeres are now no longer visible under the microscope. In the morula stage, cell division continues and eventually the cells will rearrange around a hollow sphere; a *blastocyst* is formed. The compartment enclosed within the cells is called the *blastocoel*. Already at this stage, the cells have been targeted to perform specialized tasks and this division of labor is visible in the early blastocyst with one layer of cells, the *trophectoderm*, forming the sphere, and another, the *inner cell mass*, located to one side. Eventually, the inner cell mass will develop into the fetus, and the trophectoderm will form the anchorage to the wall of the uterus.

1.3.5   Embryo Time Lapse Imaging

In Vitro Fertilization (IVF) is the process of extracting gametes donated by a couple, performing the fertilization outside the body and after a few days growth, implanting the young embryo in the uterus of a woman where it - if all goes well - attaches and grows into a healthy fetus. Despite 30 years of practice, the success rate for implantation of embryos into the uterus in IVF is still only around 30% [10, 11]. Thus, when transferring embryos from *in vitro* culture and implanting them, it is critical that only

the highest quality embryos are selected. Selecting the embryo with the highest chance of survival will not only optimize the chance of live birth, but also reduce the need for multiple embryo transfer, with the subsequent risk of twin pregnancy and the neonatal complications and maternal pregnancy-related health problems associated. Embryo selection can be performed based on a number of criteria and much research has been done in identifying morphological features correlated with embryo health. Other methods for embryo selection exist, such as genetic screening and metabolic profiles of culture media, but have not been proven to increase pregnancy rates [12–18]. Discussions concerning the relevance of embryo morphology in quality assessment exist [19], but it is likely that it will continue to play a large part in IVF embryo evaluation also in the future.

Prior to implantation, the embryo is kept in an incubator for a time period of 2-5 days. Traditionally, embryo quality assessment has been performed by manual inspection using light microscopy at intermittent time points during embryo development. Identification of key points in embryo development, along with annotation of images, is one tool commonly used by embryologists to make it possible to share and archive the information leading up to diagnosis. A combination of incubator technology with non-invasive imaging has recently made it possible to monitor embryos continuously during the course of their development without any known consequences to their health. At the same time a lot of image material is being generated in IVF clinics around the world, covering precisely a period in embryo growth of which we currently know very little. Instead of studying the embryo at fixed times, the embryo can now be monitored continuously, making it possible to study morphokinetics and assess the dynamics of embryo development in a way that has not previously been possible. It has since been shown that the timing of key occurrences within the embryo can vary greatly between embryos that have similar morphological appearance at the conclusion of the recording period and that embryo morphology can change in a matter of hours [20–23], emphasizing the fact that dynamic monitoring is preferred over intermittent monitoring of embryos.

1.4   Objectives

1.4.1   Definition of Objectives

An automated, standardized framework for embryo selection using TLI must address several questions. First it must be concluded if time lapse monitoring in itself can be regarded as without effect on embryo health, and second if dynamic monitoring will provide any new information. Regardless of whether time lapse is used or not, a number of standardized criteria must be decided upon. The current selection of embryos is primarily based on morphology and a unified evidence-based objective ranking strategy is lacking, though attempts have been made to work out such a strategy [24]. Ways to automate the embryo image analysis process and complement the manual work would save time in IVF research, and could also introduce new ways to assess embryo quality as well as providing a tool for standardization.

When validating the potential of any new method to predict embryo health, it must also be defined what is meant by embryo health and embryo quality. The embryo may appear to be of top quality and successfully develop a blastocyst during culture but still fail to implant. If we are only concerned with the effects of *in vitro* culture, the formation of a high quality morphological apperance may be enough, but ideally, successful pregnancy or live birth outcome may be a more realistic measure of the embryos potential to survive.

In fulfillment of the aim of this Thesis (Chapter 1.2) we formulate the following four objectives:

1. What choice of hardware has the best potential for automatic embryo analysis?
2. What endpoint should be used when evaluating the selected criteria?
3. What additional data is available using TLI compared to intermittent monitoring?
4. Which criteria are most relevant for automatic monitoring?

1.4.2   About this Thesis

Chapter 2 briefly addresses Objective 1-3 and describes the current state-of-the-art. The remainder of this Thesis is focused on the fulfillment of Objective 4 by evaluation of the criteria selected in Objective 3, using methods from computer vision and probabilistic theory. Chapter 3 summarizes the results when methods are applied to embryo imaging, referencing enclosed papers. The methods used are not limited to embryo monitoring, which is shown in Chapter 4 describing one additional application in toxicology as well as a possible extension to the embryology framework using a combination of non-invasive

and invasive imaging methods, with applications in developmental biology. Chapter 5 includes a discussion of the implications of the findings, an outlook on the future and suggestions for continued work.

Chapter 2

Methods

2.1 Addressing the Objectives

2.1.1 Objective 1: Choice of Methodology

We can require that a framework for long-term non-invasive cell monitoring fulfills at least the following criteria:

- **Environment:** Use a hardware and environment constructed to sustain the sample for the time of study. The imaging procedure, sample mount and light exposure must be without effect to the sample. The system should also be in need of low maintenance and capable of imaging during long periods of time without operator supervision. This includes insensitivity to focal drift, no need for frequent change of cell culture medium or other environment and an enclosed chamber to avoid contamination.

- **Imaging:** Use an imaging technique capable of non-invasive enhancement of cellular structures, producing images suited for digital analysis. The images should contain low levels of noise and optical artifacts and at the same time display a high level of detail.

- **Software:** Use a digital analytical framework capable of sufficient robustness to handle real-time image analysis of clinically produced images.

Under optimal conditions, there can be a dialogue between image acquisition and digital analysis so that the conditions of image acquisition can be improved to optimize for digital analysis. In a clinical environment this is not always possible, and it is here

beneficial if the digital process is robust enough to handle a broad range of imaging conditions. All images used in the work leading to this Thesis were acquired in either clinical or laboratory settings outside the control of the software analyzer. This limited the possibility to adjust the image acquisition at the hardware stage, and provided a situation close to that of clinical data collection. The main objective was to investigate to which extent these real world images could be used to extract meaningful information. This Chapter gives a brief overview of the methodology chosen when targeting the four objectives described in Chapter 1.4.1. Some information has been extracted from literature and in discussion with collaborators in embryology and cell biology. The methods developed have been mainly in software and are described in further detail in Chapter 2.3. Important for the successful implementation of image analysis techniques has been the understanding of the imaging hardware and its characteristics. Chapter 2.2 gives an overview of relevant hardware techniques.

2.1.2 Objective 2: Choice of Study Endpoint

The work with this Thesis aimed to develop a model to rank embryos, not to calculate chance of success, therefore considering embryo parameters only and leaving out couple and treatment characteristics. When aiming to rank embryos within the same IVF treatment cycle it is of little interest to have perfect discriminating capability with implantation potential (*i.e.* live birth outcome) and in this context, a relative score is sufficient for comparison. Many studies use blastocyst as endpoint [33–38]. It has been argued [39] that using blastocyst formation as a predictive endpoint has limitations, since it yields a high number of embryos deemed of lesser quality, which may still result in implantation. Therefore blastocyst development is a surrogate endpoint at best, but can still be useful in the selection of positive selection criteria. Targeting Objective 2, the choice of the formation of a blastocyst was chosen as a sufficient endpoint when evaluating criteria for a software analytical framework.

One benefit of this choice is the potential of being able to predict blastocyst formation. Transfer of blastocysts generally yields higher implantation rates than transfer at the cleavage stage, but the higher implantation rate must be weighed against the potential drawbacks of longer culture, so a way to predict blastocyst formation would provide a way to reduce culture time, while at the same time maximize the chance of success. If the same criteria would apply when using live birth as an outcome, remains to be seen.

2.1.3 Objective 3: Choice of Study Target
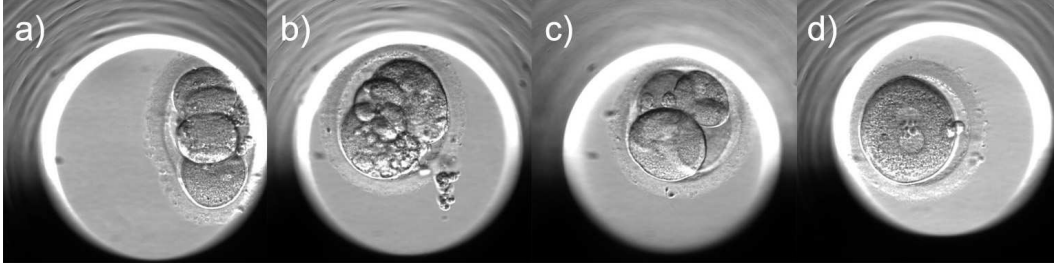
FIGURE 2.1: Examples of embryo instances posing analytical challenges. a) Deformed embryo located close to the container wall. b) Fragmented embryo. c) Unevenly sized blastomeres. d) Multinucleation is normal in the case of the zygote, but at later embryo stages can be a sign of replication problems.

Several studies [33, 34, 40] have addressed the potential role of time-lapse monitoring in clinical selection of competent embryos. The majority have investigated the correlation between the timing of mitotic divisions and embryo quality [33, 41–47]. It has been demonstrated that early cleavage is correlated with embryo survival because early cleavage embryos are likely to cleave more evenly [35, 48]. A relationship has been observed between either prolonged or delayed duration of the 2nd to 3rd mitosis and blastocyst formation [49], and further examples of morphological criteria include pronuclear appearance and orientation [50], embryo respiration [17], chromosomal abnormalities [51], degree of fragmentation [52], number, size, shape of blastomeres, degree of blastocoelic expansion, cellular composition and compactness of the inner cell mass and trophectoderm [53]. It has also been shown that uneven blastomere size (Figure 2.1b) and multinucleation (Figure 2.1d) can preclude implantation [43]. Recently, a study [54] concluded that rhythmic contractions induced by sperm entry in the zygote of mice may predict embryo viability. These vibrations took place at a scale of seconds. See [55] for an overview of criteria.

Making use of the proposed criteria, there are several proposed prediction models for embryo ranking described in literature, giving a good indication of the most promising morphological and morphokinetic criteria for embryo health, but it remains to be seen whether the new data available (embryo dynamics, timing of events *etc.*) is also relevant to embryo health. Most models were developed on small data sets [56–60]. A recent review [61] concluded no significant effects to embryo selection when based on data from TLI. A few larger studies include Loendersloot [62] and Racowsky [63] where the correlation of eight potential predictive criteria for embryo quality in 6021 embryos was studied. Ciray [24] proposes guide-lines for a standard methodology and terminology for TLI use, but before adoption into clinical practice, any framework be subject to a

number of larger, randomized clinical studies, preferably as prospective cohort, multi-center trials.

TABLE 2.1: Selected study target embryo traits

Embryo event	Target
Embryo location	Tracking of position to facilitate analysis
Syngami	Providing reference point in time
Timing of mitotic splittings and interval between mitosis	Embryo health and development
Timing of blastocyst formation	Embryo health, development and providing study endpoint
Blastomere number	Embryo health and development
Embryo movement and activity	Embryo health
Embryo fragmentation	Embryo health

Table 2.1 summarizes the embryo structures focused in this Thesis. Table 2.2 summarizes in more detail the current proposed criteria known from literature. Also included in both tables is the location of the embryo within the growth container. Though being assumed of no importance to embryo health, knowing the location makes it possible to extract image information continuously in the presence of vibrations and dislocation of embryos during the necessary procedure of replacement of growth medium. Column 3 and 4 indicates expected visibility and detectability in HMC images [64]. The chosen criteria in Table 2.1 have been selected primarily based on their expected detectability in the HMC images produced.

TABLE 2.2: Summary of embryo selection criteria for digital image cytometry

Embryo trait	Relevance to embryo health	Visibility (HMC)	Expected detectability (HMC)	Requires TLI	Required temporal resolution	Ref.
Embryo location	no	yes	high	no	$\leq$ 10 min	[33–35] [36–38]
Timing of mitosis	yes	yes	high	yes	$\leq$ 10 min	[34, 47] [43, 52]
Interval between mitosis	yes	yes	high	yes	$\leq$ 1 h	[42, 44] [43, 47]
Duration of mitosis	yes	yes	medium	yes	$\leq$ 1 min	[33]
Blastomere size, shape	yes	yes	low	no	-	[43]
Blastomere number	yes	yes	low	no	-	[52, 62]
Timing of compaction	yes	yes	high	yes	$\leq$ 1 h	[62]
Fragmentation number	yes	yes	medium	no	-	[34, 52]
Embryo respiration	yes	no	-	no	-	[17]
Embryo activity, movement	yes	yes	medium	yes	1 s–1 min	[54]
Chromosomal abnormalities	yes	no	-	no	-	[51]
Nuclei appearance	yes	yes	medium	yes	$\leq$ 1 min	[22]
Nuclei number	yes	yes	medium	no	-	[43]
Timing of pronuclear formation	yes	yes	high	yes	$\leq$ 1 min	[41, 50]
Pronuclear orientation	yes	yes	medium	no	-	[50]
Pronuclear size, shape	yes	yes	low	no	-	[41]
Timing polar body formation	yes	yes	very low	yes	-	[41, 50]
Polar body orientation, number	yes	yes	very low	no	-	[50, 51]

2.1.4 Objective 4: Evaluation of Criteria

In this Thesis, work has been focused on selecting relevant criteria in Objective 3, develop methods to detect them, and finally evaluating those methods (Objective 4) using the imaging hardware and study endpoint chosen in Objectives 1 and 2, respectively. In the evaluation of results, images have been compared manually, with the help of embryology expertise. It is noteworthy that trials which compare the computer aided quality assessment with a manual assessment should ideally be made on the same data set and using the same incubator, lest it will also indirectly compare the quality of incubators. In the studies described in this Thesis, the comparative evaluation has been done on the same images as the computer analysis, thus avoiding the problem of separate experimental set ups. Instead, it introduces a reason for caution as the manual evaluation is made on image sequences different from the manual microscopy procedure normally used by embryologists and there can have been discrepancies caused by the fact that:

1. There is a limited number of focal planes, and embryologists are confined to study the images after capture, with no manual real-time control of focus.

2. In the experimental set up used, the embryos have a tendency to migrate towards the edge of the container, causing image artifacts, and making them harder to evaluate, by both human and computer (Figure 2.1a).

3. To the human, the embryo captured with the time lapse imaging does not appear the same as under a traditional microscope due to differences in imaging technique and lighting conditions, causing uncertainty in manual evaluation due to lack of experience.

The following Chapter 2.2 gives and overview of non-invasive microscopy imaging and Chapter 2.3 describes of digital image processing methods and the implemented variations of each used in this Thesis.

2.2   Non-invasive Microscopy

As an alternative to using contrast-enhancing agents, there exist methods to optically enhance the contrast of relevant structures in the sample. This section contains a brief introduction to a few commonly used full-field non-invasive microscopy techniques. In Chapter 2.2.6 follows an analysis of the case of embryo imaging and a motivation for the choice of HMC (Chapter 2.2.4) as imaging technique.

2.2.1   Brightfield Imaging
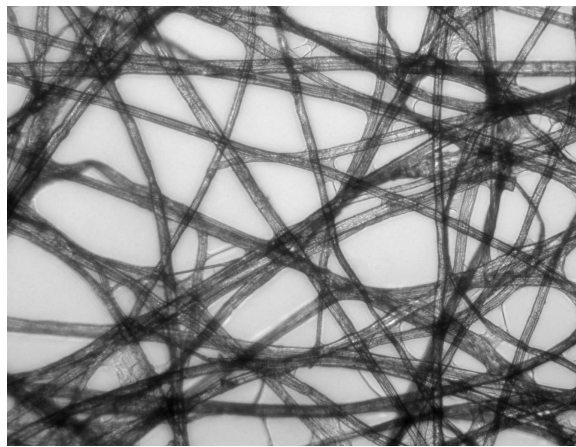
FIGURE 2.2: Texture fibers viewed by brightfield imaging [127].

Brightfield imaging is a collective term for the simplest form of microscopy where light is transmitted through a sample and recorded by a camera. A brightfield microscope is very simple to set up and use, but does not have the ability to image non-pigmented specimens such as living cells. The sample appears against a bright background and the contrast is determined by the ability of the sample to absorb the wavelength of the light transmitted. Brightfield microscopy is best used with naturally pigmented samples (Figure 2.2). Living cells are often close to transparent in white light and samples are then stained to enhance contrast. In addition, several optical techniques exist, capable of non-invasively enhacing the contrast of the sample. Some of those are so common that they are loosely referred to as "brightfield microscopy", even though each technique has very distinctly different characteristics. Two common techniques are described below; Phase Contrast microscopy (Chapter 2.2.3) and Hoffman Modulation Contrast microscopy (Chapter 2.2.4). The former is common in *in vitro* cell culture imaging and the latter often used in embryo imaging.

2.2.2   Darkfield Imaging
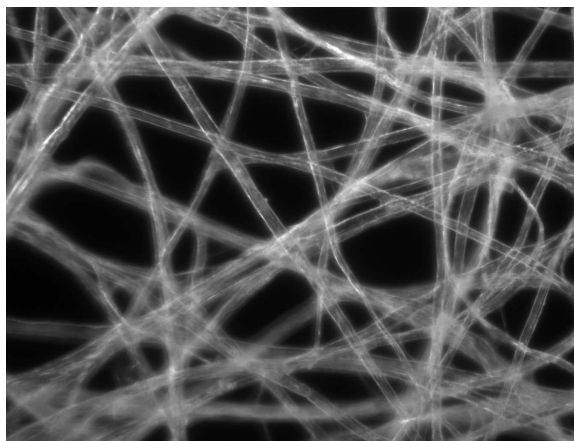
FIGURE 2.3: Texture fibers viewed by darkfield imaging [128].

Darkfield imaging are imaging methods which exclude any unscattered light, causing the sample to appear brighter on a darker background (Figure 2.3) and enhancing the contrast of any imaged (unstained) sample.

A central aperture blocks light directly from the light source and the sample is only imaged by light incident at an angle. A ring aperture is then used to collect light only from the light path along the optical axis, effectively removing the angled light from the source. The image on the sensor is formed only by light which has been scattered by the specimen. Darkfield is a simple yet effective method to non-invasively enhance sample contrast but has the disadvantage of low light levels available for collection. To compensate the sample must be strongly illuminated and the heavy light exposure can cause sample damage. However, the low light level also means the image is almost entirely free from optical artifacts.

Optically, a darkfield image may appear as the negative of a brightfield image, but this is not the case. Structures are imaged differently, and details visible with one technique may be completely invisible in the other. Darkfield microscopy is most useful for studying boundary structures with a high difference in refractive index, thus imaging cell membranes more effectively than internal cell structures. It is best suited for thin samples and with high differences in refractive index (such as for sharp edges) and thick samples, it may produce artifacts. Preparation and quality of the sample can greatly affect the accuracy of the image and sample contamination such as dust or medium debris will easily appear as structures of the sample itself. Darkfield Microscopy is generally not a reliable tool to obtain quantitative measurements of specimens due to the strong dependence on the placement of condenser and aperture and its relation to

image appearance. It has been used for embryo studies in a number of research set ups and commercial products.

2.2.3   Phase Contrast Imaging
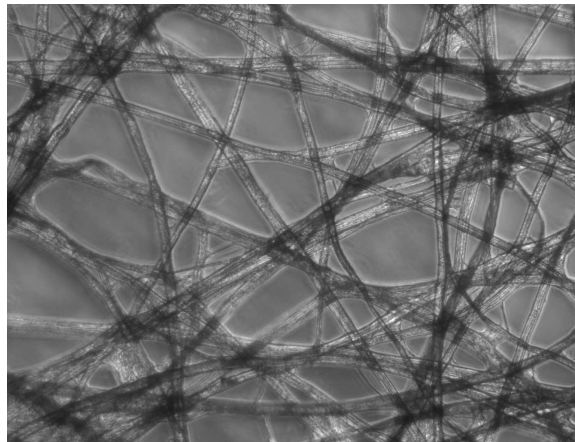
FIGURE 2.4: Texture fibers viewed by phase contrast imaging [129].

The first phase contrast microscope was built in 1938 by Frits Zernike [130]. Today, phase contrast is standard in most modern microscopes. Together with a phase contrast objective, it provides the possibility to non-invasively enhance the visibility of low contrast samples in their natural environment.

In a typical biological sample, the scattered light is phase shifted 90°due to the thickness of the sample and its refractive index difference with respect to the surrounding medium. In the phase contrast microscope the background light is phase shifted another 90°by the means of a phase ring, causing constructive interference between the scattered foreground light and the phase shifted background light, resulting in an increase in brightness in the areas of the image containing a sample. The intensity of background light is further dimmed by a gray filter. Some of the foreground light will also be dimmed, but to much less extent than the background.

Disadvantages of phase contrast microscopy include distortions to thicker specimens and a high degree of image artifacts, the most apparent observed as a halo effect around edges (Figure 2.4), obscuring details along the specimen perimeter. Modern advances exist which help to reduce the halo effect, but it cannot be completely eliminated. In the above described technique and in Figure 2.4, the sample is brighter than the background. In a variation of the set up, the background light is phase shifted by -90°. In this case, the foreground appears darker than the background and the halo is inversed. The halo is the reason for using Phase Contrast Microscopy mainly for single layered cell lines and

other thin objects, and not for embryos and other samples with strong three dimensional character.

2.2.4 Hoffman Modulation Contrast Imaging
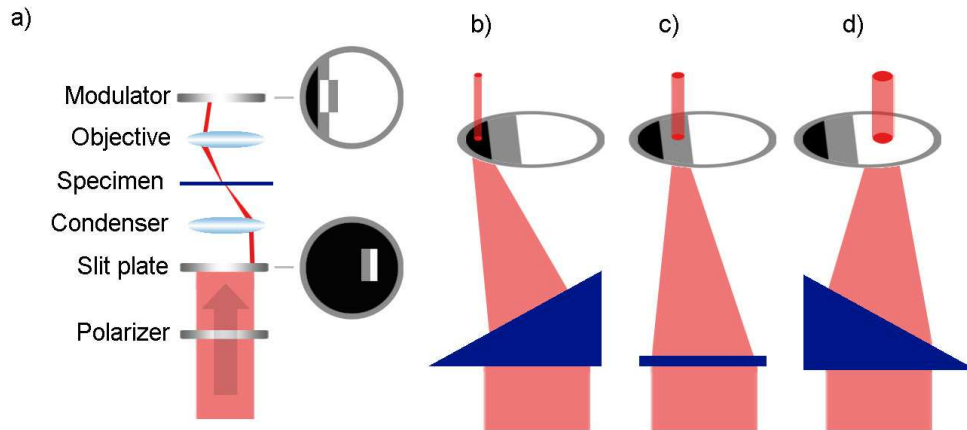
FIGURE 2.5: a) Schematic setup of Hoffman Modulation Contrast Microscopy. Polarized light passes through a semi-permeable slit (bottom), is focused using a condenser and passes at an angle through the specimen. After being collected by an objective, the light passes through a modulator (top) which is a mirror of the slit. Any gradient in refractive index in the specimen will shift the light on the modulator towards more (b) or less (d) attenuation, whereas undisturbed light will be recorded at intermediate (or gray) pixel levels (c)[8].

Hoffman Modulation Contrast (HMC) Imaging was invented by Hoffman in 1975 [131], and is today a common technique for non-invasive contrast enhancement of biological samples, available in many commercial microscopes. Light is passed through a pair of off-axis slits, converting gradients in sample optical path to bands of light and dark appearance, depending on the spatial sample direction (Figure 2.5).

Disadvantages in HMC include optical artifacts, which is particularly apparent when manually examining images, as the angular position of the slits lead to a pseudo three-dimensional effect. The orientation of the specimen has implications for the image since the system is most sensitive to gradients perpendicular to the orientation of the slit. On the other hand images are free from the halo shown in phase contrast images.

Advantages includes good contrast, low light exposure, excellent resolution and a short depth of field, giving the opportunity of focal sectioning at a resolution controllable by the numerical aperture of the objective. The ability to section is also influenced by the homogeneity of the sample. HMC is commonly used for embryology studies, including a number of commercial products.

2.2.5   Digital Holographic Microscopy
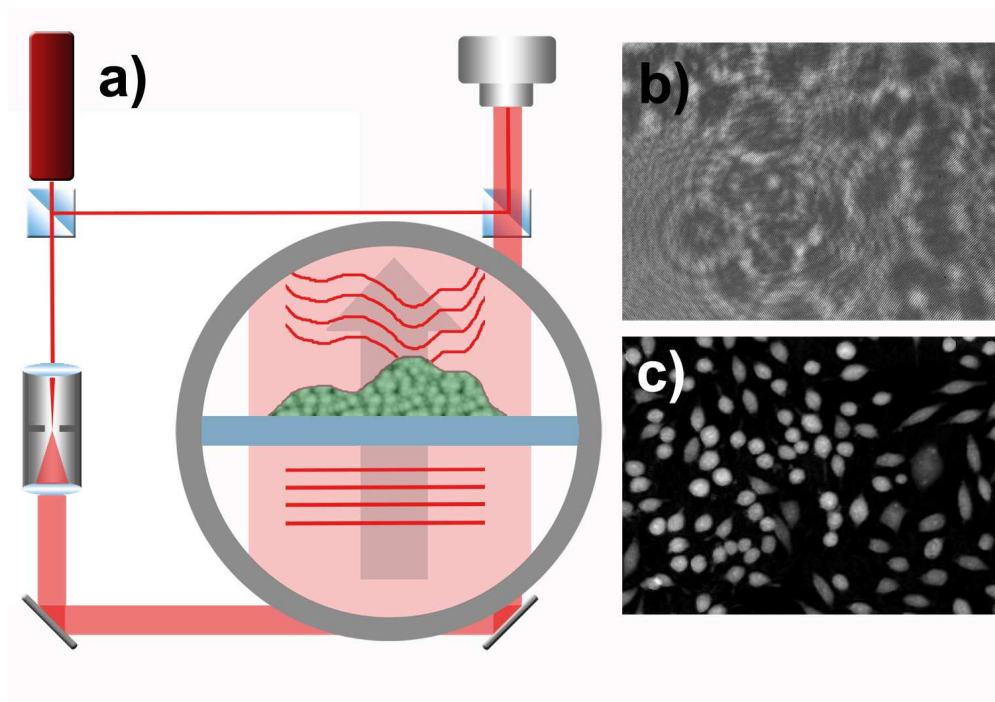
FIGURE 2.6: a) Schematic setup of Digital Holographic Microscopy. Laser light (top left) is divided into two paths; one reference beam (top) and one passing through the sample (left, enlargement). The combined beams form an interference pattern on the camera (top right). A typical pattern is shown in b). Using a succession of Fourier transforms and image adjustment an image of the sample can be computed. The computed image of cells growing in a Petri dish is shown in c) [7].

Dennis Gabor was first to invent a way to encode the phase of light in a single recording, *i.e.* a hologram [132]. The technique was later used as base for the development of digital holography [133–136]. Some years later, the first Digital Holographic (DH) microscopy images showing living cells were published [137].

In DH (Figure 2.6), low intense light from a laser is divided into two paths, one transmitted through the sample, the other providing background reference. On the sensor, light from the two interferes, constructing a pattern characteristic of the sample. An image can be calculated from the interference patters by a set of Fourier transforms, resulting in an intensity map where gray levels are related to sample thickness and refractive index difference between sample and surrounding environment.

Advantages to DH include the quantitative nature of images - gray level intensity can be directly correlated with physical traits, and the ability to visualize un-stained samples in their natural environment. Disadvantages include sensitivity to the settings of the image calculation and image artifacts for samples thick enough to scatter light more than one image wavelength, thus making this technique more suitable for thin samples.

Like Phase Contrast Microscopy, DH is better suited for thin samples such as adherent cell lines and tissue samples of only a few cell layers.

2.2.6 Embryo Imaging

In this Thesis we limit ourselves to a comparison between three systems commercially available for embryo monitoring. Two of those are based on Hoffman Modulation Contrast (HMC) imaging (Chapter 2.2.4), one of which uses scanning imaging [25] and one full-field [26]. The third is a full-field system based on Darkfield (DF) imaging (Chapter 2.2.2) [27]. All of them are combined with incubator technology to provide a beneficial environment for embryo growth. The aim of this Thesis is not to evaluate a commercial product, but rather the potential of the technique, and will therefore not take into account any differences originating from limitations posed by the technological setup, such as choice of light source and cultivation chamber. For an overview and comparison of available commercial systems, see [28]. No difference has yet been seen in growth and implantation rates of embryos grown in the standard intermittent incubator system and a time lapse incubator system [29–31]. One study [32] found a higher rate of miscarriage for the TLI group, indicating there are reasons for caution. However, the same study noted no effect on pregnancy rates or embryo health prior to implantation. At present, it is therefore safe to conclude that the imaging process and the incubation is not a liming factor in the experiment. Table 2.3 summarizes the three technological approaches.

The scanning HMC is the only system where the imaging equipment and incubator is integrated, while the other two are systems where a stand-alone microscope is placed inside a standard incubator. With a sensitive image sensor, the light exposure for both HMC systems can be significantly reduced compared to that on a standard manual HMC microscope, whereas the exposure levels for DF are inherently higher. Where the scanning technology introduces constant movement of the samples, the full-field imaging has the advantage of a completely undisturbed environment. On the other hand, the scanning makes it possible to make full use of the image field, thus providing much better resolution. For the work in this Thesis, scanning HMC has been chosen primarily for this reason. The HMC has a disadvantage in that the images are bi-gradient, which makes edge structures less apparent. On the other hand, it has the ability for focal slicing with the option to reconstruct the entire three-dimensional structure of the sample. Focal slicing with HMC suffers some bleed-through from out-of-focus material, but the defocused information is superimposed on focused image details, which makes it potentially possible to handle in image processing. The resolution of focal slicing is ultimately determined by the relationship between the depth of field and the sample size. The

TABLE 2.3: Comparison of embryo imaging systems

	Scanning HMC	**Full-field HMC**	**Full-field DF**
Imaging method	HMC	HMC	DF
Movement of sample	Yes	No	No
Focal planes	Multiple	Multiple	Single
Images	Grey background, bi-gradient	Grey background, bi-gradient	Dark background, single gradient
Limit to temporal resolution	No. of samples	Image acquisition	Image acquisition
Limit to image resolution	Numerical aperture	Samples per image	Samples per image
Most visible features	Internal cell structure, small structures	Internal cell structure, small structures	Cell membranes, large structures

scanning procedure limits the temporal resolution, but instead provides a higher spacial resolution, maximizing the ability for analysis of internal embryo structure. Table 2.4 summarizes the drawbacks and benefits of using HMC from the perspective of digital image cytometry.

2.3 Digital Image Analysis

In this Chapter is given an overview of the foundation for the methods for image analysis used in the experimental section. The methods range from standard image processing techniques to edge detection, segmentation and feature extraction. Image processing is a general term describing any kind of manipulation of image data. The purpose is usually to provide more favorable conditions for later processing steps. The digital image consists of discrete picture elements, pixels, the value of which represents the image intensity. In a color image, each image consists of three layers, one for each color channel. In this Thesis we are only concerned with gray scale images, *i.e.* all channels carry the same pixel value. For the benefit of future reference we define the original

TABLE 2.4: Advantages and disadvantages of HMC for automated image analysis

Advantages	Disadvantages
High image resolution and high level of detail of internal cellular structure	Image edge hard to detect and difficult to interpret
Low sample light exposure and the ability to adjust imaging wavelength to minimize sample damage	No color information
Focal sectioning possible	Resolution of focal sectioning limited by microscope numerical aperture

gray scale image $F(x, y)$, with dimensions X and Y respectively, as a two-dimensional matrix of gray levels i, to a maximum gray level I. Furthermore we can define mean μ and standard deviation σ of $F(x, y)$ as:

$$\mu = \frac{1}{X \cdot Y} \sum_{x=1}^{X} \sum_{y=1}^{Y} F(x, y), \tag{2.1a}$$

$$\sigma^2 = \frac{1}{X \cdot Y} \sum_{x=1}^{X} \sum_{y=1}^{Y} [F(x, y) - \mu]^2. \tag{2.1b}$$

Several image analysis steps are necessary when making measurements on digital images. In a sense, an image analysis task can be thought of as a basic recipe, which can be varied to accommodate different purposes and where a number of parameters vary a few key settings in the algorithm. For research and analysis tasks, the goal is often to develop algorithms of sufficient robustness to make the results reproducible on a large image set with as little manual adjustment of parameters as possible. Only when no parameters need to be set can the algorithm be called *automatic*, otherwise it is referred to as *semi-automatic*. The exact implementation of each step need to be adjusted depending on the nature of the images, the application and the intended outcome of the analysis. A complete overview of methods is beyond the scope of this Thesis, but a few key steps are commonly present and will be described further in the following sections. The methods can be divided into:

- Per pixel translations, where all pixel values in the image are replaced by another value. An example is the application of an image filter (Chapter 2.3.1), or a

segmentation process (Chapter 2.3.3,2.3.5).

■ The location of interest points or regions, such as edges (Chapter 2.3.2) or image features (Chapter 2.3.7).

■ Dimension reduction, where the image information is compacted using a set of descriptors (Chapter 2.3.7).

There are reasons for caution in image manipulation since even a simple transformation may affect the result of the final analysis. On the other hand, some pre-processing may be necessary, especially when comparing images obtained from different sources with highly varying lighting and capture conditions.

2.3.1   Filters

After applying a filter, the new pixel values $F'(x, y)$ consist of the weighted sum of the intensities of pixels in the surrounding area, defined by the size of the filter $g(x_g, y_g)$. The weights are stored in the filter *kernel*, a matrix with entries

$$F'(x, y) = \sum_{x_g=1}^{X_g} \sum_{y_g=1}^{Y_g} F(x - x_g, y - y_g) \cdot g(x_g, y_g) \tag{2.2}$$

By convention, a filter is applied by convolution, so that the top left weights the image pixels in the bottom right. If the filter is symmetric, this makes no practical difference. If the filter is asymmetric, it can be used to detect directions in the image. A commonly used symmetric filter is the gaussian filter used for blurring:

$$g(x_g, y_g) = \frac{1}{2\pi\sigma^2} exp(-\frac{x_g^2 + y_g^2}{2\sigma^2}) \tag{2.3}$$

The Prewitt [138] filter is a simple approximation of a first degree derivation:

$$\widehat{P_x} = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix} \tag{2.4a}$$

$$\widehat{P_y} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} \tag{2.4b}$$

while the Laplacian filter [139] is an approximation of the Laplacian operator $\widehat{L} = \nabla^2$:

$$\widehat{L} = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} \tag{2.5}$$

In practice, the Laplacian operator is often combined with a Gaussian smoothing in order to reduce noise. Due to the associative nature of convolution filters, the application of a Gaussian followed by a Laplacian filter to an image is equivalent to the application of a Gaussian to the Laplacian filter and applying the resulting *Laplacian of Gaussian* filter to the image.

2.3.2 Edge Detection

In microscopic imaging, polarization effects from the optics must be taken into account when applying asymmetric filters. In a DH or fluorescent image the foreground cells are brighter than the background. The image is unpolarized, and we will refer to this type of image as *single gradient*. In an HMC image on the other hand, foreground pixels along structure borders are both brighter and darker than the gray background. The reason is optical polarization; the angle of incident light produces a shift in edges in one half of the image compared to the other half. The direction of change in optical path length is given by the polarity of the artifacts. We will refer to this type of image as *bi-gradient*. Applying an unsymmetric filter to a bi-gradient image can have unforeseen effects and when applying a gradient-based edge detector, the bi-gradient image displays double edges, if edges are at all located.

In the HMC microscope, light is passed through a pair of off-axis slits, converting gradients in sample optical path to bands of light and dark appearance, depending on the spatial sample direction (Figure 2.7). In this example, the slit pair is positioned so that image gradients appear symmetrical around a horizontal axis. This effect is most apparent when plotting the image as an isocontour (Figure 2.7b). This will result in ambiguity when applying asymmetric image operators, such as a derivative (Figure 2.7c-d) and certain edge detectors. A Canny edge detector is based on the application of two Prewitt filters (Eq 2.4) in x and y-direction, respectively. Applied to the raw HMC image (Figure 2.7e) it is clear that the filter may find edges on both sides of the lighter and darker bands, resulting in an uncertainty when trying to determine the location of the border of the embryo or of a single blastomere. Since the direction of light depends on the azimuthal angle between slits and sample, this effect can be reduced at the hardware stage by rotating the sample around a vertical axis and combining information
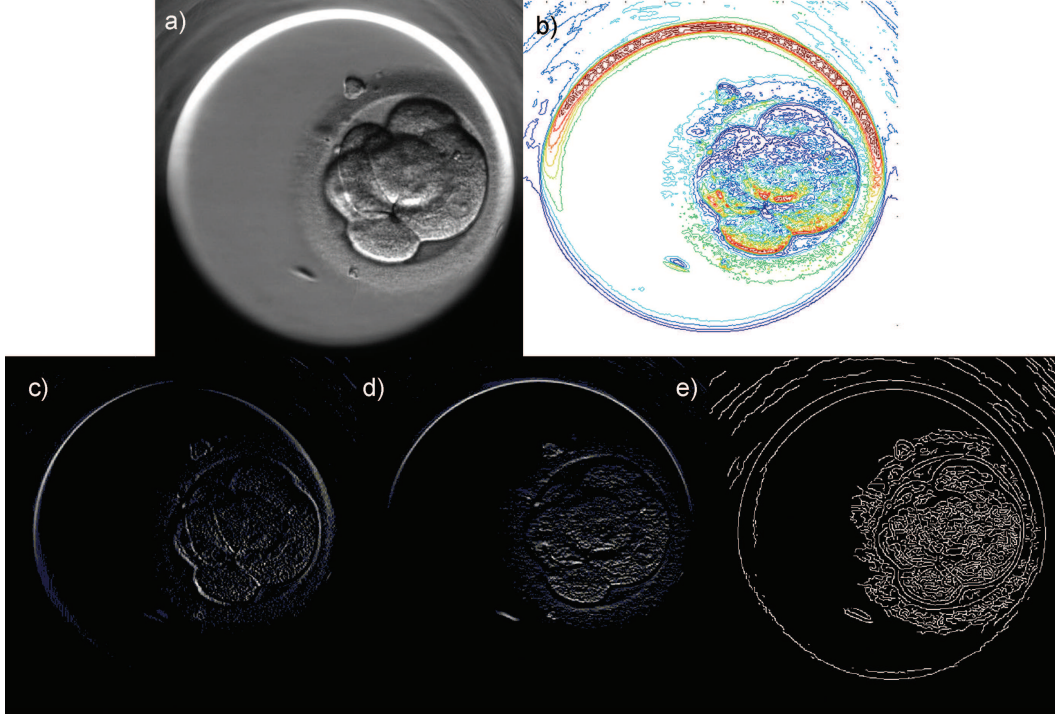
FIGURE 2.7: Edge detection in HMC images. a) Original image. Bright outlines at the top of the well show a polarity opposite to that of the embryo, where the bright outlines are at the bottom. b) Contour. c-d) Prewitt filter applied (Eq 2.4), approximating derivation in x and y direction, respectively. e) Canny edge detector applied [8]. The location of edges is shifted and broadened as a result of the optical artifacts.

from several images along the rotation. When hardware adjustment is not possible, edge detection must proceed by other means.

If foreground and background pixels have differing ranges of pixel values, we can define a suitable intensity threshold separating the two. There are several methods for determining a suitable threshold. Many of those are histogram based, taking advantage of the fact that background and foreground pixels are more abundant than indeterminate pixels. In this case, foreground and background pixels will display as peaks in a histogram, separated by a local minimum, where the threshold can be placed. This is the case in single gradient images. In bi-gradient images, the histogram will contain two minima instead of one, and the background pixels will occupy the central peak. In those cases, a double threshold can be used to extract edges (Figure 2.8). A method for extracting edge information in bi-gradient images is evaluated in Paper III [3]. Here, the edge map was constructed from the lowest and highest pixel ranges in the image I using the threshold

$$t = c \cdot median(I). \tag{2.6}$$

FIGURE 2.8: Grey scale distribution in DH (a-c) and HMC (d-f) images. a,d) Original image. b,e) Histogram over pixel values. Threshold(s) have been marked, one in DH image (b), two in HMC image (e). c,f) Thresholded image at the threshold defined in b and e, respectively.
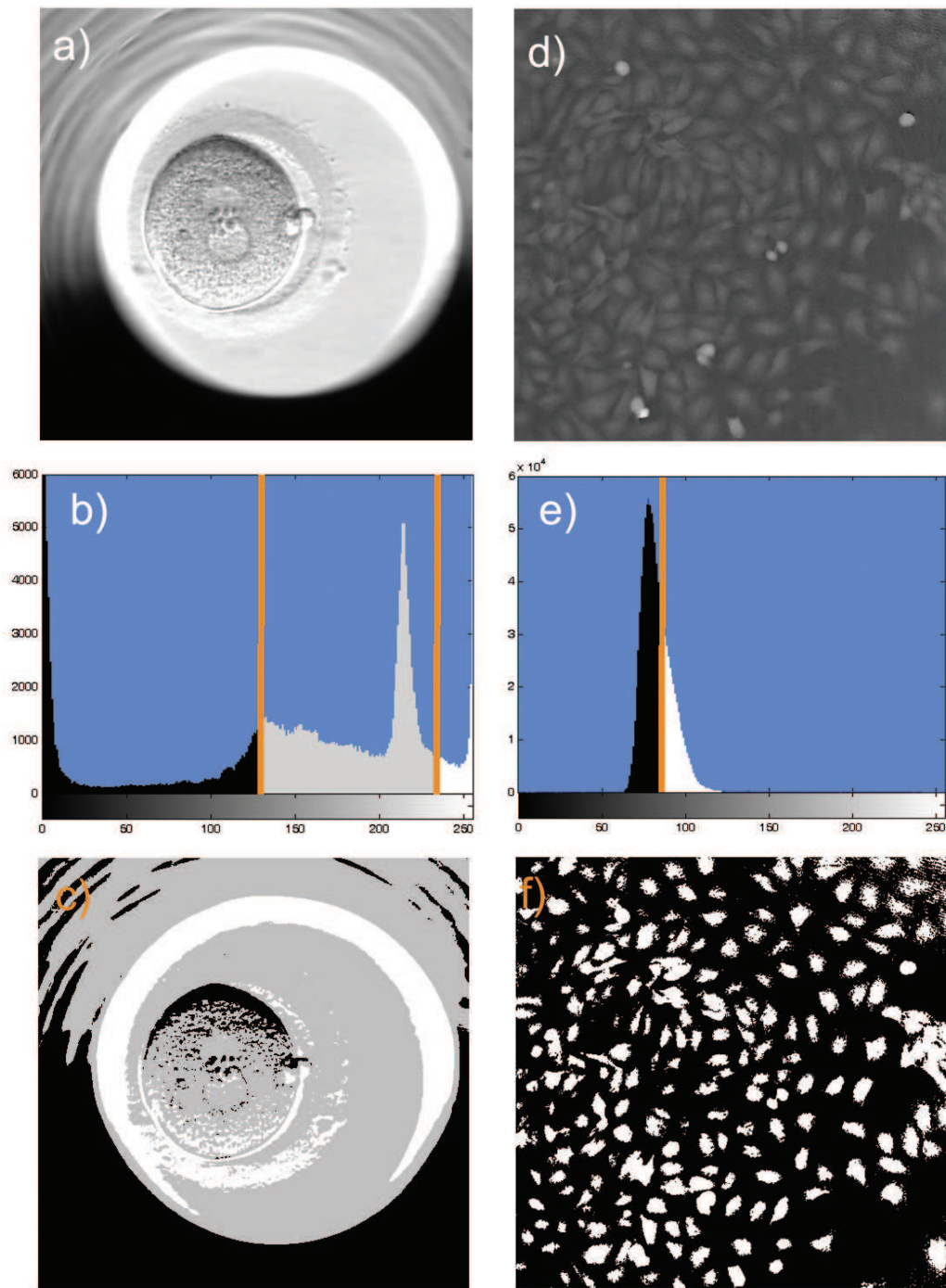
If applied to the image as it is, this results in a high number of false positive edges. To filter those out, a focus filter (Chapter 3.3.1.1) was constructed and used to compute a convex hull around regions in focus which in turn was used to exclude false positives.

2.3.3   Segmentation

A segmentation is a transformation which assigns labels to different regions of the image, identifying them as contents of separate meaning. A simple example of segmentation is a threshold, separating image pixels into foreground and background, and a more complex to separate the foreground into different segments [140]. What is defined as foreground may vary and is determined by the purpose of the study. Segmentation can be seen as a classification task in itself, but in the work described in this Thesis, segmentation is used as a pre-processing step, identifying relevant image content.

When thresholding is not sufficient to identify object contours, a common approach is the application of an edge finding filter. It exploits the fact that pixels values change more rapidly at edges. The basis for this type of filter is often the first or second image derivative, or an approximation thereof. Simple thresholds or edge detectors will not help us in separating clustered objects from each other. In the case of DH images, the brighter foreground pixels are often locally less bright at object borders, and we can use this to separate clustered cells. A robust segmentation technique for this kind of problems is the watershed segmentation [141]. It can be intuitively understood if we think of the image as a landscape, being immersed in water. In the case of a DH image, the pixel values are first inverted, so that the center of each bright object represents a local minimum. A level (the water) then rises from each local minimum, forming a basin. When water from two basins meet, the border between two objects is defined. All pixels associated with the same basin are then given the same label, and are regarded as belonging to the same object.

2.3.4   Segmentation vs. Segmentation Free

There are many examples of segmentation in microscopic images [68–75] and in non-destructive *in vitro* cell imaging, segmentation has been performed in a number of studies in Darkfield [69, 76], Digital Holography [77], Phase Contrast [68, 70, 78–81] and Bright Field [82]. In HMC, several attempts have been made to detect embryo traits using segmentation [83–86] by manual selection of a region of interest (ROI) [87–89] or by a combination [76] with varying degrees of success. Because of the bi-gradient nature of HMC, images are hard to segment accurately using traditional edge-based methods, watersheds, thresholds or level sets, and have so far been limited to the 4 cell case [83].

The segmentation is often simultaneously the most vital and the most difficult step in an analysis task. The result of the segmentation is critical for the outcome of the analysis, and a manual segmentation is impractical for large image sets. For some tasks where the image data set is highly diverse it may be very difficult to construct a segmentation algorithm which is robust enough to perform well across the entire set. Under clinical conditions, images are often hard to segment due to presence of clutter, poor lighting and imperfect focus. In these cases, manual segmentation can be used, identifying regions of interest before calculation commences. For clinical conditions, a cumbersome initialization of calculation does not necessarily speed up the workflow. Instead, a segmentation-free approach may be more robust, but instead not provide the same detailed information. In the case of cells, many important analytic end-points such as information on cell size and shape, and calculation of cell lineage require full segmentation of the image. In some cases, it may be possible to use a shape-controlled segmentation, not requiring perfect edge detection but instead exploiting the fact that the sought structure is of known shape (Chapter 2.3.5). Other study endpoints, such as the number of cells and degree of fragmentation can sometimes be extracted without segmentation using image descriptors (Chapter 2.3.6).

2.3.5 The Hough Transform

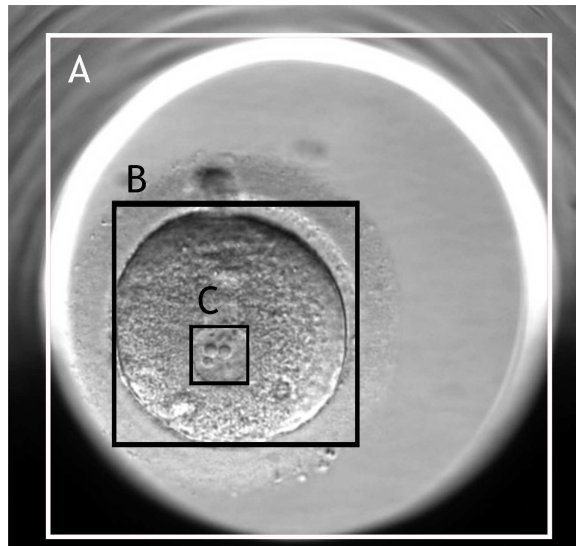FIGURE 2.9: Annotated raw image of an embryo, indicating structures of interest. A − well outline. B − zygote boundary. C − pronucleus [2].

In Figure 2.9, three typical structures in an HMC embryo image are outlined. They all have a close to circular shape, a fact which can be exploited using shape-based segmentation in the form of a Hough Transform (HT). The HT detects possible circular structures

in an image in decreasing order of strength (Chapter 2.3.5). The HT benefits from being able to detect also structures with partial or fragmentized outlines, an advantage under clinical circumstances, where high levels of noise can be expected. The HT operates on an edge image and here we can choose a gradient-based edge detector such as Canny edge, if edges are not extremely thick or with a high difference in refractive index, in which case we risk detecting double edges (Chapter 2.3.2).
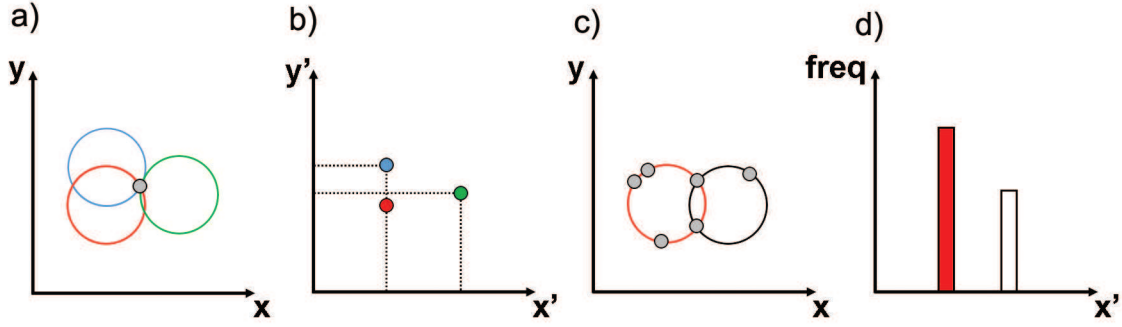

FIGURE 2.10: Principle of the circular Hough Transform. a) A set of circle candidates for one point in $\Re$ space are mapped to a set of points in Hough space (b), where the set of coordinates for each point represents the centroid position of the circle. c) Two circle candidates matched to points in $\Re$ space have different strength as seen in Hough space (d), depending on the number of points in $\Re$ space they intersect. Selecting the strongest peak in Hough space is equivalent to selecting the most probable circle candidate in $\Re$ space (red).

A Hough Transform (HT) can be seen as a type of segmentation procedure, in which the foreground is modeled as a particular shape object. Figure 2.10 shows a circular HT, where a circle is parameterized as:

$$(\frac{x - x'}{r})^2 + (\frac{y - y'}{r})^2 = 1, \tag{2.7}$$

with centroid position (x', y') and radius r as parameters, thus producing a three-dimensional Hough space. In Figure 2.10a, we reduce the dimensionality by drawing a set of circles with the same radius, all passing through a single point in $\Re$ space. Mapped to Hough space (Figure 2.10b) they will each appear as a point, determining the centroid position of each circle. In Figure 2.10c we simplify matters further by approximating two circles to a set of points in $\Re$ space, where we select the circle candidates such that they are all located at the same y. We now collect the number of points matched by each circle candidate in $\Re$ space into the corresponding bin (given by the circle's x' position) and plot it as a histogram. The relative difference in frequency represents the relative strength in Hough space for each circle candidate and the strength with which a structure appears in the image is represented by the number of points in
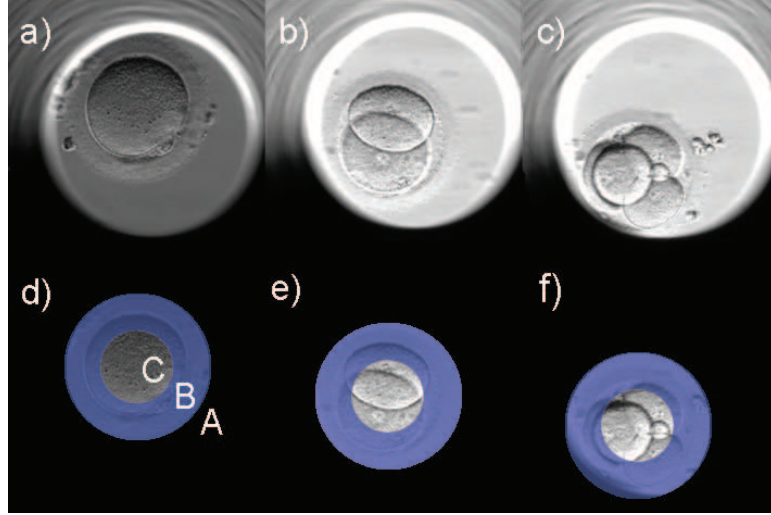
FIGURE 2.11: Example of test images with a) 1, b) 2 and c) 4 cells per embryo. All images are from separate embryos, illustrating differences in lighting conditions. d-f) Illustration of image feature extraction of the example images a-c). Region A: Excluded information, region B: Outer layer, region C: Inner embryo [6].

the corresponding bin (Figure 2.10d). The shapes most likely to exist in the image can be extracted by selecting peaks in Hough space in order of decreasing height. In the case of cell detection, we would prefer to model cells as ellipses rather than circles, and the circular HT can be readily extended to a four-dimensional elliptic space using

$$(\frac{x - x'}{a})^2 + (\frac{y - y'}{b})^2 = 1, \tag{2.8}$$

where a and b are major and minor axis, respectively. The main advantage of the HT is its tolerance to gaps in the boundary description and subsequent resistance to image noise. A disadvantage is that it sometimes returns large number of false negatives, and another that the complexity of the accumulator increases as $o(D^{\rho-2})$ [142], where D is the size of the image space and ρ is the number of parameters. Thus, by increasing the number of parameters, the noise in Hough space will increase accordingly, and it is advantageous to select as few parameters as possible to describe a shape.

2.3.6 Image Descriptors

An *image descriptor* or *feature descriptor* is a fixed size vector, with dimension $\leq X \cdot Y$ of the image, describing a set of image identifying characteristics. In an inference problem, this is a common way to reduce the amount of information and describe the image in the compact form necessary for efficient computation. A simple example of an image descriptor is a histogram. It effectively transforms the image of dimension $X \cdot Y$ to

a fixed size array, whose elements note the frequency of image gray scale values within a predefined range. Descriptors in general may be calculated manually, as response to a filter or a set of filters, or by any kind of specific algorithm. The choice of descriptor is highly dependent on the endpoint of the study and may consist of geometrical information such as area and shape, textural information such as regularities within the image, the location and appearance of specific interest points or statistical measures of intensity values (mean, max or distribution). Many descriptors have been developed over the years and it is outside the scope of this Thesis to describe them all. Instead we satisfy ourselves with defining a list of criteria relevant to cell imaging and discuss the considerations which have lead up to them. For this Thesis, the choice of descriptors has been a combination of customized descriptors defined by these criteria, together with several Texture Descriptors (Chapter 2.3.7), which gave the opportunity to describe any regularities of cell and embryo interior. The extracted image features can then be used for classification, a form of probabilistic inference (Chapter 2.4).

In the case of imaging cells randomly distributed over an image imaged with a technique where the pixel intensity is related to physical properties in the sample, a suitable descriptor should:

- be rotationally invariant (cells are located randomly and are not expected to grow directionally)

- not necessarily require a segmented contour (shape information may help, but should not be required since full object segmentation is not always possible)

- focus on characterizing small regions within the image, rather than large sections

- not be scale invariant (ideally we fix the scale at the hardware stage, and wish to use features to compare scale, rather than impose scale invariance)

- not necessarily be optimized for speed, since cells naturally grow and divide at a rate slower than the image acquisition.

2.3.7 Texture Analysis

A common way to study a regular structure, or *texture*, in an image is by means of a Gray Level Co-Occurence Matrix (GLCM). It can be thought of as an intermediate matrix, whose values indicate how often one pixel-value in an image tends to be similar to its adjacent values. A mathematical definition [96] of a GLCM G of an image $F(x, y)$ with a maximum of I gray levels describes the number of times gray level i is oriented with respect to gray level j such that $F(x+1, y+1) = j$, where $i, j \in [0, G]$, where $i, j \in [0, H]$,

and $H < I$ is the number of gray level used in the GLCM, and will determine the scale at which we look at the texture. We can normalize the GLCM, G, so that:

$$\sum_{i=1}^{H}\sum_{j=1}^{H}[G(i,j)] = 1. \tag{2.9}$$

In this way, each value in $G(i,j)$ will denote the probability of a range of gray levels to appear in each others' proximity. If we can not assume the cells and texture to be randomly oriented in the image, we may have to compute several GLCM along each texture direction. Given the G it is then straight forward to compute a number of texture features:

The energy S_{energy} describes the uniformity of the texture. In a uniform image, the number of gray-tone transitions are few, and the co-occurence matrix will have fewer entries of high magnitude, Hence a homogeneous image has high energy.

$$S_{energy} = \sum_{i=1}^{H}\sum_{j=1}^{H}[G(i,j)]^2 \tag{2.10a}$$

The entropy $S_{entropy}$ describes the randomness of the elements in the matrix. A homogeneous image has lower entropy than an inhomogeneous image. So an image with low entropy will have high energy.

$$S_{entropy} = -\sum_{i=1}^{H}\sum_{j=1}^{H}G(i,j)\cdot log[G(i,j)] \tag{2.10b}$$

Cluster shade S_{cs} and cluster prominence S_{cp} are measurements of the skewness of the image. An image with high skewness will be highly asymmetric.

$$S_{cs} = \sum_{i=1}^{H}\sum_{j=1}^{H}[i+j-\mu_x-\mu_y]^3 G(i,j) \tag{2.10c}$$

$$S_{cp} = \sum_{i=1}^{H}\sum_{j=1}^{H}[i+j-\mu_x-\mu_y]^4 G(i,j) \tag{2.10d}$$

Contrast S_{con}

$$S_{con} = \sum_{i=1}^{H}\sum_{j=1}^{H}|(i-j)|^2 G(i,j) \tag{2.10e}$$

and correlation, where S_{corr} refers to the correlation between elements. A high correlation means the image will be more complex than if the correlation is low.

$$S_{corr} = \sum_{i=1}^{H} \sum_{j=1}^{H} G(i,j) \frac{(i - \mu_x)(j - \mu_y)}{\sigma_x \sigma_y} \qquad (2.10\text{f})$$

The homogeneity S_{hom} is large if values are gathered along a diagonal.

$$S_{hom} = \sum_{i=1}^{H} \sum_{j=1}^{H} \frac{G(i,j)}{1 - |(i-j)|^2} \qquad (2.10\text{g})$$

2.4   Inference & Classification

Classification of image features used a number of standard supervised classification methods and variations thereof. This section begins with an overview of probabilistic theory with definition of concepts such as inference and logistic regression and then gives an overview of the specific methods used in this Thesis and motivation for the implemented changes for the purpose of cell classification.

In the mathematical definition of an inference problem, we take a set of observations (images) of the world $\overline{v}$ and from it infer a state from the set of states $\overline{w}$ in the world, where each component of the vector $\overline{v}$ is an observation (an image) and each component of $\overline{w}$ is a real world state. If $\overline{w}$ is continuous, $\overline{w} \in [-\infty, \infty]$, we have a special case of inference, which we define as *regression*. If $\overline{w} \in [0, K]$ is discrete, with $K \in \Re$ being a limited vector, we call it *classification*.

The inference $Q(\overline{w}|\overline{v})$ is defined as the real world state we can deduce from each image state. *E.g.* if we see an image of an object we think resembles a dog, we conclude it is a dog. There is an inherit ambiguity in visual data, *i.e.* several real world states can share the same image state $\overline{v}$ and *vice versa*. For instance, a change in light and angle can cause different objects to appear the same way (Figure 2.12).

Ideally, we would like to compute the inference $Q(\overline{w}|\overline{v})$ as a 1:1 relationship, *i.e.* combine observations with conclusions as a set of pairs (v_α, w_α), each having a definitive certainty. As a consequence of observational ambiguity, all we can do is compute a *posterior probability distribution* $P(\overline{w}|\overline{v})$ over possible states $\overline{w}$. If we see an image of an object we think resembles a dog, we conclude there is a certain probability that it actually is a dog, but if the image is not very clear, we may also deduce there is a small probability it is a cat, a garden statue or something else entirely. We have computed a limited collection of sample *posteriors* from the image: dog, cat, statue. Most often,
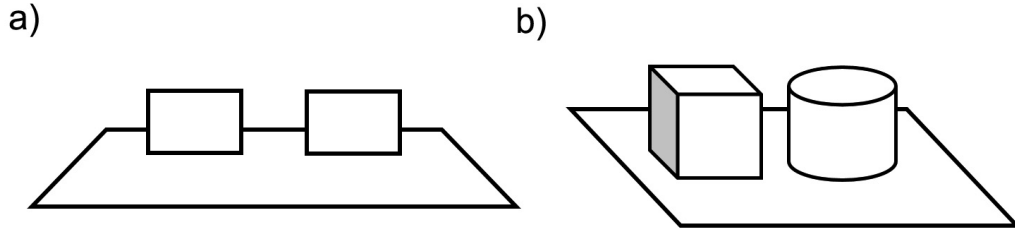
a)   b)

FIGURE 2.12: Illustration of the ambiguity on an inference problem. Two three-dimensional objects, when viewed from one angle, can lead us to infer them to share the same world state $\overline{w}$ (a). Using an observation from a different angle (b), we infer a completely different solution.

we go for the most probable solution, as long as evidence does not change to make us believe otherwise. We call the most probable solution the *maximum a posteriori* (MAP) solution. To complicate things further, it is not possible to compute $P(\overline{w}|\overline{v})$ over *all* possible states - there is still a small probability that the image of the dog is an image of something we have not been able to deduce. We learn as children what a dog is supposed to look like and our brain draws on this experience when interpreting the image, but without that experience we could not draw the same conclusions - we can not infer something we can not imagine. We may say that we can only infer probabilities that are somehow built into our *model* of the world.

We may now formulate the solution for a visual problem as composed of three components:

- A *model* to describe our expectations - our translation from world to image and *vice versa*, describing the relationship between the observation $\overline{v}$ and the world $\overline{w}$. The model is described using a set of learned parameters $\overline{\psi}$.

- A *learning algorithm* which computes the parameters for our model $\overline{\psi}$ from training examples (v_α, w_α).

- An *inference algorithm* which takefusings a new observation v_β and uses the model to return the posterior $P(\overline{w}|v_\beta, \overline{\psi})$ over world states $\overline{w}$.

The model construction can be of two types:

- Inferring the world state from the observation $P(\overline{w}|\overline{v})$ (conclusion) or

- Inferring the observation from the world $P(\overline{v}|\overline{w})$ (prediction).

The first type of model is referred to as a *discriminative* model, and in the case of imaging it is the same as drawing conclusions about the world from what is seen in the image. The second type is referred to as a *generative* model. In the case of imaging, it would be equivalent to setting up a model for how the real world object is translated into the image, *e.g.* modeling the light reflected or transmitted by the object and how it is translated by the optical setup, recorded by the camera sensor, digitized and finally displayed on a screen or as a print. The relationship between the two model types is described by Bayes' rule:

$$P(\overline{w}|\overline{v}) = \frac{P(\overline{v}|\overline{w})P(\overline{w})}{P(\overline{v})} = \frac{P(\overline{v}|\overline{w})P(\overline{w})}{\int P(\overline{v}|\overline{w})P(\overline{w})\delta\overline{w}}. \tag{2.11}$$

Commonly, we also assume that the data in our observations (the components of each observation v_α) are independent of one another, a condition known as *Naïve Bayes*. The solution to a discriminative problem is often referred to as logistic regression (not to be confused with linear regression, which is defined as the solution to a regression problem of continuous world states). Table 2.5 summarizes the different cases.

TABLE 2.5: Methods for inference

	Model discriminative $P(\overline{w}|\overline{v})$	Model generative $P(\overline{v}|\overline{w})$
Regression $\overline{v} \in [-\infty, \infty], \overline{w} \in [-\infty, \infty]$	Linear regression	Linear regression
Classification $\overline{v} \in [-\infty, \infty], \overline{w} \in [0, K]$	Logistic regression	Probability density function

All inference problems can be interpreted in the light of this abstract framework. In the context of this Thesis, we are concerned only with classification, i.e. $\overline{w} \in [0, K]$ using the discriminative models and logistic regression solutions described further in Chapter 2.4.1. For microscopy imaging, including HMC, the generative model is only of indirect interest, as it often would be very difficult to calculate exactly. Our model $\overline{\psi}$ varies depending on the problem, and so does the learning algorithm and resulting inference algorithm. In most cases, we compute the posterior directly. In all cases, we have satisfied ourselves with computing the MAP, leaving out the full posterior probability distribution (PPD). This is often a practical approach in computer aided diagnosis (CAD) where we are concerned only with detection or exclusion of a world state $\overline{w}$ which is significant of a certain cellular state, such as disease or health. Computation of the full PPD has the benefit of returning a probability confidence score to our conclusion, but also introduces a more complex computation and possibly increased computation time. In most real

world cases of CAD, a computer diagnosis is not enough to decide on patient treatment and the opinion of a human expert observer must be consulted anyway, making the confidence score returned from the PPD of less practical value. The best we can do in CAD, is returning a suggestion of disease, an *annotation* which serves as a notification for a human observer to take a closer look.

2.4.1 Logistic Regression

The goal of a logistic regression problem is to determine a posterior probability distribution $P(\overline{w}|\overline{v})$ mapping a set of observations $\overline{v}$, each taking on continuous values $\overline{v} \in [-\infty, \infty]$, to a discrete set of world states $\overline{w} \in [1, ...K]$. The discrete set may be modeled as a Bernoulli distribution using $\lambda \in [0, 1]$, where λ is our result class. To model the distribution, we construct the linear *activation* function:

$$\omega = \phi_0 + \overline{\phi} \cdot \overline{v} \tag{2.12}$$

and apply a *logistic sigmoid function* $sig(\omega)$ which maps the range $[-\infty, \infty]$ to $[0, 1]$:

$$sig(\omega) = \frac{1}{1 + exp(-\omega)} \tag{2.13}$$

The result is a PPD

$$P(\overline{w}|\phi_0, \overline{\phi}, \overline{v}) = sig(\omega) \tag{2.14}$$

where we can adjust the parameters of the activation so that $\omega = \overline{\phi} \cdot \overline{v}$ to simplify:

$$P(\overline{w}|\overline{\phi}, \overline{v}) = sig(\omega) \tag{2.15}$$

In training, the goal is to maximize the likelihood:

$$P(\overline{w}|\overline{v}, \overline{\psi}) = \sum_{\alpha=1}^{L} P(w_\alpha|v_\alpha, \overline{\psi}) = \prod_{\alpha=1}^{L} \lambda^{w_\alpha}(1 - \lambda)^{1-w_\alpha} \tag{2.16}$$

using L sample pairs in the training data set.

By computing the log probability, the expression simplifies into a sum, the *maximum log likelihood*:

$$log[P(\overline{w}|\overline{v},\overline{\psi})] = \sum_{\alpha=1}^{L} w_{\alpha} log[\lambda] + \sum_{\alpha=1}^{L}(1 - w_{\alpha}) \cdot log[1 - \lambda] \qquad (2.17)$$

2.4.2   Extension to Multiple Classes

If $K = 2$ we have only two classes, and the classification problem is binary. In the real world however, multiclass problems, $K \geq 2$ are common. A range of methods exist to extend algorithms for binary classification to multiple classes.

The optimization is readily extended to multiple classes by describing the posterior as a categorical distribution with $\overline{\lambda(x)} = [\lambda_1, ... \lambda_K]$ as functions of the data. Instead of one activation (Eq 2.12), we have K:

$$\omega_k = \overline{\phi}_k \cdot \overline{v}, \qquad (2.18)$$

K local optimizations:

$$log[P(\overline{w}|\overline{v},\overline{\psi},\lambda_k)] = \sum_{\alpha=1}^{L} w_{\alpha} log(\lambda_k) + \sum_{\alpha=1}^{L}(1 - w_{\alpha}) \cdot log(1 - \lambda_k) \qquad (2.19)$$

and K parameters for the model; $\overline{\psi} = \overline{\phi}_1, ... \overline{\phi}_K$, mapping the activation states to each λ_k as

$$\lambda_k = \Lambda[\omega_1, ... \omega_K] = \frac{exp(\omega_k)}{\sum\limits_{n=1}^{K} exp(\omega_n)} \qquad (2.20)$$

where Λ maps the set of continuous activation functions to discrete categories λ_k

$$\sum_{k} \lambda_k = 1. \qquad (2.21)$$

2.4.3   Classification

Often the task of inference is discrete; to label the image or an area within an image as belonging to some predefined class. In this case, the inference problem is commonly termed *classification*. In the simplest terms, a classification task in image analysis is

the process of assigning the correct label to an image or part of an image. Classification is an important task in computer vision. Several algorithms have been around for decades and with the arrival of cheap and efficient computer power, their use have over the years increased. Many algorithms have been applied to medical image analysis, especially for the detection of abnormalities where the task can be seductively simple; classify structures into healthy or unhealthy. This is an example of a binary problem but in many cases, the classification is multiple, such as when separating cells into the different stages of the cell cycle. The separating border between two groups in class space is referred to as a *decision boundary* (Figure 2.14). Depending on what side of the decision boundary a new data ends up on, it is classified as belonging to each respective group. Classification algorithms can roughly be divided in two groups; supervised and unsupervised techniques. A *supervised* approach uses training images labeled with known classes to define the parameters of the model $\overline{\psi}$ through reverse engineering. The model obtained can then be used to classify unlabeled images, provided the training set was representative enough. *Unsupervised* learning on the other hand, attempts to classify images directly, often by grouping images with similar characteristics together, but without assigning any labels. Unsupervised methods is the common name for methods not requiring labeled training data. They are mostly useful in image encoding, image matching and as a preprocessing step for further supervised schemes.

2.4.4 Supervised Learning

A common approach to a supervised learning problem is the generation of a descriptor, consisting of image features, which is then fed into a classifier. Images are usually preprocessed and structures sometimes segmented prior to feature extraction. The parameters for the classifier are learned in a training step where features from images are classified and parameters set to achieve the best possible match with the known result. The training step requires a representative set of labeled data - the *ground truth* - to infer the correct parameter setting of the classifier. Supervised learning methods often benefit from having a large set of labeled training data available.

2.4.5 Decision Trees

A solution to the logistic regression has a few drawbacks. It can only compute linear decision boundaries and will become inefficient or tend to overfit in higher dimensions. More complex decision boundaries can be computed if we introduce a nonlinear transformation of observations. Another way is to divide the data space $\overline{v}$ into distinct regions and use a different classifier to each region. The result is a *nested logistic regression*

model, sometimes also referred to as a *decision tree* [126].

$$\omega = (1 - \Omega(v_\alpha, \omega))\phi_0 \cdot v_\alpha + \Omega(v_\alpha, \omega)\phi_1 \cdot v_\alpha, \qquad (2.22)$$

where Ω is a *gated sigmoid*:

$$\Omega(v_\alpha, \omega) = sig(\omega \cdot v_\alpha). \qquad (2.23)$$

A decision tree has advantages over a nonlinear logistic regression in that it omits the nonlinear transformation and is thus simpler and potentially faster, since each node only produces a binary classification and each data point will only take one unique path through the node tree.

2.4.6 Multiclass Decision Trees: The Random Forest

In the multi-class case, instead of maximizing the log Bernoulli probability, we strive to optimize the log Bernoulli probability of the training class labels. In the case of more than two classes, decision trees are usually termed *random forests*. Like in the binary case, data is passed to only one child branch, and eventually ends up in a single leaf, but in this case, each node has several possible branches, and an optimization function at each node determines the chosen branch.
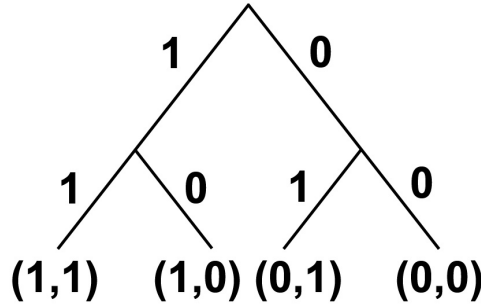

FIGURE 2.13: Principle of a two level binary Decision Tree. At each node, the descriptor is evaluated according to some optimization function separating the data set in two groups. The classification progresses step-wise downward, until all members of the data set ends up in one of the bottom nodes (the leaves).

2.4.7 Support Vector Machines

So far, we have described probabilistic algorithms for classification. Another approach is to use non-probabilistic classification algorithms. An example of these are Support Vector Machines (SVM), which have become very popular for image classification tasks

and for historical reasons they are readily available in most software packages. The simplicity in their implementation and the computational speed makes them useful for a broad range of applications, where the full probabilistic inference is not as much of interest as the class output. In solution to a binary classification problem, a SVM separates two clusters of data by means of a plane (Figure 2.14), placed to maximize the gap between the members of each cluster closest to the decision boundary (the support vectors). Originally, the SVM was designed to handle only binary problems, but a number of extensions to multiple classes exist.
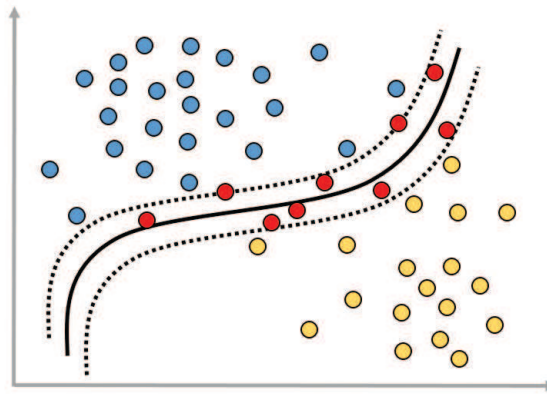

FIGURE 2.14: Principle of a Support Vector Machine. A decision boundary is defined using the fewest possible, separating descriptors (the support vectors), marked in red.

2.4.8   Multiclass SVM: The DAG-SVM

The SVM was originally designed for binary problems, but a number of extensions have been developed to handle multiple classes. Most of them progress as a conglomerate of binary classifiers in different combinations. One approach is to perform K classifications for K classes, where each class is classified versus all the others - an approach known as one-versus-all. Each classification returns a confidence score which decides the final class. In the case of embryo classification, this approach has a number of disadvantages. Each classifier must produce a real-valued confidence score, and its scale may vary between classifiers. Also, the group of positives will always be smaller than the group of negatives, which consists of all other classes. If the classification is intended to separate embryos of different number of cells, the distribution will be even more skewed since the more cells the embryo contains, the shorter the time between division and less images representative for each class are available. One way to avoid the demand to return a confidence score is to perform $K(K-1)/2$ binary classifications without confidence score and select the class voted for by the majority of classifiers. A downside is that

this method may lead to ambiguity, as the same number of votes may be cast for several classes. A third approach is to perform a step-wise exclusion, where the entire sample set is subject to $K(K-1)/2$ successive and pair-wise one-versus-one classifications. When first implemented, this approach was referred to as a Directional Acyclic Graph SVM (DAG-SVM) [99]. It avoids the problem of having to report a confidence score, as well as the problem of ambiguity with a voting scheme.

Figure 2.15 illustrates the DAG-SVM for a 4-class problem with classes A,B,C and D. For the first classification class A versus class B, it is concluded that all samples classified as belonging to class B, do not necessarily belong to class B, but can in any case not belong to class A. For the sake of argument, we call this sample cluster Ã (not-A). In the next step, the group Ã then compares class B with class C. If classified as C, it is concluded that samples from this sample cluster are not necessarily C, but in any case, do not belong to B, and may be called ÃB̃ (not-A-nor-B). Classification progresses step-wise until only one class remains and the last one-versus-one classification then finally separates class C from D.
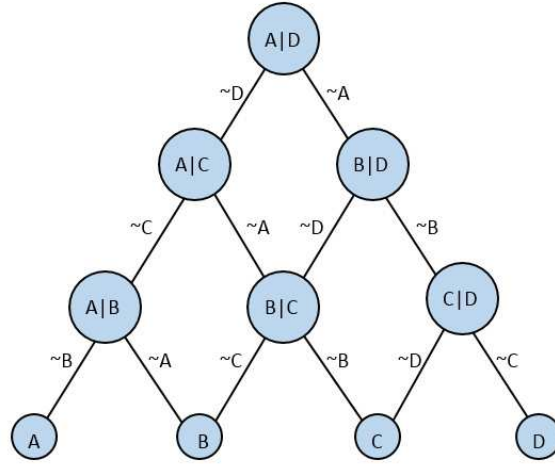

FIGURE 2.15: Principle of a DAG-SVM applied to 4 classes. Circles denotes classification nodes, lines decisions on classes and dots final classification result. An object classified as not belonging to a class k is denoted by k̃ [6].

Chapter 3

Results

In this Chapter, we focus in turn on the seven chosen criteria for embryo health (Table 2.1), describe in detail the methods developed to study them and evaluate the result. In cases where the results have been published, references are provided to the original publication. All relevant results and figures have been replicated in this Thesis, but the publications are also included as appendices.

3.1   Tracking of Embryo Location

In HMC imaging embryos are imaged in a circular well, one embryo per well, and image contents from the well surrounding the embryo and from outside of the well are of little importance to the analysis. The majority of a typical HMC image consists only of the contents and surroundings of the embryo growth chamber, and usually less than 25% of the images is actual embryo content (Figure 2.9B). Embryos are usually stationary within the well, but may shift position suddenly when medium is replaced at a few days interval. For automatic analysis, this means that the embryo position will have to be tracked during the sequence. In Paper II [2], a method is described which detects embryo location, using a HT (Chapter 2.3.5). The code was implemented in Matlab [92].

3.1.1   Methods

Figure 3.1 shows an example of how an embryo has been displaced in a well following change of medium. The time frame between Figure 3.1a and Figure 3.1b is one capture. A circular HT (Chapter 2.3.5) was applied to a Canny edge map (Chapther 2.3.2) of the raw image and the mean of the highest peaks defined the embryo position.
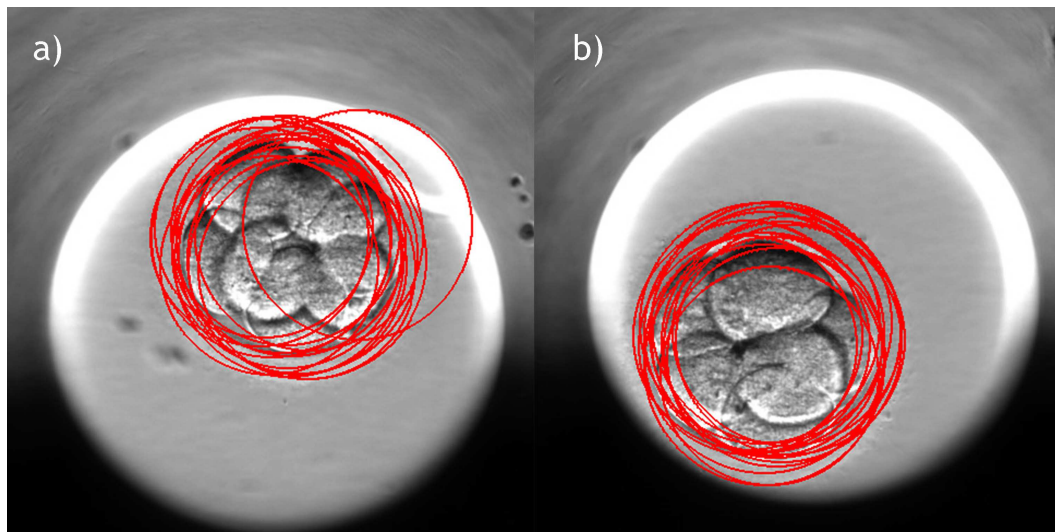
FIGURE 3.1: Tracking and automatic selection of embryo interior. The same embryo captured before (a) and after (b) dislocation. After filtering trajectories outside well boundaries (not shown), a median of several possible trajectories for inner zona pellucida boundaries was used to detect the region of the embryo [2].
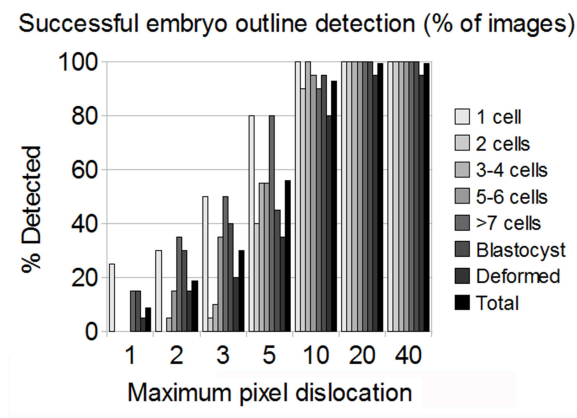
3.1.2   Results

FIGURE 3.2: Correctly detected embryo outlines for 140 images of embryos from zygote to blastocyst stage, as a function of maximum pixel dislocation. Showing calculations for each of 7 categories separately (gray) and total (black) [2].

Figure 3.2 shows the number of successful embryo outlines for embryos with 1-7 cells, plus blastocyst stage. Images considered potentially difficult were placed in a separate category. Location detection was compared using the location of the bounding box centroid. If an accurate detection was defined as within 10 pixel variation, a detection accuracy of 92.9% was achieved for all number of cells, and an accuracy of 95.0% if the class of difficult images was ignored. As a comparison, the manual selection of embryo outline had an accuracy of 70% for 1 pixel variation, when repeated on the same

picture by the same person, and 100% was within 5 pixel variation. The embryo outline detection was also tested on four image series where fresh medium was supplied at one point in the series, displacing the embryo within the well. In Figure 3.3, the positions of four embryos were automatically detected after being displaced.
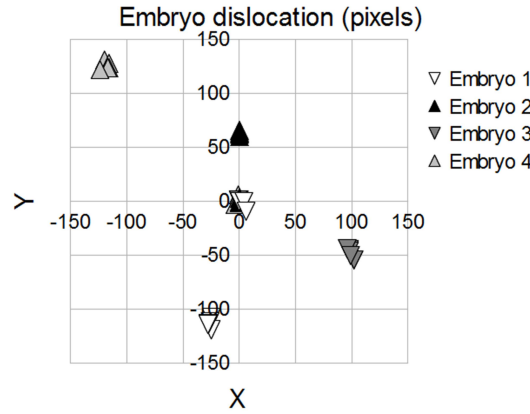

FIGURE 3.3: Embryo location in image (pixels) before and after change of medium in the containing well. The embryo position in the first image has been set to (0,0), with all other positions calculated relative to the first. Position defined as the position of the centroid detected by Hough transform, with radius 100-120 pxl [2].

3.2 Classification of Embryo Cell Number

Timing of mitotic splitting and the duration of mitosis are important cues to embryo development, as discussed in Chapter 1.3.5. Normally, it would require full segmentation of images and tracking of cells, but an indirect method to detect the timing of events in a time sequence is the classification of the sequence into groups of images belonging to different predefined states, each defined by the number of cells in the image. In Paper VI [6], the aim is to classify images in an embryo time sequence into groups with different number of cells, thus indirectly detecting the moment of mitosis between each group.

3.2.1 Methods

A set of embryos imaged by HMC were manually classified in groups of 1-8 cells for comparison. A training set was constructed using randomly selected subsets comprising 10% and 20% of the total image set, respectively. The total data set consisted of 620 8-bit gray scale images in series from 18 embryos. 12 features were used for classification. All were standard gray level image features; gray level mean, variance, maximum, minimum, root mean square, kurtosis, skewness, energy, entropy, contrast, correlation

and homogeneity[96]. To select an initial ROI, and to spatially filter out irrelevant image data, the embryo region was automatically detected (Chapter 3.1). Features were extracted from two separate regions; the entire embryo (Figure 2.11d-f: Region B) and on the inner embryo (Figure 2.11d-f: Region C only). The perimeter for region C was at half the radius of region B. Intuitively, from Figure 2.11d it is clear that this spatial filtering will have a strong effect on features since *e.g.* all cell membranes are eliminated for region C. Three state-of-the-art classifiers were compared; a Naïve Bayes classifier with a Gaussian kernel [97] (Chapter 2.4), a Random Forest (RF) [98] (Chapter 2.4.6) and a DAG-SVM [99] (Chapter 2.4.8). Also introduced was a slicing multiclass SVM (SS-SVM). Experiments were performed in three sets with the total image set containing 1-4, 1-6 and 1-8 cells, respectively. Results are given as the mean value of ten repeated identical experiments, with the training set selected randomly from the complete image set each time. The computational performance was evaluated using a standard 4 core PC with 8 GB RAM, with the code implemented in Matlab® 7.12.0.635 (R2011a) [93].

3.2.1.1   Sample Slicing SVM
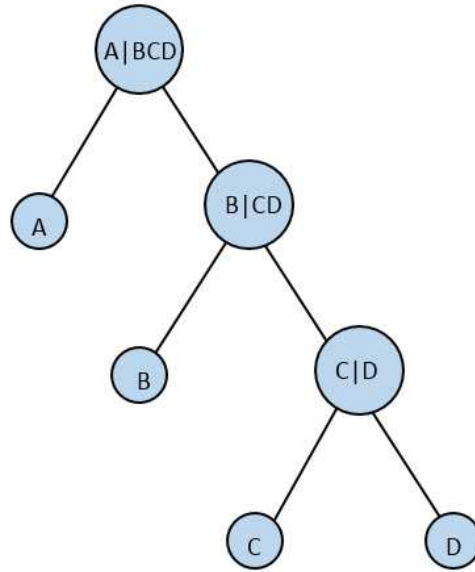
FIGURE 3.4: Principle of SS-SVM using an example of four cells. Class A has the highest cluster distance in feature space, and is excluded first. Classification progresses until all classes are determined [6].

In classification, one disadvantage of DAG-SVM (Chapter 2.4.8) is the performance reduction for higher number of K classes. Assuming some prior knowledge of the distance between sample clusters in feature space, we may improve performance and slice the sample set successively in a set of cascading one-versus-all classifications, inspired by

what has previously been described in [90, 91]. With the successive slicing SVM (SS-SVM), we improve performance by exploiting the fact that for embryos, we have prior knowledge of the hierarchy of the classification problem: The appearance of one cell compared to two cells differs more than the appearance of two cells compared to three, and so on. The test set is successively divided in smaller and smaller sets, until only one class remains for each sample. The order of classification must be previously decided upon by some general clustering method or by prior knowledge of cluster hierarchy. When classifying number of cells we assume that the optimal order of classification is from the lowest number of cells to the highest and the cluster distance can be expected to decrease for higher number of cells in the image. Simply put, the difference between an image containing one and two cells will be larger than the difference between an image containing three and four cells and so on. Exploiting this fact, we first separate all samples with one cell from the rest, progressing with classification of 2,3 cells *etc.*, until we reach the class with the maximum number of cells (Figure 3.4). Using this method, we need only K classifiers of type one-versus-all, one for each class, as opposed to $K(K-1)/2$ for the DAG-SVM.

3.2.2 Results

The accuracy for the SS-SVM was slightly higher than that of the DAG-SVM, with 89.4% correctly classified for the 1-4 cell case, 80.8% for the 1-6 cell case, and 74.9% for the 1-8 cell case (Figure 3.5). For some values, the DAG-SVM outperformed the SS-SVM, and for a few isolated values, they were both outperformed by the RF, but overall the predictive performance of the SS-SVM was comparable to the DAG-SVM. In Figure 3.6, the feature set was split into two groups, one set from each ROI, and classified using the SS-SVM. The result was improved by up to 45% by combining features from both ROI, compared to using features from one ROI only. In Figure 3.7, the Receiver Operator Characteristic (ROC) is plotted for each class and for each classification of image series comprising 1-4, 1-6 and 1-8 cells, respectively. The mean sensitivity for the classes in the 1-4 cell set was 82.2%, slightly higher than the mean sensitivity of 80.0% for the 1-6 cell set and 79.9% for the 1-8 cell set. The mean of the fallout was 4.75% for the 1-4 cell case, 4.43% for the 1-6 cell case and 5.34% for the 1-8 cell case. For a small class set (1-4 cells) the computational performance of the DAG-SVM was comparable to that of the SS-SVM, but for an increased number of classes, the SS-SVM outperformed the DAG-SVM by 63% (1-6 cells) and 73% (1-8 cells). The overall best performance was for the Naïve Bayes classifier, outperforming the SS-SVM by 86% (1-4 cells) and 88% (1-6 cells). For the 1-8 cell case, no results for the Naïve Bayes are shown, since the Gaussian kernel failed to estimate the parameters (Figure 3.8).
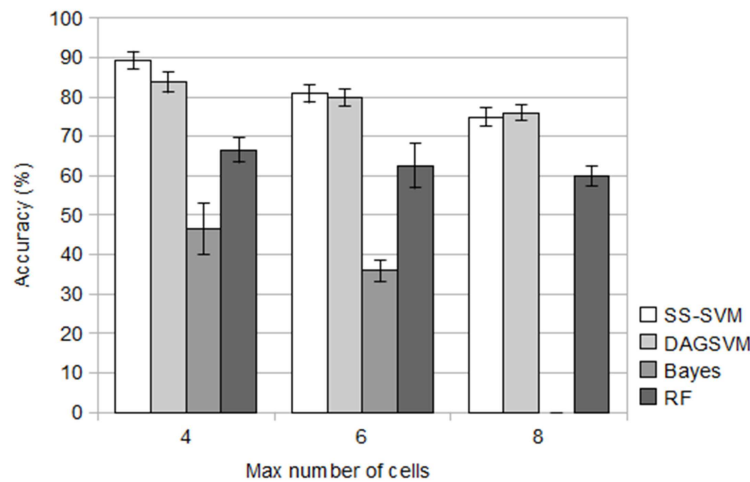
FIGURE 3.5: Accuracy as a function of maximum number of cells in the classification for all four classification methods. Mean values and confidence intervals are calculated from 10 classifications. The Bayes classifier failed to classify in the 1-8 cell case [6].
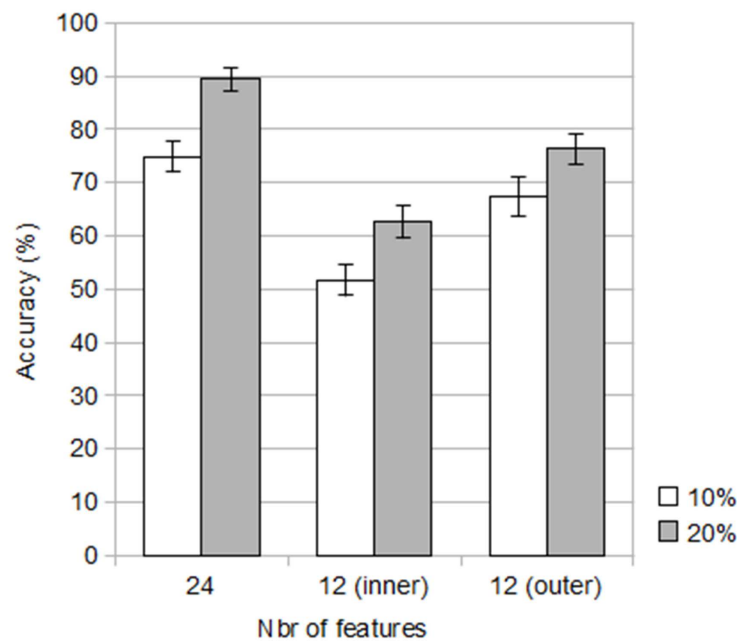
FIGURE 3.6: Accuracy as a function of number of features for the SS-SVM. The total feature set was 24 features, a combination of 12 features from the inner embryo and 12 features from the entire embryo (referred to as outer). Calculation was performed twice, with the training set selected as 10% and 20% of the total image set, respectively. The classification was performed for 1-4 cells. Mean values and confidence intervals are calculated from 10 classifications [6].

FIGURE 3.7: ROC of SS-SVM. Each data point represents one class in the image sets 1-4 cells (diamond), 1-6 cells (triangle down) and 1-8 cells (triangle up) [6].

FIGURE 3.8: Calculation time as a function of maximum number of cells for classification of the entire test set. Methods used were SS-SVM (square), DAG-SVM (diamond), Bayes (triangle down) and RF (triangle up) [6].

3.2.3   Discussion

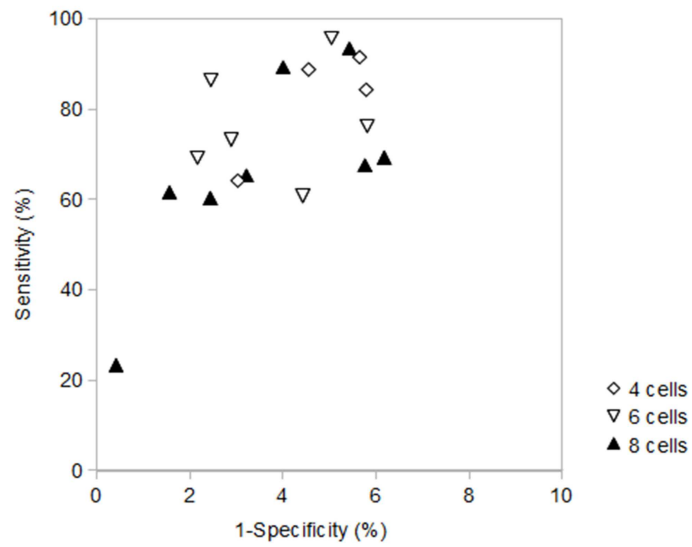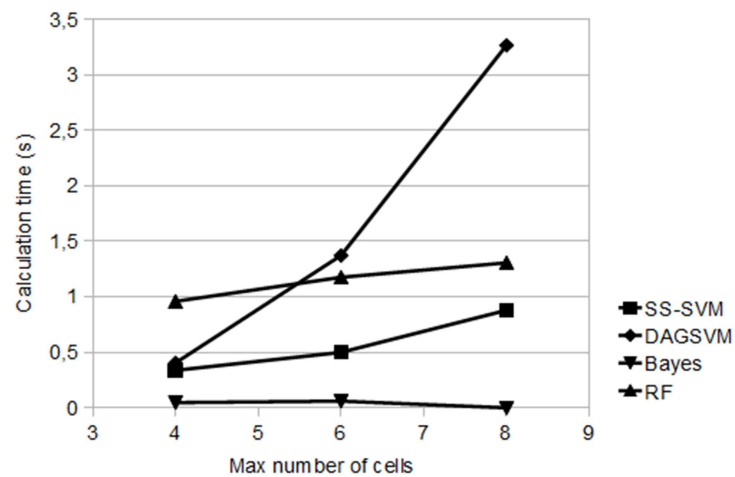The high classification accuracy is highly dependent on the fact that even if training images are separate from test images, they can still come from the same embryo, and so the training set is expected to be highly representative. In this case, the classification of number of cells is possible to a high accuracy even up to eight cells. A practical implementation of this approach would however require a manual selection of a few key frames per embryo, containing different number of cells. Ideally, a classification should be able to progress using training images from a set of training embryos, and then be applicable to images from another set of test embryos. The data set of 18 test embryos is too small to reliably train a classifier under those circumstances and this is an important continued point of investigation.

3.3   Embryo Activity Analysis

Variance is an extremely simple - thus also very robust - image characteristic. Paper IV [4] includes an example of the use of image variance to detect embryo activity, mitotic stagnation and rhythmic blastocoelic expansions. Mitotic stagnation is a sign of poor embryo quality and blastocoelic expansions conversely a state of normal embryo development. At the temporal resolution available, it is not possible to capture inter-cellular vibration at high frequencies, but intra-cellular movement and embryo movement are readily detectable.

3.3.1   Methods

Images in embryo sequences were captured with 0.2h interval between pictures. Figure 3.12 shows an example, where two embryos from the same patient are imaged simultaneously. The image was spatially filtered (Chapter 3.1) to region A, B and C (Figure 2.11) and the variance of the embryo interior (region C) relative to the variance in the cultivation well (region A) was calculated. The total image series of 448 images spanned from 4.7h (4-6 blastomere stage) to 94.1h of development, at a focal plane located approximately half-way through the embryo. For this analysis, only one of the available seven focal planes was used for analysis. The optimal plane was selected using a focus filter.
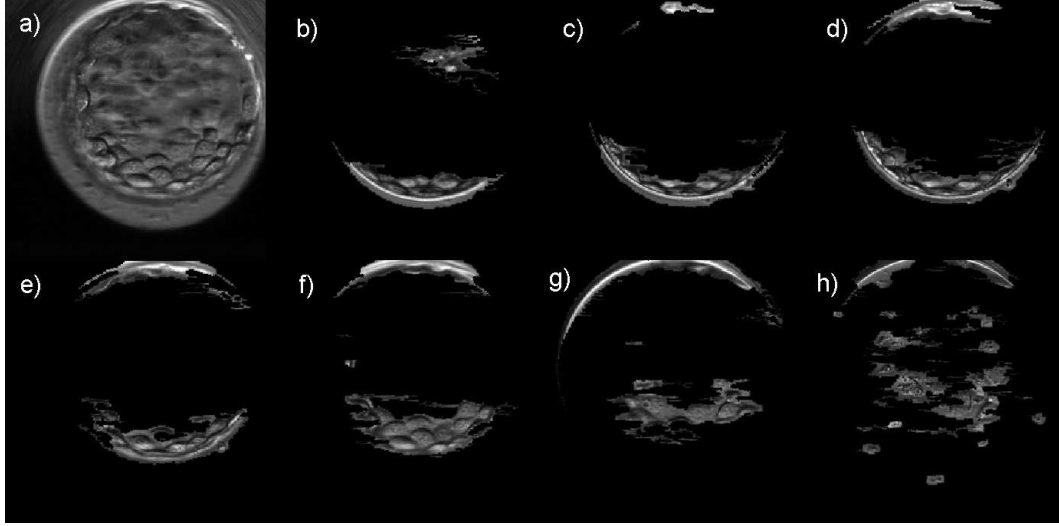
FIGURE 3.9: Hoffman Modulation Contrast image of blastocyst stage embryo at 136.2h, showing one original image (a) and seven images, captured at separate focal planes, with extracted regions in optimal focus (b-h). Focal planes cover the embryo from slightly above the embryo waist (b), down to the bottom of the containing well (h). The inner cell mass is in focus on the first slide (b). The original image corresponds to the 5th image in the series (f). Images (b)-(h) have reversed contrast for display purposes [8].

3.3.1.1 Focal Filtering

Being a focal slicing technique, HMC images are captured on several focal planes per imaging time point. Some of these images will only contain out-of-focus objects and others contain a mix of objects in-focus and out-of-focus. Out-of-focus objects contribute very little to the analysis and can even contribute negatively by providing inaccurate information, and must be removed prior to image feature extraction.

In cameras with autofocus, calculation of the image focus is often based on calculation of image contrast and several autofocus algorithms have been evaluated for microscopy [65, 66]. In optical microscopy [67], where the depth of field is usually very short, and when studying human embryos using optical microscopy with a short depth of field, the entire embryo is rarely in focus at one optical setting. Instead, the user may be required to adjust the focus back and forth in order to get a complete view of the sample. By recording images automatically at several focal planes, image contrast calculation can be used to filter images or areas of images out-of-focus. Prior to handling, captured images, $F(x, y)$, were filtered with a Gaussian filter to remove speckle noise. To detect the image contrast variation, a Laplacian filter (Eq 2.5) was applied. Note that due to the asymmetrical nature of HMC, the symmetrical Laplacian was chosen, rather than the direction dependent gradient (Chapter 2.3.2). To detect areas of high sharpness, an H-maxima transform, H, was applied to the result $\widehat{L}(F)$, and the resulting image was converted to a binary mask filtering the original image, F (Figure 3.9,3.10,3.11). The
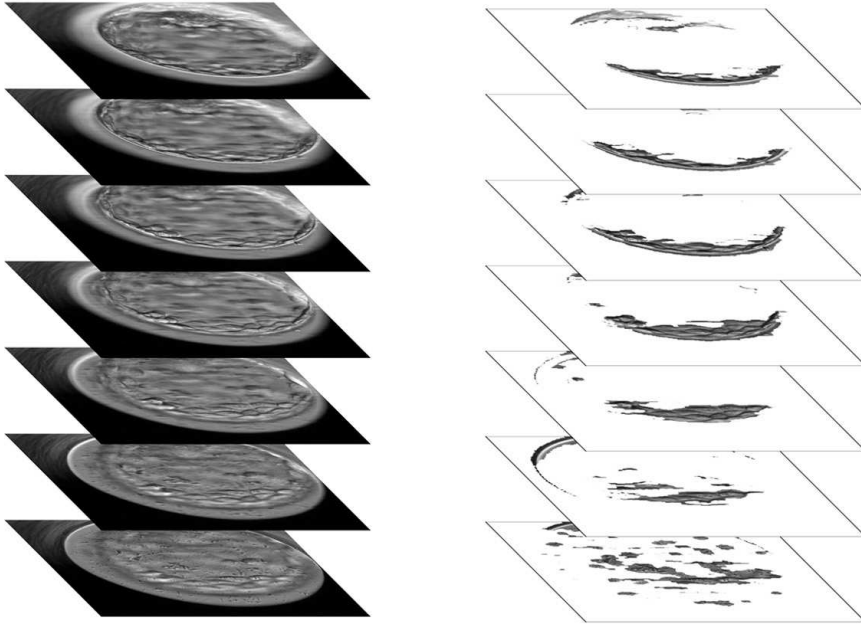
FIGURE 3.10: 7-image stack of Hoffman Modulation Contrast images of blastocyst stage embryo at 136.2h, and the extracted focused regions. The extracted regions have reversed contrast for display purposes [8].

method is evaluated in Paper I [1], and has been used in several of the enclosed papers for selection of the optimally focused image in an image stack, or for spatial filtering within an image, extracting areas in focus [94].

3.3.2   Results

The first embryo experienced several cell divisions during the first hours of the series, after which it formed a blastocoel at approximately 44.7h. After this, the embryo underwent a series of morphological changes where it reverted back and forth between a blastocyst and a tight central cell structure (Figure 3.12). These changes were also reflected in the image variance (Figure 3.13), both when looking at the internal and the entire embryo, though being less pronounced for the latter. The second embryo experienced a reduction in division activity after approximately 34h and suffered from heavy fragmentation from image 54h and forward. Within the region of interest of the embryo, this is shown as a reduced image variance when compared to the healthy embryo.
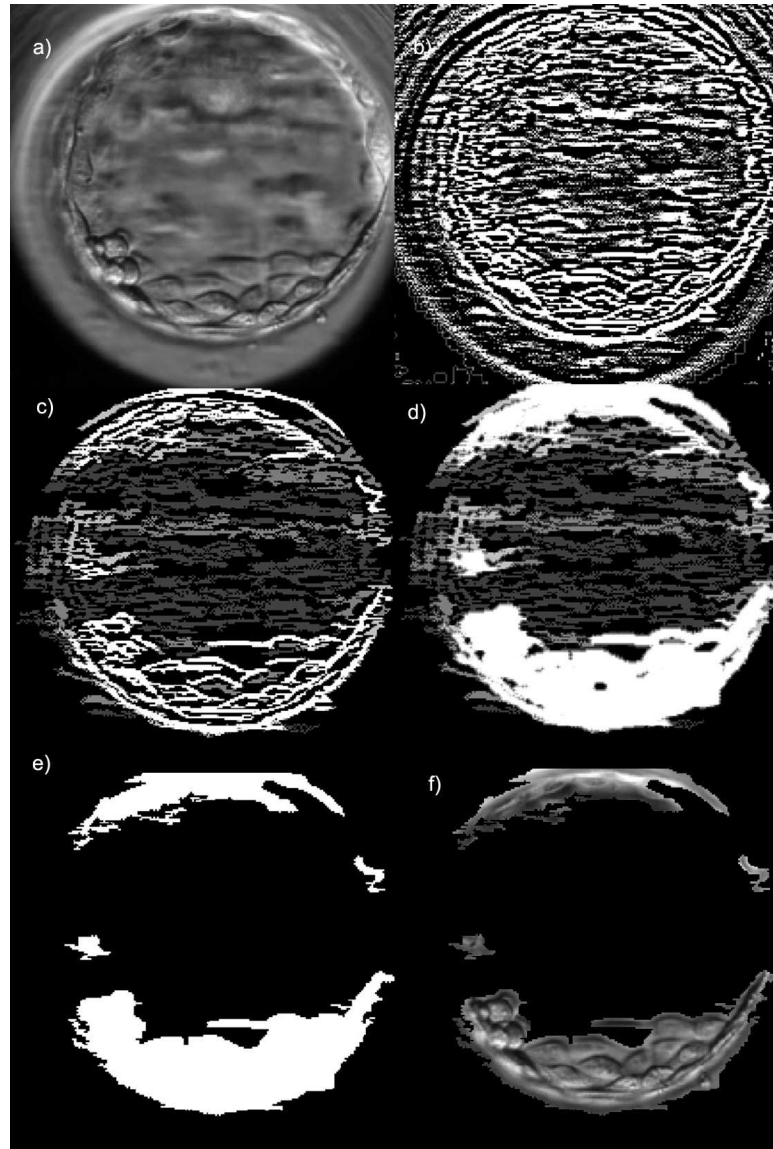
FIGURE 3.11: Step-by-step procedure for focal filtering. a) Original image. b) Laplacian of Gaussian applied. c) H-maxima transform highlights focal regions. d) A close transform further extends the selected area in focus to enclose a region, rather than isolated image patches. e) A threshold selects a mask. f) Original image filtered by mask.
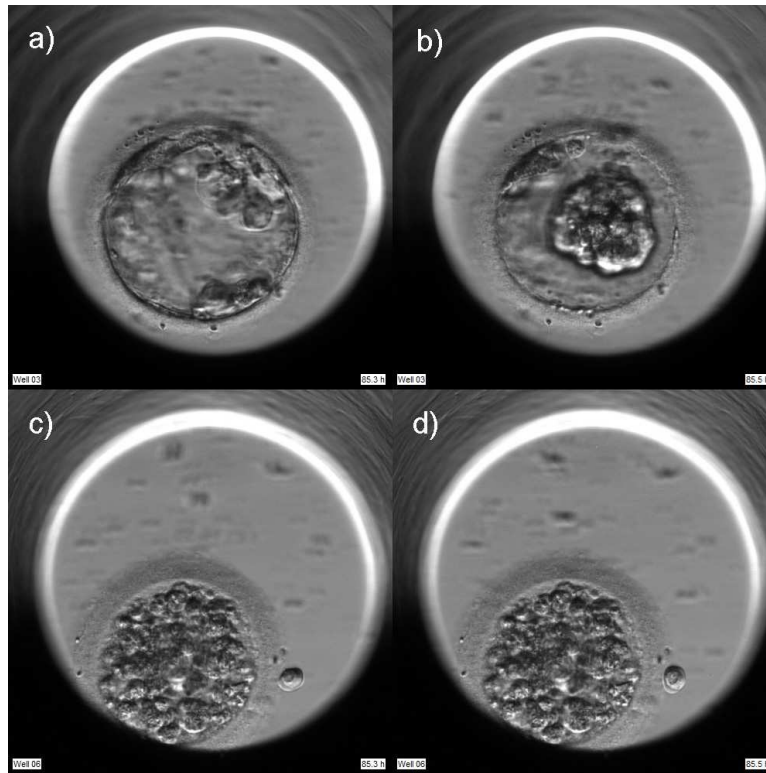
FIGURE 3.12: Two embryos at 85.3h (a,c) and 85.5h (b,d). The blastocoel starting to form for the healthy embryo (a-b). The second embryo shows heavy fragmentation (c-d), and the embryo activity is low [4].
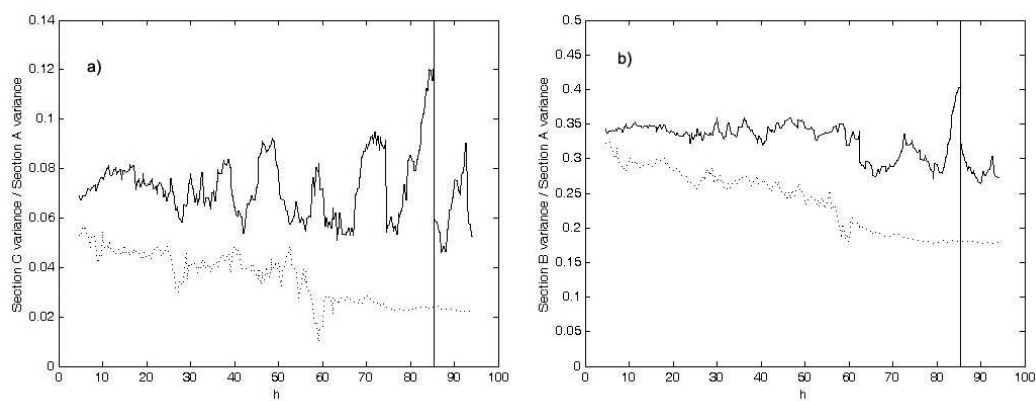
FIGURE 3.13: Variance of half embryo interior (a) and full embryo (b) relative to mean of embryo intensity. Line represents position of example images (Figure 3.12). Solid: healthy embryo, dashed: fragmented embryo [4].

3.3.3 Discussion

This examples shows the ability to use the temporal derivative of a TLI series to detect movement - or lack thereof - in living tissue. The importance of the correct spatial filtering is clear when comparing results from the full embryo and the internal parts. These results do not constitute a complete evaluation of this approach, but formed a basis for subsequent evaluations.

3.4 Embryo Fragmentation

Fragmentation can be an indicator of poor embryo health, but fragments can also be produced and reabsorbed during normal embryo growth [34, 52], making TLI a necessary tool to study the dynamics of fragmentation. It has been previously shown that the reduced activity of a fragmented embryo can be detected (Chapter 3.3). In this investigation, we evaluate a method to detect degree of fragmentation directly in images.

3.4.1 Methods

Images in embryo sequences were captured with 0.2h interval between pictures. Figure 3.12 shows an example, where two embryos from the same patient are imaged simultaneously. The image was spatially filtered (Chapter 3.1) to region A in Figure 2.11. The entire image was then divided using a square pixel grid of sizes in the range 5–100 pixels. The variance of pixel values in each grid element was calculated and normalized using the variance of Region A (entire embryo).

3.4.2 Results

Figure 3.14 shows a comparison between two embryos, one experiencing low degree of fragmentation (a-b), the other embryo higher (c-d). A heat map constructed from the normalized variance values is used to visualize fragmented regions. Using a suitable threshold, the degree of fragmentation within the embryo can be readily calculated.

3.4.3 Discussion

This examples shows the ability to use relative variance to detect embryo regions of higher fragmentation. A normalization to the image region of the embryo was selected in order to make the results comparable between images captured at different light
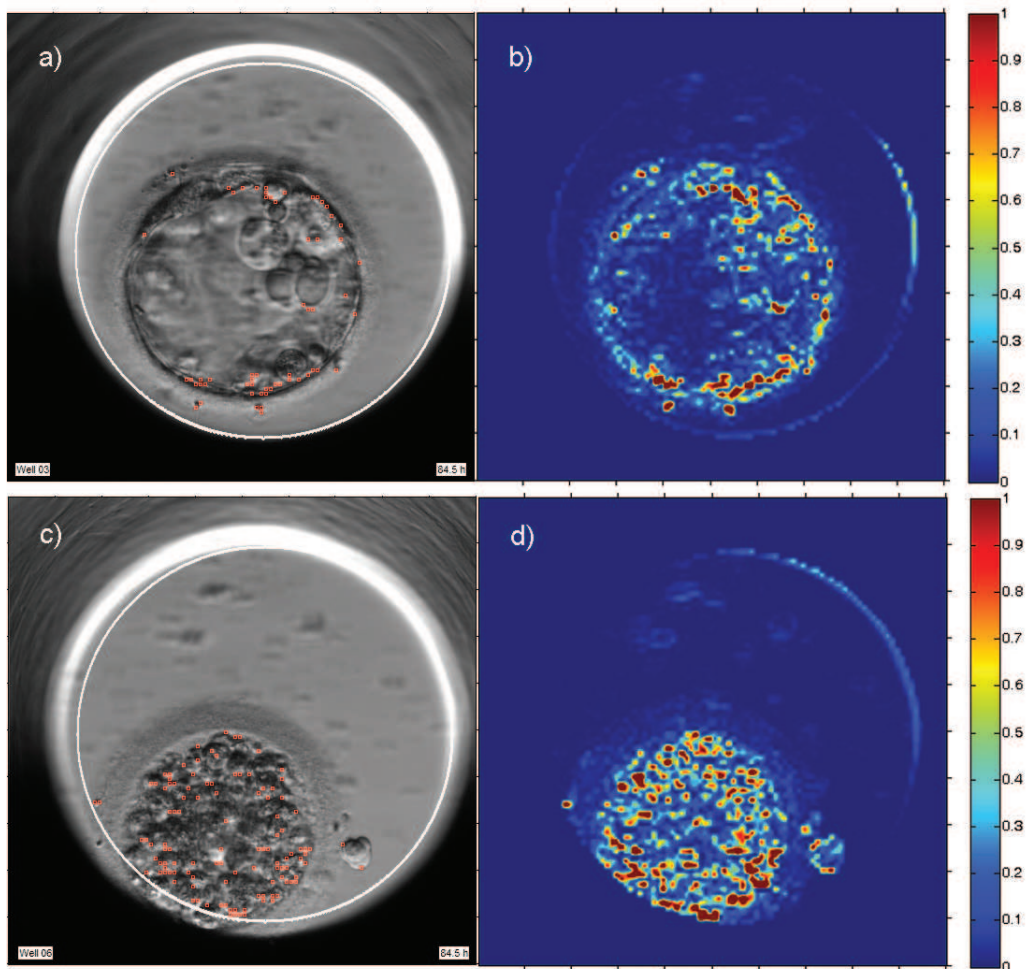
FIGURE 3.14: Comparison of variance for low fragmentation (a-b) and high fragmentation (c-d) embryo. a,d) Original image overlayed with image grid elements containing fragmentation. c-d) Heatmap of local image variance throughout image highlight heavily fragmented areas.

conditions. However, there is reason for caution when comparing images with very high degree of fragmentation, in which case the variance of the entire embryo region will also be affected. In order for this method to quantifiably measure differences in fragmentation, a consistent measure of normalization must be used. These results do not constitute a complete evaluation of this approach, and further analysis would require a comparative manual annotation of fragmentation of separate image regions.

3.5   Profiling Embryo Developmental Stages

In HMC microscopy, as in other bi-gradient images, edges are the most prominent structures. For cells, the main source of edge structure is the cell membrane. As cells divide

the number of cell membranes increase, and consequently the number of edges in the image. Conversely, cell compaction and loss of focus increase image smoothness and cause a loss of edge structures. As the number of edges in the image increases, the two-dimensional distribution of image intensity changes (Chapter 2.3.2). Image variance is a measure of the distribution of gray levels in an image and can function as an indirect measure of edge structures, and also of embryo developmental stages. Paper V [5] profiles embryo growth during a course of 5-7 days growth using image variance.

3.5.1   Methods

Image variance is computed in a filtered ROI within the embryo (Chapter 3.1) and compared to manual profiling of embryo growth (Figure 3.15). The spatial filter was set to 50% of embryo radius, and the variance calculated as fraction of the variance for the total embryo. The ratio was chosen so that during the cleavage stage, no single cell would appear completely outside the ROI, and for the blastocyst stage, the region would be small enough to exclude the trophectoderm. The result is a timeline profile, describing embryo development (Figure 3.16), effectively reducing the information of the four-dimensional image series to a one-dimensional sequence, which was then used

for further analysis detecting gradients and local minima and maxima. Six characteristics were chosen on the basis of their visbility in the graph and combined to a set of features for blastocyst detection.

- The width (duration) of the negative gradient at compaction.

- The height of the maximum variance detected at the end of the cleavage stage.

- The height of the maximum variance detected at cavitation.

- The timing of compaction.

- The timing of the maximum at the end of the cleavage stage.

- The total number of variance gradients during the entire development (a sign of strong fluctuating behavior, indicating poor quality).

A total of 39 embryos were examined from zygote to blastocyst stage, or to the equivalent point in time if no blastocyst was formed. Length of time lapse sequences varied between 350-500 images, the equivalent of 5-7 days of growth.
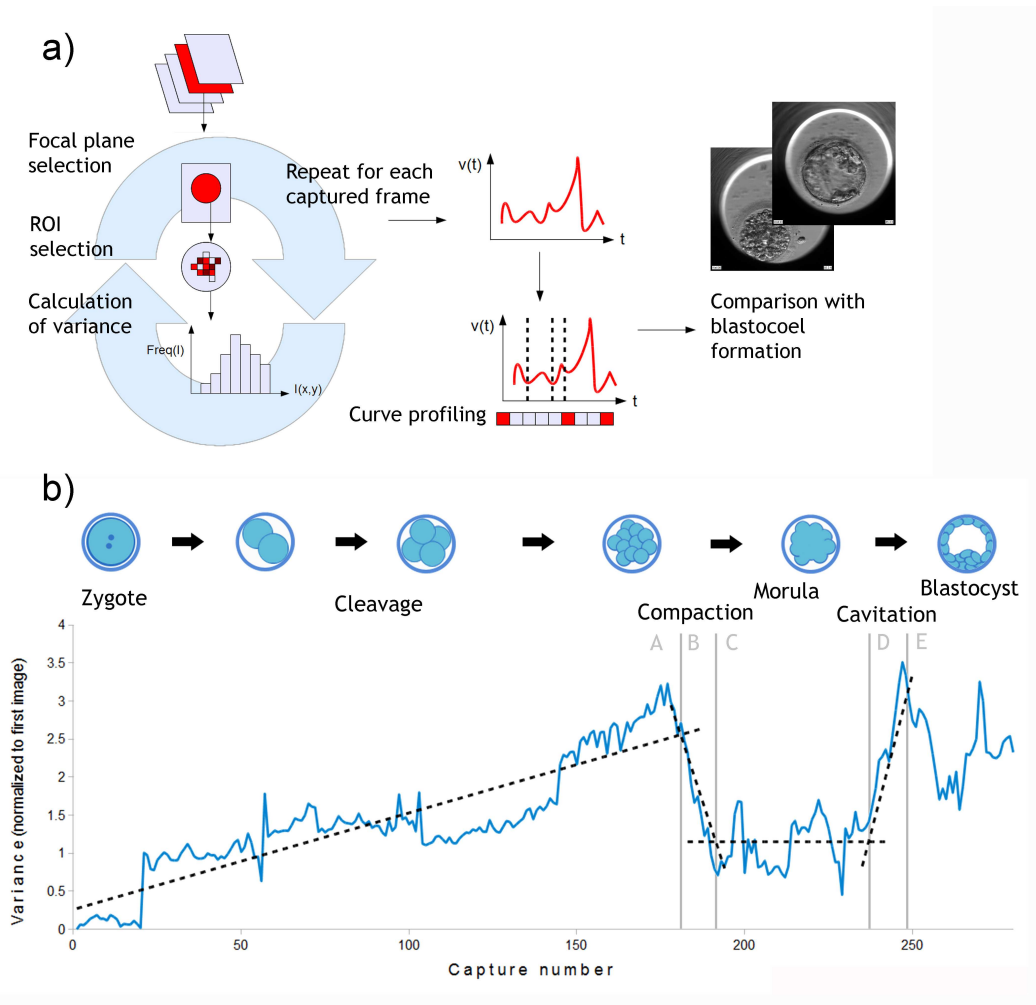
FIGURE 3.15: a) Illustration of computational pipeline of the captured image series of an embryo. The optimal focal plane from the image stack was selected. A region of interest (ROI) was selected within each individual image, and one value of the variance in image intensity was computed for each ROI. This process was repeated for each capture in the image series, resulting in a function v(t) describing the variance as a function of time. v(t) was then further analyzed for the occurrence of detectable key events, profiling the embryo development. Finally the profiles for embryos forming blastocysts and for those not forming blastocysts were compared. b) Image intensity variance of an embryo during the course of 280 frame captures, normalized to the first image in the series. Divisions during the cleavage stage are detectable as sudden increases in image variance, due to the number of increased edges in the image, as blastomeres undergo mitosis. At the onset of compaction, individual blastomere membranes are no longer distinguishable, and the variance drops and remains at a low level during the morula stage. The variance increases once more as blastocoel expansion sets in, and may fluctuate strongly during the blastocyst stage, if the embryo displays several cycles of collapse and re-expansion. The growth of the embryo has been considered in five stages. A) Initial divisions from fertilization to onset of compaction. B) Onset to completion of compaction. C) Morula. D) Cavitation. E) Blastocyst. The mean and change in variance has been calculated for each section. Dashed trend lines have been added for illustrative purpose [5].

FIGURE 3.16: Profile of three representative embryos showing decreasing quality (a-c). Variance was calculated from the image intensity at a circular region encompassing the center of the embryo. A few example images are shown at points where characteristic changes are visible in the variance profile. For a good quality embryo (a), mitotic divisions are visible as successive increases in image variance, and the morula stage as a period of lowered variance. b) illustrates a clearly expressed pronuclear breakdown, but experiences fragmentation during the cleavage stage, even though a blastocyst is eventually formed. In c), the pronuclear breakdown is also apparent, but the embryo develops early fragments, never reaching a blastocyst stage [5].
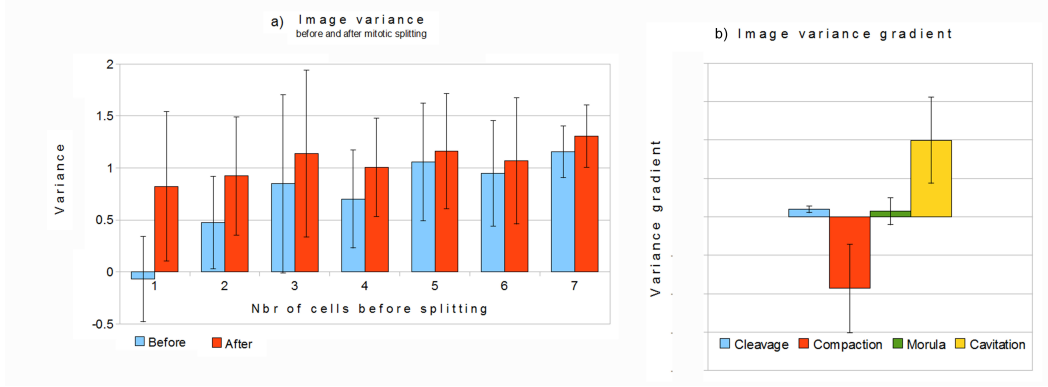
FIGURE 3.17: a) Average variance for 14 training embryos before and after mitotic division. P-values are P $\leq$ 0.05 for first and second division, P $\geq$ 0.1 for division 3-7. The negative variance before the first division is due to the drop in variance during syngamy. b) Gradient of image variance for embryo developmental stages for the 14 training embryos. P $\leq$ 0.001 for adjacent stages [5].

3.5.2   Results

In total, 37 out of 39 tested embryos had sufficient quality to detect the first five mitotic divisions. Computational detection was compared to manual detection for divisions of up to the 4 cell and 8 cell stages (Figure 3.17). The uncertainty in the exact location of division increased with the number of blastomeres. From the total image set of 37 embryos, 100% of divisions from 1 to 2 cells were detected, 73% from 2 to 3 (or 4) cells, 30% from 3 to 4 cells, and 59% from 4 to 5 (or 6) cells. The identification of the time at which a mitotic division occurred was comparative to manual detection of the same event, since also experts would sometimes disagree on the number of cells in the image (Figure 3.18).

The method of locating the timing of mitotic divisions shows a larger span between maximum and minimum deviation from the true position compared to manual detection, but on average, our method performed better. 62% of cleavages identified by automatic detection were located at the exact same capture frame as manually identified by experts. The same agreement for manual detection between different experts was only 35%. The results of the automatic method improved if only the 1-4 cell stages were considered, compared to all 1-8 cell stages. The results depended heavily on image capture frequency and at a rate of three captures per hour, some cell divisions were lost by under-sampling, most commonly the 3 and 5 cell stages. The change in variance was most prominent between the compaction and cavitation stages, one order of magnitude higher than that for the entire cleavage stage. Results also showed a high variation between patients, especially in the duration of the morula stage (Figure 3.19). Last, a set of measures from the time series profiling was used to classify the embryos into two groups, one of
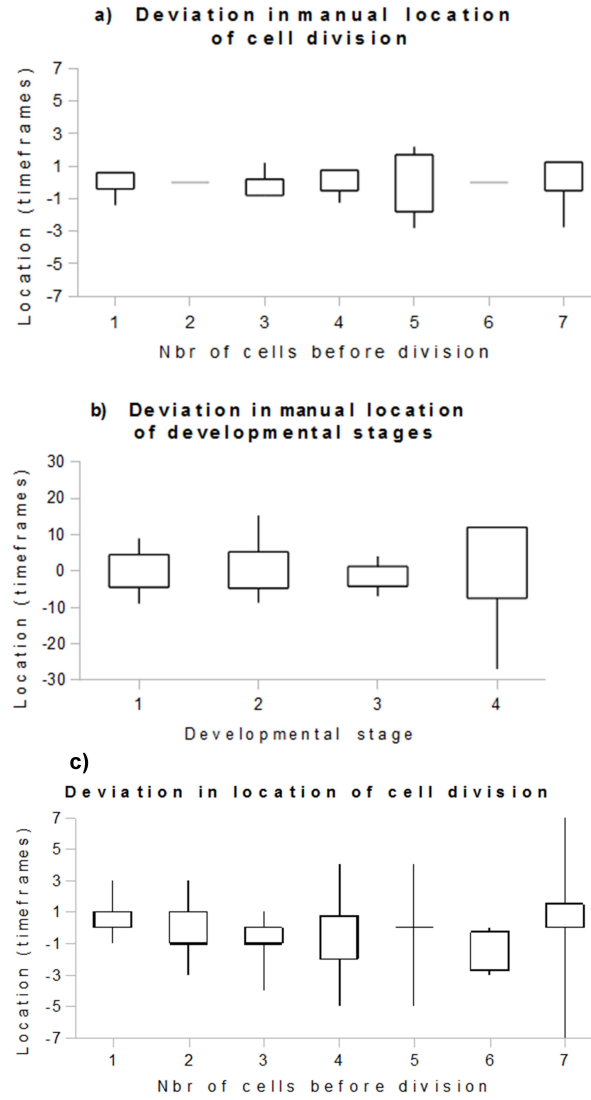
FIGURE 3.18: Standard deviation in manual annotation for the evaluation embryo in terms of a) timings of mitotic divisions up to 8 cells and b) detection of the beginning of developmental stages: 1-Compaction, 2-Morula, 3-Cavitation and 4-Blastocyst. Bars represent lower to upper quartile, whiskers minimum and maximum values. c) The deviation from expert determined location of division in terms of timeframes for the divisions that were detected, plotted vs the number of cells preceding the division. Bars represent lower to upper quartile, whiskers minimum and maximum values [5].

good quality, one of bad. The result was compared to the manually observed appearance of a blastocyst at the end of the time series, assuming the embryos forming blastocysts as being of good quality. The best overall result was correct detection of a blastocyst being formed in 71.8% of cases, but at a cost of 28.2% false positive detection with little sensitivity to parameter setting [5].
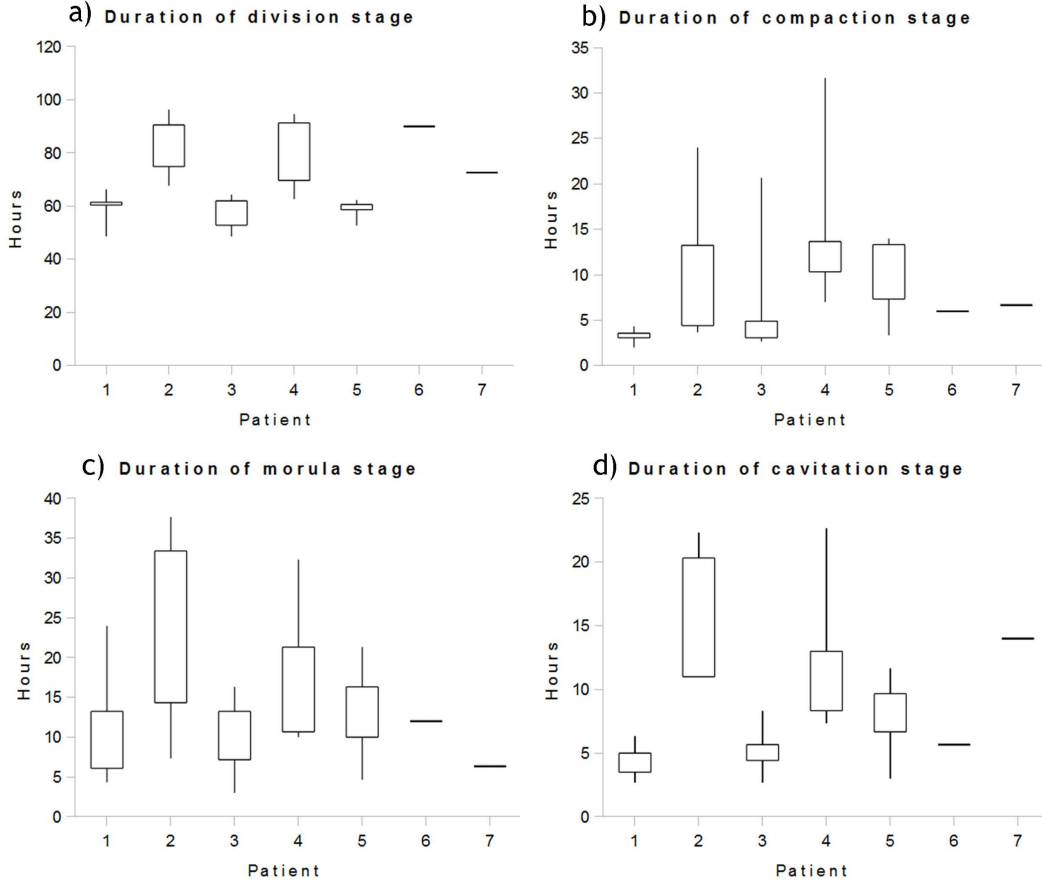
FIGURE 3.19: Duration in hours, automatically measured, of four stages of embryo development for seven patients (Total 28 embryos). a) Cleavage (the time from first frame to onset of compaction). b) Compaction (time from onset until completed). c) Duration of morula stage. d) Duration of cavitation stage (time from onset of cavitation to blastocyst). Patients 6 and 7 had only one embryo each completing all four stages. Bars represent lower to upper quartile, whiskers minimum and maximum values [5].

3.5.3   Discussion

As in Chapter 3.3, a time profiling is essential for understanding embryo development, and requires a dimension reduction from $X \cdot Y$ pixel values to something more manageable. Here it is shown that even the most extreme dimension reduction, down to a single metric, is enough to give an indication of embryo behavior through all stages from fertilization to implantation. It is theoretically possible to increase the detail of the

profile by combining several one-dimensional metrics, by combining a more fine-grained spatial filtering or using a combination of both.

3.6 Detection of Syngami
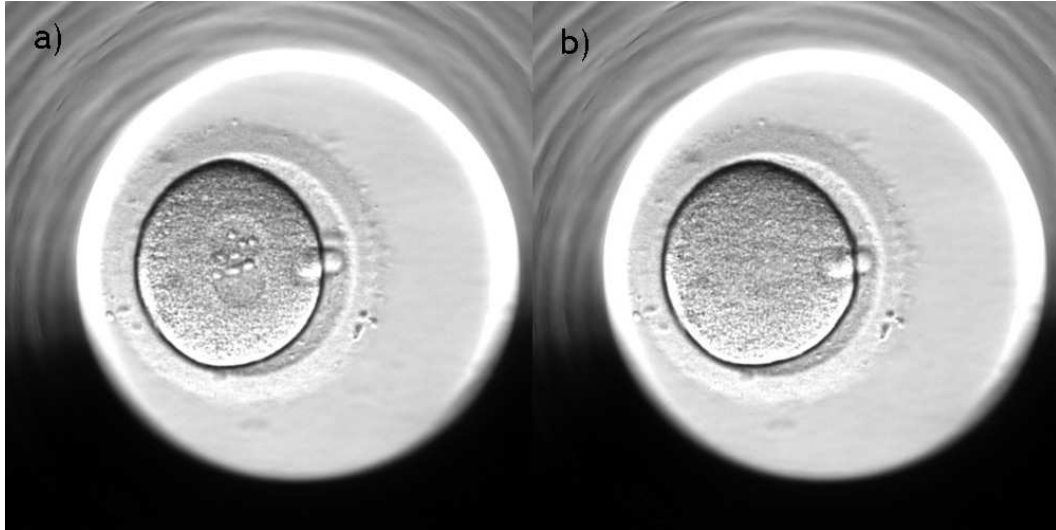
FIGURE 3.20: Example of pronuclear disappearance in preparation for the first mitotic splitting of the zygote. Before (a) and after (b) [2].

Considering the internal of the cell, there are a few structures of high diagnostic interest visible in HMC images, the nucleus being the most prominent. When comparing image time lapse series in a clinic, the moment of fertilization is often used as a reference point in time. If combining image data from several clinics, the time from fertilization to the first captured image may not be known. In this case, the moment of pronuclear appearance or disappearance (syngami) can be used as a common calibration point (Figure 3.20). As such it has the benefit of being short (less than one image capture = 12 min) and normally being the first easily detected biological event following the start of incubation. The timing of syngami at the onset of the first mitotic splitting as well as the general shape and size of nuclei may also provide clues to embryo health [50, 100, 101]. Paper II [2] describes a computerized detection of syngami using a set of different computational approaches.

3.6.1 Methods

For 20 embryos, 20 images per embryo were selected so that approximately half the time series showed pronuclei and the other half did not. After a spatial filtering of embryo
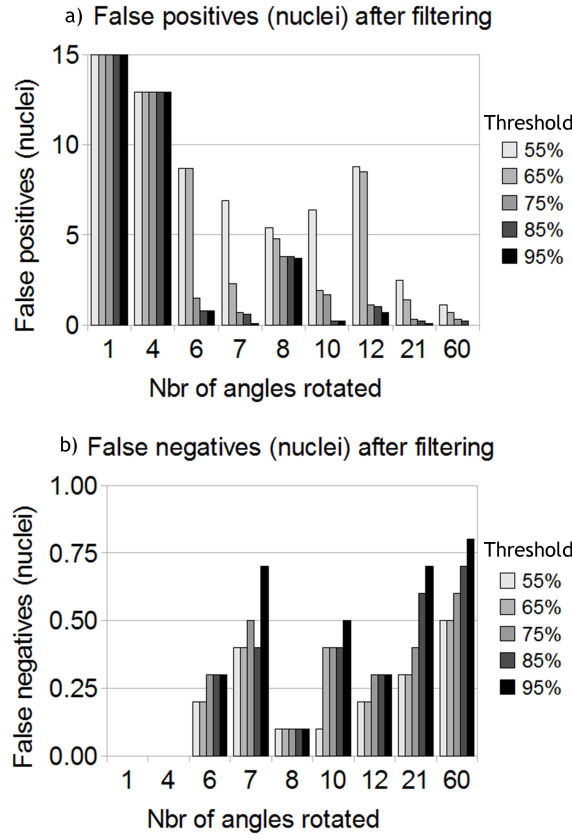
FIGURE 3.21: Number of failed detected nuclei as a) false positive (FP) and b) false negative (FN), respectively, calculated average over 10 images as function of number of angles each image has been rotated and the threshold % of images where the nucleus needs to appear in order to be counted [2].

location (Chapter 3.1) was applied. Four methods were chosen and compared: direct detection of PN using the HT (HT+PN), comparison of the image difference (ID)

$$\Sigma_{x=1}^{n}\Sigma_{y=1}^{m}(I(x,y,t_k) - I(x,y,t_{k-1}))^2, \tag{3.1}$$

the image variance (V),

$$v(t) = \frac{1}{n \cdot m} \cdot \Sigma_{x=1}^{n}\Sigma_{y=1}^{m}(I(x,y,t) - \overline{I(x,y,t)})^2, \tag{3.2}$$

and the image variance within the ROI selected as the embryo outline by the HT (HT+IV). The ability of each computational approach was evaluated for its ability to separate images from before and after the moment of syngami. For detection of circles using the HT, false positives were defined as circles displaced more than one nuclear radius. When comparing the circle location, a displacement of the centroid of at most 3 pixels resulted in the shapes being counted as identical.

3.6.1.1   Noise Reduction in the Hough Transform

The outline of the embryo itself is fairly strong, and could often be detected directly using the strongest HT circle (Figure 3.22h). In contrast, the nuclei are very weak, even when focus is optimized. For those cases the HT resulted in a high number of randomly appearing false positive hits. By introducing a slight disturbance to the image in the form of a trigonometric rotation

$$
\begin{aligned}
x_{n+1} &= x_n \cdot cos(\theta) - y_n \cdot sin(\theta) \\
y_{n+1} &= x_n \cdot sin(\theta) + y_n \cdot cos(\theta)
\end{aligned}
. \tag{3.3}
$$

textcolorgreenAs the pixels in the image are square-shaped, a rotation (except for 90°) results in slightly shifted absolute pixel values, causing the false circular shapes to appear in new locations, whereas the true shapes were slightly more consistent. This can be exploited by successively rotating the image, apply the HT to each rotation and average over the results. Figure 3.22a-g sketches the flow [2].

The described method was used in several enclosed papers. First to track the embryo throughout the image sequence, making sure the change of medium (2-3 occasions per 5 day sequence) did not affect the measurements. Second, the same spatial filter was used to exclude image content external to the embryo and as base for embryo region selection. Third it is used to detect nuclei using the rotation step described above (Chapter 2.3.5).

The number of angles to compute, and the threshold over which a structure needed to appear were optimized using a series of training images and the results are given in Figure 3.21. Here, false negative is determined as the probability of loosing the pronuclei detection. A rotation of 4 angles (90°) gave identical results, as expected. For a rotation of 6 angles (60°), an improvement was apparent for a threshold of 75% and above, giving a reduced number of false positives to below one per image at the expense of $\leq 30\%$ risk of obtaining a false negative. For an increased number of rotations, a further decrease in false positive rate was observed but at the risk of an increased number of false negatives.

3.6.2   Results

In Figure 3.23 the change in all metrics is compared before (images with pronuclei) and after (images without pronuclei) syngami. The spatially filtered variance gave the only statistically significant difference, and only at spatial filtering of $\leq 50\%$ of embryo radius. Figure 3.24 illustrates the variance shift observed using two image examples
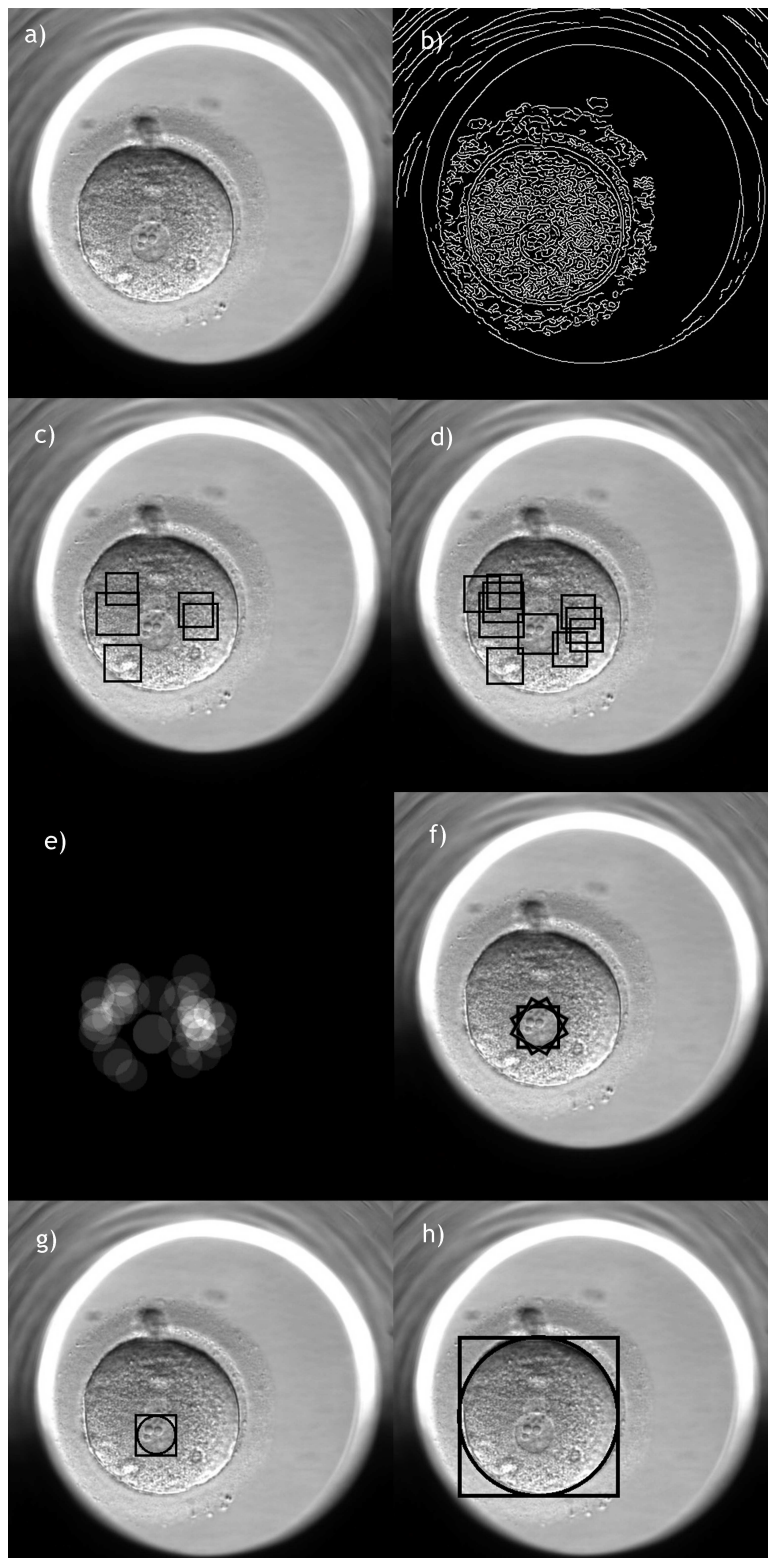
FIGURE 3.22: Detection of zygote and pronucleus in human embryo. a) Original image. b) Canny edge detection. c) 5 most significant circular structures selected. d) 10 most significant circular structures selected. e) Overlap of circular structures selected from the same image rotated $6 \cdot 60°$. f) Outline of pronucleus indicated, overlap of three calculations at separate angles. g) Outline of pronucleus selected. h) Outline of zygote selected [2].
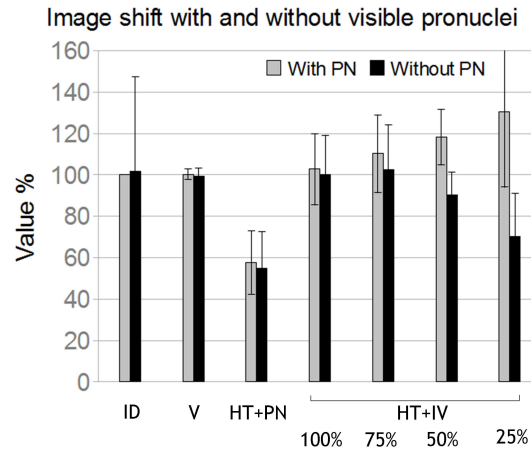
FIGURE 3.23: Values before and after syngami, averaged over 10 images each from 20 different embryos. Vales are given in % compared to the last image on which pronuclei are visible by eye. Time between images captured is 20 min. Measurements used were image difference (ID), image variance (V), detection of pronuclei using Hough Transform (HT+PN) and image variance at four different spatial filter settings (HT+IV) [2].

before and after syngami, and in Figure 3.25 the variance at 50% has been plotted (at image number 10), relative to the first image in the series. Syngami effectively lowers the variance of the embryo interior by approximately 25% over the course of at most three image captures. Time between images is 12 minutes. The increase in image variance approaching image number 15 is caused by the onset of mitosis. The best overall result was achieved at 83.0% accuracy for all 20 image series, using 20 images per series.

3.6.3   Discussion

The ability to detect syngami is of paramount importance for multiclinical trials, as it detects a common reference point in time with which to compare embryo growth. Luckily, pronuclei are the easiest detectable nuclear structures, since the embryo at this time only consists of a single zygote. Further detection of nuclei during the cleavage stage would provide important information of nuclei number, shape, size and position. It would require the same spatial filtering, but now using each blastomere as filter reference, rather than the entire embryo. Thus, a continued analysis of nuclei for later stage embryos requires a successful segmentation of embryo blastomeres.

3.7   Embryo Cell Detection

An accurate detection of cell region is one of the keys to successful digital image cytometry. As cell borders are usually defined by the edges in the image produced by the
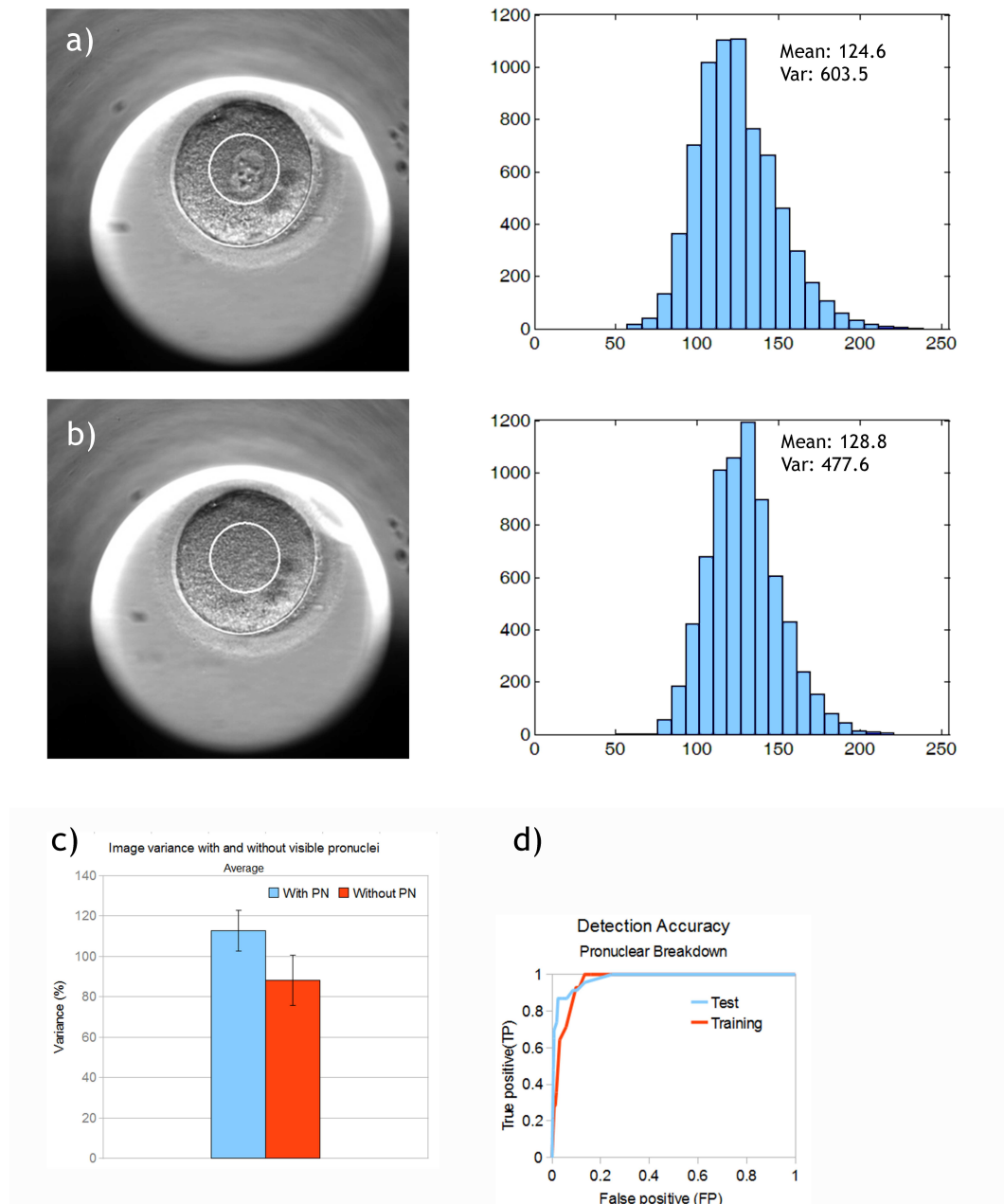
FIGURE 3.24: Calculation of variance in image intensity using pronuclei as an example. Images a) and b) were captured 20 min apart. The frequency of image gray scale values (0-255) within a selected ROI (white circle) at half embryo radius has been plotted as histograms, and the mean and variance calculated. c) Difference in image variance before (blue) and after (red) pronuclear breakdown. Standard deviation calculated as mean over the training set of 14 embryos ($P \leq 0.0001$). d) Detection accuracy of the training (14 embryos) and test (25 embryos) sets, respectively. The computation is governed by a single threshold (gradient of decreasing variance over time). Pronuclear breakdown is defined as gradients larger than some threshold, yielding an increase in true positive (TP) accuracy as the threshold decreases [5].
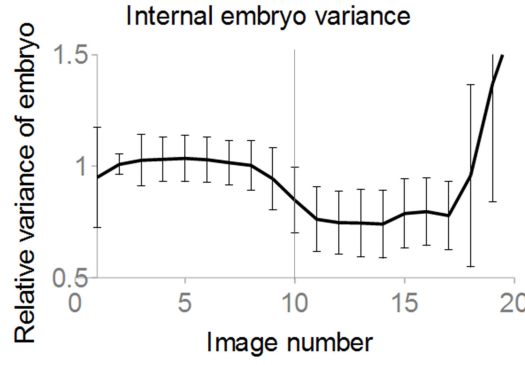
FIGURE 3.25: Image variance of ROI selected by Hough Transform at 50% of embryo radius, compared to the first image in the series, before and after syngami (marked by gray vertical line). Black line shows average over 20 embryos, bars indicate standard deviation [2].

differences in refractive index between the cell membrane and the internal and the external of the cell, edge detection is the most straight forward approach to cell detection. Although many common edge detectors are very efficient for edge detection in a single gradient image (Chapter 2.3.2), the detection of cells in an HMC embryo image set up requires a different approach. In Paper III [3] we describe an approach for detecting edges in bi-gradient images and apply it to HMC images of human embryos.

3.7.1 Methods

Images from 18 embryos containing 1-6 cells with cell outlines marked manually were used for the analysis. The image was first spatially filtered (Chapter 3.1) and a Canny edge transform applied. A set of elliptic model candidates representing cells were then located using a HT (Chapter 2.3.5), producing an initial set of cell candidates. In order to reduce computation time the candidates were computed using a circular HT, and the circular shape was then adjusted to an ellipse by stretch and rotation. Each candidate was then matched to a double thresholded edge map of the raw image. Figure 3.26 shows an example of edge detection, matching to edge map and validation to manual detection for a 2 cell and 6 cell case, respectively. The edge map was constructed from the lowest and highest pixel ranges in the image F using the threshold

$$T = m \cdot median(F). \tag{3.4}$$

where the median was selected to reduce the influence of pixel outliers. The use of two thresholds gave two image masks which were then added to a binary edge map, excluding the middle range of pixel value (*i.e.* the background). A focus filter provided a convex

hull around the regions in focus (Chapter 3.3.1.1) and was used to further exclude image background such as the zona pellucida region. In matching, the elliptic perimeter was compared to the edges, and as one cell was found, its trace was successively removed from the edge map. The matching progressed until all cells were found [95]. It is noteworthy that this approach required an initial guess of the number of cells in the image. In our case, we used the classification described in Chapter 3.2 to provide this primer. To compare results we define a correctly located cell as one where the centroid of the predicted ellipse is contained within the manually defined region and a correctly described cell outline as one where the Intersection over Union (IOU) between result and prediction is $\geq 70\%$.

3.7.2   Results

Results show a high degree of localization agreement, finding $\geq 80\%$ of cells up to the 4 cell stage and $\geq 70\%$ up to the 6 cell stage (Figure 3.27). Accuracy in detecting outlines was high only for the 1 cell stage, then at best 43% up to the 4 cell stage (the 3 cell case slightly lower). The main reason for this was deviation from a perfect elliptic shape in manually drawn outlines, where the 2 and 3 cell case often displayed more irregular outline than a case of more than 4 cells. Table 3.1 shows the accuracy as the similarity requirement is reduced to 40%. The overall score was 43.8% for IOU = 70% and 83.8% for IOU = 40%.

TABLE 3.1: Comparison of embryo outline detection

Number of cells	1	2	3	4	5	6	Total	[102]
IOU$\geq$70%	90.0	40.0	33.3	43.4	31.4	25.0	43.8	83.9
IOU$\geq$40%	100.0	87.5	76.9	85.5	74.3	78.3	83.8	-

3.7.3   Discussion

The construction of the edge map transforms the bi-gradient image into a single gradient image and the matching of candidates of fixed shape to the map eliminates the need for a complete outline. This is an advantage when handling potentially overlapping objects in a noisy image. Limitations include the inability to separate objects with a high degree of overlap, naturally limiting this method to 6-8 cell embryos. A previous study using Dark Field microscopy reports an average cell location accuracy of 83.9% using the same similarity measure (IOU $\geq$ 70%) for up to the 4 cell stage [102]. No results on each number of cells or confidence intervals are reported, making correct comparison difficult, but with the inherent capability of DF microscopy to detect edges, a discrepancy is not
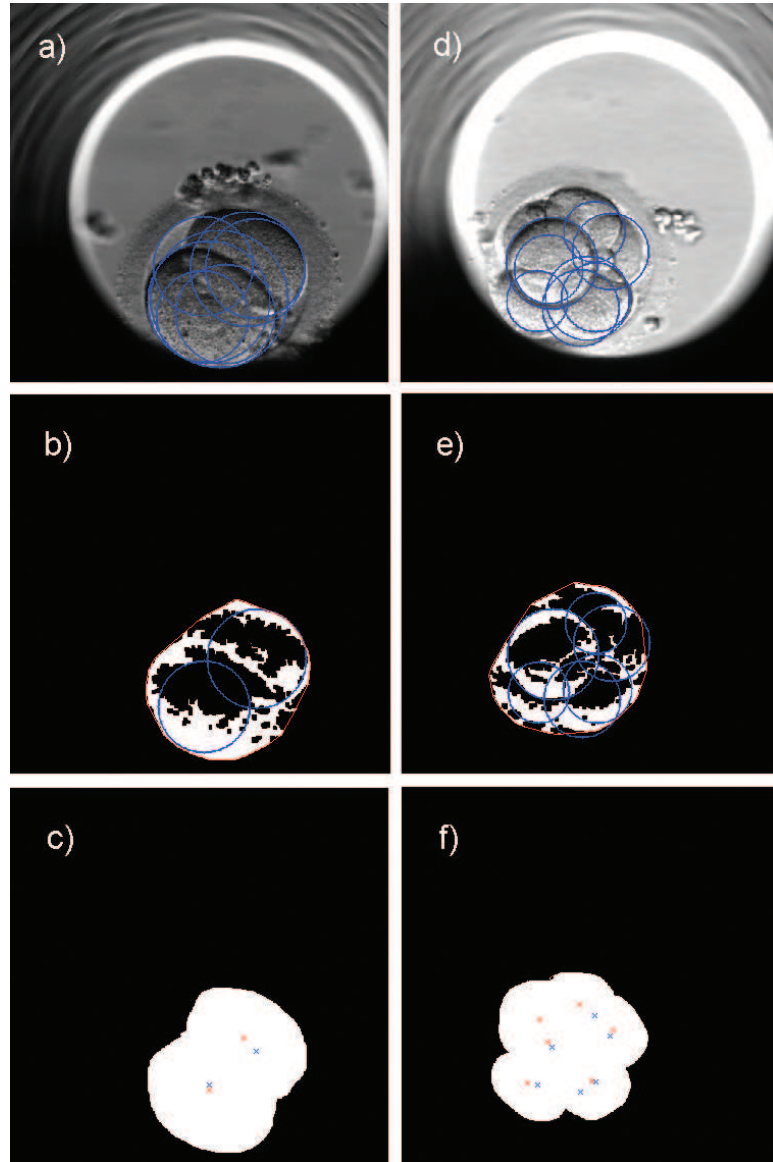
FIGURE 3.26: Step wise detection of cells in an embryo of two cells (a-c) and 6 cells (d-f). a,d) Original image, with elliptic candidates detected. b,e) Edge map with convex hull (red) and matched candidates (blue). c,f) Comparison between computed location (blue) and manual (red), x:s represent ellipse centroids [3].
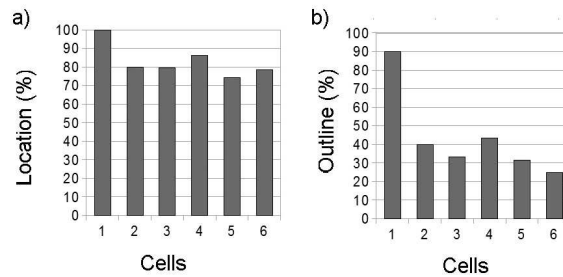
FIGURE 3.27: Localization (a) and Outline (b) detection accuracy from 1 to 6 cells [3].

surprising. However, it is difficult to extend a computational method based on DF to more than 4 cells, and DF is not capable of separating information from several focal planes. Further evaluation of this technique should therefore address this potential and include selection of cell outlines at several focal planes simultaneously, exploiting the fact that the focal filtering responds quite differently to objects in focus, compared to objects out of focus. See *e.g.* Figure 3.26d. In this image, the top cell is slightly defocused, making the dark cell membrane to appear slightly thicker than the underlying cells.

Chapter 4

Extended applications

4.1 Three-dimensional Visualization

Many physical cell models, including tumorous cells *in vitro* are two-dimensional. In their natural environment, cells grow and interact in three dimensions. Thick tissue samples containing hundreds or thousands of cells are hard to study for practical reasons, especially *in situ*. An embryo is an excellent model of three-dimensional cell interaction, containing anything from two to a few hundred cells in a self-contained natural environment. To study it in the third dimension it is necessary to perform either tomography or some kind of focal slicing in order to separate image planes. Focal slicing is the technically less complicated of the two. Paper IV [4] illusrtates this and in Figure 4.1 a 3D model is constructed using confocal microscopic images of embryos. 3D imaging is useful for a better understanding of the embryo structure and the positioning of blastomeres relative to each other. In this example, nuclei are marked for visibility but in principle any marked structure can be used. The drawback is that a fixed embryo can no longer be used for further analysis or growth. The same type of optical sectioning is also possible in some non-destructive techniques including HMC, however not with the same resolution as in confocal imaging. The majority of embryo HMC sample data used in this Thesis comprised 7-9 focal planes per time point. If a successful segmentation of this type of image could be achieved and three-dimensional visualization was possible, it would open new doors in the area of embryology, both for clinical purpose, but also for the understanding of early human development.
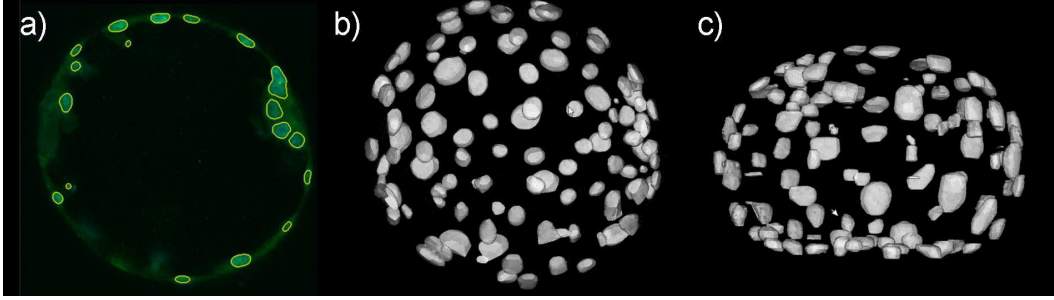
FIGURE 4.1: 3D rendering of the segmented data (a) [8] from a confocal image stack. Rendered top view (b) and side view (c) [4] of an embryo in blastocyst stage. The shapes shown are embryo nuclei (121 in total), spread along the inner wall of the blastocoel. The flattening of the embryo is due to the deformation caused by the imaging process. Blastocyst diameter is approximately 120 μm [4].

4.1.1   3D Modeling of Confocal Imaging

The short depth of field of confocal imaging makes it ideal for focal slicing and three-dimensional visualization. Paper IV [4] explores this option using confocal images of embryo nuclei marked with DAPI. Images were captured 1 μm apart. The results were segmented using a combination of region growing segmentation filters (Neighbourhood Connected Thresholding and Confidence Connected Thresholding) and Watershed segmentation [103]. The resulting bodies were then put together in a three-dimensional representation of the complete embryo [104] and the position and relative size of the blastomere nuclei were readily calculated. Note how the flattening of the embryo becomes apparent when applying a side view (Figure 4.1). This is a physical effect of being sandwiched between microscope slides. The nuclei along the perimeter form the trophectoderm. A denser pack of nuclei to the right of the top view image form the inner cell mass (Figure 4.1a-b).

4.1.2   2.5D Modeling of HMC Imaging

In Paper VIII [8], a manual segmentation of HMC images was used to compute a primitive circular embryo cell model. Blastomere outlines were selected using the software EmbryoSegmenter [105], and a simple spherical model was adjusted to the segmentation outline assuming the segmentation to be at the waist, *i.e.* on the widest part of the blastomere. The spherical shape represents a first order simplification of the true blastomere shape, and can readily be extended to a more complex model, if information from several focal planes is considered [106].
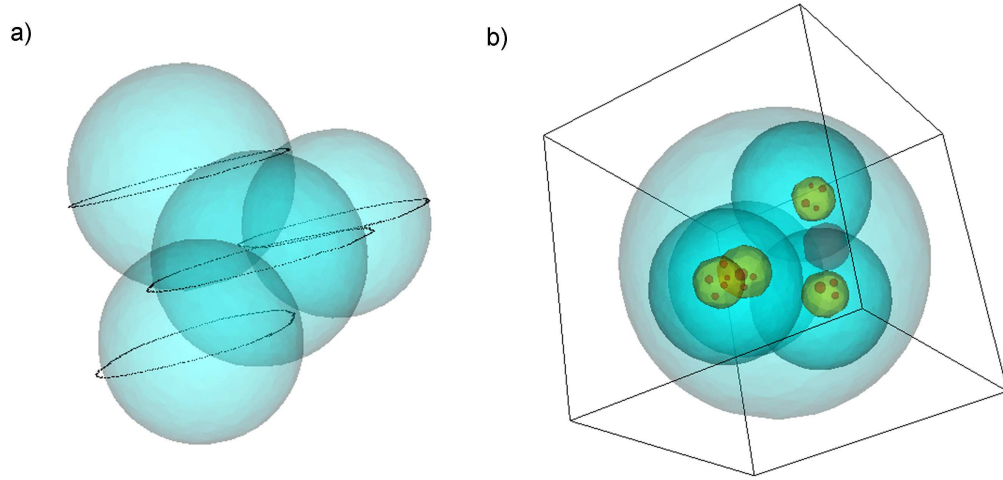
FIGURE 4.2: a) Segmentation outlines at blastomere waist are used to guide spherical cell models. b) All structures present in the original image can be visualized. Here showing nuclei and nucleoli, as well as a large fragment (visible to the right of image center). The different structures have been artificially colored for clarity [8].
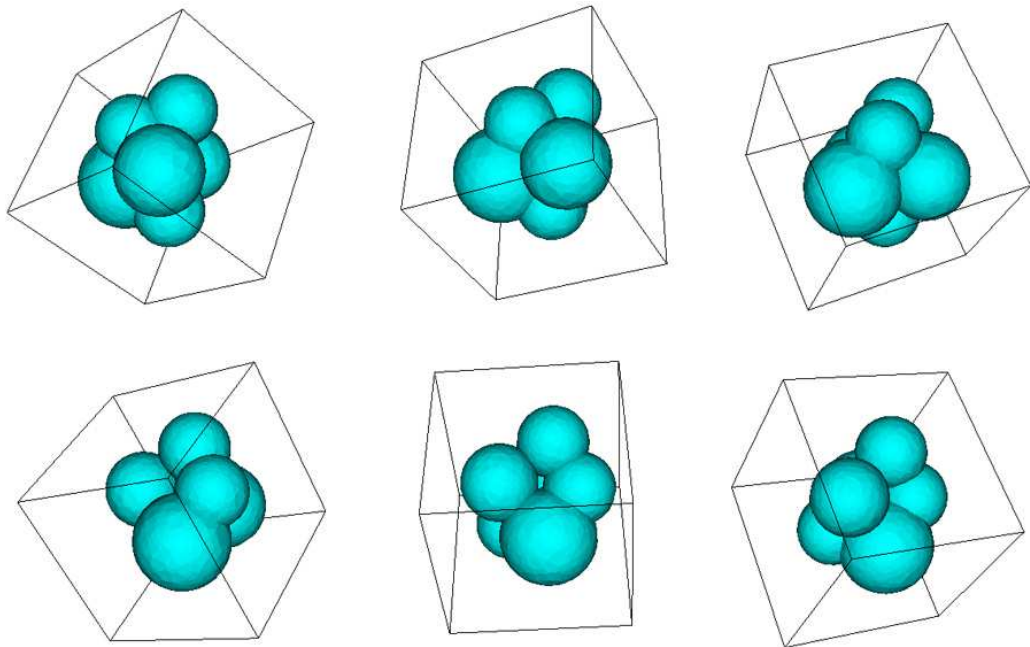
FIGURE 4.3: 6 cell stage embryo, with blastomeres modeled as spheres, rotated about its central axis. Bounding box added for clarity. Connected regions between blastomeres differ depending on blastomere size and location [8].

4.1.3   Cell Connectivity
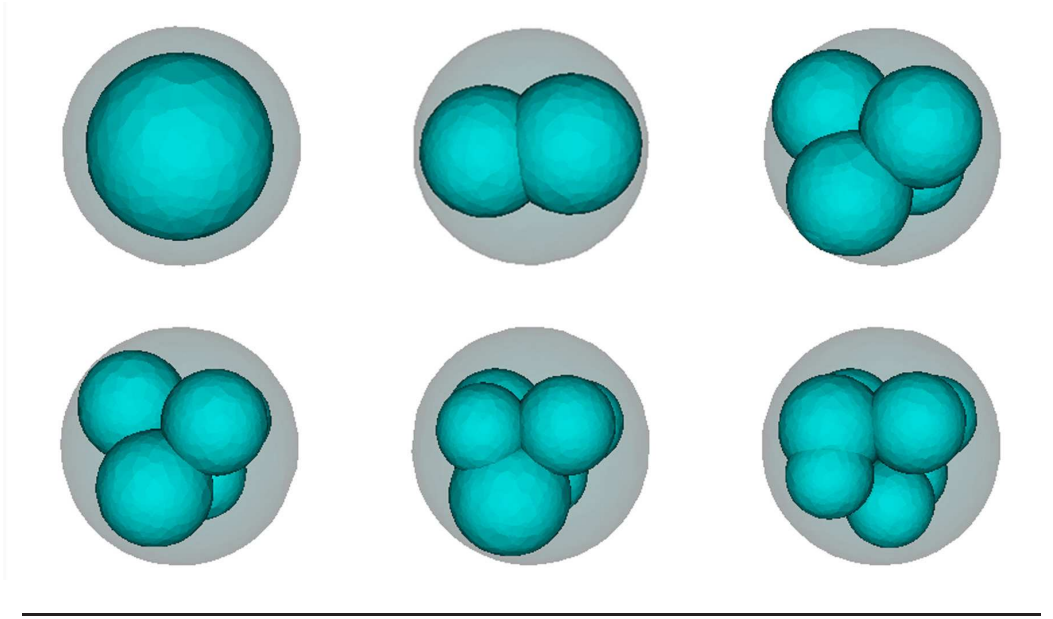
FIGURE 4.4: Cleavage stage embryo observed during development from 0 to 34h, show-ing 1, 2, 4, 5, 6 and 7 cell embryos, respectively. The boundary of the zona pellucida is indicated in gray [8].

Even with the simple model described in Chapter 4.1.2, some information on cell posi-tion, shape, size and connectivity can be extracted. During the cleavage stage, as the blastomeres undergo mitosis, contact regions between them increase when cellular com-munications are initiated. As the embryo grows, it transforms from a package of loosely aggregated cells into a complete entity. Contact regions increase between blastomeres during the cleavage stage, as cellular communications are initiated. At the onset of the compaction stage and at the following cavitation, the blastomeres adhere together more closely and start to form a single interacting multi-cellular organism. In Figure 4.4, six steps of development from one to seven blastomere stages of a cleavage stage embryo have been modeled. The increase in connectivity with higher cell number is shown in Figure 4.5.

The connectivity is computed from the spherical model

$$C = \frac{A_{con}}{(A_{con} + A_{uncon})} = \frac{r^2}{r^2 + 4p\pi R^2} \tag{4.1}$$
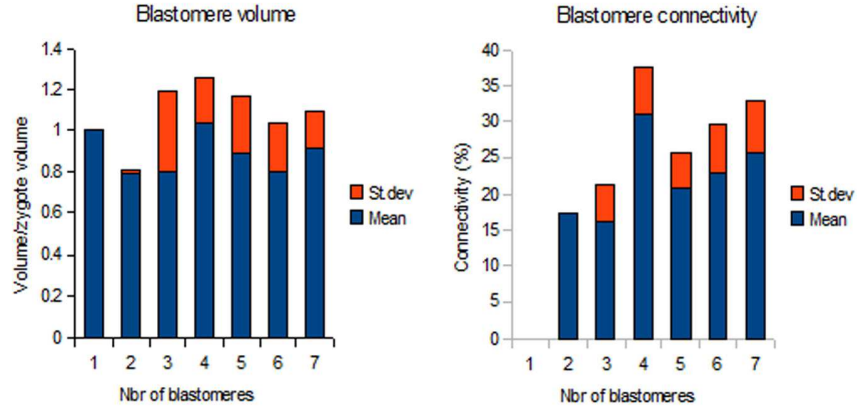
FIGURE 4.5: Blastomere volume and connectivity for the embryo in Figure 4.4. Blastomere volume has been normalized to the volume of the zygote and to number of cells in the embryo. The connectivity is displayed as percentage of surface area. Bars show both mean and standard deviation, calculated for each number of blastomeres. Apart from the 4 cell stage, which differs in both volume and connectivity, the total volume is constant, whereas the connectivity increases [8].
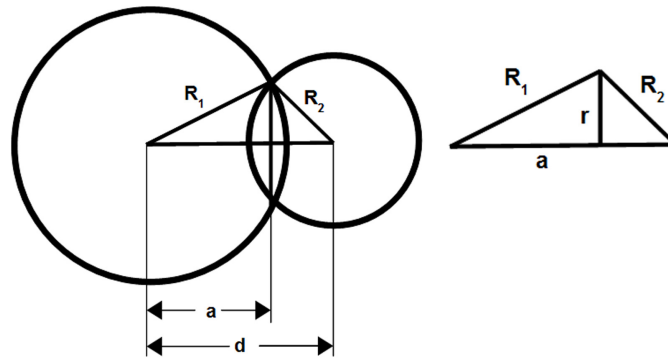
FIGURE 4.6: The intersection plane between two spheres is described by a circle, definitions [8].

where the value of p is given by integrating over the surface of the sphere across an angle φ and dividing with the total area of the sphere (Figure 4.6),

$$p = \frac{1}{4\pi r^2} \cdot \int\limits_{-\pi}^{\pi} \int\limits_{0}^{\varphi} r^2 sin(\theta)\delta\theta \cdot \delta\varphi, \qquad (4.2)$$

where

$$sin(\varphi) = \frac{r}{R_{1,2}}. \qquad (4.3)$$

4.2   *In vitro* Cell Monitoring for Drug Response Assessment

The genetic changes associated with a cancer tumor often differ between patients. As a consequence, the optimal treatment for cancer is also highly individual. Today's cancer treatments are often administered based on generalized knowledge or using trial-and-error. With a more personalized approach, it may not only be possible to significantly improve treatment and outcome for cancer patients but also to avoid the resource waste of drugs administered in vain. In addition to individual patient responses to drugs, individual tumor cells may also react differently to the same treatment, as some cancer cell types are more malignant whereas others may be more benign. By better understanding individual cell reactions to anti-cancer agents, it would be possible to target drugs better to the most malignant cancer cells.

Tumorous cell lines grown *in vitro* are a widely used model for cell studies. They can roughly be divided into suspension and adherent cells. *Suspension cells* are for instance lymphoma cell lines, growing in a medium solution where individual cells are physically separated from one another and little or no interaction occurs. *Adherent cells* grow attached to a substrate such as the bottom of a cell culture container or on top of each other. In terms of cell-cell studies, adherent cell lines provide a two-dimensional model for both cell growth, interaction and motility. When studied, adherent cells are usually detached from the growth substrate, marked with dyes or fluorophores and analyzed.

Most assays in use today study the average characteristics of large groups of cells and drug response studies have traditionally focused on concentration of a compound throughout the bloodstream and major organs. In *in vitro* analysis, the first step of drug testing, tumorous cells are prepared as parallel cultures and exposed to a range of concentrations. The effect is then measured and compared on a culture-by-culture basis. A more fine-grained analysis on cell-by-cell basis would make it possible to study how cells react when exposed to a toxin depending on their cell type, tissue type and stage in the cell cycle. Moreover, it would allow the entire analysis to take place in the same container, removing the need for parallel experiments and the need for adherent cell detachment. Single cells can be monitored over time, and different drug response to different cell types can be studied simultaneously within the same image field. This opens up a whole new perspective on cellular research, where cells are viewed not as a uniform mass, but as a spectrum of individuals. Understanding cell behavior on an individual level would not only allow for more personalized medical care, personalized studies of drug response and adverse drug effects, but also lead to a more detailed, in depth understanding of the mechanisms by which we function as biological entities.

In Paper VII [7], classification using a set of features from images of cells grown *in vitro* is used to study differences in cell morphology, health and viability while exposed to a toxic substance, etoposide, and imaged using DH (Chapter 2.2.5).

4.2.1   Background

Measure of cell viability using chemical assays is a common way to assess the effect of toxic exposure to cells. Often, reduction in viability is preceded by changes in genetic expression, metabolism and morphology [107, 108]. It can be a slow process which extends over several days depending on the concentration of toxin. In order to study morphokinetic properties at this time scale, non-invasive imaging is a necessary tool.

Etoposide is a cytotoxic chemotherapy drug, commonly used as an anti-cancer agent [109, 110], causing cell death by apoptosis. The effect to cells at sub-apoptotic concentrations and during long-term exposure is yet poorly understood, but reports of effects include drug resistance and risk of secondary leukemia [111].

4.2.2   Methods
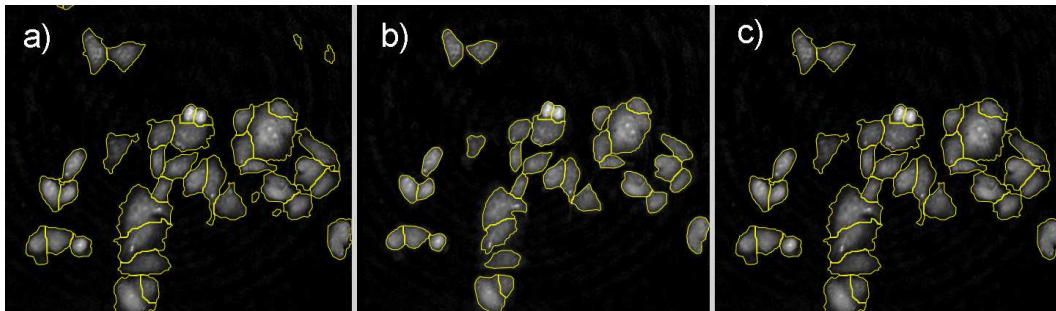
FIGURE 4.7: Cell filtering using double thresholding. The same image (control day 1) segmented using a) Minimum error thresholding, b) Otsu thresholding and c) The result of Minimum error thresholding, filtered using the Otsu thresholding to identify cells [7].

Digital holography is a full-field imaging technique where pixel values are directly related to cellular dry mass, and since the images from a mathematical point of view are similar to those from a standard epifluorescence microscope the same image analyzing algorithms can often be used. In cellular biology, DH has been used for a number of applications, including *in vivo* studies [112], sub-cellular motion within living tissue [113], migration studies [114–117] and studies of morphological changes [118–121], proliferation [77] and apoptosis [122].

A robust segmentation method for many cell image modalities is the watershed segmentation (2.3.3). Here, a threshold will define the boundary between the foreground cells and the background. The selection of threshold computation can yield highly varying results to segmentation (Figure 4.7). In this example Otsu thresholding (OT) [123] yielded a higher cut-off than Minimum Error Thresholding (MET) [124], resulting in a more accurate location detection of cells as out-of-focus debris in the growth medium were not selected. On the other hand, MET gave a more accurate cell outline. Assuming that Otsu thresholding gave a correct location detection of cells, the segmentation using OT could be used to filter the result using MET, effectively removing image debris. The filtration was done comparing the bounding boxes of MET segmentation and OT segmentation and discards all objects not represented in both [125].

Prostate cancer cells were cultivated during three days after exposure to etoposide and imaged by DH once per day. Images were segmented into regions, each describing the contents of one cell. Each cell region was then modeled using a set of gray scale features, and classified in groups according to morphology (Figure 4.8). The result was compared with the known concentration to which the cells had been exposed, and also to the result of MTS, a standard chemical assay for viability. The critical step in image processing was the segmentation of the image into cellular regions. A spatial filter was applied to adjust the description of the cell boundaries. Three methods were used to classify cells; a DAG-SVM [99], a Naïve Bayes classifier and a RF [126]. 10% of cells were used for training, the rest for validation. Each classification was repeated ten times, and the results averaged. The performance was evaluated using a standard 4 core PC with 8 GB RAM, with the code implemented in Matlab$^{\circledR}$ 7.12.0.635 (R2011a).

Cells were exposed to a total of 5 concentrations of etoposide, and we used the we use the sensitivity and specificity as indicators of classification quality:

$$Sensitivity = \frac{E(c \geq c_{min}) \cup T(c \geq c_{min})}{T(c \geq c_{min})} \tag{4.4a}$$

$$Specificity = \frac{E(c < c_{min}) \cup T(c < c_{min})}{T(c < c_{min})} \tag{4.4b}$$

Here, $E(c)$ is the set of cells being classified as concentration c and $T(c)$ is the set of cells actually exposed to concentration c and c_{min} is a threshold concentratiron. Cells were defined as treated if $c \geq c_{min}$.
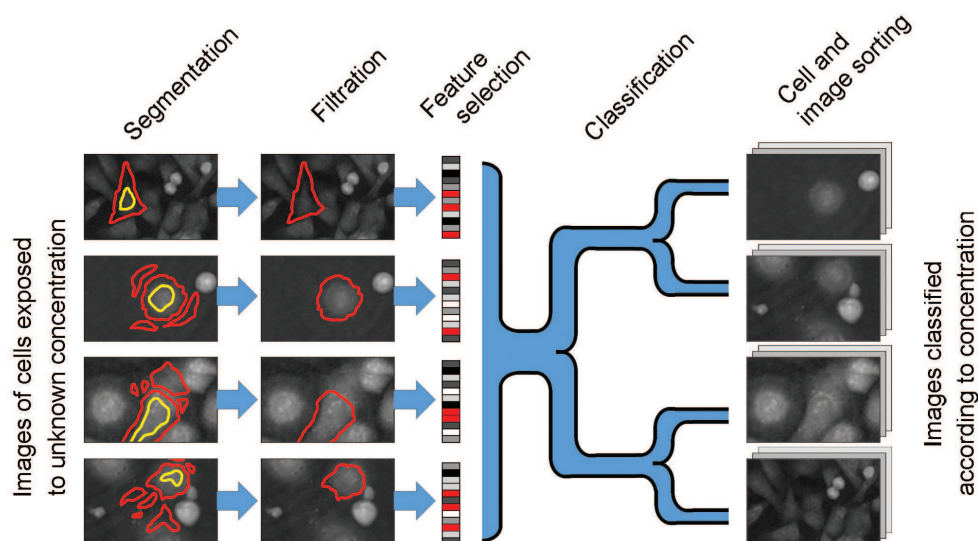
FIGURE 4.8: Algorithmic overview of the classification of cells exposed to a treatment, based on their imaged morphological characteristics. Images are segmented and filtered, features extracted and used to train a supervised classifier [7].

4.2.3 Results

First the individual cell data is combined sample-by-sample to study the morphological effects in bulk (Figure 4.9) for the ten most prominent features. For concentrations 0.25-1 μM, no effects were detectable by MTS absorbance (Figure 4.11), but clear changes occurred in morphology. The cell area was the most dominating morphological effect of long-term LC-etoposide exposure, but textural changes were also prominent. All ten features showed statistically significant differences between treated and control.

Qualitatively, treated cells differed from control in a number of ways (Figure 4.10). The control cells show normal exponential cell growth, filling the culture vessel by day 3. Healthy cells display throughout a compact spindle or prolonged elliptic shape, with uniform cell size from day 1 to day 3. The 5 μM concentration shows the most distinct case of toxicity, with immediate arrest in cell growth and reduction in cell number at day 2 and day 3. Cell shapes are slightly rounded day 1 and by day 2 and day 3 mostly spherical and detached from the bottom of the cell culture vessel. Concentration from 0.25 μM - 1 μM show intermediate stages of varying severity. Images from 1 μM day 2, 0.25 μM day 3 and 0.5 μM day 2 and day 3 show similar morphological changes. These changes, which appear to be consistent with onset of, or low level exposure to, etoposide toxicity are rounded, flattened shape and increased cell area with cells still attached to the bottom of the cell culture vessel. In some cases, an increased granularity can be seen within cells, though this may be caused by the increased area and flatness, giving a
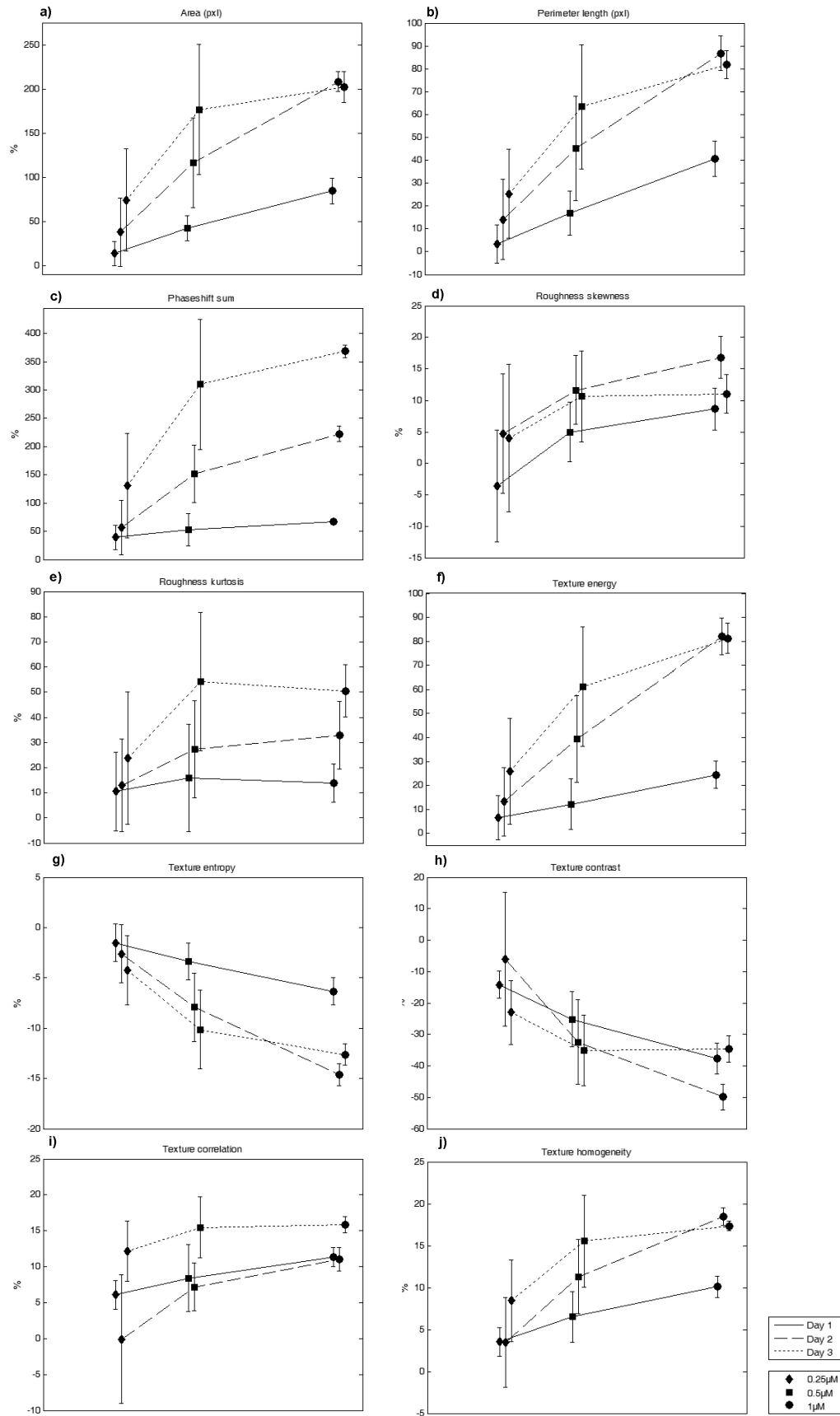
FIGURE 4.9: Variable values compared to control, calculated from sample-wise mean. Day 1 (solid), 2 (dashed) and 3 (dotted) after treatment. x-axis represents etoposide concentration; 0.25 μM (diamond), 0.5 μM (square) and 1 μM (circle), respectively. Error bars represent standard deviation, and for clarity, each series has been slightly offset in x-direction [7].
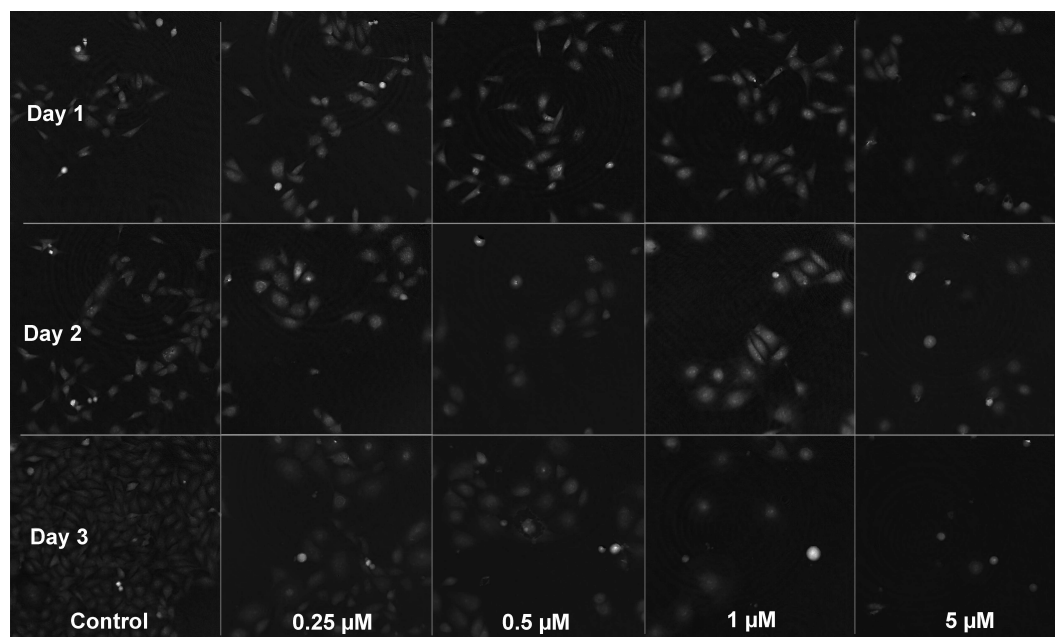
FIGURE 4.10: Typical appearance of DU-145 imaged by Digital Holography after several days exposure to LC-etoposide treatment [7].
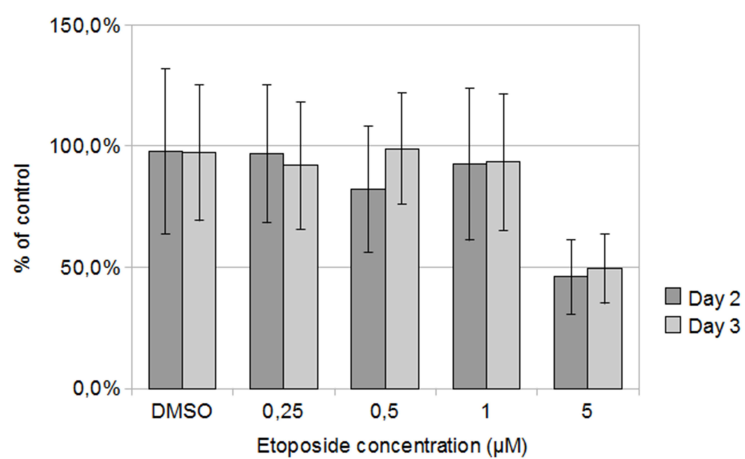
FIGURE 4.11: MTS absorbance as % of control day 2 and 3, as a function of etoposide concentration, after 2 hours of MTS incubation [7].
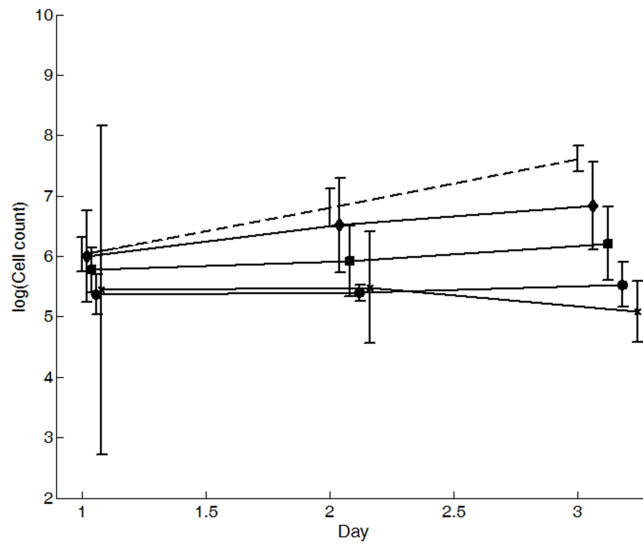
FIGURE 4.12: Proliferation curves for DU-145 under the effect of etoposide 0 μM (dashed), 0.25 μM (diamond), 0.5 μM (square), 0.1 μM (circle and , 5 μM (cross). Measurements were done at the same time and values for x-axis have been displaced slightly to display confidence intervals [7].

higher contrast to variations in internal cellular structure than can be seen in the health compact elliptical shape. It appears as if cell proliferation has been halted at this stage, but with little or no increase in cell mortality, and the cell number is approximately constant.

Measurement of cell growth was consistent with the MTS results, and close to exponential for the control cells, as could be expected (Figure 4.12). It is notable that the morphological changes were detectable already day 1 with DH for all examined concentrations of etoposide, whereas changes in MTS absorbance were only detectable for the highest concentration of etoposide (5 μM) on day 3.

There was a lower probability of separating control and the lowest concentration used (0.25 μM) as well as of separating higher concentrations from each other. When comparing the positive and negative predictive power we note a more rapid change from 0.25 μM to 0.5 μM, reflecting the fact that cell morphology changes more rapidly at these concentrations. There was a best case sensitivity in separating control from treated (etoposide $\leq$ 0.25 μM) at 88%$\pm$0.17 on Day 3, at a specificity of 94%$\pm$0.001 .

The data on 5 μM was very scarce due to very low number of cells in each image (typically $\leq$ 10 cells), giving high standard deviations and low predictive power in results. The accuracy of prediction increased throughout Day 1-3 for all classifiers, reflecting the fact that cell morphology becomes more and more distinct (Figure 4.10). In most cases, the
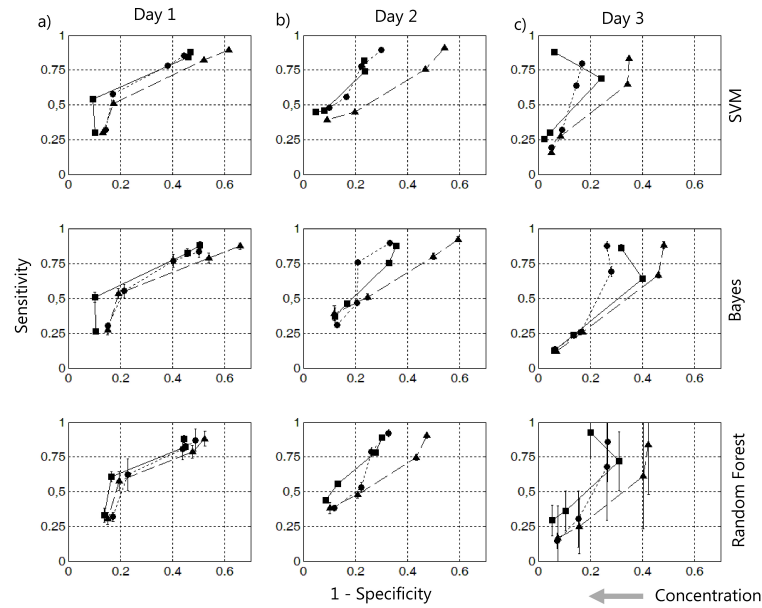
FIGURE 4.13: Positive predictive power for classification of cells exposed to a concentration or higher (x-axis). Classifications used are SVM (Black), Naïve Bayes (Gray) and RF (White). Three methods for outline detection were used in segmentation; Otsu, Minimum Error (Min.Err.) and Minimum Error filtered with Otsu (Filtered). Results have been grouped according to day of measurement taken [7].

multiclass SVM was the best classifier. It was closely matched by RF, but the errors for RF were higher, especially when cell morphology differed strongly (Figure 4.13). Day 1 there was only marginal difference between the three segmentation methods, but for Day 2-3 the Otsu thresholding gives the most reliable classification. The computational performance of the Naïve Bayes classifier was an order of magnitude better than that of the SVM, while the RF rank between the two [7].

4.2.4 Discussion

Despite the high positive predictive power, the ability to pinpoint the concentration for a single cell was low, reflecting the individual nature of cells. Thus, single cell imaging may be more convenient for individual cell monitoring than for determination of drug exposure. Variations in the cell border using different selections of threshold also had a large impact on accuracy, reflecting the fact that one of the most prominent effects of LC-etoposide exposure was the increase in cell area. For Day 2-3 the Otsu thresholding gives the most reliable classification. This can be due to the fact that cell area is larger for treated cells, but that a thresholding including the cell perimeter does not contribute significantly to cell classification.

The causes of the morphological changes observed have yet to be examined, but may be connected to cell cycle arrest, since they are accompanied by a reduction in proliferation rate. However, changes are not directly related to reduction in viability or metabolism, as detected by MTS. For a future investigation, it would be of interest to let the incubation proceed to investigate if viability changes would occur in time. Also, by interrupting the exposure to etoposide, it would be possible to examine if the effects of etoposide are reversible or not.

Chapter 5

Discussion

5.1 Conclusions

In this Thesis, the potential of non-invasive clinical imaging for analysis of embryos growing *in vitro* has been evaluated. It has also been shown that there is a potential for extension to other cell types, growth varieties and imaging hardware [7].

Methods have been developed to track the location of embryos through a time series [2] at 92.9% accuracy under noisy conditions, making it possible to perform automatic image analysis despite sample disturbance and displacement. The potential of the embryo imaging hardware selected to perform focal slicing has been demonstrated, and techniques for automatic focal filtering were developed to make use of the slicing capability [1]. Focal filtering can be used as a pre-processing step to facilitate image stack analysis or as a method for automatic focus of microscope hardware when images are captured during long periods of time or when the capturing process is automatic and the sample inaccessible inside incubator and imaging equipment. If connected back to the hardware, it can be used as an input for mechanical automatic focus. Focal filtering can also be used in image analysis for spatial filtering, selecting the regions of interest and for partial segmentation. It provides a means to quantify image sharpness and the opportunity to select different image regions at different image planes of the same stack, thus taking the first step towards non-invasive three-dimensional image analysis.

Furthermore, it has been shown how several known embryological traits can be detected in non-invasive imaging using standard computer vision techniques (or varieties thereof). Table 5.1 summarizes the results.

Features of biological interest suggested in literature (Table 2.2) are for the most part based on visual analysis, relying on the expertise of clinicians to identify structural

TABLE 5.1: Summary of evaluated target embryo traits

Property	Detection accuracy (HMC)	Reference
Embryo location	70.0% (1-6 cells)	[3]
Syngami	83.0%	[2]
Timing of mitotic splittings	80.8% (1-6 cells)	[6]
Blastocyst formation	71.8%	[5]
Blastomere number	80.8% (1-6 cells)	[6]
Embryo activity	Detectable, not quantified	
Embryo fragmentation	Detectable, not quantified	

information in images as healthy or non-healthy. A computerized evaluation introduces the possibility of extraction of non-visual metrics, with the advantage of being objective and easy to calculate, and the disadvantage of sometimes being hard to directly relate to biological characteristics. Several embryo characteristics which relate to embryo health are based solely of the timing of morphological events such as timing and duration of mitosis. During the work with this Thesis, methods were developed to explore the possibility of analysis of non-visual metrics [2, 5, 6] and it is shown that in time sequence analysis, they can be used to detect the timing of events without a more complicated full visual analysis and segmentation.

To summarize, the main outcome of this study is an evaluation of health criteria for embryo monitoring, and their potential for automatic image analysis. The evaluation has been based on HMC imaging, a commonly used method for embryo imaging, which has so far been used very little for computerized diagnostics. The methods used are tailored for embryo analysis, but can also benefit other fields of study in Life Science, a fact which has also been demonstrated in this Thesis.

5.2   Future Work

There are several reasons for caution when comparing the new TLI systems with traditional microscopy and evaluation. One obvious improvement to both manual and

TABLE 5.2: Comparison of embryo selection criteria: HMC & DF

Embryo trait	Expected detectability (HMC)	Expected detectability (DF)	Means of improvement
Embryo location	very high	unknown	
Timing of mitosis	high	very high	Temporal resolution
Interval between mitosis	high	very high	Temporal resolution
Duration of mitosis	high	high	Blastomere detection
Blastomere size, shape	low	high	3D visualization
Blastomere number	low	high	Blastomere detection
Timing of compaction	high	medium	Temporal resolution
Fragmentation	medium	medium	
Embryo respiration	-	-	
Embryo activity, movement	medium	medium	Temporal resolution
Chromosomal abnormalities	-	-	
Nuclei appearance	medium	low	Blastomere detection
Nuclei number	medium	low	Blastomere detection
Timing of pronuclear formation	high	low	Temporal resolution
Pronuclear orientation	medium	low	Spatial resolution
Pronuclear size, shape	low	low	Spatial resolution, Image detail
Timing of polar body formation	very low	very low	Spatial resolution
Polar body orientation, number	very low	very low	Spatial resolution

computer aided evaluation is the construction of the embryo container. In the circular wells used, embryos tend to migrate to the edge of the well (Figure 2.1a), making both manual and computerized analysis difficult. An optimization of well shape to facilitate imaging and avoid artifacts caused by the well edges would significantly improve the robustness of image evaluation.

The selected imaging technique HMC had the potential of revealing internal embryo and cell detail at high resolution but had limitations in the detection of embryo and blastomere outlines [3]. Studies [102] have shown the capability of other non-invasive imaging techniques (DF) to extract outline information at the expense of image detail. Table 5.2 summarizes the embryo selection criteria suggested by literature (Table 2.2), now comparing both HMC and DF as detection technique. As can be seen, the two methods largely complement one another. For nuclei detection, which is possible for HMC, a detection of blastomere outlines is required which is hard to accomplish using HMC. Using DF, blastomere outline detection is possible for up to four cells. For a detailed clinical analysis, a combination of both imaging techniques carries the potential of the most complete automatic scanning of embryo image characteristics known today.

To HMC alone, there are a number of possibilities to improve image detail and achieve higher resolution images by hardware adjustment. One is the numerical aperture of the optical set up, which would decrease the image depth of field, thus making it possible to achieve a more fine grained focal slicing and a better image stack resolution. Another is the rotation of the sample relative to image optics, which - despite being complicated to achieve mechanically - would annul the undesired effect of image polarity, making it possible to use gradient-based edge detectors and also render a tomographic image of the sample.

The spherical structure of an embryo is exploited in this Thesis by the construction of separate ROI for feature extraction (Chapter 3.1,3.5). Here, the properties for an internal region are compared to the entire embryo (see *e.g.* Figure 2.9). In Chapter 3.4, the use of a more detailed grid was demonstrated, but it is also pointed out that a square grid is not consistent with embryo geometry and that a standardized region for normalization is lacking. The zona pellucida may be the best image region to use for pixel value normalization, as it is expected to stay static until the moment of hatching and will not be as effected as the embryo exterior by sample contamination. In Chapter 3.3 it has been shown that a concentric, circular grid layout is necessary to analyze embryo rhythmic contraction and motion, such as the repeated blastocoelic expansions and the formation of the morula and in Chapter 3.6 is investigated the effect of changing the circular grid proportions. This symmetry could be exploited further using a more fine-grained grid (Figure 5.1).
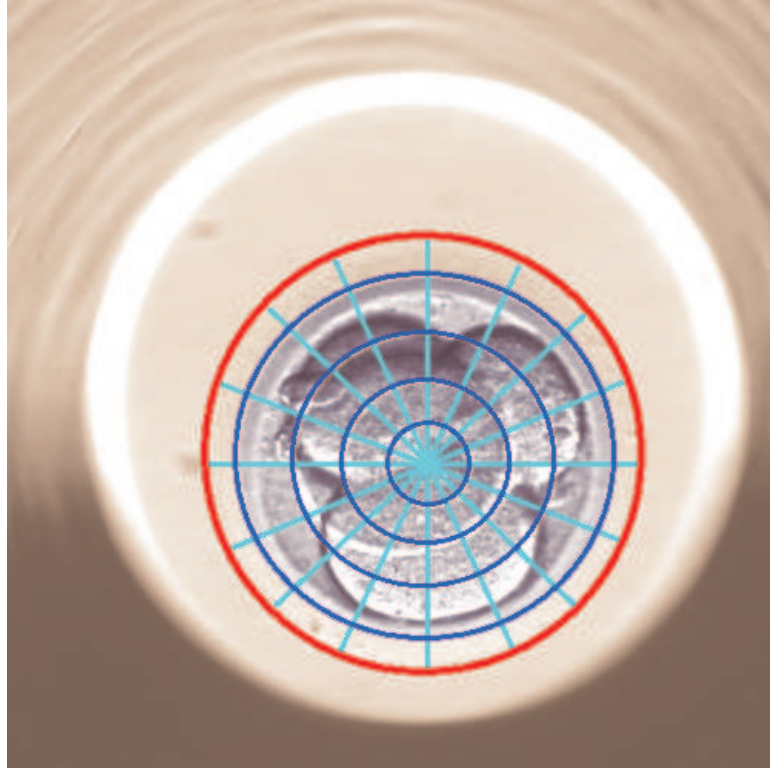
FIGURE 5.1: Suggested layout for polar embryo image grid.

A more detailed analysis of vibrational and motion events will also require a higher frequency image capture than what was available for this analysis. Table 2.2 lists expected image capture intervals for a number of embryo health related criteria. Part of the limitation was caused by the scanning procedure, limiting the temporal resolution of sequence capture. Studies using higher temporal resolution have suggested [54] that today's imaging frequency in TLI may be under-sampling, thus missing influential events in embryo morphokinetics. For high end study purposes, it would be interesting with the option to sacrifice the number of samples imaged for a higher temporal resolution. Given this, there is an interesting possibility in the study of rhythmic movement and embryo activity and motion.

If not limited to non-invasive technologies, there is the potential for combining the methods described with confocal microscopy or fluorescence, *e.g.* in the study of animal embryo models in evolutionary development. Here, a non-invasive imaging can contribute to the analysis by continuous monitoring over long periods of time and serve as a complement to a more detailed analysis of molecular targets using attached fluorophores.

There is also potential for improvement in the selection of methods and the choice of model for embryo health. In this Thesis non-visual characteristics have been targeted

by means of hand-crafted image features. Given large amounts of training data, there is a possibility to use machine learning to automatically extract features using *e.g.* deep learning. In several challenges in recent years in medical imaging, neural networks and deep learning algorithms have emerged as the most accurate. However, these methods rely on large ($\gg$1000) sets of labeled training data, and requires substantial computational power in the training stage. Medical imaging analysis is a field with high variability in the nature of the data, and transferring classifiers from one field to another may not always be possible. It may therefore take some time before suitable classifiers are available in the areas where imaging data is now being gathered in increasing extent.

5.3 Outlook: A Framework for Large Scale Analysis of Medical Imaging

[Single cell studies] "opens up a whole new perspective on cellular research, where cells are viewed not as a uniform mass, but rather as a broad spectrum of individual bodies, all contributing in its own way to the characteristics of the tissue" [7].

What is needed in order to progress computer vision to the realm of clinical medical image analysis?

Using computer vision, it is possible to extract more information from images than ever could be possible by manual means. To progress with large scale data analysis of images it is necessary to access a large amount of images and it is vital to be able to combine and compare large amounts of data in a standardized fashion. Today, large scale data analysis is not possible for small and medium sized enterprises, or for research institutions without access to their own data storage. When collaborations in data sharing take place, it often involves a cumbersome and time-consuming alignment procedure of the source material. To allow also smaller institutions to take active part in the development of drugs, medical instrumentation and diagnostic tools, data must be shared in a more efficient manner.

Supervised learning methods often benefit from having a large set of labeled training data. Crowd sourcing enables annotation of large data sets for real world images, but the application for medical imaging may require deeper understanding and expert knowledge of image content [143, 144]. The conditions under study are also sometimes very rare, resulting in less image material available. To obtain large medical data sets for training is challenging for several reasons. First, it is often difficult to obtain funding for the construction of data sets alone. Second, privacy issues make it more difficult to share

medical data than natural images. Third, the breadth of applications in medical imaging requires that many different data sets need to be collectively shared. Despite this, there is rapid progress in collection and sharing of clinical data. As images are often routinely labeled in the clinic, it is also possible that routine sharing in the future will enable access to multi-clinical data. One prerequisite is that the storing and annotation is done in a standardized format, and the earlier this standardization can be initiated, the less the risk that historical data will become inaccessible and obsolete in a future global system.

There are several obstacles causing delay in data sharing and disrupting scientific progress:

- Data is stored locally and inaccessible due to technical and ethical constraints

- A global approach to data tagging and data storage format is lacking, making data hard to compare, even if it could be accessed without restraint

- A common standard for manual data annotation is lacking, making it hard to establish a ground truth for comparison with computerized analysis, as algorithms are developed

If steps could be taken in medical imaging to remove those obstacles, and an agreement could be reached to share and store data in a standardized manner, the scientific approach to the medical and biological sciences would change forever. In order to not only supply users in the high end of image processing, but also be useful in a routine, clinical setting, work need to progress in the following:

- generalization of methods including agreement of the definition of general concepts such as "image quality" and agreement of standards for data storage, format and meta-data.

- establishment of multi-clinical cooperation around standards and data annotation for comparison

- establish cooperation around clinical trials for method evaluation.

Work in this direction has already begun. Public data sets have been released and are being routinely used in experimental validation [145, 146]. Several research groups also routinely release their raw data, and a multitude of initiatives have emerged in recent years, dedicated to the coordination of research data reporting [147–149] and more and more software for image analysis is being released as open source, such as the software used in this Thesis (Figure 5.2).

When standardizing a framework, it is not only validation of data which has to be reported in a comparable fashion, researchers themselves need to report results in a comparable fashion. An issue illustrated by three recent studies, all using the same data set and reporting results in three different forms [150–152]. The risk of different reporting metrics is a common and not unsurprising issue in a new emerging field, where research is advancing simultaneously on multiple fronts, sometimes merging from different disciplines, and sometimes diverging into new research areas. Some of this research diversity can be expected to reduce as the area matures and as methods migrate from the laboratory to the clinic. In the meantime, it is important for researchers in this field to take special care to pay attention to other research efforts and to report their own results in several ways and in accordance with previous reports to allow for comparison whenever possible.
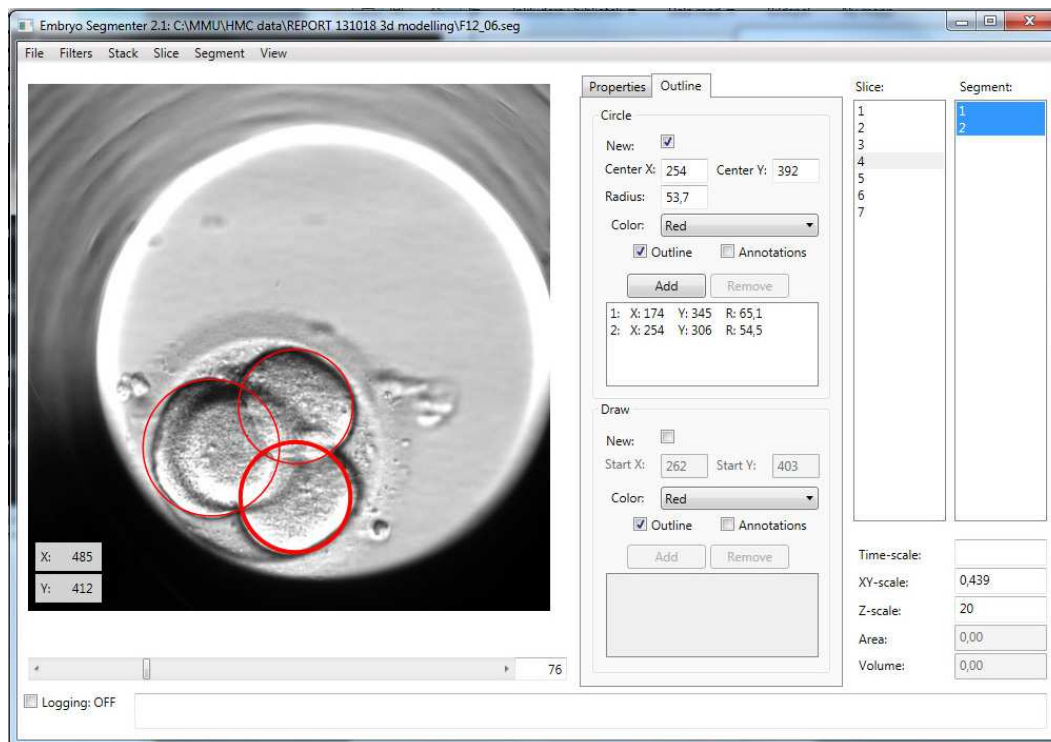
FIGURE 5.2: EmbryoSegmenter for annotation, selection of focus level and blastomere outlines [105].

Biological sciences are progressing from a qualitative field with evaluation based on experience, to a quantitative field, heavily dependent on computer guided assessment and evaluation of metrics. The task of developing computer systems capable of reliable medical image analysis will require a high degree of future cooperation between software developers and clinical scientists active in the field. The implications of making wrong decisions in a clinical environment are potentially severe and any software solution must

be implemented with great care, but a healthy skepticism for technology must not prevent us from critically evaluating it and embracing it wherever possible to improve our understanding of ourselves and the world.

Bibliography

[1] A Mölder, S Czanner and N Costen. Focal Plane Selection in Microscopic Embryo Images. *Conference on Computer Graphics & Visual Computing*, 2014.

[2] A Mölder, S Czanner, N Costen and G Hartshorne. Automatic detection of circular structures in human embryo imaging using trigonometric rotation of the Hough Transform. *22nd International Conference on Pattern Recognition*, 22:3239–3244, 2014.

[3] A L Mölder, N Costen, G M Hartshorne and S Czanner. Method for segmentation of single blastomeres in three dimensional microscopy image stacks. *In print.*

[4] A Mölder, S Czanner and N Costen. Multidimensional visualisation to improve the understanding of biological data sets. *Proceedings of Central European Seminar on Computer Graphics*, 17:121–128, 2013.

[5] A Mölder, S Drury, N Costen, G M Hartshorne and S Czanner. Semiautomated analysis of embryoscope images: Using localized variance of image intensity to detect embryo developmental stages. *Cytometry Part A*, 87(2):119–128, 2015.

[6] A L Mölder, S Drury, N Costen, G M Hartshorne and S Czanner. Efficient multi-class Support Vector Machine for the classification of embryo time-lapse imaging. *Submitted for publication*, 2016.

[7] A L Mölder, J Persson, Z El-Schich, S Czanner and A Gjörloff-Wingren. Machine learning for the study of etoposide-treated in vitro adherent cells based on non-invasive imaging growth characteristics. *Submitted for publication*, 2016.

[8] A L Mölder, S Drury, N Costen, G M Hartshorne and S Czanner. Three Dimensional Visualisation of Microscope Imaging to Improve Understanding of Human Embryo Development. *Visualization in Medicine and Life Sciences III*, Springer International Publishing, Cham, 249–264, 2016. ISBN: 978-3-319-24521-8.

[9] T W Deacon. *Incomplete Nature How Mind Emerged from Matter*. W. W. Norton & Company Inc, New York, 2013.

[10] M P Diamond, S Willman, P Chenette and M I Cedars. The clinical need for a method of identification of embryos destined to become a blastocyst in assisted reproductive technology cycles. *Journal of Assisted Reproduction and Genetics*, 29(5):391–396, 2012.

[11] M S Kupka, T DHooghe, A P Ferraretti, J de Mouzon, K Erb, J A Castilla, C Calhaz-Jorge, Ch De Geyter and V Goossens. The European IVF-Monitoring Consortium for the European Society of Human Reproduction and Embryology. Assisted reproductive technology in Europe, 2011: results generated from European registers by ESHRE. *Human Reproduction*, 31(2):233–248, 2016. PMID: 26740578.

[12] T Hardarson, A Ahlström, L Rogberg, L Botros, T Hillensjö, G Westlander, D Sakkas and M Wikland. Non-invasive metabolomic profiling of Day 2 and 5 embryo culture medium: a prospective randomized trial. *Human Reproduction*, 27(1):89–96, 2012. PMID: 22068638.

[13] S Mastenbroek, M Twisk, J van Echten-Arends, B Sikkema-Raddatz, J C Korevaar, H R Verhoeve, N E A Vogel, E G J M Arts, J W A de Vries, P M Bossuyt, C H C M Buys, M J Heineman, S Repping and F van der Veen. In Vitro Fertilization with Preimplantation Genetic Screening. *New England Journal of Medicine*, 357(1):9–17, 2007. PMID: 17611204.

[14] C G Vergouw, D C Kieslinger, E H Kostelijk, P G Hompes, R Schats and C B Lambalk. Metabolomic profiling of culture media by near infrared spectroscopy as an adjunct to morphology for selection of a single day 3 embryo to transfer in ivf: a double blind randomised trial. *Fertility and Sterility*, 96(3):S3, 2011.

[15] L D M Ottosen, J Hindkjaer and J Ingerslev. Light exposure of the ovum and preimplantation embryo during ART procedures. *Journal of Assisted Reproduction and Genetics*, 24(2–3):99–103, 2007. PMID: 17216346.

[16] G M Jones, D S Cram, B Song, G Kokkali, K Pantos and A O Trounson. Novel strategy with potential to identify developmentally competent IVF blastocysts. *Human Reproduction*, 23(8):1748–1759, 2008. PMID: 18477572.

[17] L Scott, J Berntsen, D Davies, J Gundersen, J Hill and N Ramsing. Symposium: innovative techniques in human embryo viability assessment. Human oocyte respiration-rate measurement–potential to improve oocyte and embryo selection? *Reproductive Biomedicine Online*, 17(4):461–469, 2008. PMID: 18854099.

[18] E Seli, C Robert and M-A Sirard. OMICS in assisted reproduction: possibilities and pitfalls. *Molecular Human Reproduction*, 16(8):513–530, 2010. PMID: 20538894.

[19] T Hardarson, G Caisander, A Sjögren, C Hanson, L Hamberger and K Lundin. A morphological and chromosomal study of blastocysts developing from morphologically suboptimal human pre-embryos compared with control blastocysts. *Human reproduction*, 18(2):399–407, 2003. PMID: 12571180.

[20] M Montag, J Liebenthron and M Köster. Which morphological scoring system is relevant in human embryo development? *Placenta*, 32 Suppl 3:S252–256, 2011. PMID: 21782239.

[21] J Van Blerkom, P Davis and S Alexander. A microscopic and biochemical study of fragmentation phenotypes in stage-appropriate human embryos. *Human reproduction*, 16(4):719–729, 2001. PMID: 11278225.

[22] J G Lemmen, I Agerholm and S Ziebe. Kinetic markers of human embryo quality using time-lapse recordings of IVF/ICSI-fertilized oocytes. *Reproductive biomedicine online*, 17(3):385–391, 2008. PMID: 18765009.

[23] M Alikani, J Cohen, G Tomkin, G J Garrisi, C Mack and R T Scott. Human embryo fragmentation in vitro and its implications for pregnancy and implantation. *Fertility and sterility*, 71(5):836–842, 1999. PMID: 10231042.

[24] H N Ciray, A Campbell, I E Agerholm, J Aguilar, S Chamayou, M Esbert, S Sayed and Time-Lapse User Group. Proposed guidelines on the nomenclature and annotation of dynamic human embryo monitoring by a time-lapse user group. *Human Reproduction*, 29(12):2650–2660, 2014. PMID: 25344070.

[25] URL: http://www.vitrolife.com/sv/Products/EmbryoScope-Time-Lapse-System/

[26] URL: http://www.vitrolife.com/en/Fertility/Products/Primo-Vision-Time-Lapse-System/

[27] URL: https://www.eevatest.com/

[28] P Kovacs. Embryo selection: the role of time-lapse monitoring. *Reproductive Biology and Endocrinology*, 12, 2014. PMID: 25510244.

[29] L T Polanski, M A Coelho Neto, C O Nastri, P A Navarro, R A Ferriani, N Raine-Fenning and W P Martins. Time-lapse embryo imaging for improving reproductive outcomes: systematic review and meta-analysis. *Ultrasound in Obstetrics & Gynecology*, 44(4):394–401, 2014.

[30] S Kahraman, M etinkaya, C Pirkevi, H Yelke and Y Kumtepe. Comparison of Blastocyst Development and Cycle Outcome in Patients with eSET Using Either Conventional or Time Lapse Incubators. A Prospective Study of Good Prognosis Patients. *Journal of Reproductive Biotechnology and Fertility*, 3(2):55–61, 2012.

[31] K Kirkegaard, J J Hindkjaer, M L Grøndahl, U S Kesmodel and H J Ingerslev. A randomized clinical trial comparing embryo culture in a conventional incubator with a time-lapse incubator. *Journal of Assisted Reproduction and Genetics*, 29(6):565–572, 2012. PMID: 22460082.

[32] H Park, C Bergh, U Selleskog, A Thurin-Kjellberg and K Lundin. No benefit of culturing embryos in a closed system compared with a conventional incubator in terms of number of good quality embryos: results from an RCT. *Human Reproduction*, 30(2):268–275, 2015. PMID: 25432920.

[33] C C Wong, K E Loewke, N L Bossert, B B Christopher J De Jonge, T M Baer and R A R Pera. Non-invasive imaging of human embryos before embryonic genome activation predicts development to the blastocyst stage. *Nature Biotechnology*, 28(10):1115–1121, 2010.

[34] C Pribenszky, E Losonczi, M Molnár, Z Lang, S Mátyás, K Rajczy, K Molnár, P Kovács, P Nagy, J Conceicao and G Vajta. Prediction of in-vitro developmental competence of early cleavage-stage mouse embryos with compact time-lapse equipment. *Reproductive Biomedicine Online*, 20(3):371–379, 2010. PMID: 20089456.

[35] A Arav, A Aroyo, S Yavin and Z Roth. Prediction of embryonic developmental competence by time-lapse observation and shortest-half analysis. *Reproductive Biomedicine Online*, 17(5):669–675, 2008. PMID: 18983751.

[36] P Holm, N N Shukri, G Vajta, P Booth, C Bendixen and H Callesen. Developmental kinetics of the first cell cycles of bovine in vitro produced embryos in relation to their in vitro viability and sex. *Theriogenology*, 50(8):1285–1299, 1998.

[37] B Grisart, A Massip and F Dessy. Cinematographic analysis of bovine embryo development in serum-free oviduct-conditioned medium. *Journal of Reproduction and Fertility*, 101(2):257–264, 1994.

[38] D S Gonzales, J M Jones, T Pinyopummintr, E M Carnevale, O J Ginther, S S Shapiro and B D Bavister. Trophectoderm projections: a potential means for locomotion, attachment and implantation of bovine, equine and human blastocysts. *Human Reproduction*, (11):2739–2745, 1996.

[39] K Kirkegaard, A Campbell, I Agerholm, U Bentin-Ley, A Gabrielsen, J Kirk, S Sayed and H J Ingerslev. Limitations of a time-lapse blastocyst prediction model: a large multicentre outcome analysis. *Reproductive BioMedicine Online*, 29(2):156–158, 2014. PMID: 24912418.

[40] J Conaghan, A A Chen, S P Willman, K Ivani, P E Chenette, R Boostanfar, V L Baker, G D Adamson, M E Abusief, M Gvakharia, K E Loewke and S Shen.

Improving embryo selection using a computer-automated time-lapse image analysis test plus day 3 morphology: results from a prospective multicenter trial. *Fertility and Sterility*, 100(2):412–419.e5, 2013. PMID: 23721712.

[41] D Payne, S P Flaherty, M F Barry and C D Matthews. Preliminary observations on polar body extrusion and pronuclear formation in human oocytes using time-lapse video cinematography. *Human reproduction*, 12(3):532–541, 1997. PMID: 9130755.

[42] M Cruz, N Garrido, J Herrero, I Pérez-Cano, M Muoz and M Meseguer. Timing of cell division in human cleavage-stage embryos is linked with blastocyst formation and quality. *Reproductive BioMedicine Online*, 25(4):371–381, 2012.

[43] M Meseguer, J Herrero, A Tejera, K M Hilligsœ, N B Ramsing and J Remohí. The use of morphokinetics as a predictor of embryo implantation. *Human Reproduction*, 26(10):2658–2671, 2011. PMID: 21828117.

[44] D Hlinka, B Kalatová, I Uhrinová, S Dolinská, J Rutarová, J Rezáčová, S Lazarovská and M Dudáš. Time-lapse cleavage rating predicts human embryo viability. *Physiological Research*, 61(5):513–525, 2012. PMID: 22881225.

[45] A A Chen, L Tan, V Suraj, R R Pera and S Shen. Biomarkers identified with time-lapse imaging: discovery, validation and practical application. *Fertility and Sterility*, 99(4):1035–1043, 2013. PMID: 23499001.

[46] M D VerMilyea, L Tan, J T Anthony, J Conaghan, K Ivani, M Gvakharia, R Boostanfar, V L Baker, V Suraj, A A Chen, M Mainigi, C Coutifaris and S Shen. Computer-automated time-lapse analysis results correlate with embryo implantation and clinical pregnancy: a blinded, multi-centre study. *Reproductive Biomedicine Online*, 29(6):729–736, 2014. PMID: 25444507.

[47] R Milewski, P Kuć, A Kuczyńska, B Stankiewicz, K Łukaszuk and W Kuczyński. A predictive model for blastocyst formation based on morphokinetic parameters in time-lapse monitoring of embryo development. *Journal of Assisted Reproduction and Genetics*, 32(4):571–579, 2015. PMID: 25690157.

[48] T Hardarson, C Hanson, A Sjögren and K Lundin. Human embryos with unevenly sized blastomeres have lower pregnancy and implantation rates: indications for aneuploidy and multinucleation. *Human Reproduction*, 16(2):313–318, 2001. PMID: 11157826.

[49] V Burruel, K Klooster, C M Barker, R R Pera and S Meyers. Abnormal early cleavage events predict early embryo demise: sperm oxidative stress and early abnormal cleavage. *Scientific Reports*, 4:6598, 2014. PMID: 25307782.

[50] C Garello, H Baker, J Rai, S Montgomery, P Wilson, C R Kennedy and G M Hartshorne. Pronuclear orientation, polar body placement and embryo quality after intracytoplasmic sperm injection and in-vitro fertilization: further evidence for polarity in human oocytes? *Human Reproduction*, 14(10):2588–2595, 1999.

[51] S Munné and J Cohen. Chromosome abnormalities in human embryos. *Human Reproduction Update*, 4(6):842–855, 1998. PMID: 10098475.

[52] F Guerif, A Le Gouge, B Giraudeau, J Poindron, R Bidault, O Gasnier and D Royere. Limited value of morphological assessment at days 1 and 2 to predict blastocyst development potential: a prospective study based on 4042 embryos. *Human reproduction*, 22(7):1973–1981, 2007. PMID: 17496054.

[53] W B Schoolcraft, D K Gardner, M Lane, T Schlenker, F Hamilton and D R Meldrum. Blastocyst culture and transfer: analysis of results and parameters affecting outcome in two in vitro fertilization programs. *Fertility and sterility*, 72(4):604–609, 1999. PMID: 10521095.

[54] A Ajduk, T Ilozue, S Windsor, Y Yu, K B Seres, R J Bomphrey, B D Tom, K Swann, A Thomas, C Graham and M Zernicka-Goetz. Rhythmic actomyosin-driven contractions induced by sperm entry predict mammalian embryo viability. *Nature Communications*, 2:417, 2011.

[55] K Kirkegaard, I E Agerholm and H J Ingerslev. Time-lapse monitoring as a tool for clinical embryo assessment. *Human Reproduction*, 27(5):1277–1285, 2012. PMID: 22419744.

[56] C Giorgetti, P Terriou, P Auquier, E Hans, J L Spach, J Salzmann and R Roulier. Embryo score to predict implantation after in-vitro fertilization: based on 957 single embryo transfers. *Human reproduction*, 10(9):2427–2431, 1995. PMID: 8530679.

[57] J Holte, L Berglund, K Milton, C Garello, G Gennarelli, A Revelli and T Bergh. Construction of an evidence-based integrated morphology cleavage embryo score for implantation potential of embryos scored and transferred on day 2 after oocyte retrieval. *Human Reproduction*, 22(2):548–557, 2007. PMID: 17095516.

[58] C V Steer, C L Mills, S L Tan, S Campbell and R G Edwards. The cumulative embryo score: a predictive embryo scoring technique to select the optimal number of embryos to transfer in an in-vitro fertilization and embryo transfer programme. *Human Reproduction*, 7(1):117–119, 1992. PMID: 1551945.

[59] E Van Royen, K Mangelschots, D De Neubourg, I Laureys, G Ryckaert and J Gerris. Calculating the implantation potential of day 3 embryos in women younger than

38 years of age: a new model. *Human Reproduction*, 16(2):326–332, 2001. PMID: 11157828.

[60] S Ziebe, K Petersen, S Lindenberg, A G Andersen, A Gabrielsen and A N Andersen. Embryo morphology or cleavage stage: how to select the best embryos for transfer after in-vitro fertilization. *Human Reproduction*, 12(7):1545–1549, 1997. PMID: 9262293.

[61] C Racowsky, P Kovacs and W P Martins. A critical appraisal of time-lapse imaging for embryo selection: where are we and where do we need to go? *Journal of Assisted Reproduction and Genetics*, 32(7):1025–1030, 2015.

[62] L van Loendersloot, M van Wely, F van der Veen, P Bossuyt and S Repping. Selection of embryos for transfer in IVF: ranking embryos based on their implantation potential using morphological scoring. *Reproductive Biomedicine Online*, 29(2):222–230, 2014. PMID: 24925309.

[63] C Racowsky, J E Stern, W E Gibbons, B Behr, K O Pomeroy and J D Biggers. National collection of embryo morphology data into Society for Assisted Reproductive Technology Clinic Outcomes Reporting System: associations among day 3 cell number, fragmentation and blastomere asymmetry and live birth rate. *Fertility and Sterility*, 95(6):1985–1989, 2011.

[64] E S Filho, J A Noble and D Wells. A Review on Automatic Analysis of Human Embryo Microscope Images. *The Open Biomedical Engineering Journal*, 4:170–177, 2010. PMID: 21379391.

[65] J H Price and D A Gough. Comparison of phase-contrast and fluorescence digital autofocus for scanning microscopy. *Cytometry*, 16(4):283–297, 1994.

[66] F C A Groen, I T Young and G Ligthart. A comparison of different focus functions for use in autofocus algorithms. *Cytometry*, 6(2):81–91, 1985.

[67] M W Davidson and M Abramowitz. Optical Microscopy. *Encyclopedia of Imaging Science and Technology*, John Wiley & Sons, Inc, 2002.

[68] M E Ambühl, C Brepsant, J-J Meister, A B Verkhovsky and I F Sbalzarini. High-resolution cell outline segmentation and tracking from phase-contrast microscopy images. *Journal of microscopy*, 245(2):161–170, 2012. PMID: 21999192.

[69] T A Nenasheva, T Carter and G I Mashanov. Automatic tracking of individual migrating cells using low-magnification dark-field microscopy. *Journal of microscopy*, 246(1):83–88, 2012. PMID: 22260196.

[70] J Pan, T Kanade and M Chen. Learning to Detect Different Types of Cells under Phase Contrast Microscopy. *Microscopic Image Analysis with Applications in Biology*, The Robotics Institute, Carnegie Mellon University, 2009.

[71] M R Lamprecht, D M Sabatini and A E Carpenter. CellProfiler: free, versatile software for automated biological image analysis. *BioTechniques*, 42(1):71–75, 2007. PMID: 17269487.

[72] M V Boland, M K Markey and R F Murphy. Automated recognition of patterns characteristic of subcellular structures in fluorescence microscopy images. *Cytometry*, 33(3):366–375, 1998. PMID: 9822349.

[73] A E Carpenter, T R Jones, M R Lamprecht, C Clarke, I H Kang, O Friman, D A Guertin, J H Chang, R A Lindquist, J Moffat, P Golland and D M Sabatini. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biology*, 7(10):R100, 2006. PMID: 17076895.

[74] X Chen, X Zhou and S T C Wong. Automated segmentation, classification and tracking of cancer cell nuclei in time-lapse microscopy. *IEEE transactions on biomedical engineering*, 53(4):762–766, 2006. PMID: 16602586.

[75] X Yang, H Li and X Zhou. Nuclei Segmentation Using Marker-Controlled Watershed, Tracking Using Mean-Shift and Kalman Filter in Time-Lapse Microscopy. *IEEE Transactions on Circuits and Systems I*, 53(11):2405–2414, 2006.

[76] F Moussavi, Y Wang, P Lorenzen, J Oakley, D Russakoff and S Gould. A Unified Graphical Models Framework for Automated Mitosis Detection in Human Embryos. *IEEE Transactions on Medical Imaging*, 33(7):1551–1562, 2014.

[77] A. Mölder, M Sebesta, M Gustafsson, L Gisselson, A Gjörloff Wingren and K Alm. Non-invasive, label-free cell counting and quantitative analysis of adherent cells using digital holography. *Journal of Microscopy*, 232(2):240–247, 2008.

[78] O Debeir, P Van Ham, R Kiss and C Decaestecker. Tracking of migrating cells under phase-contrast video microscopy with combined mean-shift processes. *IEEE transactions on medical imaging*, 24(6):697–711, 2005. PMID: 15957594.

[79] A J Hand, T Sun, D C Barber, D R Hose and S MacNeil. Automated tracking of migrating cells in phase-contrast video microscopy sequences using image registration. *Journal of Microscopy*, 234(1):62–79, 2009. PMID: 19335457.

[80] K Li, E Miller, L Weiss, P Campbell and T Kanade. Online Tracking of Migrating and Proliferating Cells Imaged with Phase-Contrast Microscopy. *Conference on Computer Vision and Pattern Recognition Workshop*, 65–72, 2006.

[81] I Seroussi, D Veikherman, N Ofer, S Yehudai-Resheff and K Keren. Segmentation and tracking of live cells in phase-contrast images using directional gradient vector flow for snakes. *Journal of Microscopy*, 247(2):137–146, 2012. PMID: 22591174.

[82] J Selinummi, P Ruusuvuori, I Podolsky, A Ozinsky, E Gold, O Yli-Harja, A Aderem and I Shmulevich. Bright field microscopy as an alternative to whole cell fluorescence in automated analysis of macrophage images. *PloS One*, 4(10):e7497, 2009. PMID: 19847301.

[83] A Giusti, G Corani, L M Gambardella, M C Magli and L Gianaroli. Blastomere segmentation and 3D morphology measurements of early embryos from Hoffman Modulation Contrast image stacks. In *2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 1261–1264, 2010.

[84] I E Agerholm, C Hnida, D G Cruger, C Berg, G Bruun-Petersen, S Kølvraa and S Ziebe. Nuclei size in relation to nuclear status and aneuploidy rate for 13 chromosomes in donated four cells embryos. *Journal of Assisted Reproduction and Genetics*, 25(2-3):95–102, March 2008. PMID: 18256921.

[85] E S Filho, J A Noble, M Poli, T Griffiths, G Emerson and D Wells. A method for semi-automatic grading of human blastocyst microscope images. *Human reproduction*, 27(9):2641–2648, 2012. PMID: 22736327.

[86] D A Morales, E Bengoetxea and P Larranaga. Automatic segmentation of zona pellucida in human embryo images applying an active contour model. *Proceedings of the 12th Annual Conference on Medical Image Understanding and Analysis*, 209–213, 2008.

[87] C Hnida, E Engenheiro and S Ziebe. Computer-controlled, multilevel, morphometric analysis of blastomere size as biomarker of fragmentation and multinuclearity in human embryos. *Human reproduction*, 19(2):288–293, 2004. PMID: 14747169.

[88] A Beuchat, P Thévenaz, M Unser, T Ebner, A Senn, F Urner, M Germond and C O S Sorzano. Quantitative morphometrical characterization of human pronuclear zygotes. *Human reproduction*, 23(9):1983–1992, 2008. PMID: 18540007.

[89] U D Pedersen, O F Olsen and N H Olsen. A multiphase variational level set approach for modelling human embryos. *Proceedings of the 2nd IEEE Workshop on variational, Geometric and Level Set Methods*, 2003.

[90] R Tibshirani and T Hastie. Margin Trees for High-dimensional Classification. *Journal of Machine Learning Research*, 8:637–652, 2007.

[91] V Vural and G D Jennifer. A Hierarchical Method for Multi-class Support Vector Machines. In *Proceedings of the Twenty-first International Conference on Machine Learning*, 105–112, 2004.

[92] URL: https://github.com/annaleida/embryo_location

[93] URL: https://github.com/annaleida/embryo_classification

[94] URL: https://github.com/annaleida/embryo_filter_focus

[95] URL: https://github.com/annaleida/embryo_blastomere_location

[96] R M Haralick, K Shanmugam and I Dinstein. Textural Features for Image Classification. *IEEE Transactions on Systems, Man and Cybernetics*, 6:610–621, 1973.

[97] C Bishop. *Pattern Recognition and Machine Learning.* Springer–Verlag, New York, 2006.

[98] S J D Prince. *Computer Vision: Models, Learning and Inference.* Cambridge University Press, New York, 1 ed, 2012.

[99] J C Platt, J Shawe-Taylor and N Cristianini. Large Margin DAGs for Multiclass Classification. *Proceedings of Advances in Neural Information*, 12:547–553, 1999.

[100] A Azzarello, T Hoest and A L Mikkelsen. The impact of pronuclei morphology and dynamicity on live birth outcome after time-lapse culture. *Human Reproduction*, 27(9):2649–2657, 2012. PMID: 22740496.

[101] L A Scott and S Smith. The successful use of pronuclear embryo transfers the day following oocyte retrieval. *Human reproduction*, 13(4):1003–1013, 1998. PMID: 9619562.

[102] A Khan, S Gould and M Salzmann. A Linear Chain Markov Model for Detection and Localization of Cells in Early Stage Embryo Development. *Proceedings of IEEE Winter Conference on Applications of Computer Vision*, 526–533, 2015.

[103] T S Yoo. *Insight into Images*, CRC Press, 2004. ISBN: 9781568812175.

[104] K Rafferty, S Drury, G Hartshorne and S Czanner. Use of Concave Corners in the Segmentation of Embryological Datasets. *Review of Bioinformatics and Biometrics*, 1(1):1–8, 2012.

[105] URL: https://github.com/annaleida/embryosegmenter

[106] URL: https://github.com/annaleida/embryo_3Dmodel

[107] G Kroemer, W S El-Deiry, P Golstein, M E Peter, D Vaux, P Vandenabeele, B Zhivotovsky, M V Blagosklonny, W Malorni, R A Knight, M Piacentini, S Nagata and G Melino. Classification of cell death: recommendations of the Nomenclature Committee on Cell Death. *Cell Death & Differentiation*, 12(S2):1463–1467, 2005.

[108] D V Krysko, T V Berghe, K DHerde and P Vandenabeele. Apoptosis and necrosis: detection, discrimination and phagocytosis. *Methods*, 44(3):205–221, 2008. PMID: 18314051.

[109] K R Hande. Etoposide: four decades of development of a topoisomerase II inhibitor. *European Journal of Cancer*, 34(10):1514–1521, 1998. PMID: 9893622.

[110] E L Baldwin and N Osheroff. Etoposide, topoisomerase II and cancer. *Current Medicinal Chemistry. Anti-Cancer Agents*, 5(4):363–372, 2005. PMID: 16101488.

[111] K C Stine, R L Saylors, J R Sawyer and D L Becton. Secondary acute myelogenous leukemia following safe exposure to etoposide. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, 15(4):1583–1586, 1997. PMID: 9193356.

[112] H Sun, B Song, H Dong, B Reid, M A Player, J Watson and M Zhao. Visualization of fast-moving cells in vivo using digital holographic video microscopy. *Journal of Biomedical Optics*, 13(1):014007, 2008. PMID: 18315365.

[113] K Jeong, J J Turek and D D Nolte. Fourier-domain digital holographic optical coherence imaging of living tissue. *Applied Optics*, 46(22):4999, 2007.

[114] C J Mann, L Yu and M K Kim. Movies of cellular and sub-cellular motion by digital holographic microscopy. *BioMedical Engineering OnLine*, 5(1):21, 2006. PMID: 16556319.

[115] F Dubois, C Yourassowsky, O Monnom, J-C Legros, O Debeir, P Van Ham, R Kiss and C Decaestecker. Digital holographic microscopy for the three-dimensional dynamic analysis of in vitro cancer cell migration. *Journal of Biomedical Optics*, 11(5):054032, 2006. PMID: 17092181.

[116] J Persson, A Mölder, S-G Pettersson and K Alm. Cell motility studies using digital holographic microscopy. In *Microscopy, Science, Technology, Applications and Education*, Formatex Research Center, 1063–1072, 2010.

[117] P Langehanenberg, L Ivanova, I Bernhardt, S Ketelhut, A Vollmer, D Dirksen, G Georgiev, G von Bally and B Kemper. Automated three-dimensional tracking of living cells by digital holographic microscopy. *Journal of Biomedical Optics*, 2009(1):014018. PMID: 19256706.

[118] I Moon and B Javidi. 3-D Visualization and Identification of Biological Microorganisms Using Partially Temporal Incoherent Light In-Line Computational Holographic Imaging. *IEEE Transactions on Medical Imaging*, 27(12):1782–1790, 2008.

[119] M Kemmler, M Fratz, D Giel, N Saum, A Brandenburg and C Hoffmann. Noninvasive time-dependent cytometry monitoring by digital holography. *Journal of Biomedical Optics*, 12(6):064002, 2007. PMID: 18163818.

[120] Y Emery, E Cuche, T Colomb, C Depeursinge, B Rappaz, P Marquet and P Magistretti. DHM (Digital Holography Microscope) for imaging cells. *Journal of Physics: Conference Series*, 61:1317–1321, 2007.

[121] B Kemper and G von Bally. Digital holographic microscopy for live cell applications and technical inspection. *Applied Optics*, 47(4):A52, 2008.

[122] Z El-Schich, A Mölder, H Tassidis, P Härkönen, M Falck Miniotis and A Gjörloff Wingren. Induction of morphological changes in death-induced cancer cells monitored by holographic microscopy. *Journal of Structural Biology*, 189(3):207–212, 2015.

[123] N Otsu. A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man and Cybernetics*, 9(1):62–66, 1979.

[124] J Kittler and J Illingworth. Minimum error thresholding. *Pattern Recognition*, 19(1):41–47, 1986.

[125] URL: https://github.com/annaleida/comparing_classifiers

[126] L Breiman, J Friedman, R Olshen and C Stone. *Classification and Regression Trees*. CRC Press, Boca Raton, 1984.

[127] By Zephyris - Own work, CC BY-SA 3.0. URL: https://commons.wikimedia.org/w/index.php?curid=10762658

[128] By Zephyris - Own work, CC BY-SA 3.0. URL: https://commons.wikimedia.org/w/index.php?curid=10762664

[129] By Zephyris - Own work, CC BY-SA 3.0. URL: https://commons.wikimedia.org/w/index.php?curid=10762662

[130] F Zernike. Phase contrast, a new method for the microscopic observation of transparent objects. *Physica*, 9(7):686–698, 1942.

[131] R Hoffman. The modulation contrast microscope: principles and performance. *Journal of Microscopy*, 110(3):205–222, 1977.

[132] D Gabor. A New Microscopic Principle. *Nature*, 161(4098):777–778, 1948.

[133] U Schnars and W P O Jüptner. Digital recording and reconstruction of holograms in hologram interferometry and shearography. *Applied Optics*, 33(20):4373–4377, 1994.

[134] E Cuche, P Marquet and C Depeursinge. Simultaneous amplitude-contrast and quantitative phase-contrast microscopy by numerical reconstruction of Fresnel off-axis holograms. *Applied Optics*, 38(34):6994–7001, 1999.

[135] J W Goodman and R W Lawrence. Digital Image Formation from Electronically Detected Holograms. *Applied Physics Letters*, 11(3):77–79, 1967.

[136] M A Kronrod, N S Merzlyakov and L P Yaroslavskii. Reconstruction of a hologram with a computer. *Soviet Physics-Technical Physics*, 17(2):329–332, 1972.

[137] P Yu, M Mustata, J J Turek, P M W French, M R Melloch and D D Nolte. Holographic optical coherence imaging of tumor spheroids. *Applied Physics Letters*, 83(3):575, 2003.

[138] J M S Prewitt. Object Enhancement and Extraction. *Picture processing and Psychopictorics*, New York Academic Press, 75–149, 1970.

[139] M Reuter, S Biasotti, D Giorgi, G Patanè and M Spagnuolo. Discrete Laplace–Beltrami operators for shape analysis and segmentation. *Computers & Graphics*, 33(3):381–390, 2009.

[140] R C Gonzales and R E Woods. *Digital Image Processing*. Prentice Hall, Upper Saddle River, New Jersey, 2nd ed, 2002.

[141] S E Hernandez and K E Barner. Joint region merging criteria for watershed-based image segmentation. In *2000 International Conference on Image Processing*, 2:108–111, 2000.

[142] L Shapiro and G Stockman. *Computer Vision*, Prentice-Hall, 2001. ISBN: 9780130307965.

[143] T B Nguyen, S Wang, V Anugu, N Rose, M McKenna, N Petrick, J E Burns and R M Summers. Distributed human intelligence for colonic polyp classification in computer-aided detection for CT colonography. *Radiology*, 262(3):824–833, 2012. PMID: 22274839

[144] M T McKenna, S Wang, T B Nguyen, J E Burns, N Petrick and R M Summers. Strategies for improved interpretation of computer-aided detections for CT colonography utilizing distributed human intelligence. *Medical Image Analysis*, 16(6):1280–1292, 2012. PMID: 22705287.

[145] URL: http://www.visceral.eu/

[146] URL: http://www.cancerimagingarchive.net/

[147] URL: https://rd-alliance.org/

[148] URL: https://www.elixir-europe.org/

[149] URL:
       http://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud

[150] M Anthimopoulos, S Christodoulidis, L Ebner, A Christe and S Mougiakakou.
       Lung Pattern Classification for Interstitial Lung Diseases Using a Deep Convolu-
       tional Neural Network. *IEEE Transactions on Medical Imaging*, 35(5):1207–1216,
       2016.

[151] H-C Shin, H R Roth, M Gao, L Lu, Z Xu, I Nogues, J Yao, D Mollura and
       R M Summers. Deep Convolutional Neural Networks for Computer-Aided Detec-
       tion: CNN Architectures, Dataset Characteristics and Transfer Learning. Cornell
       University Library, 2016. arXiv: 1602.03409.

[152] G van Tulder and M de Bruijne. Combining Generative and Discriminative Repre-
       sentation Learning for Lung CT Analysis with Convolutional Restricted Boltzmann
       Machines. *IEEE transactions on medical imaging*, 35(5):1262–1272, 2016. PMID:
       26886968.

Focal plane selection in microscopic embryo images

A.L. Mölder[1], S. Czanner[1], N. Costen[1] and G.Hartshorne[2,3]

[1]School of Computing, Mathematics and Digital technology, Manchester Metropolitan university, UK
[2]The Division of Reproductive Health, Warwick Medical School, University of Warwick, UK
[3]The Centre for Reproductive Medicine, University Hospitals Coventry and Warwickshire NHS Trust, Coventry, UK

Abstract

The most critical quality property for an image is its focus. Without a good sharpness, the image will be meaningless for both human eyes and for computerised analysis. As the use of camera recordings increase around us, both in our daily life, for the purpose of documentation, communication, surveillance or for recreational purposes, so it also increases in the laboratory and in the science communities. More often, images are captured as part of the scientific process. Often as a method of documentation, but more and more also as a part of the analytical process itself. The number and complexity of available algorithms for image segmentation, computer vision or pattern recognition continues to grow and is likely to play a large role in how we handle data in the future. When we hand over more and more of medical surveillance and diagnostic tasks to automation, it is crucial that we can rely on the accuracy of these automatic procedures. One way to ensure a level of quality, and to make sure we do not waste time trying to analyse material of poor standard, is to make sure that the images we introduce to an analytical pipeline hold a high enough quality. Image quality may depend on a number of properties such as focus, level of noise and a number of spherical and chromatic aberrations. Here we propose a simple, adjustable software framework for detection of image focal planes, making it possible to extract image sections for further analysis.

Categories and Subject Descriptors (according to ACM CCS): I.4.8 [Image Processing and Computer Vision]: scene analysis—Depth cues

1. Introduction

Many algorithms for automatic focus rely on computation of the power spectrum[1]. In cameras with moving lenses, it is also possible to adjust the focus based on the image contrast, a method often referred to as contrast detection autofocus. Several images are then captured in sequence, while searching for a local maxima of the image contrast or the gradient of the image contrast. This is usually not done for the entire image but, for a selected area of interest. Several autofocus algorithms have also been evaluated for microscopy[2], [3]. In optical microscopy[4], where the depth of field is usually very short, and when studying human embryos, the entire embryo is rarely in focus at one optical setting. Instead, the user may be required to adjust the focus back and forth in order to get a complete view of the sample. Here, we implement a variety of the contrast detection, adapt it to Hoffman Modulation Contrast microscopy[5], and use it as a filter to

select regions of interest within an already captured image of a human embryo.

2. Materials and methods

2.1. Sample preparation

Human tissue sample material was anonymously donated by patients and the project has been approved by Coventry Research Ethics Committee (04/Q2802/26) and the Human Fertilisation and Embryology Authority (R0155). Fresh embryos unsuitable for transfer or cryopreservation and thawed embryos no longer required for treatment were cultured using Medicult media (Origio, Redhill, UK) for up to 7 days and incubated in 37°C 5%CO_2 in air. Embryos were otherwise untreated or undisturbed during the imaging process.

2.2. Image capture

During image capture, embryos were cultured in an Embryoscope (Fertilitech, Copenhagen, Denmark), and images were captured at up to 7 focal depths and 5 recordings per hour using a Hoffman Modulation Contrast optical set-up, with a 635nm LED as light source. Image capture ran for up to 95 hours.

2.3. Algorithms

All image processing was done using standard libraries from Matlab. Prior to handling, captured images, I(x,y), were filtered with a Gaussian filter to remove speckle noise. For most images, the gaussian filter had a size of 5x5 pixels, and a standard deviation of 1. To detect the image contrast variation, a Laplacian filter,

$$L(I) = \frac{\nabla^2 I}{4} = \frac{1}{4} \cdot \left(\frac{d^2 I}{dx^2} + \frac{d^2 I}{dy^2} \right) \qquad (1)$$

was applied. Note that due to the asymmetrical nature of Hoffman Modulation Contrast imaging[6], the symmetrical Laplacian was chosen, rather than the direction-dependant gradient. To detect areas of high sharpness, an H-maxima transform, H, was applied to the result L(I), using an 8-connected neighbourhood, and suppressing maxima lower than 15% of the image maximum. The recovered maxima were then extended using a close transform with a 7 pixel diameter circular structure element, and the resulting image was thresholded and converted to a binary mask, M. Holes were removed from the mask using a filling function, and the mask was then used to extract the corresponding region from the original image, I.

3. Results

Figure 1 shows the extraction of regions in focus, based on the contrast detection algorithm, from an image of a human embryo. The embryo is in its blastocyst stage and has the shape of a hollow sphere, with cells covering its walls. Due to the varying distribution of cells, the spherical shell of the embryo may vary in tissue content. The stack moves from slightly above the horizontal embryo central plane and downwards until it reaches the embryo base, where structures on bottom of the embryo container are also encountered, and can be seen outside the embryo circumference.

4. Discussion

A reliable and robust algorithm for focus level detection may be useful not only as a pre-processing step for image analysis, but also as method for automatic focus of microscope hardware[7], when the images are being captured during
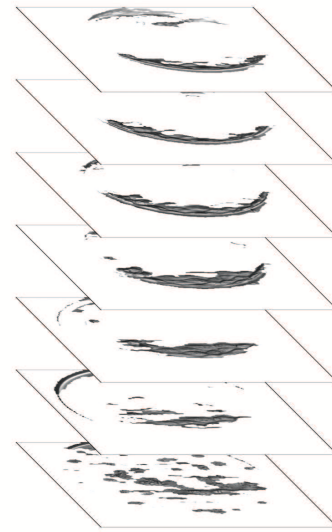

Figure 1: *A human embryo, captured at seven different focal planes: extraction of regions of interest based on image contrast.*

long periods of time, when the capturing process is automatic, or when the microscope is placed in a climate chamber, as is often the case in time-lapse sequencing. Here we present the construction of a simple algorithm for focal plane detection in Hoffman Modulation Contrast microscopic images. We intend to use this algorithm to facilitate future analysis of these images, by selecting images, or regions of images, as appropriate for segmentation, pattern recognition or other computerised analysis. With this type of quality detection as a basis, we hope to be able to quantify the image sharpness, thus being able to exclude regions of an image, or whole images, from further treatment, if their quality is deemed as being too poor. For the future, we propose to generalise the quality detection to cover several types of images. Only when more than one image set can be evaluated against the same standardised criteria can we have a true comparison of the image quality.

5. Acknowledgment

6. References

[1] P. Welch, "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," IEEE Transactions on

Audio and Electroacoustics, vol. 15, no. 2, pp. 70 – 73, Jun. 1967.

[2] J. H. Price and D. A. Gough, "Comparison of phase-contrast and fluorescence digital autofocus for scanning microscopy," Cytometry, vol. 16, no. 4, pp. 283–297, 1994.

[3] F. C. A. Groen, I. T. Young, and G. Ligthart, "A comparison of different focus functions for use in autofocus algorithms," Cytometry, vol. 6, no. 2, pp. 81–91, 1985.

[4] M. W. Davidson and M. Abramowitz, "Optical Microscopy," in Encyclopedia of Imaging Science and Technology, John Wiley & Sons, Inc., 2002.

[5] R. Hoffman, "The modulation contrast microscope: principles and performance," Journal of Microscopy, vol. 110, no. 3, pp. 205–222, 1977.

[6] A. Mölder, S. Czanner, and N. Costen, "Multidimensional visualisation to improve the understanding of biological data sets," presented at the Central European Seminar on Computer Graphics, Smolenice, 2013, vol. 17.

[7] A. Mölder, M. Sebesta, M. Gustafsson, L. Gisselson, A. G. Wingren, and K. Alm, "Non-invasive, label-free cell counting and quantitative analysis of adherent cells using digital holography," Journal of Microscopy, vol. 232, no. 2, pp. 240–247, 2008.

II

Automatic Detection of Embryo Location in Medical Imaging Using Trigonometric Rotation for Noise Reduction

Anna Mölder, Silvester Czanner, Nicholas Costen
School of Computing, Mathematics and Digital Technology
Manchester Metropolitan University
Manchester, UK
Email: Anna.Molder@mmu.ac.uk

Geraldine Hartshorne
Division of Reproductive Medicine
Warwick Medical School, University of Warwick
and Centre for Reproductive Medicine
University Hospitals Coventry and Warwickshire NHS Trust
Coventry, UK

Abstract—The assessment of embryos *in vitro* is an important tool both in In Vitro Fertilization (IVF) treatment and for research purposes. Traditionally, such assessment has been done manually and demands extensive training and expertise. Inter-observer variability limits the use of evaluation, and the manual labor increases the cost of treatment and research. To this end, feature extraction and automatic annotation using computer aided tools would both improve objectivity and save time. Due to the sensitivity of embryos, any evaluation must be performed using non-invasive imaging techniques, which sets a limit for the ability to artificially enhance image contrast, and computer aided assessment of microscopic images under clinical conditions also poses a number of difficulties due to the high level of noise and the unpredictable imaging environment. To overcome this, we propose a framework using a Hough Transform (HT) for the extraction of circular shapes in microscopic embryo images, and also use the algorithm for successful automatic profiling of the pronuclear breakdown (PNB) at syngami and detection of its timing. The best results are achieved when using a HT for embryo outline detection in combination with image variance of the embryo interior. Using our method, we achieve an overall accuracy of 95.0% detecting the position of the embryo outline, and 83.0% detecting the timing of the PNB.

I. INTRODUCTION

Despite 30 years of practice, the success rate for implantation of embryos into the uterus in IVF is still only around 30% [4]. Embryos can be selected for transfer based on a number of morphological criteria, but the sensitivity of embryos makes them difficult to monitor, as only non-destructive methods can be used. Traditionally, embryo quality assessment has been performed by manual inspection using light microscopy at intermittent time points during embryo development. Identification of key points in embryo development, along with annotation of images, is one tool commonly used by embryologists to make it possible to share and archive the information leading up to diagnosis.

Recent developments in the construction of incubators and cultivation chambers has recently made it possible to monitor embryos growing *in vitro* prior to implantation over the course of several days, without any recorded consequences to their health. This increases the amount of image material available and also increases the work load for medical personnel. In order to speed up the process and make it more objective,

tools for decision support and computer aided diagnosis are required. Ways to automate part of the image analysis process and complement the manual work would save time in IVF research, and could also introduce new ways to assess embryo quality.

Image analysis is a commonly used tool in many medical imaging applications, and often a key initial step is a successful image segmentation, assisted by artificial coloring of the sample. Despite sophisticated optics to improve image contrast the necessity to use non-invasive imaging introduces limits to the contrast enhancement possibilities in embryo imaging. A system often used is Hoffman Modulation Contrast imaging (HMC) [9], where changes in sample refractive index and thickness are translated to image amplitude, causing changes in optical path length to appear as lighter or darker bands on a grey background. The nature of the resulting image, along with the artefacts created by the image optics, makes it difficult to use global segmentation methods or methods identifying objects of arbitrary shape, such as watersheds, thresholding or level sets. Several attempts have been made to detect embryo features using segmentation [7][1][5][11] or by manual selection of a region of interest (ROI) [8][3][12]. For clinical conditions, a cumbersome initialization of calculation does not necessarily improve working conditions, and it is less likely that all images will have the quality needed for a detailed segmentation in the presence of clutter, poor lighting and imperfect focus, resulting in image objects having broken or smudged edges. One approach to solve this would be the introduction of a pre-processing step, in which images are filtered or excluded based on their level of noise. Such pre-processing can greatly improve results, but may also introduce a time-consuming step of parameter optimization, which is not desirable in a clinical context. Here we wish to attempt another approach, where the filtering is done using a simple trigonometric operation, which will allow the use of raw images 'as is'.

A. Proposed approach

Fig. 1 shows three structures of interest in embryology: The outline of the well, the embryo (here defined by the inner zona pellucida) and (pro)nuclei (with possible nucleoli), all of which appear as highly circular shapes under the microscope. The

shape and size of nuclei may provide clues to embryo health [6][13]. The appearance of two pronuclei (PN) is an early sign of successful fertilization, and for image series of embryos where the time of fertilization is unknown, the appearance of PN can be used as an early calibration point in time. Also, the timing of the pronuclear breakdown (PNB) at the onset of the first mitotic splitting (Fig. 2) has been reported to relate to live birth outcome [2].

In an attempt to automate part of the embryo evaluation process, this study investigated the possibility to detect PNB automatically using pattern recognition in series of embryo images *in vitro* with four different computational approaches, utilizing a circular HT on HMC images.

A Hough Transform (HT) allows parameterization of a known shape, transforming each structure in the image information to a corresponding bin in an accumulator, the Hough space [10]. When mapped to Hough space, each potential shape will occupy a point in the accumulator, and the strength with which it appears in the image is represented by the number of points in the corresponding bin. The parameters for the shapes most likely to exist in the image can be extracted by selecting peaks in Hough space in order of decreasing height. The main advantage of the HT is its tolerance to gaps in the boundary description and its resistance to image noise. A HT is particularly useful in attempts to locate shapes of known structures under noisy conditions and on varying backgrounds, where the position of the shape is unknown. The main disadvantage is that the complexity of the accumulator increases as $o(A^{p-2})$ [14], where A is the size of the image space and p is the number of parameters. Thus, by increasing the number of parameters, the noise in Hough space will increase accordingly, and a peak representing a shape may not appear much higher than its random neighbors. For this reason, a circular HT was used with three parameters and a three dimensional accumulator; a centroid location in two dimensions, a and b, and a radius r:

$$(x - a)^2 + (y - b)^2 = r^2 \tag{1}$$

In this case, each circular shape in an image $I(x, y)$ was mapped to a peak in Hough space, where the height of the peak represented the strongest circular structure of a particular radius. In our images, a HT applied to an embryo image resulted in a high number of randomly appearing false positive hits, but the high sensitivity also led to inconsistent results under rotational transformation. For the detection of nuclei, this sensitivity was used to filter out incorrectly detected shapes, by repeating the transform under rotation and comparing the results.

II. MATERIALS AND METHODS

A. Embryo culture and image capture

Time lapse images of human embryos fertilized *in vitro* were acquired as anonymized image sequences provided by Unisense Fertilitech. Embryos were cultured in 25 μl culture media (Origio, Redhill, UK) under mineral oil for up to 6 days, incubated in 37°C in an atmosphere of 5%CO_2 in air. Embryos were mounted in wells in an EmbryoSlide (Unisense Fertilitech, Copenhagen, Denmark), one embryo per well.
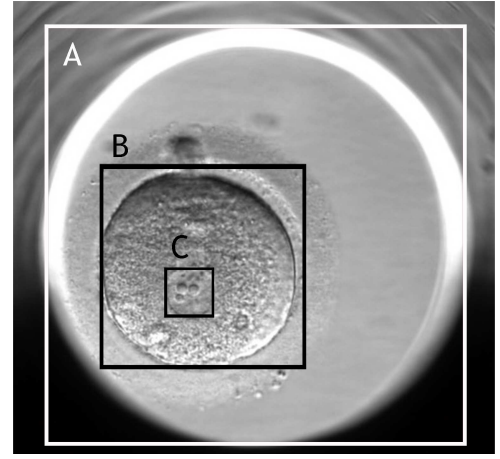

Fig. 1. Annotated raw image, indicating structures of interest. A − well outline. B − zygote boundary. C − pronucleus.
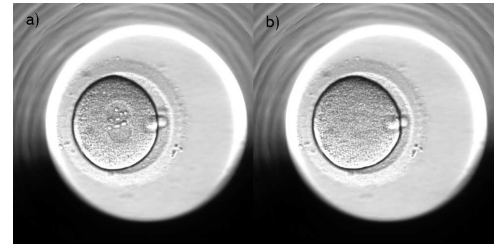

Fig. 2. Example of pronuclear breakdown (PNB) in preparation for the first mitotic splitting of the zygote. Before (a) and after (b).

The images were captured using the Embryoscope system (Unisense Fertilitech, Copenhagen, Denmark), with up to 7 focal depth planes using HMC optics and a 635 nm LED as light source. The complete test set consisted of image series of 20 embryos. Each image series consisted of 290–454 image stacks, each stack of images being recorded at 20 min intervals. Embryos were usually stationary within the well during image capture, but since at regular intervals, fresh medium was supplied to the well, the embryo was occasionally visibly displaced between captures.

B. Training sets and choice of standard

For nuclei, 10 images with 1-2 visible pronuclei were chosen as a training set to trim the detection parameters. For the embryo outlines, a training set of 140 images was selected, and up to 28 images were chosen randomly from each embryo and placed in one of seven categories by visual inspection, showing zygote, 2 cell stage, 3-4 cell stage, 5-6 cell stage, >7 cell stage and blastocyst, respectively. Images considered potentially difficult were placed in a separate category. Each category consisted of 20 images, with no more than 4 images from a single embryo per category. To provide ground truth for nuclei, a true positive (TP) or false positive (FP) classification was determined by eye, with the criterion that false positives were displaced more than one nuclear radius (Fig. 3b-c). For detection of the zygote outline, a manual annotation was performed on each image preceding computation, and the displacement of the centroid of the annotated bounding box (Fig. 1B) relative to the centroid of the detected circle was
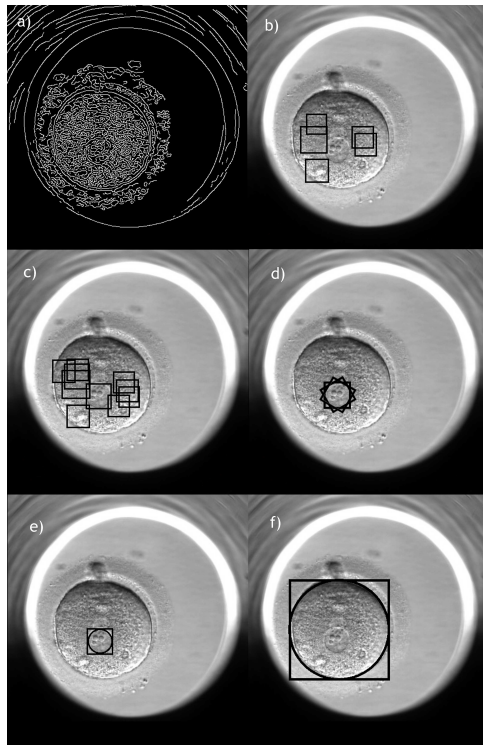
Fig. 3. a) Canny edge detection of the original image(Fig. 1).) 5 most significant circular structures selected. d) 10 most significant circular structures selected. e) Overlap of circular structures selected from the same image rotated $6 \cdot 60°$. f) Outline of pronucleus indicated, overlap of three calculations at separate angles. g) Outline of pronucleus selected. h) Outline of zygote selected.
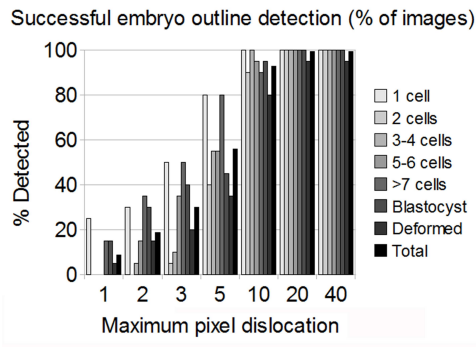

Fig. 4. Correctly detected embryo outlines for 140 images of embryos from zygote to blastocyst stage, as a function of maximum pixel dislocation. Showing calculations for each of 7 categories separately (grey) and total (black).
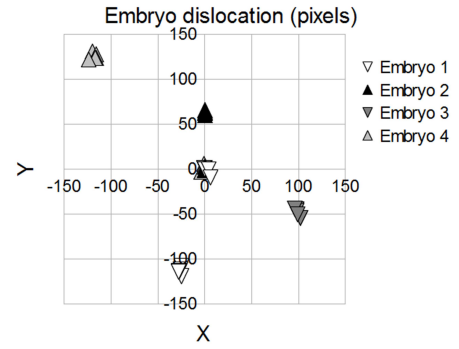

Fig. 5. Embryo location in image (pixels) before and after change of medium in the containing well. The embryo position in the first image has been set to (0,0), with all other positions calculated relative to the first. Position defined as the position of the centroid detected by Hough transform, with radius 100-120 pxl.
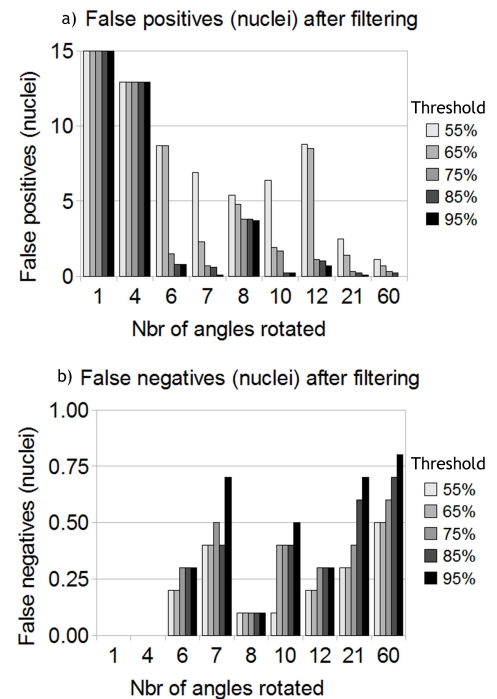

Fig. 6. Number of failed detected nuclei as a) false positive (FP) and b) false negative (FN), respectively, calculated average over 10 images as function of number of angles each image has been rotated and the threshold % of images where the nucleus needs to appear in order to be counted.

used as a measure for the accuracy of the location detection. The accuracy of the manual outline selection was measured by repeatedly selecting the embryo outline for the same image.

C. Software Implementation

A Canny edge detector was used to detect edges in raw images (Fig. 3a). Circular structures in the edge map were then parameterized using a circular HT (1). Circular structures were selected based on relative peak height in Hough space. For the parameterization, the span of a and b corresponded to the image width in pixels, and r was chosen as $(20 \ldots 25)$

pixels for nuclei, and as $(100 \ldots 120)$ pixels for embryo outlines. In the latter case, r was constructed using increments of 10 pixels, reducing the complexity of the accumulator, and thus reducing computational time. When extracting the EmbryoSlide or zygote outline (Fig. 1A), the desired outline was given by the strongest peak, without further filtration. When detecting nuclei or embryos of less circular shape (Fig. 1B,C), the correct outline only appeared as one in the group of 5-10 strongest Hough peaks, being weaker than a number of randomly appearing circular structures in the edge map. For these cases, the raw image was transformed using a trigonometric rotation,

$$\begin{aligned} x_{n+1} &= x_n \cdot cos(\theta) - y_n \cdot sin(\theta) \\ y_{n+1} &= x_n \cdot sin(\theta) + y_n \cdot cos(\theta) \end{aligned} \quad (2)$$

before repeating the Canny edge detection and HT. Unless the image was rotated $n \cdot 90°$, image values were shifted enough due to resampling of the image, to cause the random Hough peaks to appear in new locations. The result was a collection, C, of j circular structures per image, for i images,

$$C_{i,j} = [c_{i,j}(x,y), r_{i,j}], \quad (3)$$

where $c(x,y)$ is the centroid position, and r is the radius. After rotation, structures where $\mid c_{i,j} - c_{i+k,j} \mid \leq 3, \forall k \neq 0, j$ were counted as the same. Circular structures appearing in less than a number T of images could then be excluded, leaving only the real circular structures.

D. Automatic identification of pronuclear breakdown

Image series of all 20 embryos, with 20 images per embryo were analyzed and compared in an attempt to profile the PNB as shown in Fig. 2. Four methods were chosen and compared: direct detection of PN using the HT (HT+PN), comparison of the image difference (ID)

$$\Sigma_{x=1}^n \Sigma_{y=1}^m (I(x,y,t_k) - I(x,y,t_{k-1}))^2, \quad (4)$$

the image variance (V),

$$v(t) = \frac{1}{n \cdot m} \cdot \Sigma_{x=1}^n \Sigma_{y=1}^m (I(x,y,t) - \overline{I(x,y,t)})^2, \quad (5)$$

and the image variance within the ROI selected as the embryo outline by the HT (HT+IV). The ROI within the embryo was selected in four varieties; the complete embryo - referred to as HT+IV(100%) and a circular ROI with successively smaller radius, giving a measurement of the variance of the embryo interior. They are referred to as HT+IV(75%), HT+IV(50%) and HT+IV(25%) (with radius 75%, 50% and 25% of the radius of the embryo outline), respectively. Using the gradient of the variance of the ROI

$$g(t_k) = \frac{v(t_k) - v(t_{k-1})}{v(t_k)}. \quad (6)$$

Cases where $g(t_k) < g_{lim}$ are defined as PNB, where g_{lim} is some gradient threshold. For the two last computations; HT+IV(50%) and HT+IV(25%), the number of positive and negative hits were computed and compared. A positive signal, *i.e.* the PN disappearance was defined as the image where the PN visibly began to disappear, and the image immediately following, whereas all the rest were considered negative for PNB.

III. RESULTS

A. Embryo outline detection

The zygote outline was detectable without rotational filtering (Fig. 3f) for all cases tested, if the HT could be done using a radius span of < 20 pixels. If the radius span was larger, several false positives were detected also for the zygote outline, and were in that case filtered by six 60°angle rotations of the image. Rotational filtering was also necessary for the 2
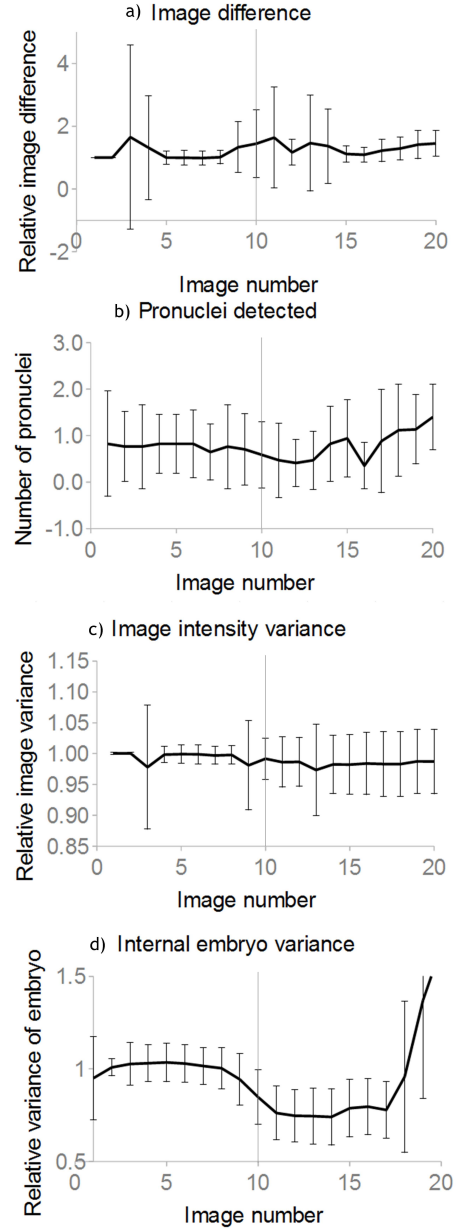

Fig. 7. Value of four calculations on images before and after PNB (marked by grey vertical line). Black line shows average over 20 embryos, bars indicate standard deviation. a) The numerical difference between each image and the previous image in the series (ID). b) Number of pronuclei detected by Hough Transform (HT+PN). c) Variance of total image (V). d) Variance of ROI selected by Hough Transform at 50% of embryo radius (HT+IV(50%)). All values are shown relative to the first image in the series.

cell and 3-4 cell stage, where the overall shape of the embryo appeared less circular. Fig. 4 shows the number of successful embryo outlines for the seven growth categories. If an accurate detection was defined as within 10% of the manually annotated values (and if the embryo is assumed to have a 100 pixel radius), a detection accuracy of 92.9% was achieved for all seven classes combined, and an accuracy of 95.0% if the class of especially difficult images was ignored. (Fig. 10 shows some examples of images in this category.) As a comparison, the manual selection of embryo outline had an accuracy of 70%
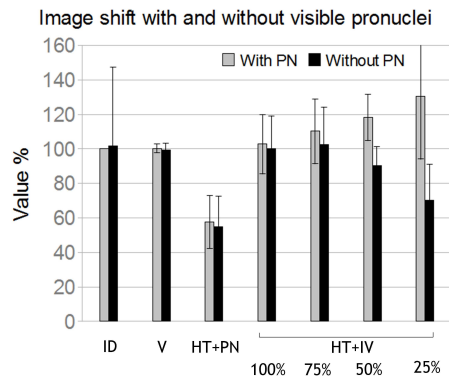
Fig. 8. Image values before and after PNB, averaged over 10 images each from 20 different embryos. Vales are given in % compared to the last image on which pronuclei are visible by eye. Time between images captured: 20 min.
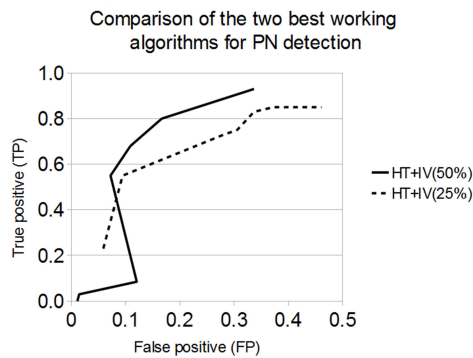

Fig. 9. Comparison of number of detected transitions from PN to no PN in 20 series of embryo images, consisting of 20 images each for the two algorithms showing the best potential. A true positive signal is determined by eye as the image where the PN start to break down, and the image immediately following (20 min later).

for 1 pixel variation, when repeated on the same picture by the same person, and 100% were within 5 pixel variation. The embryo outline detection was also tested on four image series where fresh medium was supplied at one point in the series, severely displacing the embryo. In Fig. 5, the positions of four embryos were automatically detected after being displaced.

B. Pronuclei detection

Pronuclei were detectable only after rotational filtering, because the outline of the pronucleus was suppressed by several stronger, but randomly located circular structures in the edge map. After repeating the circle detection on a number of rotated images, the circular structures were summed up and thresholded. The several randomly appearing circles had only partial overlap, in contrast to the pronucleus, which was detected in closely the same position in a large part of the images (Fig. 3a-e). The two parameters; number of angles to compute, and the threshold over which a structure needed to appear in order to be counted as a true structure, given in percent, were optimized using a series of 10 training images containing 1-2 PN each. The results are given in Fig. 6. The parameters for the original HT implementation, without rotation, was set to give 15 false positives initially. A rotation of 4 angles (90°) gave identical results, and an increased
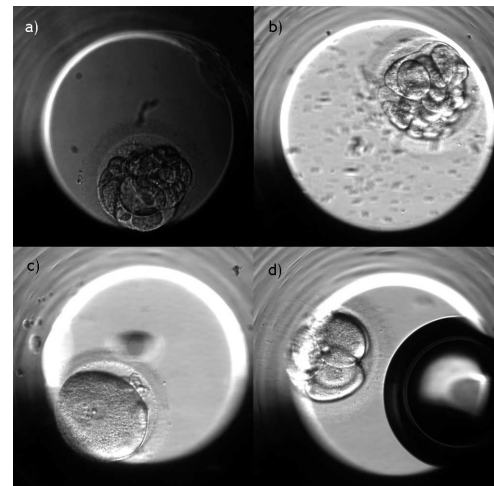

Fig. 10. Examples of difficult images, including imperfections in lighting conditions (a,b), deformed embryo placed close to the imaging well (c), and visible foreign objects (b,d).

number of false positives can be seen for 8 and 12 angles, as expected. For a rotation of 6 angles (60°), an improvement was apparent for a threshold of 75% and above, giving a reduced number of false positives to below 1/image at the expense of < 30% risk of obtaining a false negative. For an increased number of rotations, a further decrease in false positive rate was observed. However, at the same time the risk of experiencing a false negative, *i. e.* losing the detection, increased.

C. Detection of pronuclear breakdown

The mean of four of the six parameters in percent of the last image containing PN is shown in Fig. 7, where the first mitotic splitting is initiated at image 10 with the breakdown of the PN. For comparison, all values have been normalized to 1 for the first image. Fig. 8 shows a comparison of the ability of the four methods to distinguish images before and after PNB for all 20 test embryos. The ID in Fig. 7a shows an increase at the PNB, as expected, but the signal was often masked by slight displacements of the well, causing large random jumps in the signal. The detection of PN (HT+PN) in Fig. 7b shows a decrease at the PNB, but at a high variation of randomly appearing circular structures, especially toward the end of the time series (> image 15), where the cleavage began for several of the test series. The V in Fig. 7c was not as affected by image field dislocations as the ID, but on the other hand was not able to detect the small changes in image detail at the PNB. The HT+IV(50%) and HT+IV(25%) were the only measurements with a distinct difference before and after PNB (Fig. 8), showing a decrease in ROI variance at the PNB, and also an increase at the onset of cleavage, as expected (Fig. 7d). Fig. 9 shows a comparison of the HT+IV(50%) and HT+IV(25%) for varying thresholds g_{lim} for the gradient of the image variance. The best overall result was achieved by the HT+IV(50%),with 83.0% (80% TP, 17% FP) accuracy for all 20 image series, using 20 images per series.

IV. Conclusions

We have shown a method which automatically locates the embryo position from microscopic images at 92.9% accuracy even under noisy conditions. The detection accuracy was lower than for manual detection, but on the other hand required no user interaction with each image. We have also shown that the detection of the breakdown of pronuclei can be done automatically using the image variance, when an appropriate ROI is selected by a Hough Transform.

Our method of Hough detection performed best when the images contained clear circular structures. Embryo shapes showing edges differing strongly from a circular shape, such as the 2 cell stage and the blastocyst stage were more difficult to detect. One solution is to handle more complex shapes by introducing more parameters in the Hough Transform. However, care must be taken as the complexity of the transform increases rapidly.

For pronuclei and nuclei detection, there was a trade-off between falsely detected structures, and failing to detect the object we were looking for (false negatives). Here we have shown, that by using rotational filtering, we were able to reduce the number of false hits. For our purpose, automatic annotation, it was desirable to reduce the number of false positives at the expense of having a larger probability to drop the detection of nuclei and the overall best results were < 1/image at the expense of $< 30\%$ risk of obtaining a false negative for a $60°$ rotation.

Other options of nuclei detection, such as pre-processing to allow the edge detector to select only pronuclear structures, selection of a different edge detector, or different method of rotation are possible, and for future work, we will attempt to compare this with our method of rotational filtering.

The accuracy for embryo detection was 95% within 10% margin, but dropped dramatically to 59.2% within 5% margin. This can be compared to the accuracy of manual detection, which was 100% for both cases. (The manual accuracy may have been lower, if a comparison had been possible between several users.) However, the main reason for mismatch in the detection was the difference in outline definition between manual and computed. Whereas the manual selection was consistently done along the inner outline of the zona pellucida, the computation could select a stronger circular structure if the embryo was slightly displaced within the zona. To correct this for future work, it would be of interest to relax the demand for a circular shape, use the HT circle only as a starting point, and fine-tune the embryo outline detection by also using other image clues. Lastly, we have shown that when the ROI is restricted to the embryo interior, the achieved accuracy of detecting location was enough to detect the pronuclear breakdown at 83% accuracy.

References

[1] I. E. Agerholm, C. Hnida, D. G. Cruger, C. Berg, G. Bruun-Petersen, S. Klvraa, and S. Ziebe, *Nuclei size in relation to nuclear status and aneuploidy rate for 13 chromosomes in donated four cells embryos*, J. Assist. Reprod. Genet., vol. 25, no. 23, pp. 95−102, Mar. 2008.

[2] A. Azzarello, T. Hoest, och A. L. Mikkelsen, *The impact of pronuclei morphology and dynamicity on live birth outcome after time-lapse culture*, Hum. Reprod., vol 27, num 9, ss 2649−2657, Jan 2012.

[3] A. Beuchat, P. Thvenaz, M. Unser, T. Ebner, A. Senn, F. Urner, M. Germond, and C. O. S. Sorzano, *Quantitative morphometrical characterization of human pronuclear zygotes*, Hum. Reprod. Oxf. Engl., vol. 23, no. 9, pp. 1983−1992, Sep. 2008.

[4] M. P. Diamond, S. Willman, P. Chenette, och M. I. Cedars, *The clinical need for a method of identification of embryos destined to become a blastocyst in assisted reproductive technology cycles*, J. Assist. Reprod. Genet., vol 29, num 5, ss 391−396, Maj 2012.

[5] E. S. Filho, J. A. Noble, M. Poli, T. Griffiths, G. Emerson, and D. Wells, *A method for semi-automatic grading of human blastocyst microscope images*, Hum. Reprod. Oxf. Engl., vol. 27, no. 9, pp. 2641−2648, Sep. 2012.

[6] C. Garello, H. Baker, J. Rai, S. Montgomery, P. Wilson, C. R. Kennedy, and G. M. Hartshorne, *Pronuclear orientation, polar body placement, and embryo quality after intracytoplasmic sperm injection and in-vitro fertilization: further evidence for polarity in human oocytes?*, Hum. Reprod., vol. 14, no. 10, pp. 2588−2595, Oct. 1999.

[7] A. Giusti, G. Corani, L. M. Gambardella, M. C. Magli, and L. Gianaroli, *Blastomere segmentation and 3D morphology measurements of early embryos from Hoffman Modulation Contrast image stacks*, in 2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2010, pp. 1261−1264.

[8] C. Hnida, E. Engenheiro, and S. Ziebe, *Computer-controlled, multilevel, morphometric analysis of blastomere size as biomarker of fragmentation and multinuclearity in human embryos*, Hum. Reprod. Oxf. Engl., vol. 19, no. 2, pp. 288−293, Feb. 2004.

[9] R. Hoffman, *The modulation contrast microscope: principles and performance*, J. Microsc., vol. 110, no. 3, pp. 205−222, 1977.

[10] P. V. Hough, *Method and means for recognizing complex patterns*, US3069654 A18-Dec-1962.

[11] D. A. Morales, E. Bengoetxea, and P. Larranaga, *Automatic segmentation of zona pellucida in human embryo images applying an active contour model*, Proc. 12th Annu. Conf. Med. Image Underst. Anal., vol. 2008, pp. 209−213.

[12] U. D. Pedersen, O. F. Olsen, and Olsen, N.H., *A multiphase variational level set approach for modelling human embryos* Proc. 2nd IEEE Work. Var. Geom. Level Set Methods, vol. 2003.

[13] L. A. Scott and S. Smith, *The successful use of pronuclear embryo transfers the day following oocyte retrieval*, Hum. Reprod. Oxf. Engl., vol. 13, no. 4, pp. 1003−1013, Apr. 1998.

[14] L. Shapiro and G. Stockman, *Computer Vision*, vol. 2001. Prentice-Hall.

III

Robust detection of overlapping embryonic cells in microscopic bi-gradient images

Anna Mölder, Silvester Czanner, Nicholas Costen
School of Computing, Mathematics and Digital Technology
Manchester Metropolitan University
Manchester, UK
Email: Anna.Molder@mmu.ac.uk

Geraldine Hartshorne
Division of Reproductive Medicine
Warwick Medical School, University of Warwick
and Centre for Reproductive Medicine
University Hospitals Coventry and Warwickshire NHS Trust
Coventry, UK

Abstract—Detecting cells in a growing embryo is necessary for *in vitro* evaluation of embryo health, cell development and computation of cell lineage. Because of the sensitive nature of embryos, there are technical and ethical limitations imposed on the hardware available to examine them. Non-invasive imaging in combination with digital image cytometry is an appealing option, but faces a number of computational challenges, including disturbance from sample contamination, out of focus image regions, samples of low contrast and overlapping sample objects. Through optical modulation it is possible to enhance contrast, but the resulting images are bi-gradient in nature, further complicating cell detection. We present an algorithm for cell detection in bi-gradient images with overlapping objects, able to handle cell detection up to the 6 cell stage.

I. Introduction

Most common edge detectors are based on gradient detection, such as Canny, Sobel and Laplacian of Gaussian [Par97]. They are all very efficient for edge detection in a single gradient image, *i.e.* an image where pixel values increase or decrease continuously from image background to foreground. Examples include common microscopy techniques such as fluorescent imaging, brightfield and darkfield microscopy and confocal microscopy. The images from many optical contrast enhancing techniques are bi-gradient, *i.e.* the background pixels span the intermediate range of pixels values, and the pixels within the minimum and maximum range both belong to the foreground. Examples include common microscopy techniques such as phase contrast imaging and Hoffman Modulation Contrast (HMC) imaging. When applying a gradient based edge detector, the bi-gradient image displays double edges, if edges are at all located. In this study we describe a novel pathway for detecting edges in bi-gradient images, apply it to HMC images of human embryos and validate the result.

II. Materials and methods

A. Sample preparation

Human embryos were donated to research by patients undergoing IVF treatment, with informed consent. The project was approved by Coventry Research Ethics Committee (04/Q2802/26) and the Human Fertilisation and Embryology Authority (R0155). Embryos no longer required for treatment were cultured using Medicult media (Origio, Redhill, UK) for up to 7 days and incubated in 37°C 5%CO_2 in air.

B. Image capture

During image capture, embryos were cultured using Embryoscope (Fertilitech, Copenhagen, Denmark), and images were captured at up to 7 focal depths and 6 recordings per hour using a HMC [Hof77] optical set-up, with a 635nm LED as light source. Images from 18 embryos containing 1-6 cells, with cell outlines marked manually, were used for the analysis.

C. Image analysis

A set of elliptic model candidates representing cells were computed and combined for the best match with an edge map of the image. In matching, the elliptic perimeter was compared to the edges, and as one cell was found, its trace was successively removed from the edge map. The matching progressed until all cells were found.

1) Cell candidate location: Elliptic bodies in each image were first located using Hough Transform of a Canny edge version of the raw image [MCCH14]. The size range of elliptic structures to detect was set by the number of cells expected in the image. Before progressing to matching with the edge map, ellipses with excessive overlap were removed.

2) Edge map: The edge map was constructed from the lowest and highest pixel ranges in the image I using the threshold

$$t = c \cdot median(I). \tag{1}$$

If applied to the image as it is, this results in a high number of false positive edges. To filter those out, we constructed a focus filter [MCC14], computed a convex hull around the regions in focus and used this to exclude false positives. In the case of embryos, we know there is only one embryo per image, so we also add a pre-processing spatial filtering, where we detect the embryo region using a Hough Transform [Hou62].

III. Results

Figure 1 shows the result of edge detection, matching to edge map and validation to manual detection for a 2 cell and 6 cell case, respectively. To compare results (Figure 2) we define a correctly located cell as one where the centroid of the predicted ellipse is contained within the manually defined region, and a correctly described cell outline as one where the Intersection over Union (IOU) between result and prediction is $\geq 70\%$.
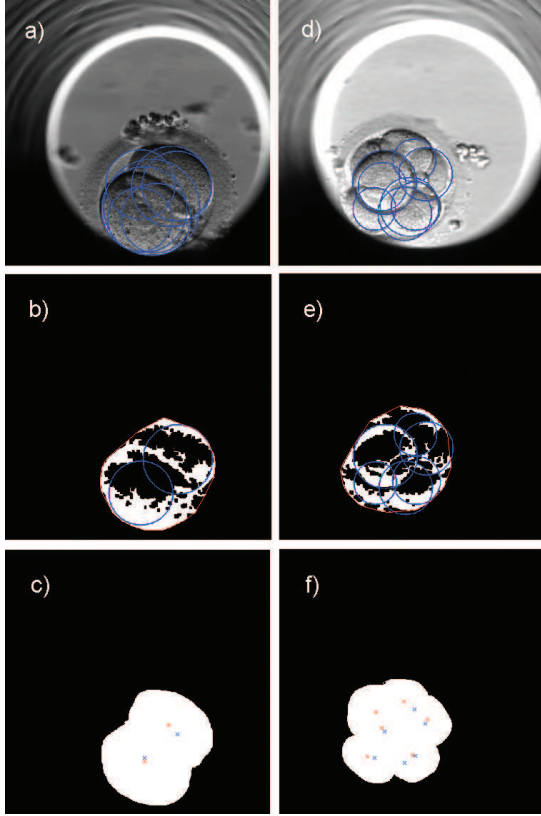
Fig. 1. Stepwise detection of cells in an embryo of two cells (a-c) and 6 cells (d-f). a,d) Original image, with elliptic candidates detected. b,e) Edge map with convex hull (red) and matched candidates (blue). c,f) Comparison between computed location (blue) and manual (red), x:s represent ellipse centroids.
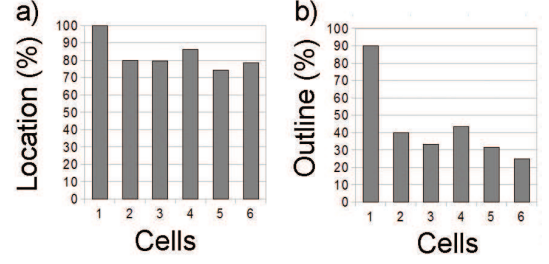
(Figure 1d, top cell).

Fig. 2. Localization (a) and Outline (b) detection accuracy from 1 to 6 cells.

V. REFERENCES

[Hof77] Hoffman R.: The modulation contrast microscope: principles and performance. *J. Microsc. 110*, 3 (1977), 205222.

[Hou62] Hough P. V.: Method and means for recognizing complex patterns. *US3069654 A*, (Dec. 1962).

[MCCH14] Mölder A., Czanner S., Costen N., Hartshorne G.: Automatic detection of circular structures in human embryo imaging using trigonometric rotation of the Hough Transform. *International Conference on Pattern Recognition 22*, (2014), 32393244.

[MCC14] Mölder A., Czanner S., Costen N.: Focal Plane Selection in Microscopic Embryo Images. *CGVC*, (2014).

[Par97] Parker, James R.: *Algorithms for Image Processing and Computer Vision*. John Wiley & Sons, Inc., 1997.

IV. DISCUSSION

The benefit of the described approach is two-fold. First, the construction of the edge map transforms the bi-gradient image into a single gradient image. This operation would have been trivial, if it were not for the need for focal filtering. The second advantage lies in the matching of candidates to the edge map. As long as the edge map contains some pixels along the candidate perimeter, it will be detected and there is no need for a complete outline. This is an advantage when handling potentially overlapping objects, which is the case in a three dimensional embryo. Limitations include the inability to separate objects with a high degree of overlap, naturally limiting this method to 6-8 cell embryos.

Results show a high degree of localization agreement, finding ≥ 80 % of cells up to the 4 cell stage and $\geq 70\%$ up to the 6 cell stage. Accuracy in detecting outlines was high only for the 1 cell stage, then at best 43% up to the 4 cell stage (the three cell case slightly lower). The main reason for this was deviation from a perfect elliptic shape in manually drawn outlines, where the 2 and 3 cell case often displayed more irregular outline than a case of more than 4 cells.

For future work we intend to adjust the outline more accurately, and also examine the possibility of extracting cells at different focal levels simultaneously. To do this, we wish to further exploit the fact that the focal filtering responds quite differently to objects in focus, compared to objects out of focus

IV

Methods to improve the understanding of microscopic embryo data sets using image analysis

Anna Mölder[*]

Supervised by: Silvester Czanner[†] and Nicholas Costen

School of Computing, Mathematics and Digital Technology
Faculty of Science and Engineering
Manchester Metropolitan University

Abstract

In this work we propose a framework for visualisation of semi-automatically segmented microscopic images of human embryos. A large part of the education of biologists consists of learning to interpret the output from a variety of analytical methods and medical imaging modalities, which can be more or less abstract in nature. Even in visual microscopy, the optical setup and the different ways to increase contrast between the sample and background produce image artefacts which have to be taken into account when interpreting the image. In *in vitro* fertilisation, the correct evaluation of the quality of the embryo is crucial for successful future development of the implanted foetus. Embryos are selected for transfer based on a number of characteristics, such as blastomere symmetry, degree of fragmentation and number and size of blastomeres. Traditionally, this evaluation has to a large extent been done by manual observation through visual microscopy, and obtaining the necessary expertise takes years of training. Here we show how the output from different analytical methods may be combined and how creative visualisation and improved user interaction with large data sets may improve the understanding of the sample under study. We show how existing computer-aided tools can be used in embryo selection and discuss automation as a way to quantify the subjective bias of manual embryo selection. We use data from human embryos as a case study, but the methods may be applied to any type of biological or microscopic material.

Keywords: Human-computer interaction, medical visualisation, embryology, image segmentation, pattern recognition, computer-aided diagnosis

1 Introduction

In several standard computer applications, computer vision algorithms are readily available, and in many fields of research there is a vast collection of tools for imaging and plotting, allowing the user to produce different visualisations of the data. In many cases these tools allow for multidimensional plotting, rescaling, the application of cross sections, the use of transfer functions, the addition of annotations, and the sorting and rescaling of data through various gates or histogram-based offsets. In the area of medicine, where the output from the method of detection is not visualised *a priori*, such as MRI, techniques for presenting the data have evolved for some time. Also, novel technique has made it possible to gather more and more data simultaneously in several dimensions and modalities. With this trend, the understanding of this data through a limited set of cross-sections has also become more difficult, and the interpretation of medical images has come to require more and more time in the education of the physician or the biologist. In microscopy, where data is directly visualised in two dimensions, little efforts have been done to present this data in any other way. Although the use of images from microscopic data has increased, the presentation tools available are usually two dimensional in nature, and lack interactivity. In microscopy, the same sample may look very different under different optical set-ups and the interpretation of microscopic images requires a high level of expertise. The aim of this paper is to give some insight into how standard computer vision techniques can be applied to microscopic data and how creative visualisation can help in the interpretation. We discuss how the physical process of image capture may influence the final image, and how knowledge of this process in some cases can be used to further improve the computer analysis. It is our hope that this article may further improve the microscopists understanding of the world of computer vision, and how they may use it for their benefit. We also believe that this article may be of interest to the computer scientists working in the field of computer vision, developing or studying algorithms for automated image analysis in the medical field.

---
[*]anna.molder@stu.mmu.ac.uk
[†]s.czanner@mmu.ac.uk

The background section of this paper (2) first gives a short overview of a few in embryology commonly used, yet very different microscopic techniques (2.1), then gives an overview of the embryo selection process (2.2), time-lapse microscopy (2.3) and segmentation (2.4) and how these can be applied to embryology. Section 3 describes the methods used and in the results section (4) we describe data extraction using two separate imaging techniques; HMC and confocal imaging. We apply three different methods of analysis to the data sets, using successively more refined methods of segmenting the image. Last, we discuss future work where we intend to extend the knowledge of human embryo development by combining information from several techniques. The beginning of each result section briefly describes the methods used.

2 Background and related work

2.1 Microscopic Imaging

Microscopic techniques can roughly be divided into quantitative or non-quantitative imaging, and destructive or non-descructive techniques. Non-descructive techniques is preferable in many cases, where there is a need to keep interference with the sample at a minimum. In *in vitro* fertilisation (IVF), the sample under observation cannot be manipulated or disturbed in any way, but must be observed "as is", if it is to be used for implantation. With a few exceptions, most quantitative imaging is in some way destructive, so for research purposes, destructive techniques can sometimes be desirable.

The two destructive techniques most commonly used in the study of embryos are fluorescence microscopy [1] and confocal microscopy [2]. Microscopical techniques which can be counted as non-destructive include bright- and dark-field microscopy, phase-contrast microscopy (PC) [3], Hoffman Modulation Contrast (HMC) microscopy [4], Differential Interference Contrast (DIC) [5] microscopy and digital holography (DH) [6], [7]. For uses in cellular biology, see [8], [9]. Of these, bright- and dark-field microscopy produces an image of the amplitude of the transmitted (or reflected) light, as we are used to seeing it. However, cellular material is usually highly transparent, and for such objects, we can get a better sample-to-background contrast, if we study the phase of light instead of the amplitude. PC microscopy and HMC imaging are techniques where the phase information of diffracted light is optically converted to amplitude information. Microscopic imaging techniques such as these are very good for visualisation, but cannot be directly translated to quantitative data. DIC and DH are techniques where the sample phase-shift is imaged directly, and can therefore count as quantitative imaging techniques.
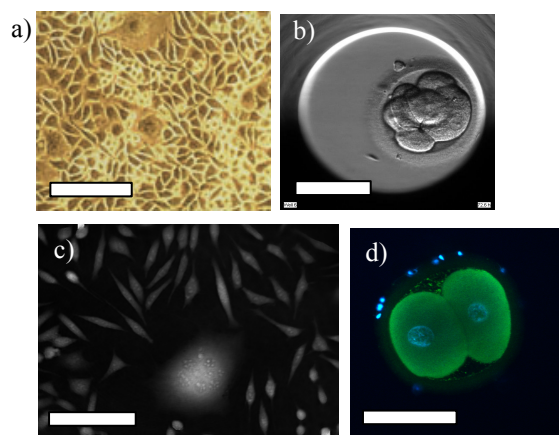

Figure 1: Examples of cellular images with common microscopic techniques. Phase contrast image of L-929 mouse fibroblast (a), HMC image of a human embryo (b), Digital holographic image (a) of mouse fibroblast L-929 (c) and Confocal image of a human embryo (d). Scale bars: 100µm.

2.2 The Embryo selection process

When selecting an embryo suitable for implantation, the embryologist may look at a number of criteria, such as pronuclear appearance and orientation [10], [11], number, size, shape of blastomeres, degree of fragmentation [12], degree of blastocoelic expansion, cellular composition and compactness of the inner cell mass and trophectoderm [13]. Discussions concerning the relevance of embryo morphology in quality assessment exists [14], but it is likely that embryo morphology will continue to play a large part in IVF embryo evaluation.

Traditionally, embryos have been studied using a microscope (commonly HMC) only at certain time points during the course of their development. It has been shown in time-lapse studies that the timing of key occurrences within the embryo can vary greatly between embryos that have similar morphologic appearance at the conclusion of the recording period, and correlation has been shown between the timing of key developmental events and embryo quality [15]. Some features, such as embryo fragmentation, which is usually connected with poor prognosis, have shown a high degree of variation in time-lapse studies [16], and embryos have also shown the capacity to reabsorb fragments [17]. It may be that the spatial and temporal pattern of fragmentation has higher impact on embryo quality than merely the presence of fragmentation [16], [18]. Such indication, in combination with new possibilities for time-lapse imaging of human embryos for an extended period of time with less negative effects to their health, makes it likely that the use of time-lapse recordings is going to increase in the future.

2.3 Time Lapse Microscopy

Time-lapse microscopy is the recording of an image sequence at intervals during a continuous period of time [17]. The length of the period and the time between intervals is determined as a trade-off between temporal resolution and potential sample deterioration. In fluorescence microscopy and confocal microscopy, which both count as destructive imaging techniques, the sample is usually fixed and no longer evolving, and it is rather fluorophore bleaching than potential damage to the sample itself, which limits sampling frequency. In non-destructive light microscopy on the other hand, imaging of live samples may be possible over several days, or even weeks. Long-term time-lapse imaging does not only require that the imaging technique causes low stress to the specimen. It also requires that the sample can be kept undisturbed in a favourable atmosphere for an extended period of time. Novel construction of incubators and cultivation chambers has recently made it possible to monitor embryos during the course of several days, without any registered severe consequences to their health.

There are difficulties other than the pure technical when combining automatic long term time-lapse imaging and microscopy. When examining embryos under the microscope, the three dimensional structure is very much of interest. In a traditional, manually handled microscope, much information can be gained by making proper use of the microscope controls, moving the sample around, scanning the focus, adjusting strength of illumination or making use of various filters and apertures in order to scan the three dimensional object in real time. In an automated time-lapse set-up, the possibility to manipulate optics is reduced when the optical set-up must incorporate a climate chamber to accommodate the living cellular material. If the microscope is instead meant to sit inside an incubator or other external chamber, the possibility to manipulate the optics is equally reduced, either because its operation requires the doors of the chamber to be opened, or because the optics is again shielded, to protect it from the high humidity of the chamber. In many time-lapse set-ups, the possibility to adjust image quality in real time has vanished, and the biologist is now limited to study the images some time after they are captured. This calls for new techniques to visualise this already captured data in creative ways, and possibly to regain some of the interactivity which was lost to the user in the process. Also, with the increased use of cameras and automated microscopic equipment, the amount of image data obtained has increased. Here is a possibility for more analytical material, but lots of data also means that time has to be spent interpreting the data. It would be beneficial to automatically point out features of interest in order to decrease the user workload.

2.4 Segmentation

There exists several examples of the segmentation of microscopic images in general [19]–[26], and segmentation of embryos in particular [27]–[31], but so far few attempts have been made to apply fully- or semi-automatic image treatment to the problem of selecting embryos. There are a number of potential benefits of automated image processing: Sampling time can be used for image processing, and the large amounts of stored image data available after capture will make the images available for further analysis and for validation by other experts. Automatic procedures will make the system less subjective, and the evaluation process will be more transparent, given that the automation process itself is made transparent. However, the differences between a standard camera image and microscopic images have a number of pitfalls, when applying standard image processing algorithms.

3 Materials and methods

The experimental section is divided into three parts. In the first, we apply a number of simple full image field transformations to a set of embryo images, to illustrate a common problem when working with the entire image. In the second, we restrict the region of analysis to regions of interest, and show how analysis of non-quantitative data still can give useful information. In the third and final section we illustrate how an embryo can be visualised in three dimensions given enough scans and a complete segmentation.

3.1 Asymmetric Imaging

Images of a human embryo at 72.6h after fertilisation were captured with the Embryoscope® system (Fertilitech, Copenhagen, Denmark), using HMC imaging at 635 nm. The raw images were plotted using Matlab, and a Canny edge and a one dimensional gradient was computed.

3.2 Embryo activity

The images in Figure 4 and 5 were captured in a 90h time-lapse series using the Embryoscope® system (Fertilitech, Copenhagen, Denmark), with a 0.2h interval between pictures. The embryos were mounted in wells in an EmbryoSlide® (Fertilitech, Copenhagen, Denmark) (Figure 3), one embryo per well, and the imaging of both wells (3 and 6 respectively) was done simultaneously, using a 635 nm LED. Both embryos are from the same patient. Three circular regions of interest were selected per image, one representing the total image field of the well (A), one selecting the body of the embryo within the zona pellucida (B), and one selecting the embryo centre at half the diameter of the embryo outline B (C) (Figure 4). The regions of interest have been manually

chosen from one image slide, and then applied to the rest of the image series. The variance of each region of interest was computed and plotted for each image in the series. The total image series of 448 images spans from 4.7h (4-6 blastomere stage) to 94.1h of development, at a focal plane located approximately half-way through the embryo.

3.3 Three dimensional Visualisation

Embryo nuclei were marked with DAPI, fixed on microscope slides and images were captured using the LSM510 confocal system (Zeiss, Hertfordshire, UK), using 400X magnification. The scans were captured with 1 μm between scans. The results were segmented using a combination of region growing segmentation filters (Neighbourhood Connected Thresholding and Confidence Connected Thresholding) and Watershed segmentation [32]. The segmented outlines where plotted as point clouds, and a Delauney triangulation was used to compute the surfaces. The resulting bodies were then put together in a three dimensional representation of the complete embryo. The complete experimental process has been described in [33].

4   Results

4.1 Asymmetric Imaging

In HMC imaging, light is passed through a pair of off-axis slits, converting gradients in sample optical path to bands of light and dark appearance, depending on the spatial sample direction (Figure 1b (raw image) and 2). Here, the slit pair is positioned so that image gradients appear symmetrical around a horizontal axis. This effect is most apparent when plotting the image as an isocontour (Figure 2a). When performing any kind of non-symmetrical image operations, this break in symmetry must be taken into account. Compare for instance the output of a derivative of the raw image taken along the horizontal and vertical axis, respectively (Figure 2c-d). In computer vision, a common approach is the application of an edge finding filter. Here too, the anti-symmetry of the HMC must be taken into account, because the angle of incident light will produce a shift in edges in the upper half of the image compared to the lower half. In Figure 2b, a Canny edge detection filter has been applied to the raw image. It is clear that the filter may find edges on both sides of the lighter and darker bands, resulting in an uncertainty when trying to determine the exact border of the embryo, or of a single embryo blastomere. However, since the direction of light depends on the azimuthal angle between splits and sample, this effect can be reduced by rotating the sample around a vertical axis, and combining information from several images along the rotation. Note also that the direction of change in optical path length is given by the polarity of the artefacts: Compare for instance the bright

upper half of the sample well with the dark upper edges of the blastomeres themselves.
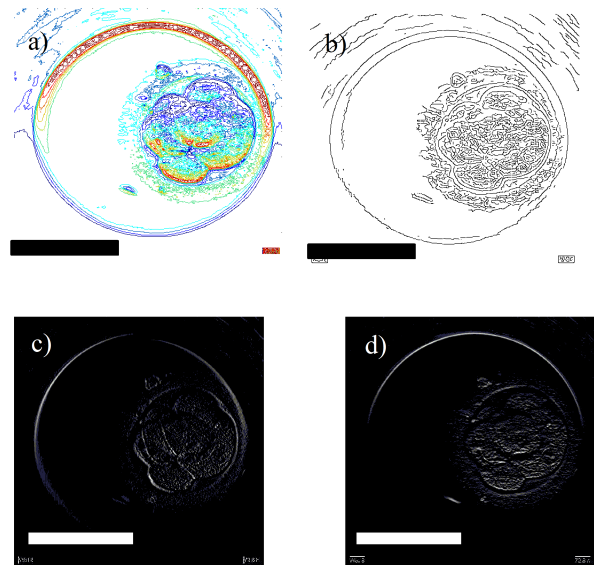

Figure 2: Contour plot of the raw HMC image given in fig 1b (a). Canny edge detector applied (b). Gradient of raw image taken in X (c) and Y (d) direction, respectively.

4.2 Embryo activity

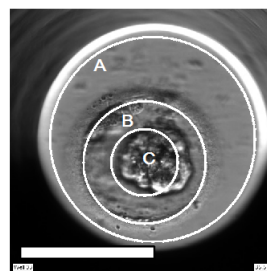
Figure 3: Cultivation slide used for embryos.

Figure 4: Regions of interest are a well (A), an embryo (B), and an internal region (C) of the embryo. The region C has been taken to be a region with centre equal to that of region B, but with half the radius. Scale bar: 100μm.
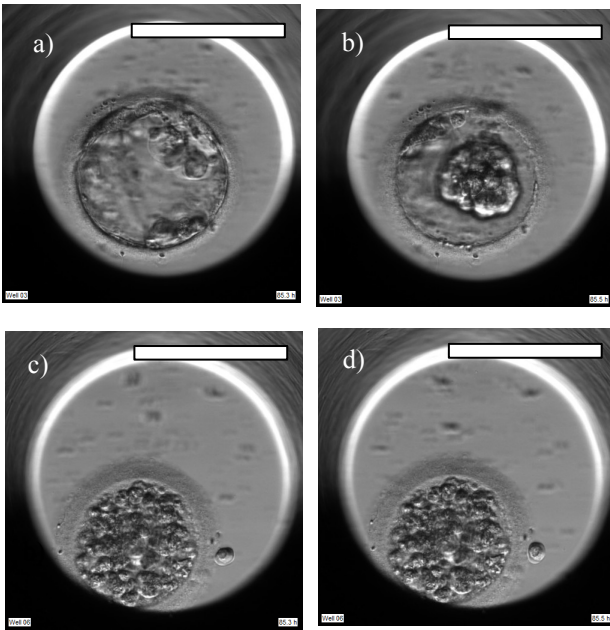
Figure 5: The embryo of well 3 at 85.3h (a) and 85.5h (b). The blastocoel starting to form. The embryo of well 6 at 85.3h (c) and 85.5h (d). Heavy fragmentation is visible, and the embryo activity is low. Scale bar: 100 μm.
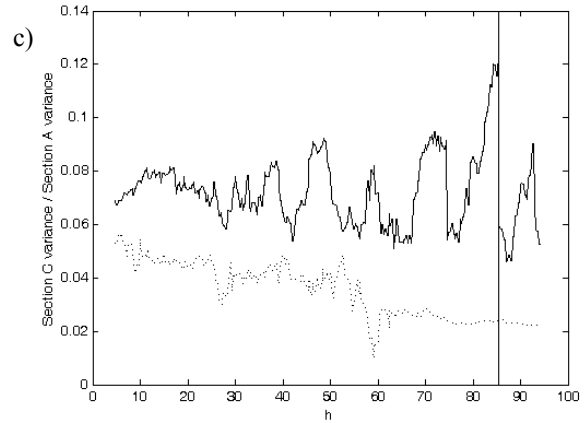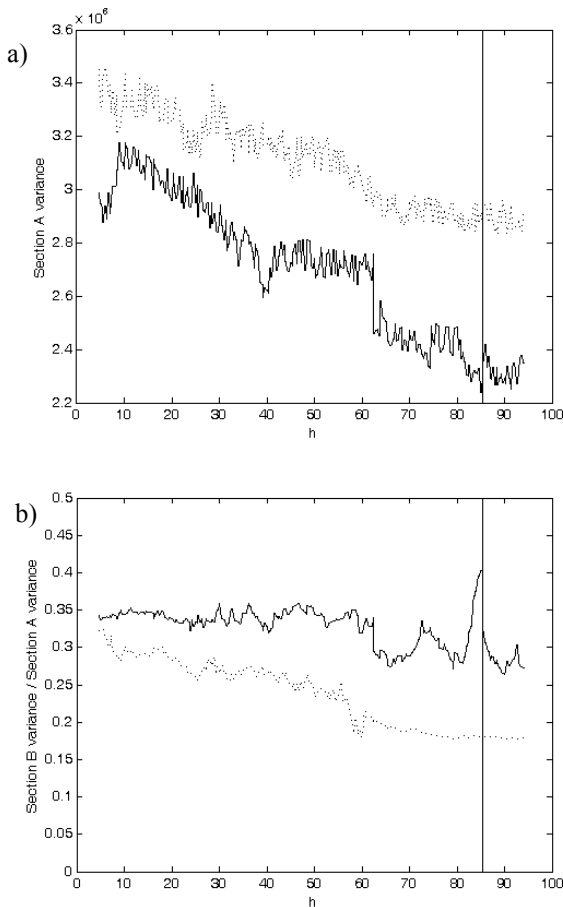

Figure 6: Variance (a) of region A. Arbitrary units. Image variance of region B (b) and region C (c) relative to mean of region A. Line represents position of example images (Figure 5). Solid: well 3, dashed: well 6.

The first embryo in well 3 experiences several cell divisions during the first hours of the series, after which it forms its first indication of a blastocoel at approximately 44.7h. After this, the embryo undergoes a series of morphological changes where it reverts back and forth between a blastocyst and a tight central cell structure (Figure 5). These changes are clearly reflected in the image amplitude and variance (Figure 6, solid).

The second embryo in well 6 experiences a reduction in division activity after approximately 34h, and clearly suffers from heavy fragmentation from image 54h and forward (Figure 5). Within the region of interest of the embryo, this is shown as a decreased image intensity, as well as a reduced image variance, when compared to the first embryo, even though for the image as a whole, the conditions are reversed. The effects are even more marked when the more restricted region of interest C is chosen (Figure 6, dashed).

4.3 Three dimensional Visualisation

Here we show that microscopic data from confocal images can generate enough data to form the basis for a three dimensional plotting with very little undersampling artefacts (Figure 7). The position and relative size of the blastomere nuclei are readily calculated. Note how the flattening of the embryo (a physical effect of being sandwiched between microscope slides) becomes apparent when applying a side view (Figure 7b).
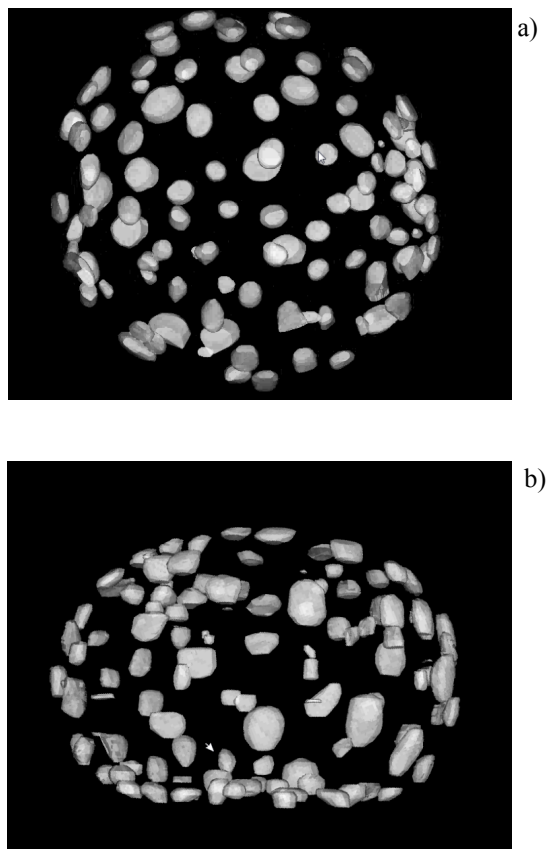

a)

b)

Figure 7: 3D rendering of the segmented data from a confocal image stack. Top view (a) and side view (b) of an embryo in blastocyst stage. The shapes shown are embryo nuclei (121 in total), spread along the inner wall of the blastocoel. The flattening of the embryo is due to the deformation caused by the imaging process. Blastocyst diameter is approximately 120 μm.

5   Conclusion and future work

Here we have shown how existing methods for image analysis may be combined to extract additional data from embryological data sets, and how computer analysis may be used to quantify results. Also, using 3D plotting, not only can we get a much more intuitive understanding of the embryo structure and the

positioning of blastomeres relative to each other – it is also possible to get a measurement of cellular or nuclear volume, which is not possible with a single scan. With a three dimensional display, it is possible to view the sample from different directions, thus getting a clearer view of its spatial layout, and gaining a better interactivity with the sample. Here we show that the methods for a complete analytical chain from raw image to three dimensional vector plotting exists, and for future work we intend to put these methods together into a working one-piece semi-automatic framework for embryo evaluation, simulation and visualisation. Clearly, the accuracy of the 3D model depends on the amount of available data, the xy-resolution and the number of scans in z-direction. Confocal microscopy has been chosen here because the low depth of field allows us to separate the signal between images in the stack, thus obtaining cleaner data. Optical sectioning is also possible in some non-destructive techniques, in particular in HMC which has a relatively low depth of field compared to other types of light microscopy. If a successful segmentation of this type of image could be achieved, it would open new doors in the area of embryology, both for clinical purpose, but also for the understanding of early human development. An improved embryo selection can in turn result in a greater number of successful implantations, less need for multiple embryo transfer, which will in turn increase the chance of survival for the foetus and reduce the risk to both the foetus and the mother.

A computerised model of an embryo is useful for embryologists for training purposes, and would also in many ways be of great help when understanding the three dimensional dynamics of the embryological content, and may bring further insight into the early stages of human embryo formation. In the future we intend to further investigate the possibilities to extract data from microscopic images, in particular focusing on the non-destructive modalities, and using other techniques such as confocal microscopy as an endpoint and method for comparison.

6   Acknowledgements

8 References

[1] J. W. Lichtman and J.-A. Conchello, "Fluorescence microscopy", *Nature Methods*, vol 2, issue 12, pp 910–919, Jan 2005.

[2] Y.-C. Liu and A.-S. Chiang, "High-resolution confocal imaging and three-dimensional rendering", *Methods*, vol 30, issue 1, pp 86–93, May 2003.

[3] M. W. Davidson and M. Abramowitz, "Optical Microscopy", in *Encyclopedia of Imaging Science and Technology*, John Wiley & Sons, Inc., 2002.

[4] R. Hoffman, "The modulation contrast microscope: principles and performance", *Journal of Microscopy*, vol 110, issue 3, pp 205–222, 1977.

[5] G. Nomarski, "Microinterferometrie differentiel a ondes polarisees", *J. Phys. Radium*, vol 16, s 9S–11S, 1955.

[6] U. Schnars and W. P. O. Jüptner, "Digital recording and reconstruction of holograms in hologram interferometry and shearography", *Appl. Opt.*, vol 33, issue 20, pp 4373–4377, Jul 1994.

[7] E. Cuche, P. Marquet, and C. Depeursinge, "Simultaneous amplitude-contrast and quantitative phase-contrast microscopy by numerical reconstruction of Fresnel off-axis holograms", *Appl. Opt.*, vol 38, issue 34, pp 6994–7001, Dec 1999.

[8] C. L. Curl, C. J. Bellair, P. J. Harris, B. E. Allman, A. Roberts, K. A. Nugent, and L. M. Delbridge, "Quantitative phase microscopy: A new tool for investigating the structure and function of unstained live cells", *Clinical and Experimental Pharmacology and Physiology*, vol 31, issue 12, pp 896–901, 2004.

[9] K. Alm, H. Cirenajwis, L. Gisselsson, A. Gjorloff, B. Janicke, A. Molder, S. Oredsson, and J. Persson, "Digital Holography and Cell Studies", in *Holography, Research and Technologies*, J. Rosen, Red InTech, 2011.

[10] L. A. Scott and S. Smith, "The successful use of pronuclear embryo transfers the day following oocyte retrieval", *Hum. Reprod.*, vol 13, issue 4, pp 1003–1013, Apr 1998.

[11] C. Garello, H. Baker, J. Rai, S. Montgomery, P. Wilson, C. R. Kennedy, and G. M. Hartshorne, "Pronuclear orientation, polar body placement, and embryo quality after intracytoplasmic sperm injection and in-vitro fertilization: further evidence for polarity in human oocytes?", *Hum. Reprod.*, vol 14, issue 10, pp 2588–2595, Jan 1999.

[12] F. Guerif, A. Le Gouge, B. Giraudeau, J. Poindron, R. Bidault, O. Gasnier, and D. Royere, "Limited value of morphological assessment at days 1 and 2 to predict blastocyst development potential: a prospective study based on 4042 embryos", *Hum. Reprod.*, vol 22, issue 7, pp 1973–1981, Jul 2007.

[13] W. B. Schoolcraft, D. K. Gardner, M. Lane, T. Schlenker, F. Hamilton, and D. R. Meldrum, "Blastocyst culture and transfer: analysis of results and parameters affecting outcome in two in vitro fertilization programs", *Fertil. Steril.*, vol 72, issue 4, pp 604–609, Oct 1999.

[14] T. Hardarson, G. Caisander, A. Sjögren, C. Hanson, L. Hamberger, and K. Lundin, "A morphological and chromosomal study of blastocysts developing from morphologically suboptimal human pre-embryos compared with control blastocysts", *Hum. Reprod.*, vol 18, issue 2, pp 399–407, Feb 2003.

[15] D. Payne, S. P. Flaherty, M. F. Barry, and C. D. Matthews, "Preliminary observations on polar body extrusion and pronuclear formation in human oocytes using time-lapse video cinematography", *Hum. Reprod.*, vol 12, issue 3, pp 532–541, Mar 1997.

[16] J. Van Blerkom, P. Davis, and S. Alexander, "A microscopic and biochemical study of fragmentation phenotypes in stage-appropriate human embryos", *Hum. Reprod.*, vol 16, issue 4, pp 719–729, Apr 2001.

[17] J. G. Lemmen, I. Agerholm, and S. Ziebe, "Kinetic markers of human embryo quality using time-lapse recordings of IVF/ICSI-fertilized oocytes", *Reprod. Biomed. Online*, vol 17, issue 3, pp 385–391, Sep 2008.

[18] M. Alikani, J. Cohen, G. Tomkin, G. J. Garrisi, C. Mack, and R. T. Scott, "Human embryo fragmentation in vitro and its implications for pregnancy and implantation", *Fertil. Steril.*, vol 71, issue 5, pp 836–842, May 1999.

[19] M. E. Ambühl, C. Brepsant, J.-J. Meister, A. B. Verkhovsky, and I. F. Sbalzarini, "High-resolution cell outline segmentation and tracking from phase-contrast microscopy images", *J Microsc*, vol 245, issue 2, pp 161–170, Feb 2012.

[20] T. A. Nenasheva, T. Carter, and G. I. Mashanov, "Automatic tracking of individual migrating cells using low-magnification dark-field microscopy", *J Microsc*, vol 246, issue 1, pp 83–88, Apr 2012.

[21] T. K. Jiyan Pan, "Learning to Detect Different Types of Cells under Phase Contrast Microscopy", 2009.

[22] M. R. Lamprecht, D. M. Sabatini, and A. E. Carpenter, "CellProfiler: free, versatile software for automated biological image analysis", *BioTechniques*, vol 42, issue 1, pp 71–75, Jan 2007.

[23] M. V. Boland, M. K. Markey, and R. F. Murphy, "Automated recognition of patterns characteristic of subcellular structures in fluorescence microscopy images", *Cytometry*, vol 33, issue 3, pp 366–375, Nov 1998.

[24] A. E. Carpenter, T. R. Jones, M. R. Lamprecht, C. Clarke, I. H. Kang, O. Friman, D. A. Guertin, J. H. Chang, R. A. Lindquist, J. Moffat, P. Golland, and D. M. Sabatini, "CellProfiler: image analysis software for identifying and quantifying cell phenotypes", *Genome Biology*, vol 7, issue 10, s R100, Okt 2006.

[25] X. Chen, X. Zhou, and S. T. C. Wong, "Automated segmentation, classification, and tracking of cancer cell nuclei in time-lapse microscopy", *IEEE Trans Biomed Eng*, vol 53, issue 4, pp 762–766, Apr 2006.

[26] X. Yang, H. Li, and X. Zhou, "Nuclei Segmentation Using Marker-Controlled Watershed, Tracking Using Mean-Shift, and Kalman Filter in Time-Lapse Microscopy", *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol 53, issue 11, pp 2405 –2414, Nov 2006.

[27] E. S. Filho, J. . Noble, and D. Wells, "A Review on Automatic Analysis of Human Embryo Microscope Images", *Open Biomed Eng J*, vol 4, pp 170–177, Oct 2010.

[28] E. Santos Filho, J. A. Noble, M. Poli, T. Griffiths, G. Emerson, and D. Wells, "A method for semi-automatic grading of human blastocyst microscope images", *Hum. Reprod.*, vol 27, issue 9, pp 2641–2648, Sep 2012.

[29] A. Giusti, G. Corani, L. Gambardella, C. Magli, and L. Gianaroli, "Blastomere segmentation and 3D morphology measurements of early embryos from Hoffman Modulation Contrast image stacks", in *2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 2010, pp 1261 –1264.

[30] A. Giusti, G. Corani, L. Gambardella, C. Magli, and L. Gianaroli,, *Segmentation of Human Zygotes in Hoffman Modulation Contrast Images. .*

[31] E. B. Dinora A Morales, "Automatic segmentation of zona pellucida in human embryo images applying an active contour model".

[32] T. S. Yoo, Red, *Insight into Images - CRC Press Book. .*

[33] K. Rafferty, S. Drury, G. Hartshorne, and S. Czanner, "Use of Concave Corners in the Segmentation of Embryological Datasets", *Review of Bioinformatics and Biometrics*, vol 1, issue 1, pp 1–8, 2012.

V

Cytometry
PART A
Journal of the International Society for Advancement of Cytometry

Semiautomated Analysis of Embryoscope Images: Using Localized Variance of Image Intensity to Detect Embryo Developmental Stages

Anna Mölder,[1]* Sarah Drury,[2] Nicholas Costen,[1] Geraldine M. Hartshorne,[2] Silvester Czanner[1]

International Society for Advancement of Cytometry

● **Abstract**
Embryo selection in in vitro fertilization (IVF) treatment has traditionally been done manually using microscopy at intermittent time points during embryo development. Novel technique has made it possible to monitor embryos using time lapse for long periods of time and together with the reduced cost of data storage, this has opened the door to long-term time-lapse monitoring, and large amounts of image material is now routinely gathered. However, the analysis is still to a large extent performed manually, and images are mostly used as qualitative reference. To make full use of the increased amount of microscopic image material, (semi)automated computer-aided tools are needed. An additional benefit of automation is the establishment of standardization tools for embryo selection and transfer, making decisions more transparent and less subjective. Another is the possibility to gather and analyze data in a high-throughput manner, gathering data from multiple clinics and increasing our knowledge of early human embryo development. In this study, the extraction of data to automatically select and track spatio-temporal events and features from sets of embryo images has been achieved using localized variance based on the distribution of image grey scale levels. A retrospective cohort study was performed using time-lapse imaging data derived from 39 human embryos from seven couples, covering the time from fertilization up to 6.3 days. The profile of localized variance has been used to characterize syngamy, mitotic division and stages of cleavage, compaction, and blastocoel formation. Prior to analysis, focal plane and embryo location were automatically detected, limiting precomputational user interaction to a calibration step and usable for automatic detection of region of interest (ROI) regardless of the method of analysis. The results were validated against the opinion of clinical experts. © 2015 International Society for Advancement of Cytometry

● **Key terms**
Key terms: automated image analysis; image-based embryo classification; computer-aided diagnosis; automated annotation; time-lapse microscopy; embryoscope; embryology

IN vitro fertilization (IVF) has been in clinical use for more than 30 years. Nevertheless, there is scope for improvement of the embryo selection procedure. By refining selection based on a greater understanding of embryo quality, we could not only reduce multiple births but also save patients the cost and distress of multiple failed attempts. Time-lapse imaging of embryos offers the prospect of such improvements and recent advances in incubator and imaging technology have enabled frequent observation and image capture of individual embryos at intervals of a few minutes. However, with the increased amount of generated imaging data, it is essential to find quality markers suitable for automated detection via computer-aided diagnostic tools. This technology has also opened up a new area of research studying the impact

of timing of key occurrences in embryo development. Currently, key events require to be identified and annotated manually, which is time consuming and limits the usefulness of the instrumentation. Noninvasive markers suitable for computer-aided diagnosis are being sought to standardize embryo selection procedures, speed up the annotation process, and provide diagnostic support.

Embryo quality is well known to relate to embryo morphology (1–3) but is not sufficiently precise an indicator to predict outcome reliably in individual patients. Embryo evaluation today is commonly undertaken using annotations of relevant features by experts at intermittent time points during development. Attempts have been made to standardize manual selection (4) and decision support systems exist for evaluating embryos (5–9). However, manual annotation is time consuming, the evaluation will vary according to the observer, and the different clinical conventions used (10). Automatic procedures to aid annotation would make the analysis less subjective and greatly reduce the manual workload involved.

Many reports have highlighted the need to observe embryo development dynamically (11–15). To thoroughly assess the benefits and drawbacks of time-lapse embryo imaging, large scale randomized clinical studies need to be performed, but before they can be done, at least two vital steps remain: The identification of the most promising markers to use and the setup of a system capable of collecting and analyzing large amounts of embryo data in a standardized and robust manner and consistency in evaluation is crucial to the usefulness of results. When migrating from a manual to an automated system, it may not be possible to require 1:1 correspondence between measurements, if the two selections are made based on distinctly different criteria. In these cases, a comparison must be made between manual and automatic evaluation, to establish the presence and size of any offset. Such comparative studies will need not only expertise in current methodology in both current embryo selection procedures and manual annotation but also a firm understanding of computerized image analysis and the nature of the image material used.

Several systems for automated embryo analysis using various approaches have previously been reported. One time-lapse system available uses an image-based decision tool analyzing cleavages to the four-cell stage using dark field optics (16). A few systems rely on direct modeling of physical conditions (17,18), requiring a highly controlled environment as well as detailed knowledge of the optical setup, something which is not always possible under clinical conditions. Other systems perform pattern recognition on microscopic images. Usually, a correctly performed segmentation (18–21) provides the most detailed information on blastomere position, shape, and outline, but this can be prone to errors, especially when used under clinical circumstances where complete and accurate segmentation may not always be possible. Using a semi-automatic approach, where a region of interest (ROI) has been selected manually (22–24); it is often possible to perform various computer vision and pattern recognition tasks even in a clinical setting. However, with a manual input required to initialize computation, this approach may instead increase user interaction with images, making it more suitable for indepth research purposes than for routine clinical work or large scale studies.

As an alternative, this study investigates the possibility of accessing relevant information using variations in image gray level in bright field images. The result is a framework for the detection of key events in embryo development without requiring samplewise initialization. At the same time, a graphical interpretation of embryo development as viewed in vitro is presented, serving as a complement to manual inspection of images.

MATERIALS AND METHODS

Embryo Culture and Image Capture

Time-lapse image series of human embryos fertilized in vitro were acquired as anonymized sequences of human embryos donated to research with ethical approval from Coventry Research Ethics Committee (04/Q2802/26) and the Human Fertilisation and Embryology Authority (R0155). Embryos were cultured in 25 μL culture media (Origio, Redhill, United Kingdom) under mineral oil for up to 6 days, incubated at 37 °C in an atmosphere of 5% CO_2, 5% O_2, and 90% N_2. The images were captured using the Embryoscope® system (Unisense Fertilitech, Copenhagen, Denmark), with up to seven focal depth planes, 15–25 μm apart, recorded at 20-min intervals using a Hoffman modulation contrast (HMC) optical setup (25) and a 635 nm LED as light source. Fresh medium was supplied at intervals, but embryos were otherwise undisturbed during imaging. The total dataset consisted of image series of 39 embryos from seven different couples, of which 28 developed into blastocysts. Fourteen series of embryos (of which nine developed to blastocysts) were used in an initial study (referred to as training set) to optimize algorithm parameters, and the analysis was repeated using the same parameters for the remaining 25 embryos (of which 18 developed blastocysts). The latter is referred to as the test set.

Software Implementation

Series of stacks of HMC images with a gray scale ranging from 0 to 255 provided the raw material for this study. In HMC microscopy, changes in optical path length are optically converted to light and dark gradients on an even gray background, resulting in an image where edges are the most prominent structures. As the number of edges in the image increases, the two dimensional distribution of image intensity changes. Objects in embryo development expected to result in an increased number of edges are visible nuclei and pronuclei as well as an increased number of blastomeres. Conversely, compaction and loss of focus are expected to increase image smoothness, following a loss of edge structures. Image variance is a measure of the distribution of gray levels within a specified region of the image and will increase with an increased number of edge structures. It is the hypothesis of this study that variance as a measure of edge structures can be used as an indirect method to identify the timing of embryo developmental stages.

To detect fluctuations in variance with sufficient sensitivity to distinguish changes caused by, for example, the appearance
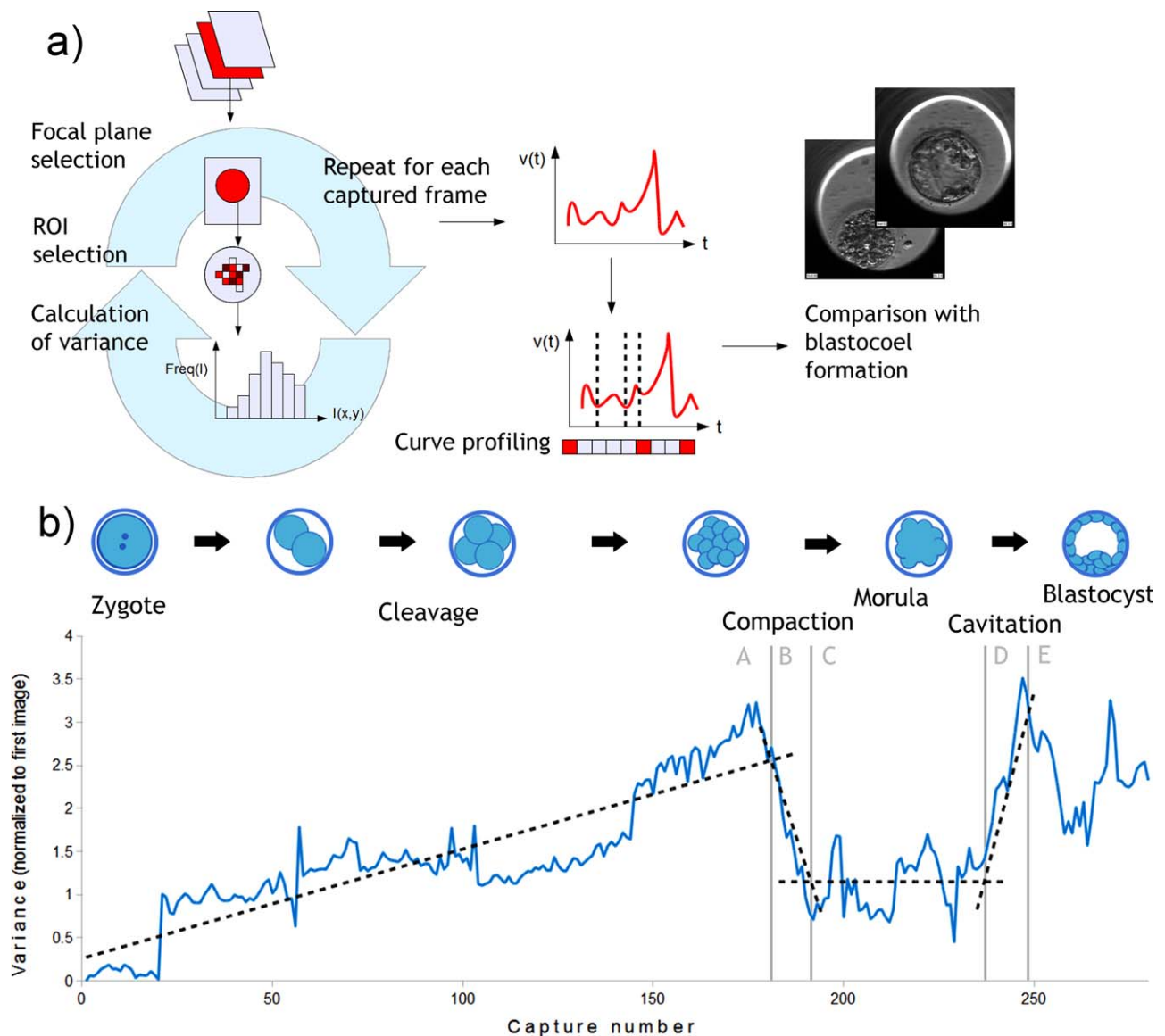
Embryo Developmental Stages Using Localized Variance

Figure 1. (**a**) Illustration of computational pipeline of the captured image series of an embryo. The optimal focal plane from the image stack was selected. A ROI was selected within each individual image, and one value of the variance in image intensity was computed for each ROI. This process was repeated for each capture in the image series, resulting in a function *v(t)* describing the variance as a function of time. *v(t)* was then further analyzed for the occurrence of detectable key events, profiling the embryo development. Finally, the profile for embryos forming blastocysts and not forming blastocysts were compared. (**b**) Image intensity variance of an embryo during the course of 280 frame captures, normalized to the first image in the series. Divisions during the cleavage stage are detectable as sudden increases in image variance, due to the number of increased edges in the image, as blastomeres undergo mitosis. At the onset of compaction, individual blastomere membranes are no longer distinguishable, and the variance drops and remains at a low level during the morula stage. The variance increases once more as blastocoel expansion sets in, and may fluctuate strongly during the blastocyst stage, if the embryo displays several cycles of collapse and re-expansion. The growth of the embryo has been considered in five stages: (A) Initial divisions from fertilization to onset of compaction. (B) Onset to completion of compaction. (C) Morula. (D) Cavitation. (E) Blastocyst. The mean and change in variance has been calculated for each section. Dashed trend lines have been added for illustrative purpose. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

of a nucleus, two prefiltering steps were necessary. The first step, which was used for every image series, selected one focal level in the stack as containing the optimal focus. This resulted in a sequence of single captures (Fig. 1a). The process is described further in the Supporting Information (Appendix A). The second step, performed on each remaining capture, automatically detected the outline of the embryo using a circular

Hough Transform. From the outline, the internal region of the embryo was selected as a circular ROI at half the embryo radius, as described in (26). The localized variance in image intensity was then calculated for the selected ROI of each image. Figure 2 shows an example of the breakdown of pronuclei and its effect on image variance. For the duration of the cleavage stage, it was assumed that no entire blastomere would
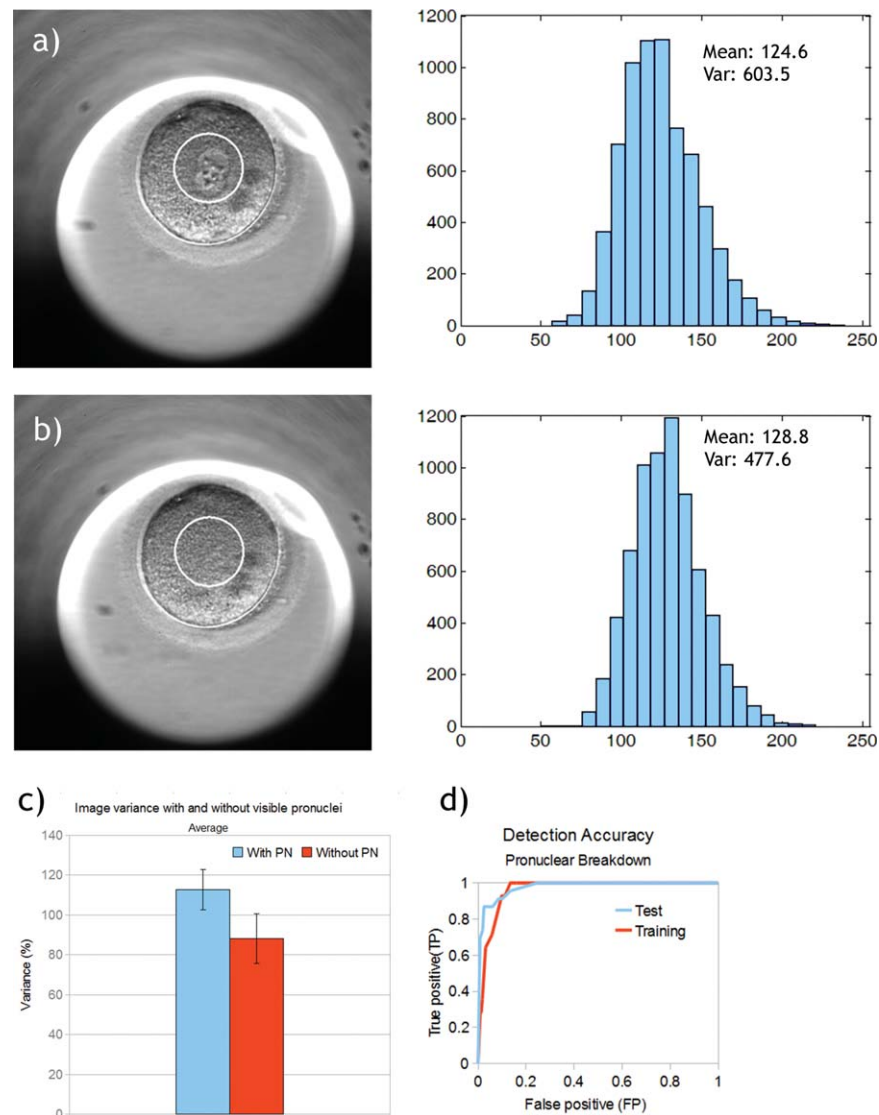
Figure 2. Calculation of variance in image intensity using pronuclei as an example. Images (**a**) and (**b**) were captured 20 min apart. The frequency of image gray scale values (0-255) within a selected ROI (white circle) at half embryo radius has been plotted as histograms, and the mean and variance calculated. (**c**) Difference in image variance before (blue) and after (red) PNB. Standard deviation calculated as mean over the training set of 14 embryos ($P < 0.0001$). (**d**) Detection accuracy of the training (14 embryos) and test (25 embryos) sets, respectively. The computation is governed by a single threshold (gradient of decreasing variance over time). PNB is defined as gradients larger than some threshold, yielding an increase in TP accuracy as the threshold decreases. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

appear completely outside the ROI. For the blastocyst stage, the choice of region proved useful as the formation of the trophectoderm removed infocus blastomeres from the embryo interior to the outline of the blastocyst (outside the ROI), making the finished blastocyst appear with a characteristic drop in image variance, once the cavity was formed. Figure 1b shows an example of an embryo growing in vitro, as viewed with the image intensity variance of the embryo interior.

Next, images were examined visually for key occurrences in embryo development, and the same events were evaluated using the image variance, constructing two characteristic profiles of a growing embryo, one obtained by manual observation, and one by mathematical inspection. The accuracy of the

hypothesis is defined by the correlation between the two profiles. The following details were included in the profile: the timing of the pronuclear breakdown (PNB) preceding syngamy, the timing of the first mitotic divisions up to eight-cell stage, and the transitions between a chosen set of main developmental stages. The details of the profiling are explained further in Supporting Information (Appendices B–D). A brief summary is given below.

Detection of Syngamy

For automatic detection of the PNB, a single threshold was optimized using the 14 training embryos. The timing of the PNB was computed for a number of thresholds, and the

Embryo Developmental Stages Using Localized Variance

minimum value giving 10% true positive (TP) detection (when comparing to visual inspection of the images of the training embryos) was selected and used for the testing embryos.

Cleavage Divisions

Mitotic divisions were also detected using a single threshold.

Compaction and Blastocyst Formation

Five stages were selected as being of interest: cleavage (A), compaction (B), morula (C), cavitation (D), and blastocyst (E) (Fig. 1b). The timing of transition between stages was defined as:

AB: main local maxima in variance, located before the main negative gradient.
BC: main negative gradient in variance.
CD: main positive gradient in variance, located after the main negative gradient.
DE: main local maxima in variance, located after the main negative gradient.

The computationally obtained stages and the transitions between stages were given letters to distinguish them from the visually defined embryo stages. For instance, the stage "B" is defined mathematically as the main negative gradient in variance, and it is part of the hypothesis that this relates to the formation of the compaction stage of the embryo. Finally, six traits for the developmental stages were combined and used simply to detect the presence or absence of a blastocoel. The six characteristics used were:

– The width (duration) of the negative gradient at compaction (B).
– The height of the maximum variance detected at the end of the cleavage stage (AB).
– The height of the maximum variance detected at cavitation (DE).
– The timing of compaction (B).
– The timing of the maximum at the end of the cleavage stage (AB).
– The total number of variance gradients during the entire development (a sign of strong fluctuating behavior, indicating poor quality).

The six traits were combined into four parameter sets, and the threshold for each one varied, while measuring the number of detected blastocyst formations.

Expert Validation and Statistical Analysis

Last, a total of 15 time-lapse image series from four different patients were used for validation. The timing of cell divisions and embryo stages was validated against the opinion of five expert clinical embryologists, each with at least 6 years of clinical embryology experience. The rest of the image series were annotated by the experimenters to the best ability using the same criteria as the embryologists. One image series was evaluated by all five experts, to allow direct comparison of their assessments. The annotation of timing in images was
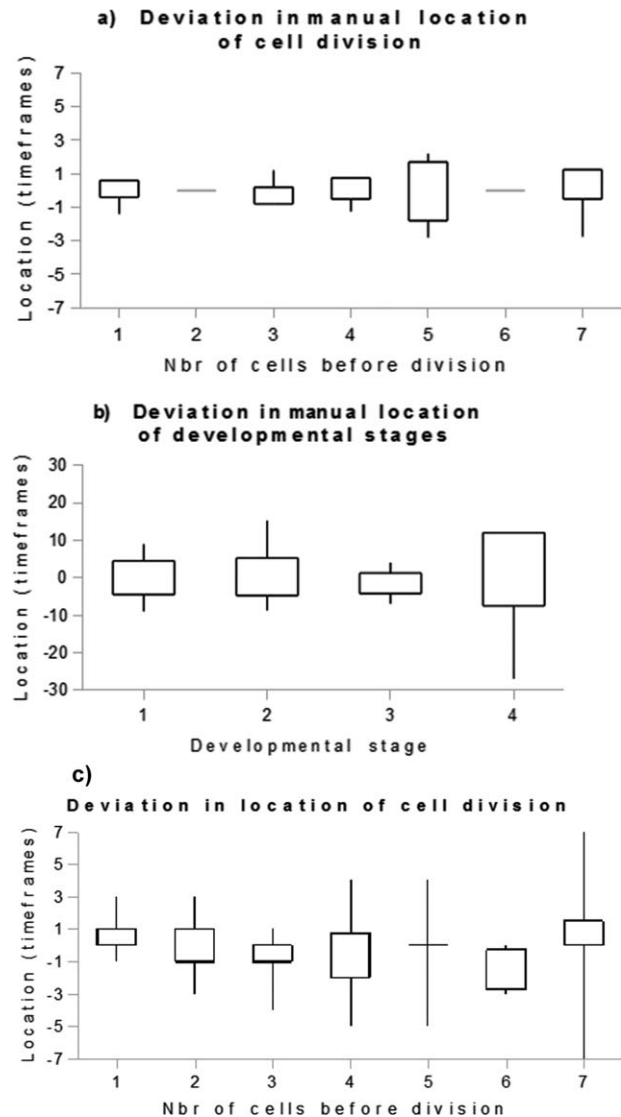

Figure 3. Standard deviation in manual annotation for the evaluation embryo in terms of (**a**) timings of mitotic divisions up to eight cells and (**b**) detection of the beginning of developmental stages: (1) Compaction, (2) Morula, (3) Cavitation, and (4) Blastocyst. Bars represent lower to upper quartile, whiskers minimum and maximum values. (**c**) The deviation from expert determined location of division in terms of timeframes for the divisions that were detected, plotted versus the number of cells preceding the division. Bars represent lower to upper quartile, whiskers minimum and maximum values.

consistently within 1–3 time frames up to an eight-cell stage, and the overall quality of the embryo in 100% agreement (Fig. 3). *P* values equal or inferior to 0.05 was used for statistical significance. Intervals in graphs and for values are given as means $\pm$ SD unless otherwise stated.

RESULTS

For embryos where developmental stages were visible in images, they were also reflected in the variance profile (Fig. 4). Both by manual observation and as measured by variance, large differences were apparent between individual embryos.
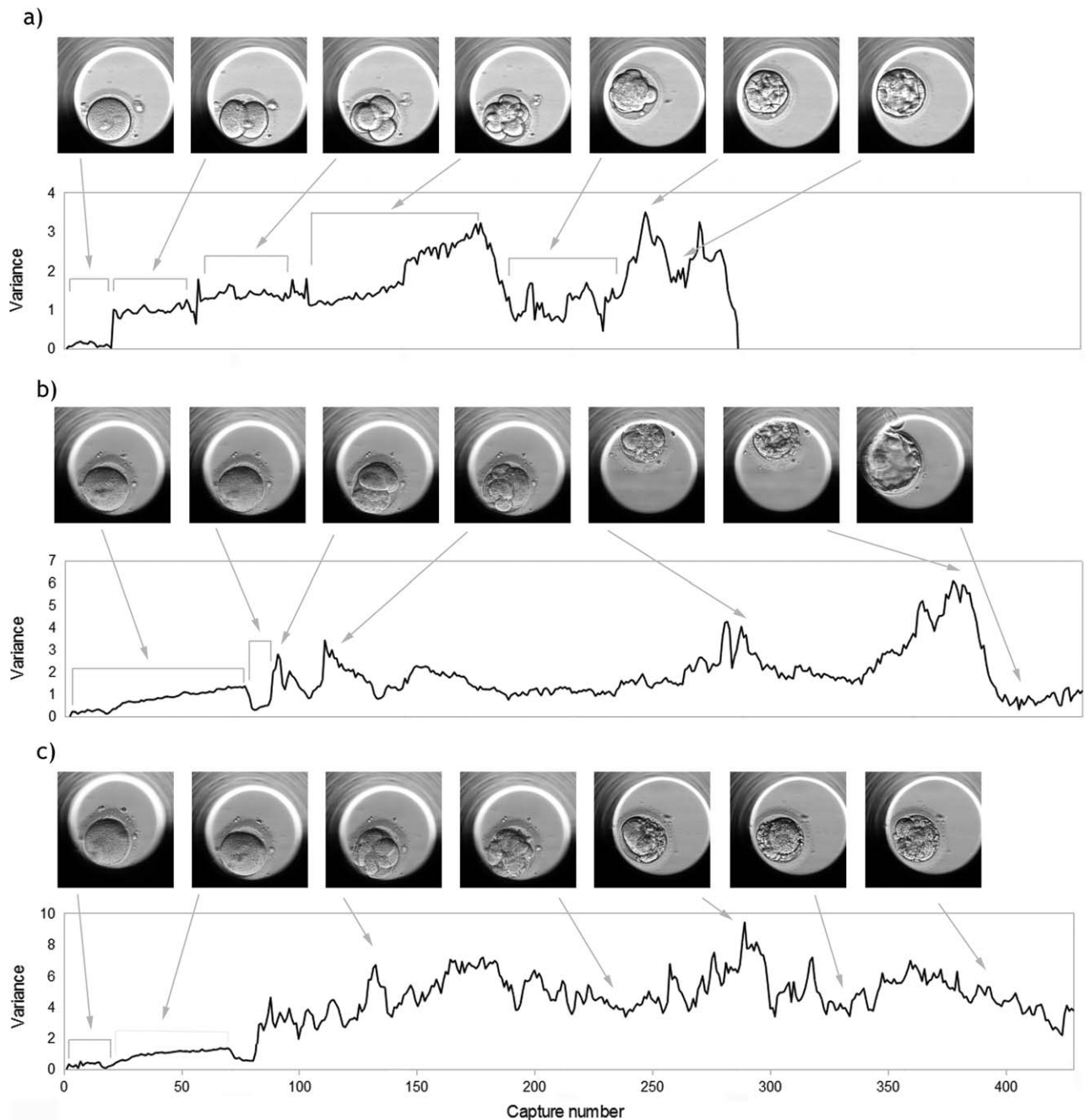
Figure 4. Profile of three representative embryos showing decreasing quality (**a–c**). Variance was calculated from the image intensity at a circular region encompassing the center of the embryo. A few example images are shown at points where characteristic changes are visible in the variance profile. For a good quality embryo (a), mitotic divisions are visible as successive increases in image variance, and the morula stage as a period of lowered variance. (b) illustrates a clearly expressed PNB, but experiences fragmentation during the cleavage stage, even though a blastocyst is eventually formed. In (c), the PNB is also apparent, but the embryo develops early fragments, never reaching a blastocyst stage.

Detection of Syngamy

In Figure 4b, the PNB is visible in the plot of variance versus image capture time as a sudden negative gradient over the course of 1–2 frames. Twenty images per embryo for the training set of 14 embryos were selected before and after PNB and used to profile the change of state. The difference in variance before and after the breakdown was large enough to be detectable, despite individual variation between images (Fig. 2c). The breakdown usually took less than one or two captures, giving an uncertainty of the timing of detection of at most 40 min at the current capture frequency. Requiring a 100% TP detection of the PNB for all 14 training embryos, the best overall result was 88% accuracy for the training embryos, and the inaccuracy being caused by false-positive
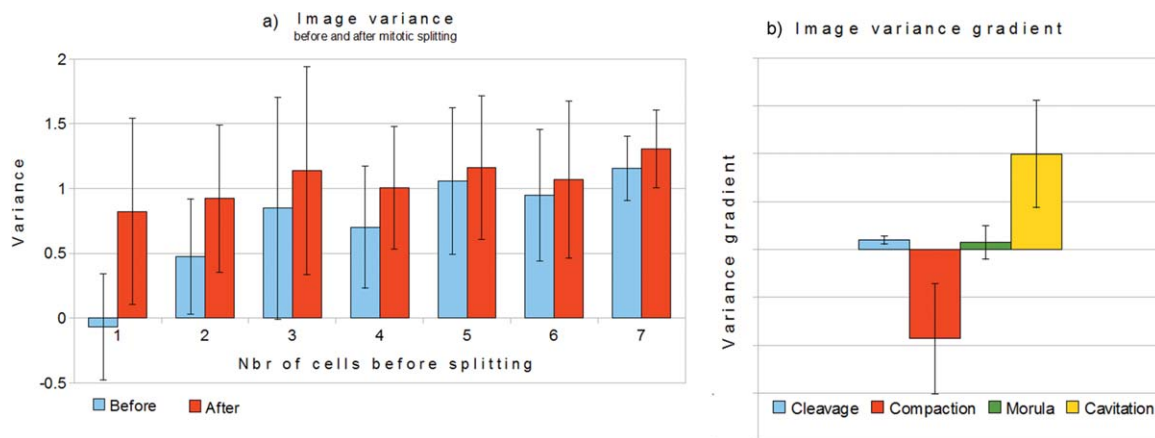
Figure 5. (**a**) Average variance for 14 training embryos before and after mitotic splitting. P values are $P < 0.05$ for first and second splitting, $P > 0.1$ for splitting 3–7. The negative variance before the first splitting is due to the drop in variance during syngamy. (**b**) Gradient of image variance for embryo developmental stages for the 14 training embryos. $P < 0.001$ for adjacent stages. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

detection. Using the same settings for the 25 test embryos, an overall detection accuracy of 90% was achieved, with 91% TP detection (Fig. 2d).

Cleavage Divisions

In total, 37 of 39 image series had sufficient quality to detect the first five mitotic divisions. Two embryos were excluded because of heavy optical interference. For most time series, it was possible to use the first automatically detected cell divisions, but manual adjustments were made in a few cases where both the division between 2 and 3 cells and between 3 and 4 cells appeared within the 20-min gap between captures. Computational detection was compared to manual detection for divisions of up to the four- and eight-cell stages, as shown in Figure 5. For embryos at the 1–8-cell stage, there was a clear bias toward divisions being under detected when using the automated procedure. For embryos at 1–4 cells, no more than two false positives (detection of divisions that were not present) or false negatives (failure to detect divisions) occurred per time series. From Figure 3c, it is apparent that the uncertainty in the exact location of division increases with the number of blastomeres. From the total image set of 37 embryos, 100% of divisions from 1 to 2 cells were detected, 73% from 2 to 3 (or 4) cells, 30% from 3 to 4 cells, and 59% from 4 to 5 (or 6) cells. The three and five-cell stages were not always distinguishable using a image capture frequency of 20 min. Of all divisions between 1 and 6 cells, 62% of divisions were located at the same captured frame index using both computer and manual detection, and 76% of divisions were located within one captured frame index from the manually noted position. The same values for the manual detection, as compared to the mean of the expert annotation, were 35% exact location and 74% within one time frame. For the 28 embryos which eventually formed a blastocyst, a measurement was also made of the time elapsed between the automatically detected cell divisions from 2 to 3 cells and from 3 to 4 cells on the total set of 39 image series, resulting in 10.27 ± 2.66 h (2–3 cells), and 1.11 ± 1.34 h (3–4 cells), respectively.

Compaction and Blastocyst Formation

Manual annotation by experts showed less agreement on timing of transitions between developmental stages (Fig. 4b), compared to detection of division. For automatic detection, the mean and gradient of the variance for each of the Stages A–E was computed for each embryo. The results are shown in Figure 5. The change in variance per unit time during the compaction and the cavitation stage was one order of magnitude higher than that for the entire cleavage stage, typically 0.3 h^{-1}. All values showed a high degree of variation (commonly with standard deviations in the range of 60–80%) between embryos. Interestingly, there was a distinct difference between embryos from different patients, when the duration of the four stages A–D was measured (Fig. 6). The duration of the morula stage showed high variability among embryos from the same patient, whereas the duration of the cleavage stage and the cavitation stage had a higher interpatient than intrapatient variability. The compaction stage, morula stage, and cavitation stage had approximately the same duration, about 1/7 that of the cleavage stage. However, the duration of the cleavage stage was only approximately determined since the exact time of fertilization was unknown for the series analyzed. Finally, the detected transitions and relative height of variance local maxima and gradients were combined and used to classify each embryo in two groups, those forming blastocysts and those failing to do so. The results were evaluated by visual inspection of the captured image series. The best overall result was correct detection of a blastocyst being formed in 71.8% of cases but at a cost of 28.2% false-positive detection (computational indication of a blastocyst without actual blastocyst formation), with little sensitivity to parameter setting (Fig. 7).

DISCUSSION

The method of locating the timing of mitotic divisions shows a larger span between maximum and minimum deviation from the true position compared to manual detection,
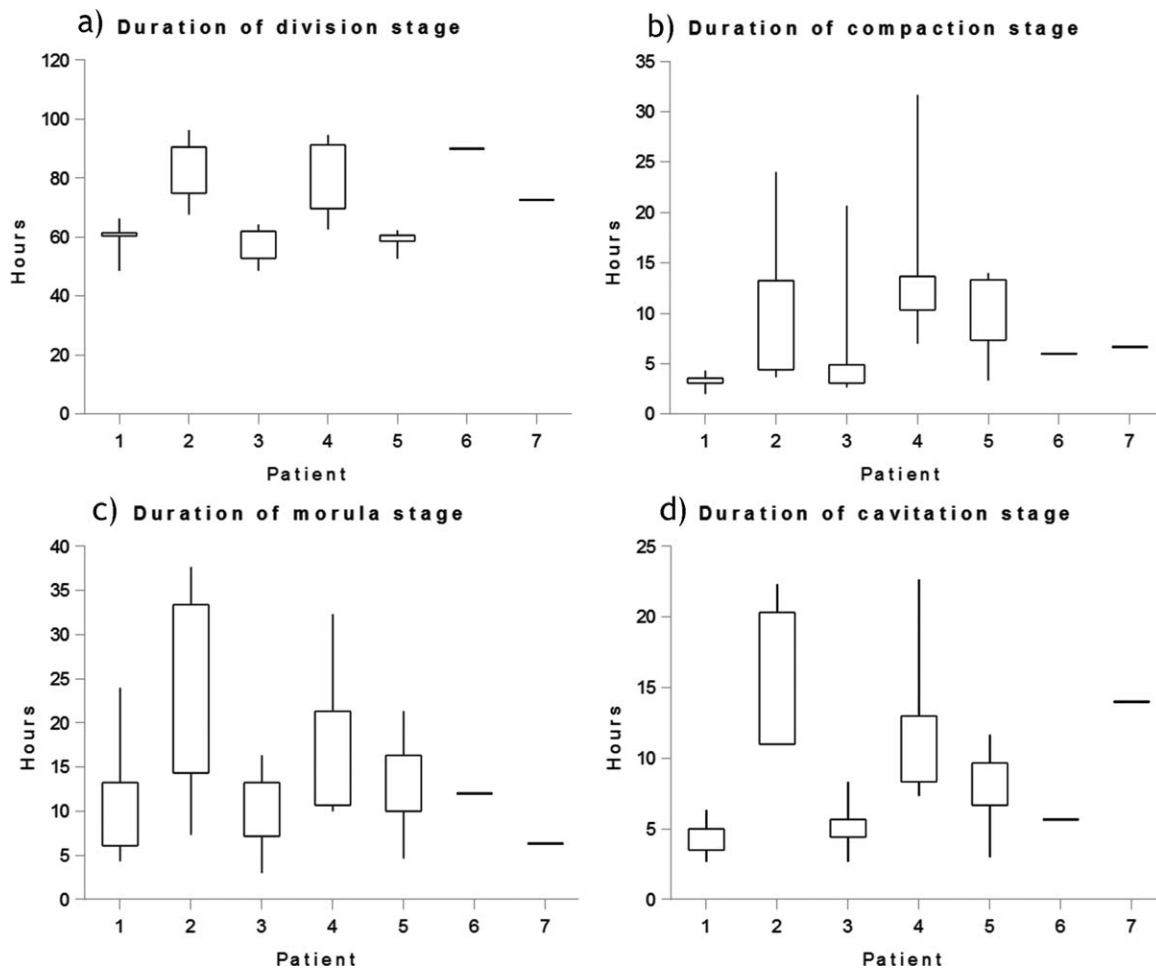
Figure 6. Duration in hours, automatically measured, of four stages of embryo development for seven patients (total 28 embryos). (**a**) Cleavage (the time from first frame to onset of compaction). (**b**) Compaction (time from onset until completed). (**c**) Duration of morula stage. (**d**) Duration of cavitation stage (time from onset of cavitation to blastocyst). Patients 6 and 7 had only one embryo each completing all four stages. Bars represent lower to upper quartile, whiskers minimum and maximum values.

but on average, our method performed better. 62% of cleavages identified by automatic detection were located at the exact same capture frame as manually identified by experts. The same agreement for manual detection between different experts was only 35%, showing the potential of automated image analysis to increase objectivity and consistency of embryo analysis. If instead, we define a correct detection to be within one captured frame of the control (corresponding to a timing inaccuracy of 20 min); the manual and computed accuracy were both approximately 75%. The results of the automatic method improved if only the 1–4-cell stages were considered, compared to all 1–8-cell stages. The results depend heavily on the frequency of image capture—20 min for this study—which was long enough for most cell divisions to take place over the course of several captured frames, but we experienced difficulty in distinguishing the three, five, and seven blastomere stages at this capture frequency. With more images captured and analyzed per unit time, it is possible that the uncertainty in location in terms of image index may increase, while at the same time decrease if computed for

clock time. In measuring the timing of the first few mitotic divisions, the results overlap but have higher standard deviations than a previously reported study (27). However, the results for (27) were obtained with visual counting of mitotic divisions, whereas the timing of divisions in this study were automatically computed.

In detecting embryonic developmental stages, there were large variations between individual embryos, as expected from clinical experience. In spite of this, a clear trend in the variance profile was apparent, and we have shown that it was possible to identify the formation of a blastocyst by automated image analysis in >70% of cases. It was also apparent that the definition of stages and transitions using the localized variance was different from that of manual detection, indicating that this way of visualizing blastocyst development may serve best as a complement to inspection of images by eye.

The parameters for the detection of blastocyst formation depend on the frequency of image capture and the hardware settings, adding a requirement for a calibration stage before analysis. For future work, a comparison between different
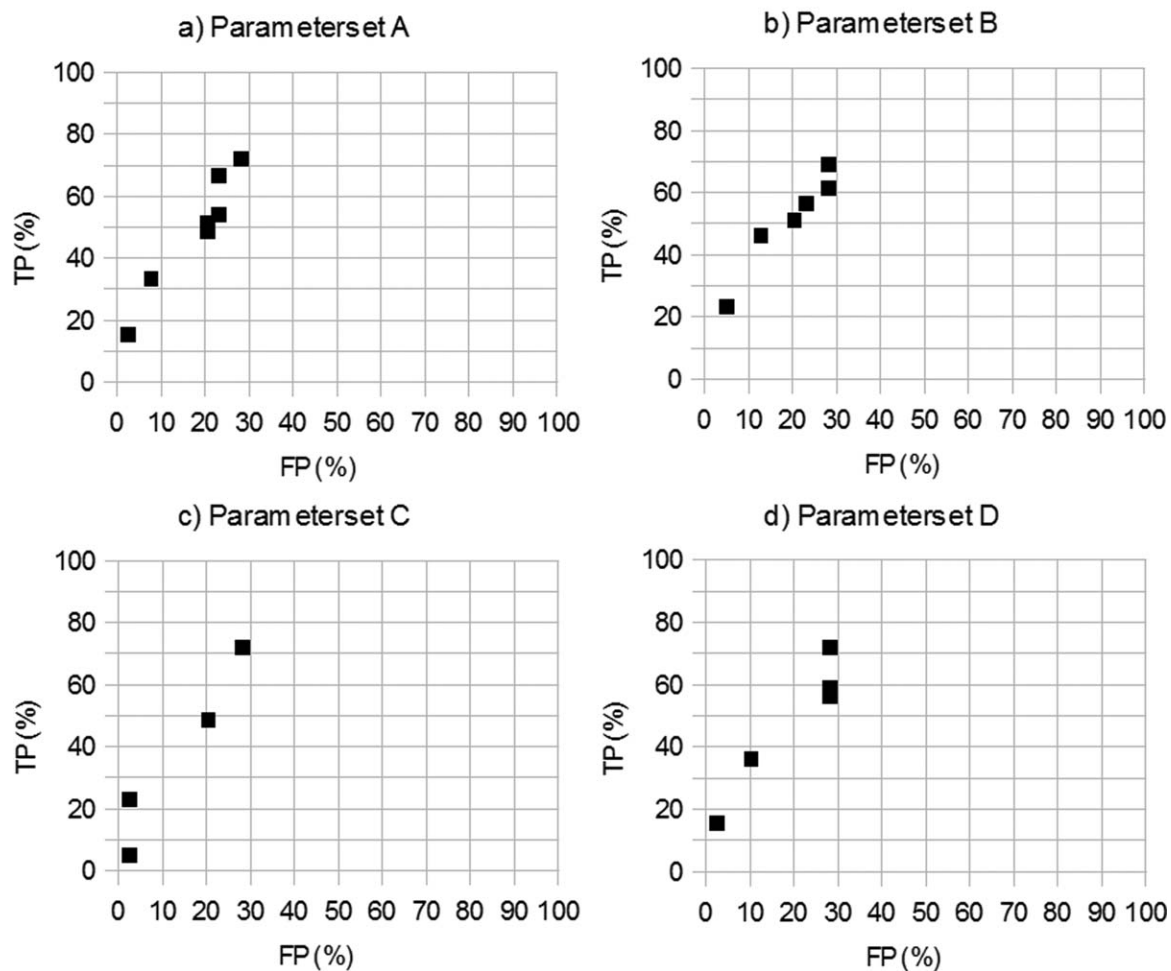
Figure 7. Example of the results in blastocoel detection using four different parameter sets. The parameters were: A: q1 = 20, q2 = 50, q3 = 0.65, q4 = 5, B: q1 = 20, q2 = 10, q3 = 0.65, q4 = 5, C: q1 = 20, q2 = 50, q3 = 0.65, q4 = 1, and D: q1 = 20, q2 = 50, q3 = 0.65, q4 = 10, where q1 is relative location of first main gradient, q2 is width of main gradient, q3 is height of the main maxima, and q4 is maximum number of gradients. Weights were w1 = 0.15, w2 = 0.15, w3 = 0.3, and w4 = 0.4 for all cases.

image capture frequencies would be desirable. It is also evident that the exact appearance of the variance function $v(t)$ depended on the choice of the ROI. In this study, selecting a circular region of four different radii were investigated. Also, parameters for this study (Appendix D) were chosen as the most feasible using our current knowledge of embryo development. The implication of any choice of parameters should be further evaluated before taking on a larger scale study, as it is possible that new technical tools to study embryo will also require new methods to define embryo health. Furthermore, there is a trade- off between minimizing false-negative and false-positive detection. For our purpose of automating annotation, it was decided that false negatives were undesirable, whereas false positives could be acceptable and handled in a future manual or automatic filtering step. This decision may change depending on the intended purpose of detection. Also, to improve the accuracy, this framework could be expanded using more extractable image cues. For example, local image texture or measurements based on direct recognition of blas-

tomere outlines could be used. Last, the use of an automatic image analysis is dependent on initial image quality, and for larger studies it will be necessary to establish robustness under clinical conditions.

There are reasons for caution in evaluating embryo quality as all studies of embryos before implantation will per definition only be able to assess embryo quality, not taking into account the uterine component of implantation. In IVF treatment, one or more embryos are normally selected for transfer to the uterus on Days 2–3 or 5–6 of development, when those developing in a normal and timely manner are usually at the four-cell, eight-cell or blastocyst stages, respectively. However, many embryos harbor abnormalities that render them incapable of prolonged development, and some of these abnormalities become manifest during preimplantation development as abnormal, delayed, or arrested growth. Thus, embryos transferred at the blastocyst stage are more likely to result in pregnancy than those transferred at earlier stages. Blastocyst transfer is, therefore, associated with a higher chance of

pregnancy and is the latest stage at which preimplantation selection can be carried out. Recent results (15) show that time-lapse studies of earlier embryonic stages can predict blastocyst development, but that the formation of a blastocyst is not necessarily an indication of a live birth outcome. In this study, we defined blastocyst formation to be evidence of a good quality embryo, but for future work we shall extend this study to clinical data where success in terms of initiating pregnancy and resulting in a live birth is known. Still, there is a need for prediction of blastocyst formation (28), and the ability for negative prediction, that is, deselection of unsuitable embryos, has the potential to save resources and allow for a more robust selection of single successful embryos for transfer. This could be achieved using automated analysis of previously identified parameters, such as immediate cleavage.

In conclusion, it is shown here that key events in preimplantation embryo development can be detected using a simple automated approach to embryo time-lapse image analysis, offering the potential of semiautomated annotation of clinical images on a large scale. The skills of the embryologist may then be better focused on checking and correcting a reduced number of uncertain computations, rather than performing routine manual annotation of the complete image set.

LITERATURE CITED

1. Schoolcraft WB, Gardner DK, Lane M, Schlenker T, Hamilton F, Meldrum DR. Blastocyst culture and transfer: Analysis of results and parameters affecting outcome in two in vitro fertilization programs. Fertil Steril 1999;72:604–609.
2. Sakkas D, Percival G, D'Arcy Y, Sharif K, Afnan M. Assessment of early cleaving in vitro fertilized human embryos at the 2-cell stage before transfer improves embryo selection. Fertil Steril 2001;76:1150–1156.
3. Guerif F, Le GA, Giraudeau B, Poindron J, Bidault R, Gasnier O, Royere D. Limited value of morphological assessment at days 1 and 2 to predict blastocyst development potential: A prospective study based on 4042 embryos. Hum Reprod 2007;22:1973–1981.
4. Racowsky C, Vernon M, Mayer J, Ball GD, Behr B, Pomeroy KO, Wininger D, Gibbons W, Conaghan J, Stern JE. Standardization of grading embryo morphology. J Assist Reprod Genet 2010;27:437–439.
5. Jurisica I, Mylopoulos J, Glasgow J, Shapiro H, Casper RF. Case-based reasoning in IVF: Prediction and knowledge mining. Artif Intell Med 1998;12:1–24.
6. Manna C, Patrizi G, Rahman A, Sallam H. Experimental results on the recognition of embryos in human assisted reproduction. Reprod Biomed Online 2004;8:460–469.
7. Patrizi G, Manna C, Moscatelli C, Nieddu L. Pattern recognition methods in human-assisted reproduction. Int Trans Oper Res 2004;11:365–379.
8. Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: A systematic review of trials to identify features critical to success. Br Med J 2005;330:765.
9. Morales DA, Bengoetxea E, Larrañaga P, García M, Franco Y, Fresnada M, Merino M. Bayesian classification for the selection of in-vitro human embryos using morphological and clinical data. Comput Methods Programs Biomed 2008;90:104–116.
10. Sundvall L, Ingerslev HJ, Knudsen UB, Kirkegaard K. Inter- and intra-observer variability of time-lapse annotations. Hum Reprod 2013;28:3215–3221.
11. Azzarello A, Hoest T, Mikkelsen AL. The impact of pronuclei morphology and dynamicity on live birth outcome after time-lapse culture. Hum Reprod 2012;27:2649–2657.
12. Basile N, Meseguer M. Time-lapse technology: Evaluation of embryo quality and new markers for embryo selection. Expert Rev Obstet Gynecol 2012;7:175–190.
13. Cruz M, Garrido N, Herrero J, Pérez-Cano I, Muñoz M, Meseguer M. Timing of cell division in human cleavage-stage embryos is linked with blastocyst formation and quality. Reprod Biomed Online 2012;25:371–381.
14. Kirkegaard K, Agerholm IE, Ingerslev HJ. Time-lapse monitoring as a tool for clinical embryo assessment. Hum Reprod 2012;27:1277–1285.
15. Kirkegaard K, Kesmodel US, Hindkjær JJ, Ingerslev HJ. Time-lapse parameters as predictors of blastocyst development and pregnancy outcome in embryos from good prognosis patients: A prospective cohort study. Hum Reprod 2013;28:2643–2651.
16. Wong CC, Loewke KE, Bossert NL, Behr B, De Jonge CJ, Baer TM, Pera RAR. Non-invasive imaging of human embryos before embryonic genome activation predicts development to the blastocyst stage. Nat Biotechnol 2010;28:1115–1121.
17. Olsen NH. Morphology and Optics of Human Embryos from Light Microscopy. Department of Innovation, IT University of Copenhagen, 2004.
18. Giusti A, Corani G, Gambardella LM, Magli MC, Gianaroli L. Blastomere segmentation and 3D morphology measurements of early embryos from Hoffman Modulation Contrast image stacks. Rotterdam: IEEE Int Symp Biomed Imaging 2010;1261−1264.
19. Agerholm IE, Hnida C, Cruger DG, Berg C, Bruun-Petersen G, Kølvraa S, Ziebe S. Nuclei size in relation to nuclear status and aneuploidy rate for 13 chromosomes in donated four cells embryos. J Assist Reprod Genet 2008;25:95–102.
20. Morales DA, Bengoetxea E. Automatic segmentation of zona pellucida in human embryo images applying an active contour model. Dundee, UK: Proceedings of the 12th Annual Conference on Medical Image Understanding and Analysis 2008;209–213.
21. Filho ES, Noble JA, Poli M, Griffiths T, Emerson G, Wells D. A method for semi-automatic grading of human blastocyst microscope images. Hum Reprod 2012;27:2641–2648.
22. Pedersen UD, Olsen, OF, Olsen, NH. A multiphase variational level set approach for modelling human embryos. In: Proceedings of the 2nd IEEE Workshop on variational, Geometric and Level Set Methods, 2003.
23. Hnida C, Engenheiro E, Ziebe S. Computer-controlled, multilevel, morphometric analysis of blastomere size as biomarker of fragmentation and multinuclearity in human embryos. Hum Reprod 2004;19:288–293.
24. Beuchat A, Thévenaz P, Unser M, Ebner T, Senn A, Urner F, Germond M, Sorzano COS. Quantitative morphometrical characterization of human pronuclear zygotes. Hum Reprod 2008;23:1983–1992.
25. Hoffman R. The modulation contrast microscope: Principles and performance. J Microsc 1977;110:205–222.
26. Mölder A, Czanner S, Costen N, Hartshorne G. Automatic detection of embryo location using rotational variance of the Hough Transform. 22nd International Conference of Pattern Recognition ICPR 2014, Stockholm, Sweden.
27. Chavez SL, Loewke KE, Han J, Moussavi F, Colls P, Munne S, Behr B, Reijo Pera RA. Dynamic blastomere behaviour reflects human embryo ploidy by the four-cell stage. Nat Commun 2012;3:1251.
28. Diamond MP, Willman S, Chenette P, Cedars MI. The clinical need for a method of identification of embryos destined to become a blastocyst in assisted reproductive technology cycles. J Assist Reprod Genet 2012;29:391–396.

VI

An efficient multiclass SVM for the classification of embryo time-lapse imaging

Leida Mölder, A.*, Hartshorne, G.M.**, Costen, N.*, Czanner, S.*

* School of Computing, Mathematics and Digital Technology, Faculty of Science and Engineering, Manchester Metropolitan University, Manchester, UK

**Division of Reproductive Medicine, Warwick Medical School, University of Warwick, and Centre for Reproductive Medicine, University Hospitals Coventry and Warwickshire NHS Trust, Coventry, UK.

Abstract

Novel technique for imaging human embryo growth continuously during IVF treatment, has brought on the possibility to monitor dynamic traits during *in vitro* culture. The timing of mitotic divisions is one such trait with importance for embryo health, the understanding of individual cell development, cell tracking and cell lineage computation. One way to detect timing of divisions is to classify images in each image sequence according to number of cells. A challenging computational task, especially when the sequences used are captured using non-invasive imaging techniques. Difficulties include variable light conditions, sample contamination, high sample variability, lack of contrast enhancing agents and sometimes overlapping objects in the sample. We pose a method for feature extraction using spatial filtering which allows us to use only 12 basic features for classification, yet with retained classification accuracy of 89.4% up to the 4 cell stage and 74.9% up to the 8 cell stage. We also introduce a variant of a multiclass SVM, giving classification performance comparable to the best compared alternative method, but significantly faster, by up to 73% for higher class spaces.

Keywords

Embryo Classification, Multiclass SVM, Medical Image Analysis, Embryology

Anna Leida Mölder
School of Computing, Mathematics and Digital Technology
Manchester Metropolitan University
John Dalton Building
Chester Street
Manchester M1 5GD
UK

mail@annaleida.com

Introduction

Human embryos are hard to study because of ethical and technical constraints. Especially the early stages of development are of high interest both clinically and for basic research [1,2,3,4,5]. Several laboratory set-ups exist and are routinely used in IVF clinics, which allows non-destructive time-lapse imaging of embryos in the early stages of growth. As a consequence, large amounts of image data is being gathered in clinics around the world, providing potentially new research material, but also an analytical challenge. Manual evaluation of images is time consuming and prone to errors and observer variability [6,7]. Automated computer aided image analysis has the potential to overcome both the workload bottleneck and to provide a standardized method of evaluation, but the images are of highly varying quality and captured under a multitude of conditions. On top of this, the sample material itself consists of living tissue with a high degree of morphological variation and the lack of fluorescent markers or any kind of contrast enhancing agents makes the task even more difficult.

An important endpoint for embryo studies is the timing of mitotic events, which has been shown to correlate to embryo health [8,3,9,10]. Embryo classification has been previously attempted using a variety of techniques. These can roughly be divided in segmentation based (requiring an identification of embryonic cell outlines) and segmentation free [11,12,13,14], or a combination of those [15]. Usually, a correctly performed segmentation [16,13,17,18] provides the most detailed information on cell position, shape and outline, but the segmentation process can itself be prone to errors, especially when used under clinical circumstances where complete and accurate segmentation may not always be possible. Here we propose a method for supervised learning and compare the performance of four different methods on a set of embryos imaged by Hoffman Modulation Contrast (HMC) imaging [19]. For each of these four methods we compare different methods to extract feature data from the images, using features from two different layers of the embryo; the complete embryo and the inner section, and finally we combine features from both

regions. We have previously illustrated this technique in [20] where a single feature was used as a cue to embryo development, but not enough for complete classification.

The purpose of this paper is two-fold. First we describe a process for image filtering and feature extraction, tailored to this specific kind of embryo images. Second we evaluate a set of simple and robust supervised machine learning techniques for the purpose of segmentation-free embryo classification. We test the results on a set of 18 embryos, and propose a tailored variant of a Support Vector Machine (SVM) as being superior in this context. Our method utilizes the fact that for the case of embryos, we have prior knowledge of the hierarchy of the classification problem. This method is not restricted to embryo images, but is applicable to any classification scheme where this is true.

Materials and Methods

Training and test set

Time lapse image series of human embryos fertilized *in vitro* were acquired as anonymized sequences donated to research with ethical approval from Coventry Research Ethics Committee (04/Q2802/26) and the Human Fertilisation and Embryology Authority (R0155). Embryos were cultured in 25µl culture media (Origio, Redhill, UK) under mineral oil for up to 6 days, incubated at 37°C in an atmosphere of 5%CO_2, 5%O_2 and 90%N_2. The images were captured using the Embryoscope® system (Unisense Fertilitech, Copenhagen, Denmark), with up to 7 focal depth planes, 15-25 µm apart, recorded at 20 minute intervals using a HMC optical set up [19] and a 635nm LED as light source. Fresh medium was supplied at intervals, but embryos were otherwise undisturbed during imaging. The total dataset consisted of 620 8-bit grey scale images in series of 18 embryos. Images from the test embryos where manually classified in groups of 1-8 cells for comparison. Table 1 provides an overview of images for each class. A training set was constructed using 10% and 20% of the total image set, respectively. The rest of the images (the test set) was used for validation.

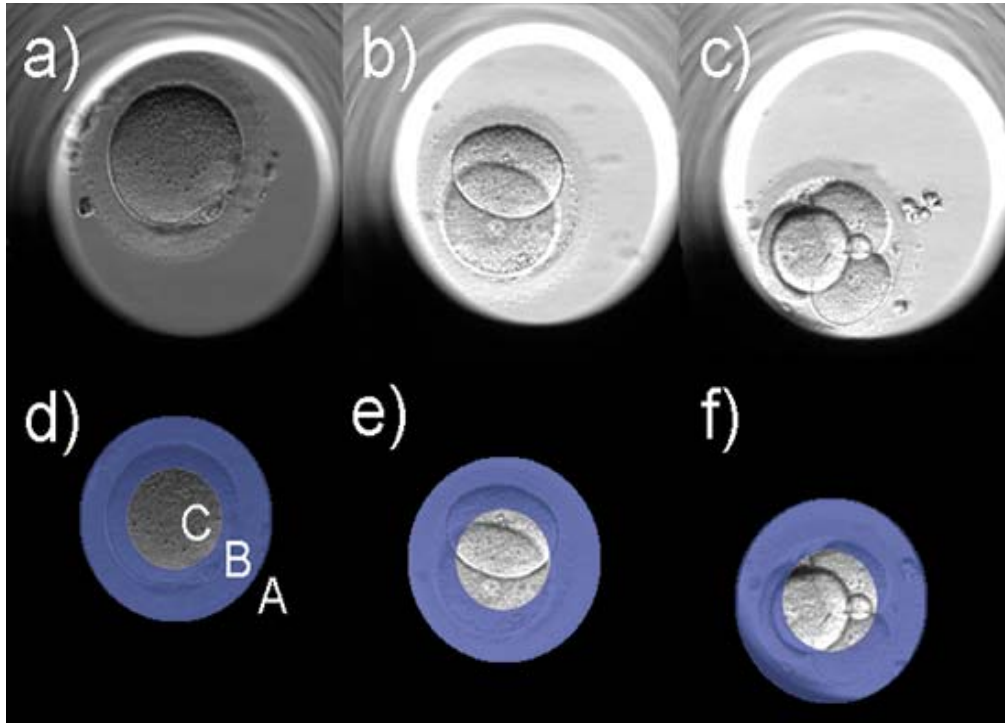| Testset	1 cell	2 cell	3 cell	4 cell	5 cell	6 cell	7 cell	8 cell
Nbr of images	130	131	44	89	56	86	14	70

Figure 1: Example of test images with a) 1, b) 2 and c) 4 cells per embryo. All images are from separate embryos, illustrating differences in lighting conditions. d-f) Illustration of image feature extraction of the example images a-c). Region A: Excluded information, region B: Outer layer, region C: Inner embryo.

Feature extraction

12 features were used for classification. All were standard grey level image traits; grey level mean, variance, maximum, minimum, root mean square, kurtosis, skewness, energy, entropy, contrast, correlation and homogeneity. A mathematical definition of all features is provided in Appendix A. All features were extracted for two separate regions of interest (ROI) and combined to a complete feature vector comprising 24 features. To select a ROI, and to spatially filter out irrelevant image data, the embryo region was automatically detected using a Hough Transform [21]. The removal of the background (Figure 1d-f: Region A) was necessary to get enough sensitivity in extracted grey scale features, and also to help eliminate noise from sample contamination. Feature extraction was then performed on the entire embryo (Figure 1d-f: Region B) and on the inner embryo (Figure 1d-f: Region C only). The perimeter for region C is at half the radius of region B. Intuitively, from Figure 1d it is clear that this spatial filtering will have a strong effect on features since all cell membranes are eliminated for region C. The optimal choice for region selection was investigated previously in [21].

Classification methods

A supervised machine learning classification is a computer algorithm which learns by example to assign labels to objects. For example it can learn to recognize handwritten digits by examining

a large collection of scanned images of digits, and has also been successfully applied to a wide variety of biological applications. In this report we used two standard supervised machine learning approaches for comparison, and also a simple Naïve Bayes classifier with a Gaussian kernel [22], where a prior probability was estimated from the relative frequencies of the classes in the training set. We also used a random forest with a binary decision tree where the optimization criterion was given by Gini's diversity index [23]. Last, we implemented two variations of an SVM [24]. An SVM is a general term for the method of separating two clusters of data by means of a plane, placed to maximize the gap between the members of each cluster closest to the border (the support vectors). Originally, the SVM was designed to handle only 2-class problems, but a number of extensions to multiple classes exist. One of the most common is the one-versus-all approach, where N classifiers are used to separate the same sample set, the final choice of class being the one with a highest confidence score. In our case, this method has a number of disadvantages. Each classifier must produce a real-valued confidence score, and its scale may vary between classifiers. Also, each classifier will see an imbalanced distribution, since the group of positives will be smaller than the group of negatives (consisting of all other classes). In the case of mitotic division, this problem is emphasized, since the distribution can be expected to be skewed in the entire sample set – the more cells in the embryo, the shorter the time between mitotic divisions, leading to less images to use as samples. Another approach is to perform N(N-1)/2 classifications and by vote deciding on the class identified by the majority of classifiers. This method has the disadvantage that the same number of votes may be cast for several classes, leading to decision ambiguity. In this study we implement the Directional Acyclic Graph SVM (DAGSVM) [25]. The DAGSVM improves on these simple multiple class SVMs in several ways, such as not having to report confidence scores for each one-versus-one classification, and avoiding the ambiguity which may arise with a voting scheme. The algorithm can be intuitively understood as a stepwise exclusion, where the entire test set is successively divided into N classes by performing a pair-wise one-versus-one classification using a total of N(N-1)/2 classifiers. Classification progresses

step-wise until only one class remains. For example, consider a 4-class problem, with classes A,B,C and D. For the first classification class A versus class B, we conclude that all samples classified as belonging to class B, does not necessarily belong to class B, but we can in any case be sure they do not belong to class A. For the sake of argument, call this sample cluster "not-A". In the next step, "not-A" then compares class B with class C. If classified as C, we conclude that this sample cluster is not necessarily C, but in any case, does not belong to B, so we may call this new cluster "not-A-nor-B". In a last one-versus-one classification we can then separate C from D.

Successive Slicing SVM

One disadvantage of DAGSVM is the performance reduction for higher number of N. In our case we are looking for a quick and efficient classification method, where we are prepared to trade accuracy for speed. The computational performance may be increased, if we can assume we have some prior knowledge of which sample clusters are more easily separated. If that is the case, we may pairwise slice the entire sample set by first removing the most obvious classes. We refer to this method as a successive slicing SVM (SS-SVM). The SS-SVM performs a set of a cascading one-versus-all SVM classification, similar to what has been described in [26,27]. Here the test set is successively divided in smaller and smaller sets, until only one class remains for each sample. The choice of classifier and thus order of classification is decided by some general clustering method. In our case we know that the optimal order of classification is from lowest number of cells to highest cell number, so we can further improve on performance by avoiding the clustering step. Simply put, the difference between an image containing one and two cells will be larger than the difference between an image containing three and four cells, so the cluster distance can be expected to decrease for higher number of cells in the image. Using this method, we need only N classifiers of type one-versus-all, one for each class, as opposed to N(N-1)/2 for the DAGSVM. We then iterate through the test set, first separating all samples with one cell from the rest. These samples are then removed from the test set, and the classification is progressed by separating

the samples with two cells from the remaining test set, until we reach the class with the maximum number of cells.

Statistics and metrics of results

Results are given as the mean value of ten repeated identical experiments, with the training set selected randomly from the complete image set each time. 95% confidence intervals are given, calculated using a 2-tailed Student's t-test, unless otherwise stated [28]. Experiments were performed in three sets, with image sets containing 1-4, 1-6 and 1-8 cells, respectively. We repeated each classification 10 times and compare the accuracy, given by

$$A = \frac{\sum_{n=1}^{N} C_{nn}}{\sum_{n=1}^{N}\sum_{m=1}^{N} C_{nm}},$$

(1)

where C_{nm} is the *(n,m)th* element in the confusion matrix of size *NxN,* with columns representing prediction, and rows representing manually classified values. The predictive performance was also measured using the sensitivity of class *k*

$$Se(k) = \frac{C_{kk}}{\sum_{n=1}^{N} C_{nm}},$$

(2)

specificity

$$Sp(k) = 1 - \frac{\sum_{m=1}^{N} C_{im}(m \neq k)}{\sum_{n=1}^{N}\sum_{m=1}^{N} C_{nm}(n \neq k)},$$

(3)

(where 1-specificity is referred to as the fallout), positive predictive value

$$Pp(k) = \frac{C_{kk}}{\sum_{m=1}^{N} C_{km}},$$

(4)

and negative predictive value

$$Np(k) = \frac{\sum_{n=1}^{N}\sum_{m=1}^{N} C_{nm}(n \neq k, m \neq k)}{\sum_{n=1}^{N}\sum_{m=1}^{N} C_{nm}(m \neq k)}.$$

(5)

Results

Accuracy

The accuracy for the SS-SVM was slightly higher than that of the DAGSVM, with 89.4% for the 1-4 cell case, 80.8% for the 1-6 cell case, and 74.9% for the 1-8 cell case (Figure 2). The complete result set is given by Table B.1-3 in Appendix B. For some values, the DAGSVM outperformed the SS-SVM, and for a few isolated values, they were both outperformed by the Random Forest, but overall the predictive performance of the SS-SVM was comparable to the DAGSVM.
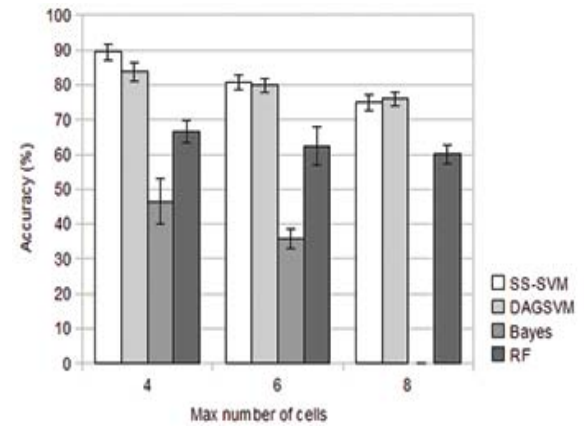

Figure 2: Accuracy as a function of maximum number of cells in the classification for all four classification methods. Mean values and confidence intervals are calculated from 10 classifications. The Bayes classifier failed to classify in the 1-8 cell case.

The best case positive predictive power of the SS-SVM is given in Table 2. The values for the classes 3, 5 and 7 cells were significantly lower than the rest, reflecting the scarcity of images in the test set from these classes. Figure 3-4 illustrates the predictive performance of the SS-SVM alone. In Figure 3, the feature set was split in two groups, one set from each ROI. The accuracy of each classification was compared to the classification when both ROI are combined. The result was improved by up to 45% by

combining features from both ROI, compared to using features from one ROI only. In Figure 4, the Receiver Operator Characteristic (ROC) is plotted for each class (representing a cell number) and for each classification of image series comprising 1-4, 1-6 and 1-8 cells, respectively. The mean sensitivity for the classes in the 1-4 cell set was 82.2%, slightly higher than the mean sensitivity of 80.0% for the 1-6 cell set and 79.9% for the 1-8 cell set. It is clear from Figure 4, that the spread of values was larger for the 1-8 cell set. The mean of the fallout was 4.75% for the 1-4 cell case, 4.43% for the 1-6 cell case and 5.34% for the 1-8 cell case.

Table 2: Positive predictive value (PPV) and negative predictive value (NPV) for each class of cells in classifications comprising images with 1-4, 1-6 and 1-8 cells, respectively.

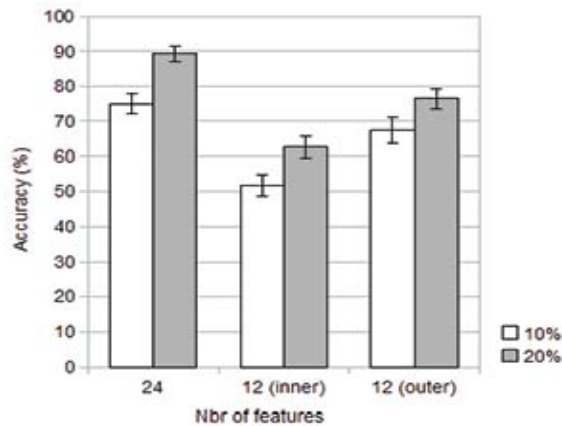| Nbr of cells:	1-4 cells				1-6 cells						1-8 cells							
	1	2	3	4	1	2	3	4	5	6	1	2	3	4	5	6	7	8
PPV	89	91	72	81	86	92	55	72	79	83	82	86	46	65	80	77	56	76
NPV	96	95	96	95	99	96	96	95	96	95	98	97	98	95	96	94	98	95

Figure 3: Accuracy as a function of number of features for the SS-SVM. The total feature set was 24 features, a combination of 12 features from the inner embryo and 12 features from the entire embryo (referred to as "outer"). Calculation was performed twice, with the training set selected as 10% and 20% of the total image set, respectively. The classification was performed for 1-4 cells. Mean values and confidence intervals are calculated from 10 classifications.
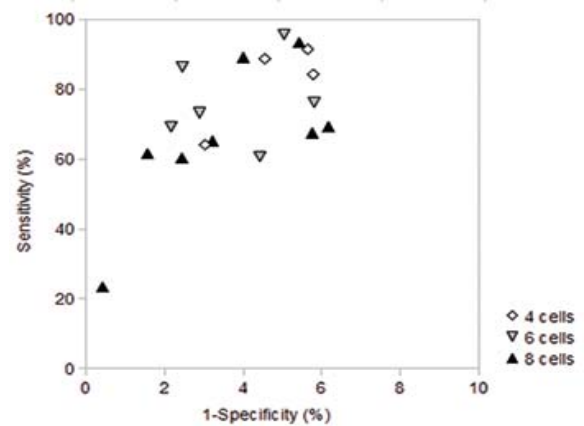

Figure 4: ROC of SS-SVM. Each data point represents one class in the image sets 1-4 cells (diamond), 1-6 cells (triangle down) and 1-8 cells (triangle up).

Performance

The performance was evaluated using a standard 4 core PC with 8 GB RAM, with the code implemented in Matlab® 7.12.0.635 (R2011a). Figure 5 illustrates performance results. The computation time for one classification (including the separation of the image set in training and test set) is plotted for each calculation comprising 1-4, 1-6 and 1-8 cells, respectively. For a small class set (1-4 cells) the DAGSVM was comparable to the SS-SVM, but for an increased number of classes, the SS-SVM outperformed the DAGSVM by 63% (1-6 cells) and

73% (1-8 cells). The overall best performance was for the Naïve Bayes classifier, outperforming the SS-SVM by 86% (1-4 cells) and 88% (1-6 cells). For the 1-8 cell case, no results for the Naïve Bayes are shown, because the Gaussian kernel failed to estimate the parameters.
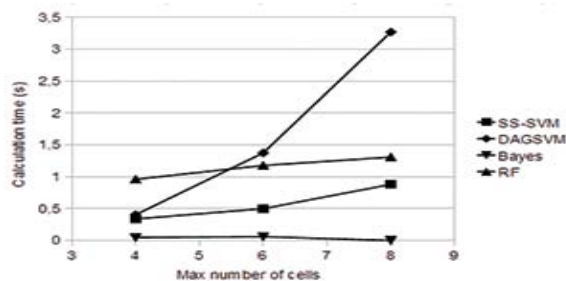

Figure 5: Calculation time as a function of maximum number of cells for classification of the entire test set. Methods used were SS-SVM (square), DAGSVM (diamond), Bayes (triangle down) and Random Forest (triangle up).

Discussion

We have demonstrated the usefulness of spatial image filtering in embryo image classification, showing an improvement in classification accuracy of up to 45% by combining features from two embryo regions. We further describe the implementation of a new supervised learning method, a successive slicing SVM, for fast and reliable embryo classification from a non-invasive imaging technique with no need for image segmentation. Using this method, we demonstrate a successful classification of embryo images up to the 4 cell stage with an accuracy of 89.4% and up to the 8 cell stage with an accuracy of 74.9%. To our knowledge, this is the highest cell number in embryos ever reported to be classified in the literature. Our method is shown to be faster than the fastest and most accurate method of comparison for classification of 8 classes by 73%.

For the purpose of this study it has been enough to illustrate this method using a limited image range, but the test set for this study has been too small to evaluate how well this method would perform with a more diverse sample set from a variety of clinics with different conditions for imaging and embryo growth. For future work, we intend to expand this study to a larger data set.

References

1. Azzarello A, Hoest T, Mikkelsen AL. The impact of pronuclei morphology and dynamicity on live birth outcome after time-lapse culture. Hum Reprod 2012;27:2649–2657.
2. Basile N, Meseguer M. Time-lapse technology: evaluation of embryo quality and new markers for embryo selection. Expert Rev Obstet Gynecol 2012;7:175–190.
3. Cruz M, Garrido N, Herrero J, Pérez-Cano I, Muñoz M, Meseguer M. Timing of cell division in human cleavage-stage embryos is linked with blastocyst formation and quality. Reprod Biomed Online 2012;25:371–381.
4. Kirkegaard K, Agerholm IE, Ingerslev HJ. Time-lapse monitoring as a tool for clinical embryo assessment. Hum Reprod 2012;27:1277–1285.
5. Kirkegaard K, Kesmodel US, Hindkjær JJ, Ingerslev HJ. Time-lapse parameters as predictors of blastocyst development and pregnancy outcome in embryos from good prognosis patients: a prospective cohort study. Hum Reprod 2013;28:2643–2651.
6. Paternot G, Wetzels AM, Thonon F, Vansteenbrugge A, Willemen D, Devroe J, Debrock S, D'Hooghe TM, Spiessens C. Intra- and interobserver analysis in the morphological assessment of early stage embryos during an IVF procedure: a multicentre study. Reprod Biol Endocrinol 2011;9:127.
7. Sundvall L, Ingerslev HJ, Knudsen UB, Kirkegaard K. Inter- and intra-observer variability of time-lapse annotations. Hum Reprod 2013;2013:3215–3221.
8. Chen AA, Tan L, Suraj V, Reijo Pera R, Shen S. Biomarkers identified with time-lapse imaging: discovery, validation, and practical application. Fertil Steril 2013;99:1035–1043.
9. Hlinka D, Kaľatová B, Uhrinová I, Dolinská S, Rutarová J, Rezáčová J, Lazarovská S, Dudáš M. Time-lapse cleavage rating predicts human embryo viability. Physiol Res Acad Sci Bohemoslov 2012;61:513–525.
10. Meseguer M, Herrero J, Tejera A, Hilligsøe KM, Ramsing NB, Remohí J. The use of morphokinetics as a predictor of embryo implantation. Hum Reprod Oxf Engl 2011;26:2658–2671.
11. Beuchat A, Thévenaz P, Unser M, Ebner T, Senn A, Urner F, Germond M, Sorzano COS. Quantitative morphometrical characterization of human pronuclear zygotes. Hum Reprod Oxf Engl 2008;23:1983–1992.
12. Filho ES, Noble J., Wells D. A Review on Automatic Analysis of Human Embryo Microscope Images. Open Biomed Eng J 2010;4:170–177.
13. Giusti A, Corani G, Gambardella LM, Magli MC, Gianaroli L. Blastomere segmentation and 3D morphology measurements of early embryos from Hoffman Modulation Contrast image stacks. 2010 IEEE Int Symp Biomed Imaging Nano Macro 2010:1261–1264.
14. Wong CC, Loewke KE, Bossert NL, Behr B, Jonge CJ De, Baer TM, Pera RAR. Non-invasive

imaging of human embryos before embryonic genome activation predicts development to the blastocyst stage. Nat Biotechnol 2010;28:1115–1121.

15. Moussavi F, Wang Y, Lorenzen P, Oakley J, Russakoff D, Gould S. A Unified Graphical Models Framework for Automated Mitosis Detection in Human Embryos. IEEE Trans Med Imaging 2014;33:1551–1562.

16. Agerholm IE, Hnida C, Cruger DG, Berg C, Bruun-Petersen G, Kølvraa S, Ziebe S. Nuclei size in relation to nuclear status and aneuploidy rate for 13 chromosomes in donated four cells embryos. J Assist Reprod Genet 2008;25:95–102.

17. Morales DA, Bengoetxea E, Larranaga P. Automatic segmentation of zona pellucida in human embryo images applying an active contour model. Proc 12th Annu Conf Med Image Underst Anal 2008:209–213.

18. Filho ES, Noble JA, Poli M, Griffiths T, Emerson G, Wells D. A method for semi-automatic grading of human blastocyst microscope images. Hum Reprod Oxf Engl 2012;27:2641–2648.

19. Hoffman R. The modulation contrast microscope: principles and performance. J Microsc 1977;110:205–222.

20. Mölder A, Drury S, Costen N, Hartshorne GM, Czanner S. Semiautomated analysis of embryoscope images: Using localized variance of image intensity to detect embryo developmental stages. Cytometry A 2015;87:119–128.

21. Mölder A, Czanner S, Costen N, Hartshorne G. Automatic detection of circular structures in human embryo imaging using trigonometric rotation of the Hough Transform. 2014 22nd Int Conf Pattern Recognit ICPR 2014;22:3239-3244.

22. Bishop C. Pattern Recognition and Machine Learning. Springer-Verlag New York; 2006.

23. Prince SJD. Computer Vision: Models, Learning, and Inference. Cambridge University Press: New York; 2012.

24. Cristianini N, Shawe-Taylor J. An Introduction to Support Vector Machines: And Other Kernel-based Learning Methods. Cambridge University Press: New York, NY, USA; 2000.

25. Platt JC, Shawe-Taylor J, Cristianini N. Large Margin DAG's for Multiclass Classification. Proc. Advances in Neural Information Processing Systems 1999;12:547-553.

26. Tibshirani R, Hastie T. Margin Trees for High-dimensional Classification. J Mach Learn Res 2007;8:637–652.

27. Vural V, Dy JG. A Hierarchical Method for Multi-class Support Vector Machines. Proc Twenty-First Int Conf Mach Learn 2004;21:105–112.

28. Lange K. Numerical Analysis for Statisticians. 2nd ed. Springer Verlag: New York; 2010.

VII

Supervised classification of etoposide-treated in vitro adherent cells based on non-invasive imaging growth characteristics

Anna Leida Mölder[a,*]**, Johan Persson**[b]**, Zahra El-Schich**[b]**, Silvester Czanner**[a]**, Anette Gjörloff Wingren**[b]

[a]School of Computing, Mathematics and Digital Technology, Faculty of Science and Engineering, Manchester Metropolitan University, Manchester, UK
[b]Department of Biomedical Science, Health and Society, Malmö University, Malmö, Sweden

Abstract. Single cell studies using non-invasive imaging is a challenging, yet appealing way to study cellular characteristics over extended periods of time. In some cases, *e. g.* transplantation culturing, real-time cellular monitoring, stem cell studies, *in vivo* studies and embryo growth studies, it is also crucial to keep the sample intact and invasive imaging using fluorophores or dyes is not an option. From the perspective of digital image analysis, non-invasive imaging poses several challenges, such as low contrast, high noise levels and weak image gradients, often impeding the outcome of the computer analysis.

In this study, adherent cell cultures of DU-145 were treated with low concentration etoposide and imaged during three days. Single cells were identified by image segmentation and subsequently classified on morphological image features, extracted for each cell. In parallel with image analysis, an MTS assay was performed to allow comparison between metabolic activity and morphological changes after long term drug response.

Results show a decrease in proliferation rate for low concentration etoposide, accompanied with changes in cell morphology, primarily leading to an increase in cell area and textural changes. It is shown that changes detected by image analysis are visible already day 1 for 0.25 micromolar etoposide, whereas effects on MTS are detected only day 3 for higher etoposide concentrations, leading to the conclusion that the morphological changes observed are not directly connected to reduction in cell metabolic activity or viability. Three classifiers are compared and we report a best case positive predictive power of 97.3 per cent for classification of cells as treated/untreated.

Keywords: single cell studies, non-invasive microscopy, low-concentration etoposide, digital image cytometry.

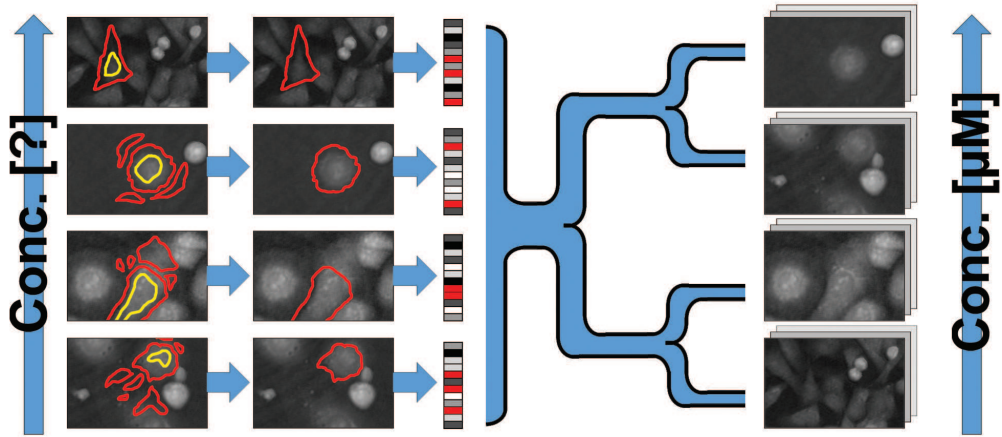*Anna Leida Mölder, mail@annaleida.com

Fig 1 Algorithmic overview of the classification of cells exposed to a treatment, based on their imaged morphological characteristics. Images are segmented and filtered, features extracted and used to train a supervised classifier.

1 Introduction

Etoposide is a cytotoxic chemotherapy drug, commonly used as an anti-cancer agent, causing DNA damage and eventually triggering cell death by apoptosis at high concentrations. The process at sub-apoptotic concentrations and during long term exposure is poorly understood, but reports of effects include drug resistance and risk of secondary leukemia. To study the effects of long term low level exposure, it is necessary to use an imaging set up where the same cells can be monitored over time.

When treating an organism with a substance *in vivo*, cells will typically be exposed to a variety of concentrations, depending on their location within the tissue and the body. Cells will also react differently when exposed to the same concentration depending on their cell type, tissue type, and sometimes stage in the cell cycle. Most assays in use today study the average characteristics of large groups of cells, but under some conditions, individual cell studies as well as bulk studies could be beneficial.

Single cells can be studied using time-lapse imaging, with or without fluorescent markers or dyes or using non-imaging marker-based techniques such as flow cytometry. The use of markers in cell imaging has several disadvantages. It is indirect, invasive, and also makes the cells unusable for further tests. In some cases this is not preferable, such as in studies of tissue or cells for transplantation, the study of embryonic cells and stem cells, or when the sample needs to be kept in continuous viable condition for a longer study. Also, when performing time-lapse studies over several days, the need to adjust focus and supervise the imaging equipment requires a time consuming surveillance by laboratory personnel.

This study was motivated by the desire to be able to detect the treatment concentration of a toxin by means of non-destructive image analysis of adherent cells *in situ* for long periods of time, on a cell-by-cell basis by means of morphological cues (Fig. 1). We use a combination of Digital Holographic Microscopy (DHM) and image analysis to profile morphological changes in a common *in vitro* cell type, which we induce by exposure to low concentration (LC) etoposide, and model them as a set of feature parameters. Principal component analysis (PCA) and clustering is used to determine the most prominent set of features. We then use the model to classify the cells

to the known treatment concentrations and in the end we compare the result to that of a standard bulk assay, the MTS, for reference. An essential step of the analysis is segmentation of the images into sub-regions, each describing a single cell. The result of segmentation is in turn affected by a threshold describing the cell border. We modify a state-of-the-art segmentation procedure by a spatial filtering using two thresholds, and compare the effect on the resulting classification.

The classification is made on a cell-by-cell basis with cells from cultures exposed to different concentrations of etoposide. Cells are also combined culture-by-culture in order to make it possible to compare the results with an established technique (MTS assay), where the evaluation is made on a culture-by-culture basis. One application for such a method is the study and classification of cells continuously on a cell-by-cell basis, even within the same culture. This possibility opens up a whole new perspective on cellular research, where cells are viewed not as a uniform mass, but rather as a broad spectrum of individual bodies, all contributing in its own way to the characteristics of the tissue.

1.1 Cell culture and treatment

In our investigation, we used prepared DU-145 cells from cultures.[1,2] When cells are dying, regardless of mechanism, their morphology changes.[3,4] Etoposide has a long history as an anti-cancer agent[5] and has been reported to cause cell cycle arrest, which has been tested previously using DHM.[6,7] It functions by interaction with the enzyme Topoisomerase II, causing breaks in the DNA-strand, ultimately leading to apoptosis[8,9] and is usually administered intravenously or orally in capsule form. It is one of the most widely used chemotherapeutic agents for several types of cancer,[10–12] but its effect is limited by toxicity.[10] It is shown[13] that both the α and β type of Topoisomerase II is targeted by etoposide, but only at concentrations above or equal to 1 μM, although some growth inhibition was observed at lower concentrations. Bleibel *et. al.*[14] reported half maximal inhibitory concentrations (IC_50) slightly below 1 μM, suggesting that effects are present also at low concentrations of etoposide. These conclusions are confirmed by Liu *et. al.*,[15] who also found that a lower limit for etoposide induced proliferation and viability reduction occurs at approximately 0.02 μM, and also shows effects on the cell cycle at concentrations ≤ 1 μM (although for different cell lines). Even lower concentrations were reported to cause changes in phenotype, and potentially also cause drug resistance or cause secondary leukemia.[16]

1.2 Digital Holography

Dennis Gabor was first to invent a way to encode the phase of the light in a single recording, *i.e.* the hologram.[17] The technique was later used as base for the development of digital holography.[18–21] Some years later, the first DHM images showing living cells were published.[22]

Digital holography is a full-field imaging technique where pixel values are directly related to cellular dry mass, and since the images from a mathematical point of view are similar to those from a standard epifluorescence microscope, the same image analyzing algorithms can often be used. In cellular biology, DHM has been used for a number of applications, including *in vivo* studies,[23] sub-cellular motion within living tissue,[24] migration studies[25–28] and studies of morphological changes,[29–32] proliferation[33] and apoptosis.[34] Fig.2 shows an example of typical morphological changes in DU-145 cells treated with etoposide, compared to control.
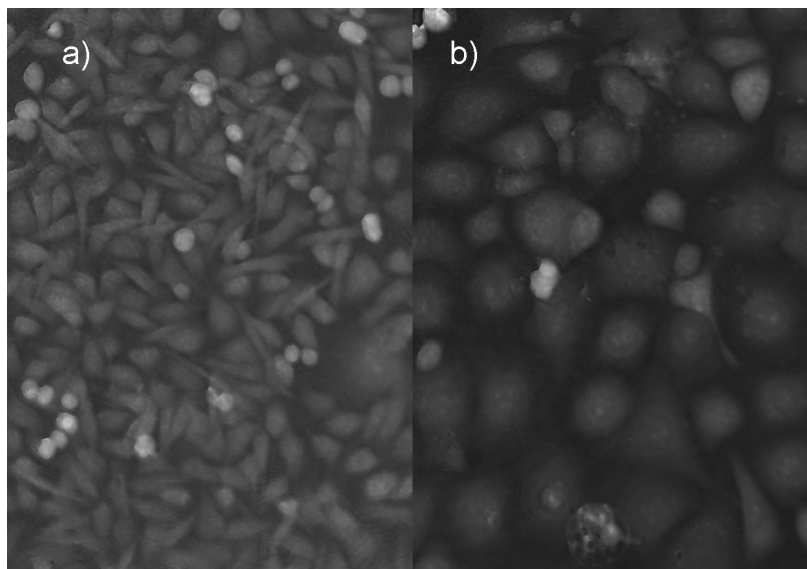
Fig 2 Example of DU-145, a) control and b) 72h after treatment with 0.5 μM etoposide, imaged with digital holography. Image intensity is related to an optical phase shift, which is in turn related to cellular dry mass.

2 Materials & Methods

2.1 Cell culture

DU-145 cells (ATCC®HTB-81, ATCC LGC Standards, Teddington, UK) were grown in 75 cm^2 flasks in Dulbeccos Modified Eagle Medium (DMEM, Invitrogen, Carlsbad, CA, USA) cell culture media with an added 10% fetal calf serum. Cells were sub-cultivated twice per week to keep the culture non-confluent.

2.1.1 DHM image capture

For DHM tests, $2 \cdot 10^4$ cells per well were seeded in 6 well plates (day -1). After 24 h incubation to allow the cells to attach, fresh media with the desired concentration etoposide (Sigma-Aldrich Co., St. Louis, USA), solved in DMSO (Sigma-Aldrich) was added (day 0). Images of cells were captured day 1-3. The experiment was repeated for etoposide concentrations 0.25 μM, 0.5 μM, 1 μM and 5 μM. Since DMSO was used to dissolve etoposide, two controls were used; untreated cells and cells in only DMSO. DHM images were captured using Holomonitor M4 (Phase Holographic Imaging AB, Lund), using a 635 nm 0.2 mW/cm^2 laser as light source. The 6 well plate was taken out of the incubator, the lid removed and 15 images were captured of each well, after which the lid was replaced and the plate was returned to the incubator. All DHM images are 1024x1024 pixels, at a scale of 0.51 μm/pixel.

2.1.2 MTS Assay

Cells were seeded day -1 into a 96-well plate at a density of $5 \cdot 10^3$ cells per well in 100 μl of cell culture medium. Six wells per test condition were used. All unused wells were filled with media including a blank column. After incubating for 24 h to allow the cells to attach, the media was substituted for new media, with the desired test concentration (day 0). After 24, 48, and 72 h the plate was then analyzed using the MTS assay. Etoposide was tested at 0.25 μM, 0.5 μM, 1 μM
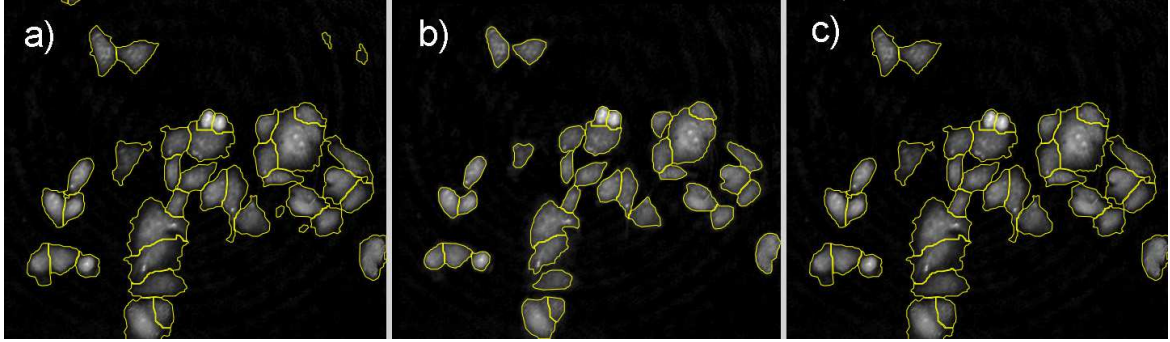
Fig 3 Cell filtering using double thresholding. The same image (control day 1) segmented using a) Otsu thresholding, b) Minimum error thresholding and c) The result of Minimum error thresholding, filtered using the Otsu thresholding to identify cells.

and 5 μM. As the etoposide was solved in DMSO, a column with only DMSO equal to the amount in the wells with the most DMSO (*i. e.* the wells with the highest concentration of etoposide) was used as secondary control. Twenty micro-liters of MTS (Promega Corporation, Madison, USA) was added to each well and carefully mixed by tapping the side of the plate. After incubating for 1 and then again after 2 hours, the plate was read by using BIO-TEK micro plate reader at 490 nm.

2.2 Image analysis

To identify single cells as regions of interest (ROI), images were segmented with the microscope software HoloStudio^{TM} (Phase Holographic Imaging AB, Lund, Sweden), using a seeded watershed segmentation.[35] Segmented objects smaller than 20 pixels were considered noise and discarded. A crucial step of the segmentation is the threshold calculation defining cell borders. Otsu thresholding (OT)[36] yielded a higher cut-off than Minimum Error Thresholding (MET),[37] resulting in a more accurate pin-pointing of cells as out-of-focus debris in the growth medium was not selected. On the other hand, MET gave a more accurate cell outline (Fig.3). In order to maintain the accuracy of the border segmentation given in MET, while at the same time restrict the selection to actual cells (and not sample contamination or image artifacts), we also introduced a pre-processing step, where we use both segmentations in combination. The location of cells given by the segmentation using OT was used to filter the non-cellular bodies from the segmentation using MET, thus gaining the advantages of both methods.

We then proceeded to extract a set of grey level image features to use as input for a random forest classification. The purpose of this was to reduce the information in the image (given as pixel by pixel), to only relevant variables, $v = [v_1, v_2, , v_m, ..., v_{M-1}, v_M]$, where M is the number of relevant variables gathered. Here, the term "relevant variable" is a matter of choice depending on the experiment. We first collected a set of 22 standard grey scale variables, which we then characterized using several unsupervised learning techniques (hierarchic clustering, principal component analysis) and cross correlated them with the known result. The variables which showed no morphological change under treatment were excluded from the supervised classification in order to reduce the risk of over-fitting. A mathematical definition of the selected variables is given in Section 2.3.

2.3 Image features

Define the original image $F(x, y)$, with dimensions X and Y respectively, as a 2-dimensional matrix of gray levels i, to a maximum gray level I. Let ψ be the subset of F belonging to the ROI of one cell selected in the segmentation (*i.e.* the area), with perimeter ρ.

Define the Gray Level Co-Occurence Matrix (GLCM) P of F as the number of times gray level i is oriented with respect to gray level j such that where $F(x, y) = i$ and $F(y) = j$ then $y = x + 1$, where $i, j \in [0, G]$, and $G < I$ is the number of gray level used in the GLCM. In our case, $G = 8$. The result is an intermediate matrix, whose values indicate how often one pixel-value in $F(x, y)$ tend to be similar to its adjacent values. We normalize the GLCM so that:

$$\sum_{i=1}^{G}\sum_{j=1}^{G}[P(i,j)] = 1, \tag{1}$$

in this way, each value in $P(i, j)$ will denote the probability of a range of gray levels to appear in each others proximity. We assume the cells to be randomly oriented, so we satisfy ourselves with computing only one GLCM along the horizontal (x) direction.

Further define μ_x, μ_y, σ_x, σ_y as the mean and standard deviations on image $F(x, y)$ in x- and y-direction, respectively.

$$\mu_x = \sum_{i=1}^{G} i \sum_{j=1}^{G} P(i,j), \tag{2a}$$

$$\mu_y = \sum_{i=1}^{G}\sum_{j=1}^{G} jP(i,j), \tag{2b}$$

$$\sigma_x{}^2 = \sum_{i=1}^{G}[i - \mu_x]^2 \sum_{j=1}^{G} P(i,j), \tag{2c}$$

$$\sigma_y{}^2 = \sum_{i=1}^{G}[j - \mu_y]^2 \sum_{j=1}^{G} P(i,j), \tag{2d}$$

For calculation of the roughness distribution $R(x, y)$, for each ψ we compute:

$$R = F(x, y) - \Omega[F(x, y)], \tag{3}$$

where Ω is a Gaussian smoothing filter, and $(x, y) \in \psi$.

The following features are used for classification:

Cell area:

$$A = \sum_{x}\sum_{y} 1, (x, y) \in \psi, \tag{4a}$$

Perimeter length:

$$P = \sum_x \sum_y 1, (x, y) \in \rho, \tag{4b}$$

Phaseshift integral:

$$\Phi_{tot} = \sum_x \sum_y F(x, y), \tag{4c}$$

Roughness kurtosis:

$$R_k = \frac{1}{A} \frac{\sum_x \sum_y R(x, y) - R_{av}^{\;\;4}}{\sigma^4}, \tag{4d}$$

Roughness skewness:

$$R_k = \frac{1}{A} \frac{\sum_x \sum_y R(x, y) - R_{av}^{\;\;3}}{\sigma^3}, \tag{4e}$$

Texture energy:

$$S_{energy} = \sum_{i=1}^{G} \sum_{j=1}^{G} [P(i, j)]^2, \tag{4f}$$

Texture entropy:

$$S_{entropy} = -\sum_{i=1}^{G} \sum_{j=1}^{G} P(i, j) \cdot log[P(i, j)], \tag{4g}$$

Texture contrast:

$$S_{con} = \sum_{i=1}^{G} \sum_{j=1}^{G} |(i - j)|^2 P(i, j), \tag{4h}$$

Texture correlation:

$$S_{corr} = \sum_{i=1}^{G} \sum_{j=1}^{G} P(i, j) \frac{(i - \mu_x)(j - \mu_y)}{\sigma_x \sigma_y}, \tag{4i}$$

Texture homogeneity:

$$S_{hom} = \sum_{i=1}^{G} \sum_{j=1}^{G} \frac{P(i, j)}{1 - |(i - j)|^2}, \tag{4j}$$

2.4 Data analysis

2.4.1 Principal Component Analysis and Clustering

Initially, an analysis of principal components was performed and a continuous modeling using Partial Least Squares on the 10 most prominent variables.[38] The PCA indicate that not all variables were significant for the classification, so v was cross-correlated with the known result and the correlation of each component v_m of variables was evaluated. The variables for the segmented cells which showed differences from control with a confidence of 95%, $P \leq 0.05$, were then used to construct a decision tree.
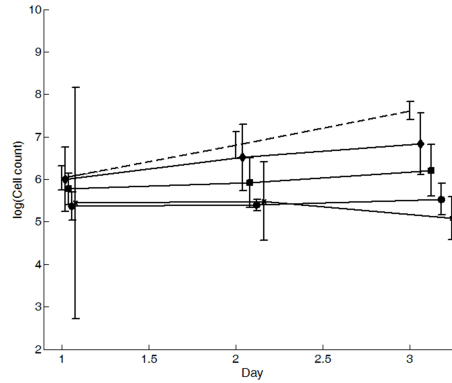
Fig 4 Proliferation curves for DU-145 under the effect of etoposide 0 μM (dashed), 0.25 μM (diamond), 0.5 μM (square), 0.1 μM (circle and , 5 μM (cross). Measurements were done at the same time and values for x-axis have been displaced slightly to display confidence intervals.

2.4.2 Classification

In order to classify single cells as treated or control, we used three different classifiers; a multiclass Support Vector Machine implementing the Directional Acyclic Graph,[40] a Naïve Bayes classifier and a classification tree using random forests, implementing the CART algorithm.[39] The optimization criterion was minimization of node impurity as defined by Gini's diversity index. In order to reduce tree complexity and avoid over-fitting, we request that the tree produced (for 5 bin classification) should have maximum 20, and minimum 9 nodes. The model cells (rows of $v(n)$) used to construct the tree were selected randomly as 10% of the total number of available N cells for that sample (the training set). The rest of the cells (the validation set) were used to test the accuracy of the result by cross correlation. The process of training and testing was repeated ten times, and the results averaged. The performance was evaluated using a standard 4 core PC with 8 GB RAM, with the code implemented in Matlab® 7.12.0.635 (R2011a).

2.5 Statistics

Errors are given as 95% confidence intervals, unless otherwise stated. To compute the P-values, a two-tailed t-test was performed. Experiment were repeated three times for concentration 0.25 μM, four times for 0.5 μM, twice for 1 μM and once for 5 μM. In 90 samples, a total of 66495 cells were imaged, averaging 738 cells per sample.

3 Results

3.0.1 MTS Correlation and Proliferation

From the daily segmentation of images, we obtain a count of cell proliferation (Fig.4). The cell growth for the control was close to exponential, and a reduction in cell growth rate was detectable for etoposide concentrations less than 5 μM, consistent with the MTS results (Fig.5). It is notable that the morphological changes were detectable already day 1 with DHM for all examined concentrations of etoposide, whereas changes in MTS absorbance are only detectable for the highest concentration of etoposide (5 μM) on day 3.
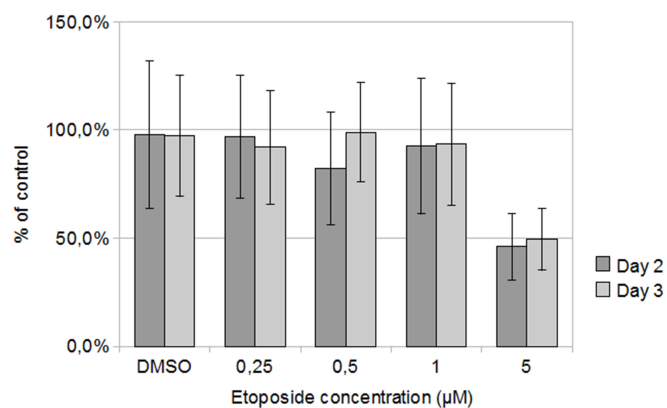
8

Fig 5 MTS absorbance as % of control day 2 and 3, as a function of etoposide concentration, after 2 hours of MTS incubation.
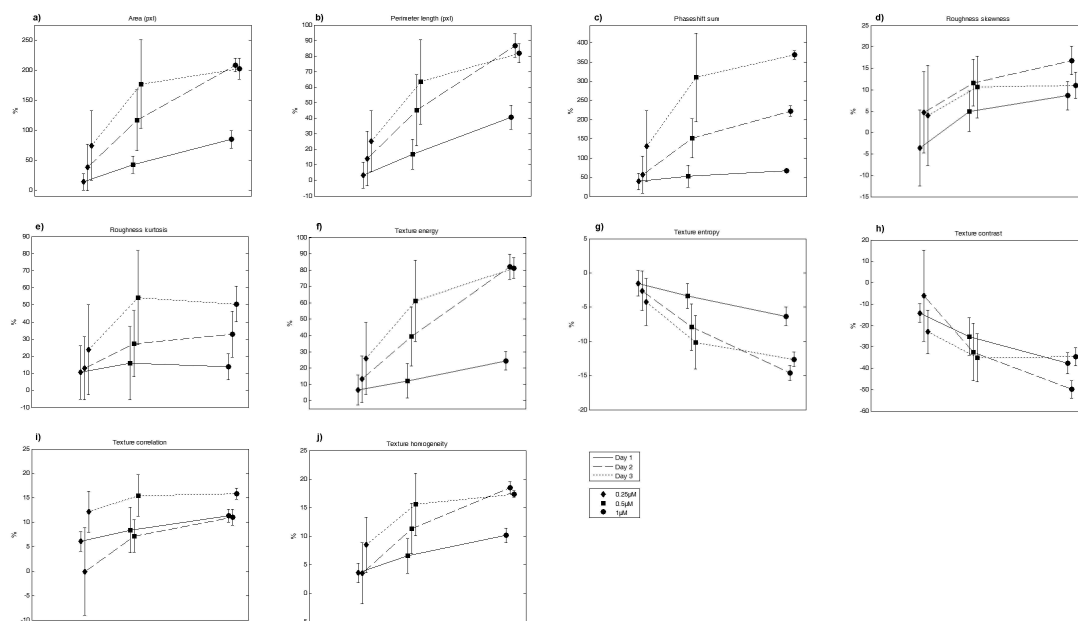

Fig 6 Variable values compared to control, calculated from sample-wise mean. Day 1 (solid), 2 (dashed) and 3 (dotted) after treatment. x-axis represents etoposide concentration; 0.25 μM (diamond), 0.5 μM (square) and 1 μM (circle), respectively. Error bars represent standard deviation, and for clarity, each series has been slightly offset in x-direction.
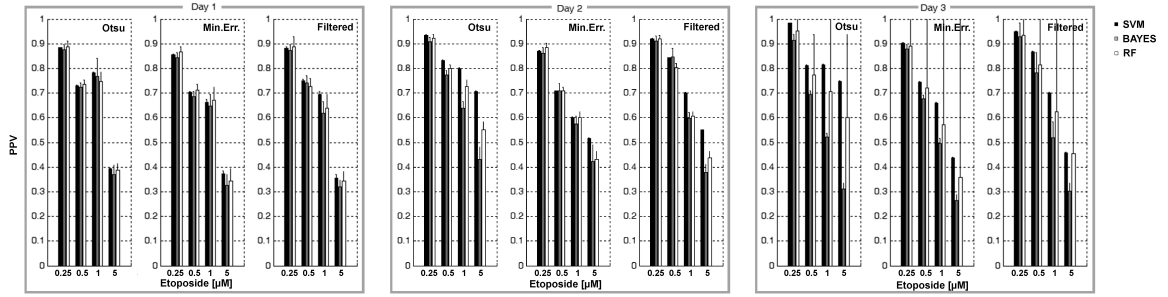
Fig 7 Positive predictive power for classification of cells exposed to a concentration or higher (x-axis). Classifications used are SVM (Black), Naïve Bayes (Grey) and Random Forest (White). Three methods for outline detection were used in segmentation; Otsu, Minimum Error (Min.Err.) and Minimum Error filtered with Otsu (Filtered). Results have been grouped according to day of measurement taken.
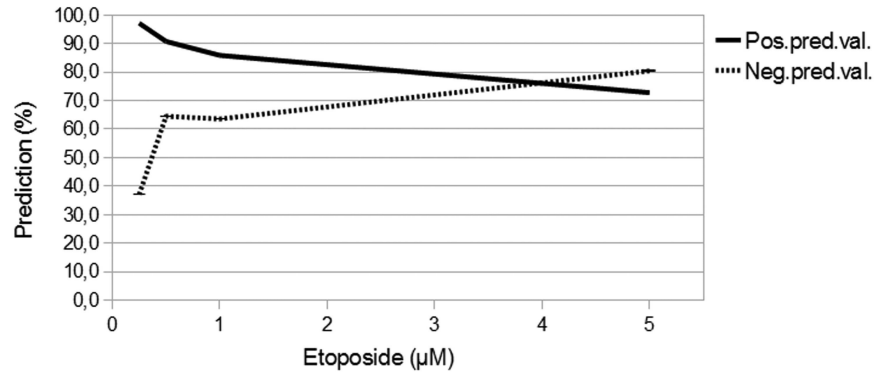

Fig 8 Trade-off between negative and positive predictive power for filtered segmentation for the case of sorting cells as being treated with more than x μM etoposide.

3.0.2 Principal Component Analysis and Clustering

PCA showed that the first principal component in almost all cases could account for more than 90% of sample variability. This was a linear combination of primarily (in decreasing order of strength) Area, Phaseshift integral and Perimeter length. Physically, these components are all related to the cell size. This is consistent with these variables showing the highest correlation also when studying cell average. An initial Partial Least Squares fit revealed the same thing, but the predictive power of the model was low (results not shown). For a subsequent discrete analysis, we filtered out the most useful components of the feature vector. After combining all cells from one sample, we could compare the variables for the sample means for control cells to the variable means for the control. Most prominent were the differences for area, perimeter length and the integral of the phase shift. We measured $P \leq 0.05$ from day 2 after treatment for 10 of the variables (Fig.6), and use these for the rest of the analysis, while the others are deemed not significant. A mathematical description of these entities is given in Section 2.3.
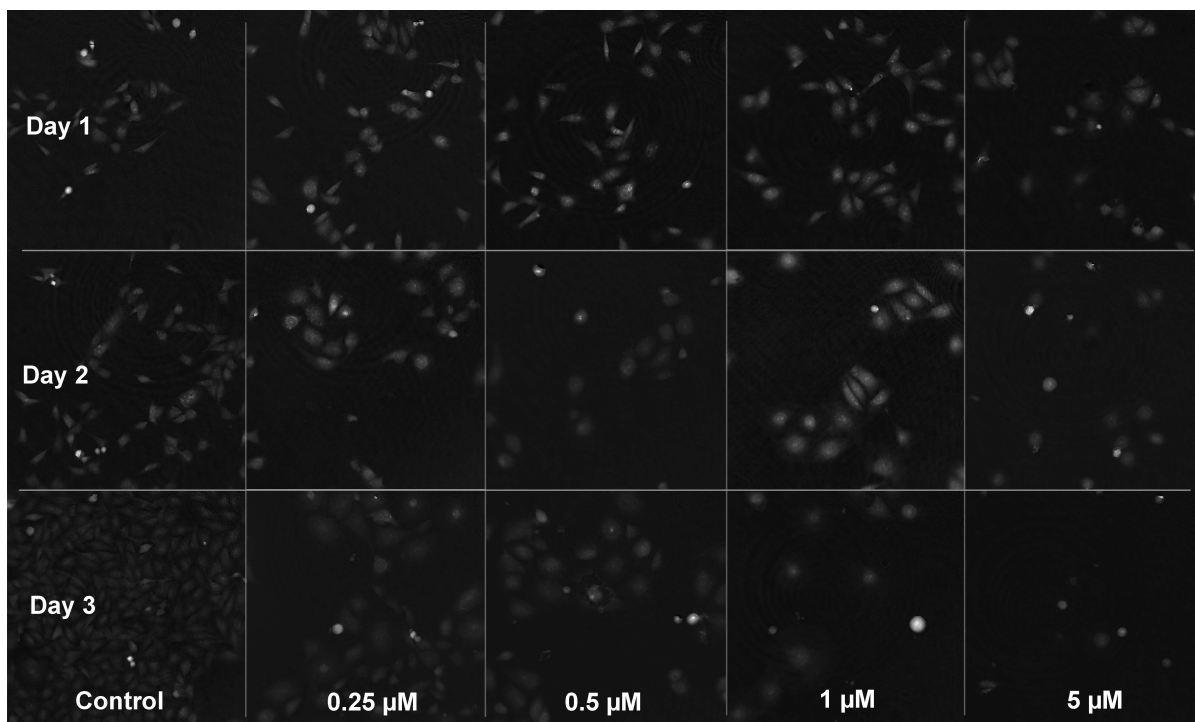
10

Fig 9 Typical appearance of DU-145 imaged by Digital Holography after several days exposure to LC-etoposide treatment.
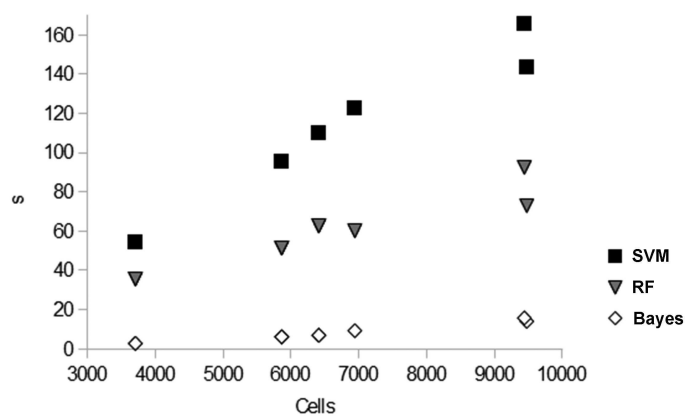

Fig 10 Performance (computation time in seconds for three classifiers; SVM, Naïve Bayes and Random Forest (RF) as a function of number of cells classified.

3.0.3 Cell Classification

With the 5 separate concentrations of etoposide, we had 5 bins available for sorting, but in a real life situation, we do not expect to know the concentration of the sample. In these cases, it is of more interest to ask whether a sample is exposed to some concentration c higher than some minimum value, $c \geq c_{min}$. In this case, the task is reduced to the binary problem of classifying cells as being treated ($c \geq c_{min}$) or not treated ($c \leq c_{min}$). We use the positive predictive power

$$PPV = \frac{P(c \geq c_{min}) \cup R(c \geq c_{min})}{P(c \geq c_{min})} \tag{5}$$

of treated cells as the comparative metric for the classification. Here, $P(c)$ is the set of cells being classified as concentration c and $R(c)$ is the set of cells actually exposed to concentration c. There was a tendency for concentrations to be underestimated, leading to high values of positive prediction (Fig 7). There was also a lower probability of separating control and 0.25 μM, as well as of separating higher concentrations from each other. We show a best case positive predictive rate in separating control from treated (etoposide $\leq$ 0.25 μM) at 97.3%$\pm$1.24 on Day 3, but at a negative predictive rate of 36.3%$\pm$0.01. Results are similar Day 1-2 (results not shown). The best overall classification was obtained for separating etoposide concentrations above and below 0.5 μM, and a theoretical trade-off between negative and positive prediction occurs at 3.3 μM (Fig 8). When comparing the positive and negative predictive power we note a more rapid change from 0.25 μM to 0.5 μM, reflecting the fact that cell morphology changes more rapidly at these concentrations.

Fig. 7 shows a comparison of classification based on the two original segmentations, compared to the version with spatial filtering of cells for all three days and methods of classification. The data on 5 μM was very scarce due to very low number of cells in each image (typically $\leq$ 10 cells), giving high standard deviations and low predictive power in results. The accuracy of prediction is increased throughout Day 1-3 for all classifiers, reflecting the fact that cell morphology becomes more and more distinct (Fig. 9). In most cases, the multiclass SVM is the best classifier. It is closely matched by a random forest, but the errors for the random forest are higher, especially when cell morphology differs strongly. Day 1 there is only marginal difference between the three segmentation methods, but for Day 2-3 the Otsu thresholding gives the most reliable classification.

The computational performance of the Naïve Bayes classifier is an order of magnitude better than that of the SVM (Fig. 10), while the RF rank between the two.

4 Discussion

We have demonstrated a method for image analysis, single-cell segmentation and cell classification for the analysis of drug response of adherent *in vitro* cells over long periods of time. The key elements are a non-destructive optical set-up and a robust image analysis, segmentation and classification. We chose a seeded watershed segmentation because of its simplicity and availability, but other segmentations may have worked equally well. However, a common trait for all is the correct location of cell border, which had a large impact on classification performance. In our case, this was influenced by the fact that one of the most prominent effects of LC-etoposide exposure was the increase in cell area. For Day 2-3 the Otsu thresholding gives the most reliable classification. This can be due to the fact that cell area is larger for treated cells, but that a thresholding including the cell perimeter does not contribute significantly to cell classification. Given the low overall probability to pinpoint the exposure concentration of a single cell, the SVM seems to be the best

option, despite its longer computational time. For future studies, we will investigate other options to optimize the classifier, and expand the set of chosen morphological features.

The classification of concentrations ≥ 0.25 μM and control was the most reliable. For future experiments, we would like to further examine this interval, studying a range of lower concentrations. In our investigation, we focused on lower concentrations of etoposide where cell growth is inhibited, but cell number not severely reduced and this is reflected in the low cell number available for analysis of the higher concentrations etoposide (≥ 1 μM). The predictive power of higher concentrations etoposide could be readily increased by obtaining more images per cell culture vessel, at the cost of longer capture time.

The most significant biological finding of this study is the characterization of morphological changes at low concentrations in the 0.25-1 μM range of etoposide, where no effects are detectable by MTS absorbance. Using a combination of computer vision techniques for image registration, followed by standard pattern recognition and a supervised classification, we are able to detect treated cells at a positive predictive value of 97.3%$\pm$1.24, but at a negative predictive rate less than 40%. The best trade-off we obtained between positive and negative predictive value occurred for etoposide concentrations above 0.5 μM. The cell area is the most dominating morphological effect of long term LC-etoposide exposure, but textural changes are also prominent. The reasons for these changes have yet to be examined, but may be connected to cell cycle arrest, since they are accompanied by a reduction in proliferation rate. It is shown here, that changes are not directly related to reduction in viability or metabolism, as detected by MTS. For the three lowest concentrations, no reduction in viability was detected for the three days examined. For all concentrations, a reduction in proliferation rate accompanies that of morphological changes. For a future investigation, it would be of interest to let the incubation proceed to investigate if viability changes would occur in time. Also, by interrupting the exposure to etoposide, it would be possible to examine if the effects of etoposide are reversible or not. Most importantly we have demonstrated the usefulness of non-invasive image analysis when examining long term processes. In this case, there was no need to set up parallel experiments and any analysis could be done on the same sample container.

References

1 K. R. Stone, D. D. Mickey, H. Wunderli, G. H. Mickey, D. F. Paulson, "Isolation of a human prostate carcinoma cell line (DU-145)," *Int J Cancer*. **21**, 274–281 (1978).

2 M. M. Webber, D. Bello, S. Quader, "Immortalized and tumorigenic adult human prostatic epithelial cell lines: Characteristics and applications part I," *Prostate*. **29**, pp. 386–394 (1996).

3 G. Kroemer, W. S. El-Deiry, P. Golstein, M. E. Peter, D. Vaux, P. Vandenabeele, B. Zhivotovsky, M. V. Blagosklonny, W. Malorni, R. A. Knight, M. Piacentini, S. Nagata, G. Melino, "Classification of cell death: recommendations of the Nomenclature Committee on Cell Death," *Cell Death Differ*. **12**, 1463–1467 (2005).

4 D. V. Krysko, T. Vanden Berghe, K. D'Herde, P. Vandenabeele, "Apoptosis and necrosis: detection, discrimination and phagocytosis," *Methods*. **44**, 205–221 (2008).

5 K. R. Hande, "Etoposide: four decades of development of a topoisomerase II inhibitor," *Eur J Cancer*. **34**, 1514–1521 (1998).

6 Z. El-Schich, A. Mölder, M. Sebesta, L. Gisselsson, K. Alm, A. Gjörloff Wingren, "Digital holographic microscopy innovative and non-destructive analysis of living cells," in *Microscopy, Science, Technology, Applications and Education*, Méndez-Vilas, A, and Díaz, J., Eds., Formatex Research Center, 1055-1062 (2010).

7 M. Falck Miniotis, A. Mukwaya, A. Gjörloff Wingren, "Digital Holographic Microscopy for Non-Invasive Monitoring of Cell Cycle Arrest in L929 Cells," *PLoS ONE*. **9**, e106546 (2014).

8 L. A. Zwelling, "DNA topoisomerase II as a target of antineoplastic drug therapy," *Cancer Metastasis Rev*. **4**, 263–276 (1985).

9 S. W. Lowe, H. E. Ruley, T. Jacks, D. E. Housman, "p53-dependent apoptosis modulates the cytotoxicity of anticancer agents," *Cell*. **74**, 957–967 (1993).

10 E. L. Baldwin, N. Osheroff, "Etoposide, topoisomerase II and cancer," *Curr Med Chem Anticancer Agents*. **5**, 363–372 (2005).

11 H.-G. Kopp, M. Kuczyk, J. Classen, A. Stenzl, L. Kanz, F. Mayer, M. Bamberg, J. T. Hartmann, "Advances in the treatment of testicular cancer," *Drugs*. **66**, 641–659 (2006).

12 J. R. Rigas, P. N. Lara, "Current perspectives on treatment strategies for locally advanced, unresectable stage III non-small cell lung cancer," *Lung Cancer*. **50**, 17–24 (2005).

13 E. Willmore, A. J. Frank, K. Padget, M. J. Tilby, C. A. Austin, "Etoposide targets topoisomerase IIalpha and IIbeta in leukemic cells: isoform-specific cleavable complexes visualized and quantified in situ by a novel immunofluorescence technique," *Molecular Pharmacology*. **54**, 78–85 (1998).

14 W. K. Bleibel, S. Duan, R. S. Huang, E. O. Kistner, S. J. Shukla, X. Wu, J. A. Badner, M. E. Dolan, "Identification of genomic regions contributing to etoposide-induced cytotoxicity," *Human genetics*. **125**, 173–180 (2009).

15 W. M. Liu, P. R. Oakley, S. P. Joel, "Exposure to low concentrations of etoposide reduces the apoptotic capability of leukaemic cell lines," *Leukemia*. **16**, 1705–1712 (2002).

16 K. C. Stine, R. L. Saylors, J. R. Sawyer, D. L. Becton, "Secondary acute myelogenous leukemia following safe exposure to etoposide," *J Clin Oncol*. **15**, 1583–1586 (1997).

17 D. Gabor, "A New Microscopic Principle," *Nature*. **161**, 777–778 (1948).

18 U. Schnars, W. Jptner, "Direct recording of holograms by a CCD target and numerical reconstruction," *Appl Opt*. **33**, 179 (1994).

19 E. Cuche, P. Marquet, C. Depeursinge, "Simultaneous amplitude-contrast and quantitative phase-contrast microscopy by numerical reconstruction of Fresnel off-axis holograms," *Appl Opt*. **38**, 6994–7001 (1998).

20 J. W. Goodman, "Digital Image Formation From Electronically Detected Holograms,", *Appl Phys Lett*. **11**, 77–79 (1967).

21 M. Kronrod, N. Merzlyakov, L. Yaroslavskii, "Reconstruction of a hologram with a computer," *Sov Phys-Tech Phys*. **17**, 333–334 (1972).

22 P. Yu, M. Mustata, J. J. Turek, P. M. W. French, M. R. Melloch, D. D. Nolte, "Holographic optical coherence imaging of tumor spheroids," *Appl Phys Lett*. **83**, 575 (2003).

23 H. Sun, B. Song, H. Dong, B. Reid, M. A. Player, J. Watson, M. Zhao, "Visualization of fast-moving cells in vivo using digital holographic video microscopy," *J Biomed Opt*. **13**, 014007 (2008).

24 K. Jeong, J. J. Turek, D. D. Nolte, "Fourier-domain digital holographic optical coherence imaging of living tissue," *Applied Optics*. **46**, 4999 (2007).

25 C. J. Mann, L. Yu, M. K. Kim, "Movies of cellular and sub-cellular motion by digital holographic microscopy," *BioMedical Engineering OnLine*. **5**, 21 (2006).

26 F. Dubois, C. Yourassowsky, O. Monnom, J.-C. Legros, O. Debeir, P. Van Ham, R. Kiss, C. Decaestecker, "Digital holographic microscopy for the three-dimensional dynamic analysis of in vitro cancer cell migration," *J Biomed Opt*. **11**, 054032 (2006).

27 J. Persson, A. Mölder, S.-G. Pettersson, K. Alm, "Cell motility studies using digital holographic microscopy," in *Microscopy, Science, Technology, Applications and Education*, Méndez-Vilas, A, and Díaz, J., Eds., Formatex Research Center, 1063-1072 (2010).

28 P. Langehanenberg, L. Ivanova, I. Bernhardt, S. Ketelhut, A. Vollmer, D. Dirksen, G. Georgiev, G. von Bally, B. Kemper, "Automated three-dimensional tracking of living cells by digital holographic microscopy," *J Biomed Opt*. **1**, 014018 (2009).

29 I. Moon, B. Javidi, "3-D Visualization and Identification of Biological Microorganisms Using Partially Temporal Incoherent Light In-Line Computational Holographic Imaging," *IEEE Trans Med Imag*. **27**, 1782–1790 (2008).

30 M. Kemmler, M. Fratz, D. Giel, N. Saum, A. Brandenburg, C. Hoffmann, "Noninvasive time-dependent cytometry monitoring by digital holography," *J Biomed Opt*. **12**, 064002 (2007).

31 Y. Emery, E. Cuche, T. Colomb, C. Depeursinge, B. Rappaz, P. Marquet, P. Magistretti, "DHM (Digital Holography Microscope) for imaging cells," *Journal of Physics: Conference Series*. **61**, 1317–1321 (2007).

32 B. Kemper, G. von Bally, "Digital holographic microscopy for live cell applications and technical inspection," *Appl Opt*. **47**, A52 (2008).

33 A. Mölder, M. Sebesta, M. Gustafsson, L. Gisselson, A. G. Wingren, K. Alm, "Non-invasive, label-free cell counting and quantitative analysis of adherent cells using digital holography," *J Microscopy*. **232**, 240–247 (2008).

34 Z. El-Schich, A. Mölder, H. Tassidis, P. Härkönen, M. Falck Miniotis, A. Gjörloff Wingren, "Induction of morphological changes in death-induced cancer cells monitored by holographic microscopy," *Journal of Structural Biology*. **189**, 207–212 (2015).

35 S. Hernandez, K. Barner, "Joint region merging criteria for watershed-based image segmentation," in *2000 International Conference on Image Processing, 2000. Proceedings*. **2**, 108–111 (2000).

36 N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Transactions on Systems, Man and Cybernetics*. **9**, 62–66 (1979).

37 J. Kittler, J. Illingworth, "Minimum error thresholding," *Pattern Recognition*. **19**, 41–47 (1986).

38 S. de Jong, "SIMPLS: An Alternative Approach to Partial Least Squares Regression," *Chemometrics and Intelligent Laboratory Systems*. **18**, 251–263 (1993).

39 L. Breiman, J. Friedman, R. Olshen, C. Stone, *Classification and Regression Trees*, CRC Press, Boca Raton (1984).

40 J.C. Platt, J. Shawe-Taylor, N. Cristianini, "Large Margin DAGs for Multiclass Classification," *Proc. Advances in Neural Information Processing Systems*. **12** 547–553 (1999).

Anna Leida Mölder received a MSc in Technical Engineering in 2006 and a parallel BSc in Biology in 2007. She specializes in computer vision and machine learning for life science applications, with a special interest in microscopy. She is currently engaged as a Ph.D. student at Manchester Metropolitan University.

Johan Persson has a Master of Science in Molecular Biology from Lund University 2008 and is a guest researcher at Malmö University. He also works in the private medical device and drug industry as a research associate in micobiology.

Zahra El-Schich is a Ph.D. student in biomedical science at Malmö University since 2011 and plans to defend her thesis in May 2016. Zahra has been performing research in different projects in immunology, cell biology, cancer and imaging. She is also teaching immunology and supervises thesis work.

Dr. Silvester Czanner is Principal Lecturer in Computing at MMU. He is doing research in areas of Computer Graphics/Medical Imaging/Digital Technologies, especially their applications in education such as using graphics, images and animations to enable and enhance learning. He obtained a MSc degree in Mathematics with specialization in Computer Graphics and Computer Science for Mathematicians, RNDr degree in Mathematics and PhD degree in Mathematics from Comenius University in Slovakia. He is a senior member of the Association for Computing Machinery (ACM) and the director of Eurographics, UK chapter.

Dr. Anette Gjörloff-Wingren received her Ph.D. in tumor immunology at Lund University 1995 and is Associate professor in immunology since 2002. She is performing research and education at Malmö University in immunology, morphology, cell biology, cancer and imaging, as well as supervision and examination of thesis work at basic and advanced level.

Biographies and photographs of the other authors are not available.

List of Figures

List of Tables

VIII

Three dimensional visualisation of microscope imaging to improve understanding of human embryo development

Anna Mölder, Sarah Drury, Nicholas Costen, Geraldine Hartshorne
and Silvester Czanner

1 Introduction

Microscopic methodologies can roughly be divided into destructive or non-destructive techniques. Non-destructive techniques are preferable in many cases, where there is a need to keep interference with the sample at a minimum. In *in vitro* fertilisation (IVF), the embryo under observation cannot normally be manipulated or disturbed in any way, but must be observed "as is", if it is to be used for implantation. For research purposes, destructive techniques can sometimes be necessary in order to perform a particular measurement. Two destructive techniques commonly used in the study of embryos are fluorescence microscopy [19] and confocal microscopy [20]. Microscopic techniques which can be counted as non-destructive include bright- and dark-field microscopy, phase-contrast microscopy (PC) [5], Hoffman Modulation Contrast (HMC) microscopy [13], Differential Interference Contrast (DIC) [22] microscopy and Digital Holography (DH) [2]. Of these, bright- and dark-field microscopy produce an image of the amplitude of the transmitted (or reflected) light. However, cellular material is usually highly transparent, and for such objects, a better sample-to-background contrast can be obtained by recording the phase of light instead of the amplitude. PC microscopy and HMC imaging are techniques where the phase information of diffracted light is optically converted to amplitude information. Techniques such as these are very good for visualisation, but the image grey scale cannot be directly translated to quantitative data. DIC and DH are techniques where the sample phase-shift is imaged directly, yielding information on sample thickness and refractive index.

An increasing amount of image material available due to lower cost of hardware and increased ease of data storage has made it more and more cumbersome to perform image analysis manually. An increasing amount of work has been done in both open source and commercial projects, developing computational solutions for

Anna Mölder

Manchester Metropolitan University, e-mail: anna.molder@mmu.ac.uk

bioimaging problems. Generalised software, with a large number of tools as well as plugins for specific tasks include ImageJ [1], Icy [4], BioImage XD [17], Cell-Profiler [18] and Fiji [27]. Each imaging modality often requires its own computational approach, demanding high technical skills from the user, in order to choose the correct algorithm for the analytical problem at hand. In addition to the multi-purpose solutions, there exists a variety of software solutions, particularly designed for a narrower purpose [26, 33]. A majority of the available computational tools for microscopic imaging focus on fluorescent microscopy imaging - a reflection of its common use in research. However, for stem cell research, when working with tissues intended for transplantation or in IVF, non-destructive imaging modalities are the only option. Interesting work has been done in modelling fixed embryos [3, 12, 31], but little has been done so far on modelling and analysing growing embryos in the early stages after fertilisation using non-destructive imaging.

1.1 The Embryo Selection process

When selecting an embryo suitable for implantation, the embryologist may look at a number of criteria, such as pronuclear appearance and orientation [6, 29], number, size, shape of blastomeres, degree of fragmentation [10], degree of blastocoelic expansion, cellular composition and compactness of the inner cell mass and trophectoderm [28]. Discussions concerning the relevance of embryo morphology in quality assessment exist [11], but it is likely that such evaluations will continue to play a large part in IVF embryo evaluation also in the future. Traditionally, embryos have been studied using a microscope (commonly HMC) only at certain time points during the course of their development. It has been shown in time-lapse studies that the timing of key occurrences can vary greatly between individual embryos with similar morphologic appearance at the conclusion of the recording period, and correlation has been shown between the timing of key developmental events and embryo quality [23]. Such indications, in combination with new possibilities for time-lapse imaging of human embryos for an extended period of time with fewer negative effects to their health, make it likely that the use of time-lapse recordings will continue to increase in the future.

1.2 Embryo Imaging using Hoffman Modulation Contrast

HMC imaging is a popular optical set-up for non-destructive microscopic imaging, routinely used in embryology. Here, light is passed through a pair of off-axis slits, converting gradients in sample optical path to bands of light and dark appearance, depending on the spatial sample direction (Fig. 1). This effect is most apparent when performing any kind of non-symmetrical image operations. Compare for instance the output of a derivative of the raw image taken along the horizontal and vertical
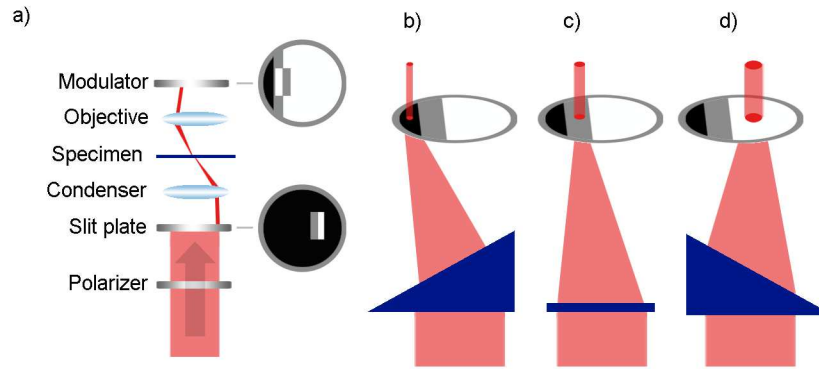
Fig. 1 a) Layout of Hoffman Modulation Contrast optical set-up. Undiffracted light falls on the grey portion of the modulator (c), and the background of the image appears light grey. Refractive index gradients in the sample result in deflection of the light to either the black (b) or the white (d) section of the modulator, so that variations are imaged darker on one side, and brighter on the other, producing a pseudo three dimensional shadowing effect.
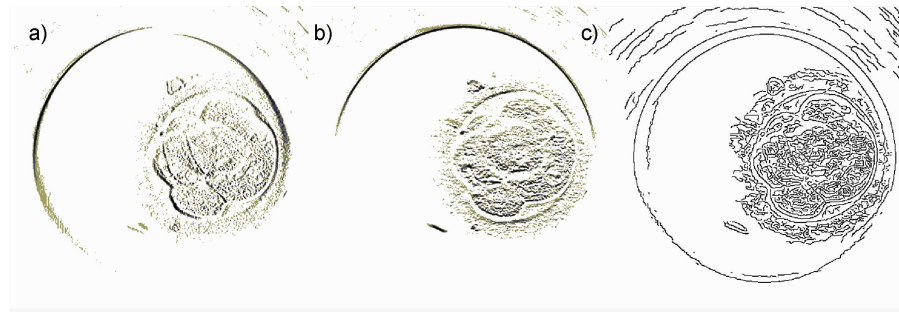

Fig. 2 Gradient of Hoffman Modulation Contrast image in X-direction (a) and Y-direction (b), respectively, and the direction invariant Canny Edge transformation (c). Note that the contrast of these images have been reversed for display purposes.

axis, respectively (Fig. 2). In computer vision, a common approach is the application of an edge finding filter. Here too, the anti-symmetry of the HMC must be taken into account, because the angle of incident light produces a shift in edges in the upper half of the image compared to the lower half. In Fig. 2c), a Canny edge detection filter has been applied to the raw image. It is clear that the filter may find edges on both sides of the lighter and darker bands, resulting in an uncertainty when trying to determine the location of the border of the embryo, or of a single blastomere. However, since the direction of light depends on the azimuthal angle between splits and sample, this effect can be reduced by rotating the sample around a vertical axis, and combining information from several images along the rotation.

Long-term imaging of sensitive material requires the process of capture to be non-destructive, but also requires that the sample can be kept undisturbed in a favourable atmosphere for an extended period of time. Novel construction of incubators and cultivation chambers has recently made it possible to monitor embryos over the course of several days, without any known consequences to their health. However, there are difficulties other than the purely technical when combining automatic long term time-lapse imaging and microscopy. When examining embryos under the microscope, the three dimensional structure is very much of interest. In a traditional, manually handled microscope, much information can be gained by making proper use of the microscope controls: moving the sample around, scanning the focus, adjusting strength of illumination or making use of various filters and apertures in order to scan the three dimensional object in real time. In an automated time-lapse set-up, the possibility to manipulate optics is reduced when the optical set-up must incorporate a climate chamber to accommodate the living cellular material. If the microscope is instead meant to sit inside an incubator or other external chamber, the possibility of manipulating the optics is equally reduced, either because its operation requires the doors of the chamber to be opened, or because the optics are shielded to protect them from the high humidity of the chamber. In many time-lapse set-ups, the possibility of adjusting image quality in real time has vanished, and the user is now limited to studying the images some time after they are captured. This calls for new techniques to visualise this previously captured data in creative ways, possibly of regaining some of the interactivity lost to the user.

2 Materials and Methods

The anonymised embryos used in these studies were donated by consenting patients and the study of them was approved by Coventry Research Ethics Committee (04/Q2802/26) and the Human Fertilisation and Embryology Authority (R0155). Fresh embryos unsuitable for transfer or cryopreservation and frozen stored embryos no longer required for treatment were cultured using Medicult media (Origio, Redhill, UK) for up to 7 days and incubated in 37°C CO_2 in air. Embryos were otherwise untreated or undisturbed during culture, while some were fixed and stained at the end of culture in preparation for the imaging process.

For confocal imaging, embryo nuclei were marked with DAPI, fixed on microscope slides and images were captured using the LSM510 confocal system (Zeiss, Hertfordshire, UK), using $400\times$ magnification, with 1μm between scans. Confocal images were segmented using a combination of region growing segmentation filters (Neighbourhood Connected Thresholding and Confidence Connected Thresholding) and Watershed segmentation [25]. A Delauney triangulation was used to compute the surfaces from the segmented outlines, and the body of each nucleus was then put together in a three dimensional representation of the complete embryo.

For HMC imaging, time-lapse series were captured using the Embryoscope® system (Fertilitech, Copenhagen, Denmark), with a 20 minute interval between
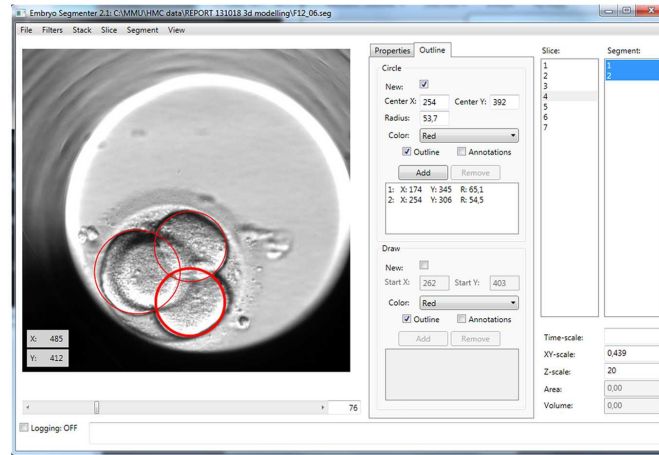
Fig. 3 EmbryoSegmenter for annotation, selection of focus level and blastomere outlines.

Fig. 4 The H-maxima transform extracts peaks higher than h. a) Original image I, subtracted by a scalar h. b) Greyscale reconstruction of I using I-h as marker. c) Subtraction of grey scale reconstruction from original image.

images. The embryos were mounted in wells in an EmbryoSlide® (Fertilitech, Copenhagen, Denmark), one embryo per well, and the imaging used a 635nm LED, and a Hoffman Modulation Contrast Microscope. Up to seven focal planes, 20μm apart, were captured simultaneously, resulting in an image stack, where the image in optimal focus was selected for each embryo using a variation of contrast detection adapted to Hoffman Modulation Contrast microscopy. Blastomere outlines were carefully selected manually using the software EmbryoSegmenter, as shown in Fig. 3, and then further processed using Python scripting, where a spherical model was adjusted to the segmentation outline as described in [8], assuming the segmentation to be at the waist, *i. e.* on the widest part of the blastomere. The spherical shape represents a first order simplification of the true blastomere shape, and can readily be extended to a more complex model, if information from several focal planes is considered. However, due to the longer field of view in HMC imaging, compared to confocal microscopy, in combination with a short depth of field, there is often bleed-through from objects out-of-focus and a higher uncertainty in the extension of objects in the focal (z) direction, compared to the xy-direction.

The image at each focal distance can be described by the two dimensional function $I(x,y)$. Prior to handling, captured images were filtered with a Gaussian filter to remove speckle noise. To detect image sharpness, *i. e.* the focus, a Laplace filter $L(I)$ was used,

$$L(I) = \frac{\nabla^2 I}{4} = \frac{1}{4} \cdot \left(\frac{d^2 I}{dx^2} + \frac{d^2 I}{dy^2}\right), \tag{1}$$

where the Laplacian has been applied to a 4-connected neighbourhood,

$$l_{i,j} = \frac{1}{4}(I_{i+1,j} + Ii-1, j + Ii, j+1 + Ii, j-1) - I_{i,j}. \tag{2}$$

Note that due to the asymmetrical nature of Hoffman Modulation Contrast Imaging (Fig. 2), the symmetrical Laplace was chosen, rather than the direction-dependent gradient.

To detect areas of high sharpness, an H-maxima transform, H, was applied to the result $L(I)$, using an 8-connected neighbourhood. The H-maxima transform consists of a morphological grey-scale reconstruction of a marker image using the original image as mask, followed by a subtraction of the result from the original image (Fig. 4). The marker image is obtained from the original image, subtracted with a constant value h. The result is a set of regional maximas lower than h. For this study, $h = 15\%$ of the image max was chosen. The detected maxima were extended using a Close transform with a 7 pixel diameter circular structure element, and the resulting image was thresholded at 99% of maximum image amplitude and converted to a binary mask. Holes were removed from the mask using a filling function, and the mask was then used to extract the corresponding region from the original image, I.

3 Modelling using Confocal Microscopy

The short field of view of a confocal microscope makes it possible to section a sample volume into image slices and several scans may be used to compile a complete three dimensional visualisation of embryo structure. In Fig. 5, semi-automatically segmented outlines of nuclei from several slices have been combined to show the complete nucelar shape. Fig. 6 shows the three dimensional layout of a human blastocyst, consisting of 121 blastomere nuclei. The nuclei along the perimeter form the trophectoderm. A denser pack of nuclei to the right of the top view image form the inner cell mass. A flattening effect due to the mounting of the embryo between microslides is clearly visible in the side view image. The accuracy of the 3D model depends on the amount of available data, the xy-resolution and the number of scans in the z-direction. Confocal microscopy has been chosen here because the low depth of field allows us to separate the signal between images in the stack, thus obtaining cleaner data. Optical sectioning is to some extent also possible in some non-destructive techniques, in particular in HMC, which has a relatively low depth of field compared to other types of light microscopy.
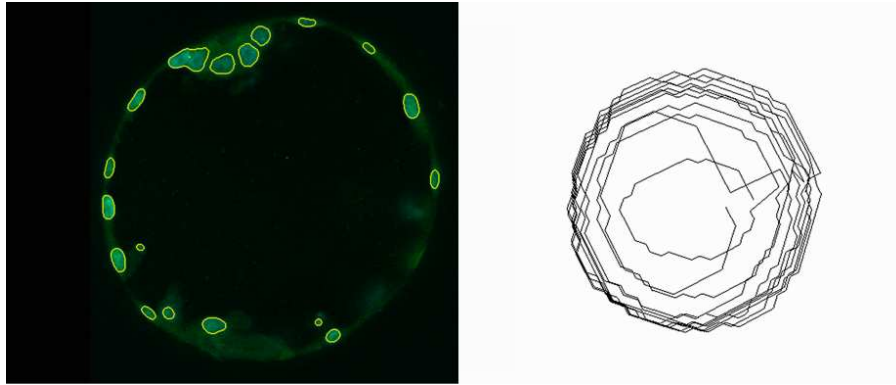
Fig. 5 Confocal image of blastocyst stage human embryo. Nuclei stained with DAPI have been segmented. Line plot of eleven segmented outlines, forming the body of a single blastomere nucleus; top view.
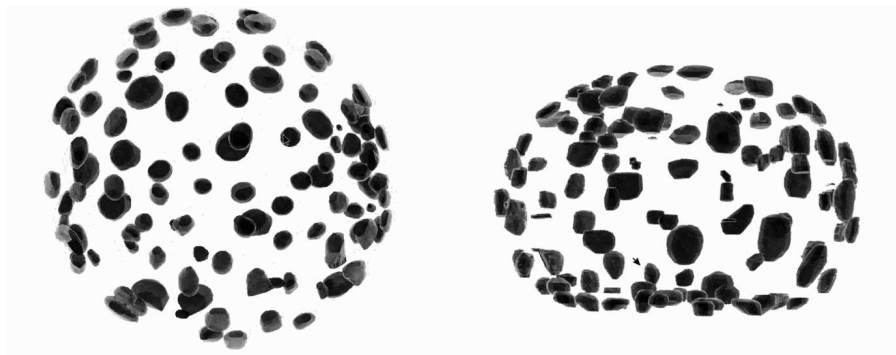

Fig. 6 Nuclei of blastocyst stage embryo, modelled after confocal microscope image scans. Front view and side view, respectively. A flattening effect is visible, where the embryo has been deformed between the microscope slide and cover slip [21].

4 Three dimensional modelling of Hoffman Modulation Contrast images

When images are unsuited for automatic segmentation, or when the clinical nature of the material calls for manual methods to be applied, creative visualisation can still add value to the analysis. Fig. 7 shows a basic spherical model of a 4-cell stage embryo, using 4 selected outlines, and a more detailed model, including the zona pellucida, and internal structures of the embryo. Blastomeres may be located at different positions along the focal axis and the selection of the correct position must be performed separately for each individual blastomere (Fig.8). The crucial step in the modelling is the selection of the appropriate focal plane from the stack of available images.
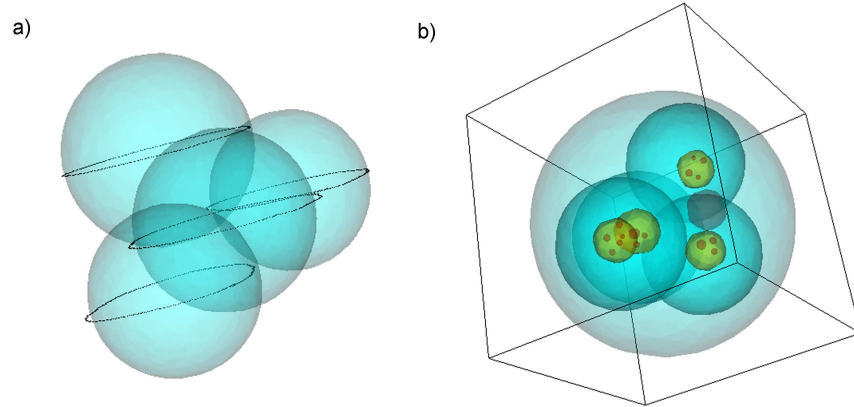
Fig. 7 a) Segmentation outlines at blastomere waist are used to guide spherical cell models. b) All structures present in the original image can be visualised. Here showing nuclei and nucleoli, as well as a large fragment (visible to the right of image centre). The different structures have been artificially coloured for clarity.

4.1 Focal selection

The most critical property of quality for an image is its focus. Without good sharpness, the image will be meaningless for both human eyes and for computerised analysis. As the use of camera recordings increases around us, both in our daily life, for the purpose of documentation, communication, surveillance or for recreational purposes, so it also increases in the laboratory. More and more often, images are captured as part of the scientific process, as a method of documentation or as a part of the analytical process itself. The number and complexity of available algorithms for image segmentation, computer vision or pattern recognition continues to grow and is likely to play a large role in how we handle data in the future. When we hand over more of medical surveillance and diagnostic tasks to automation, it is crucial that we can rely on the accuracy of these automatic procedures. One way to assure a level of quality, and to make sure we do not waste time trying to analyse material of poor standard, is to make sure that the images introduced to an analytical pipeline are captured at an optimal focus.

Many algorithms for automatic focus rely on computation of the power spectrum [32]. In cameras with moving lenses, it is also possible to adjust the focus based on the image contrast, a method referred to as contrast detection autofocus. Several images are then captured in sequence, while searching for local maxima of the image contrast or the gradient of the image contrast. This is usually not done for the entire image but for a selected area of interest. Several autofocus algorithms have also been evaluated for microscopy [7, 9, 24, 30]. In optical microscopy, where
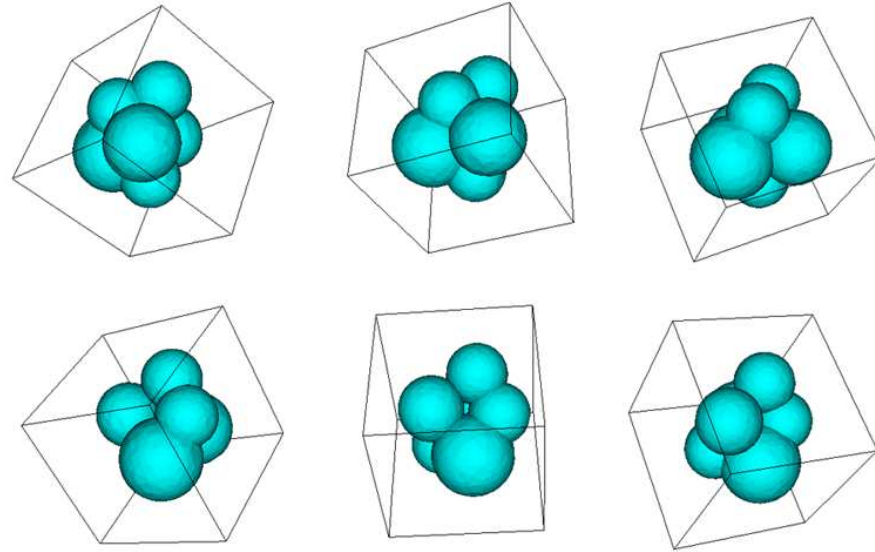
Fig. 8 6 cell stage embryo, with blastomeres modelled as spheres, rotated about its central axis. Bounding box added for clarity. Connected regions between blastomeres differ depending on blastomere size and location.

the depth of field is usually very short, the entire embryo is rarely in focus at one optical setting.

Fig. 9 and Fig. 10 illustrates the extraction of regions in focus, based on the contrast detection algorithm, from the image stack of a human embryo. The embryo is at the blastocyst stage and has the shape of a hollow sphere with cells covering its walls. Due to the varying distribution of cells, the spherical shell of the embryo may vary in tissue content. The stack moves from slightly above the horizontal embryo central plane and downwards until it reaches the embryo base, where also structures on the bottom of the embryo container are encountered.

4.2 Calculation of blastomere connectivity

During the cleavage stage, as the blastomeres undergo mitosis, contact regions between them often increase, as cellular communications are initiated. At the onset of the compaction stage and at the following cavitation, the blastomeres adhere together more closely and start to form a single interacting multicellular organism. In
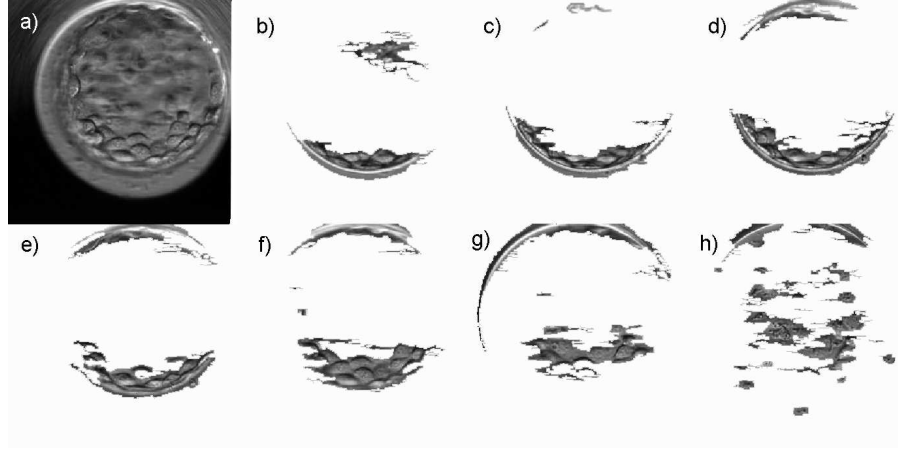
Fig. 9 Hoffman Modulation Contrast image of blastocyst stage embryo at 136.2h, showing one original image (a) and seven images, captured at separate focal planes, with extracted regions in optimal focus (b-h). Focal planes cover the embryo from slightly above the embryo waist (b), down to bottom of the containing well (h). The inner cell mass is in focus on the first slide (b). The original image corresponds to the 5th image in the series (f). Images (b)-(h) have reversed contrast for display purposes.

this way, the developing embryo eventually transforms into a complete entity, rather than a package of loosely aggregated cells.

In Fig. 11, six steps of development from one to seven blastomere stages of a cleavage stage embryo have been modelled.

When modelling each blastomere as a sphere, it is possible to compute an estimate of the blastomere connectivity, C. Let R_1 and R_2 be the radius of two intersecting model spheres, located a distance $d \leq R_1 + R_2$ from each other (Fig. 12). We assume the two spheres follow the physics of a double bubble [16]. Let a be the distance from the centre of the sphere with radius R_1 to the central point of intersection between the two bodies. The intersection will take the shape of a circle, with radius r, giving the contact area,

$$A_{con} = \pi r^2. \tag{3}$$

The connectivity

$$C = \frac{A_{con}}{(A_{con} + A_{uncon})} \tag{4}$$

where the part of the spherical surface outside the intersection, A_{uncon} can be calculated numerically from the model;

$$A_{uncon} = 4p\pi R^2 \tag{5}$$

where p is the part of the spherical surface of the model not contained within the neighbouring sphere. The value p is given by integrating over the surface of the
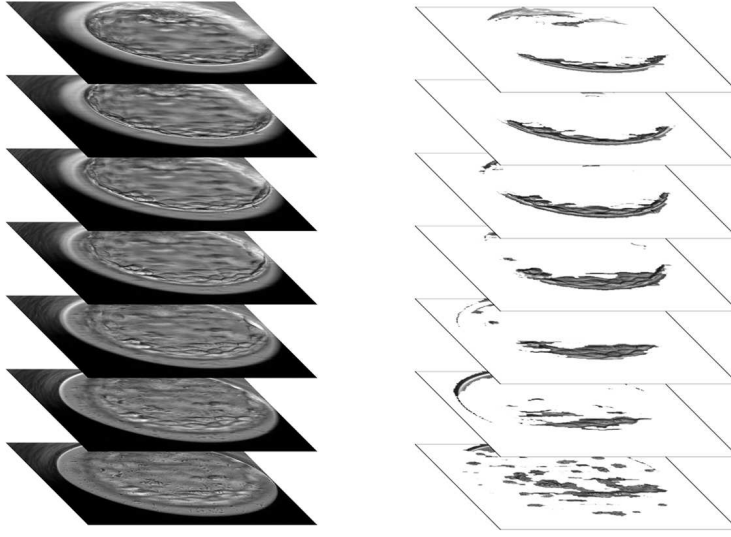
Fig. 10 7-image stack of Hoffman Modulation Contrast images of blastocyst stage embryo at 136.2h, and the extracted focused regions. The extracted regions have reversed contrast for display purposes.

sphere across an angle φ and dividing with the total area of the sphere,

$$p = \frac{1}{4\pi r^2} \cdot \int\limits_{-\pi}^{\pi} \int\limits_{0}^{\varphi} r^2 sin(\theta)\delta\theta \cdot \delta\varphi, \qquad (6)$$

where

$$sin(\varphi) = \frac{r}{R_{1,2}}. \qquad (7)$$

In Fig. 13, the blastomere volume and connectivity have been calculated for the embryo in Fig.11.

5 Conclusion

We have shown how two dimensional image data can be used to construct a simple three dimensional model, which in turn can be used to extract additional data from embryological data sets and also how computer analysis may be used to supplement two dimensional imaging. With a three dimensional display, it is possible to view the sample from different directions, thus getting a clearer view of its spatial layout,
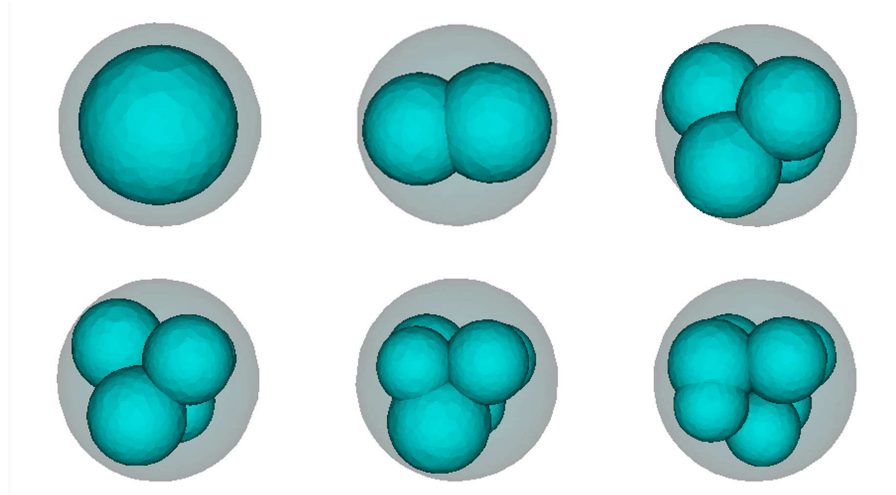
Fig. 11 Cleavage stage embryo observed during development from 0 to 34 h, showing 1, 2, 4, 5, 6 and 7 cell embryos, respectively. The boundary of the zona pellucida is indicated in grey.
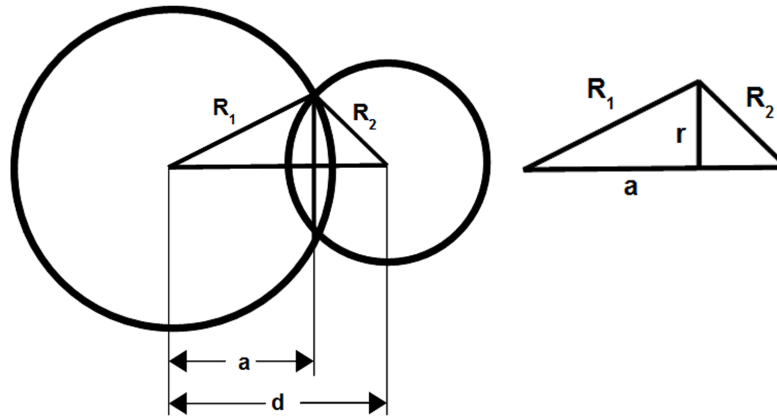

Fig. 12 The intersection plane between two spheres is described by a circle, definitions

and improving interaction with the sample. Using 3D plotting, not only can we get a much more intuitive understanding of the embryo structure and the positioning of blastomeres relative to each other, it is also possible to take measurements of cellular or nuclear volume and blastomere connectivity, which are not possible with a single scan.

When optically sectioning microscopy data, it is not always possible to rely on a short illumination path along the focus axis, as is possible in confocal microscopy,
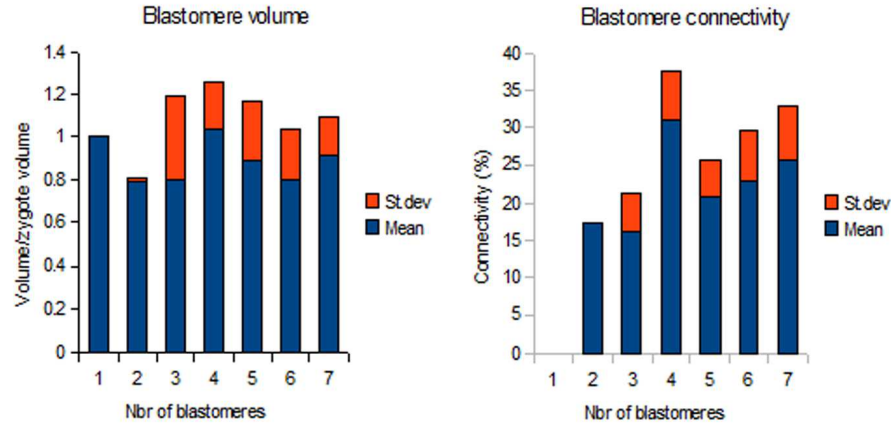
Fig. 13 Blastomere volume and connectivity for the embryo in Fig.11. Blastomere volume has been normalised to the volume of the zygote and to number of cells in the embryo. The connectivity is displayed as percentage of surface area. Bars show both mean and standard deviation, calculated for each number of blastomeres. Apart from the 4 cell stage, which differs in both volume and connectivity, the total volume is constant, whereas the connectivity increases

and the image will always suffer bleed-through from out-of-focus areas. We illustrate here how this problem to some extent may be overcome using pattern recognition and image treatment. A reliable and robust algorithm for focus level detection may be useful not only as a pre-processing step for image analysis, but also as a method for automatic focus of microscope hardware, when the images are being captured during long periods of time, when the capturing process is automatic, or when the microscope is placed in a climate chamber, as is often the case in time-lapse sequencing.

It is a challenging task for future developers, researchers and entrepreneurs to emulate real life events and actions by visualisation tools used to carry out analytical medical tasks. Generally speaking, the more technical, detailed and descriptive a function is, the longer, harder and more time consuming it is to replicate. Microscopic images are among the most detailed to be found, and the implications of making wrong decisions in a clinical environment are potentially severe. Hence, any software solution must be implemented with great care. However, it is not always necessary to simulate every aspect of reality, and a healthy scepticism for technology must not prevent us from embracing it in areas where it can increase knowledge and help reduce human suffering. The task of designing a model which simulates all the relevant features of a growing human embryo in a realistic manner will without doubt require a high degree of future cooperation between software developers and clinical scientists active in the field.

References

1. Abramoff MD, Magalhaes PJ, Ram SJ (2004) Image Processing with ImageJ. Biophotonics International 11,7:36–42.
2. Alm K, Cirenajwis H, Gisselsson L, Gjorloff A, Janicke B, Molder A, Oredsson S, Persso J (2011) Digital Holography and Cell Studies. In: Rosen J (ed) Holography, Research and Technologies. InTech
3. Baldock RA, Bard JB, Burger A, Burton N, Christiansen J, Feng G, Hill B, Houghton D, Kaufman M, Rao J, et al. (2003). EMAP and EMAGE: a framework for understanding spatially organized data. Neuroinformatics 1:309–325.
4. Chaumont, F de, Dallongeville S, Chenouard N, Hervé N, Pop S, Provoost T, Meas-Yedid V, Pankajakshan P, Lecomte T, Montagner Y Le, Lagache T, Dufour A, Olivo-Marin J-C (2012) Icy: an open bioimage informatics platform for extended reproducible research. Nature Methods 9:690–696
5. Davidson MW, Abramowitz M (2002) Optical Microscopy. Encyclopedia of Imaging Science and Technology. John Wiley & Sons, Inc.
6. Garello C, Baker H, Rai J, Montgomery S, Wilson P, Kennedy CR, Hartshorne GM (1999) Pronuclear orientation, polar body placement, and embryo quality after intracytoplasmic sperm injection and in-vitro fertilization: further evidence for polarity in human oocytes? Hum Reprod 14:2588—2595
7. Geusebroek JM, Cornelissen F (2000) Robust autofocusing in microscopy. Cytometry 39:1—9
8. Giusti A, Corani G, Gambardella LM, Magli MC, Gianaroli L (2010) Blastomere segmentation and 3D morphology measurements of early embryos from Hoffman Modulation Contrast image stacks. IEEE Int Symp Biomed Imaging Nano Macro 1261–1264
9. Groen FCA, Young IT, Ligthart G (1985) A comparison of different focus functions for use in autofocus algorithms. Cytometry 6:81-–91
10. Guerif F, Gouge A Le, Giraudeau B, Poindron J, Bidault R, Gasnier O, Royere D (2007) Limited value of morphological assessment at days 1 and 2 to predict blastocyst development potential: a prospective study based on 4042 embryos. Hum Reprod Oxf Engl 22:1973-–1981
11. Hardarson T, Caisander G, Sjögren A, Hanson C, Hamberger L, Lundin K (2003) A morphological and chromosomal study of blastocysts developing from morphologically suboptimal human pre-embryos compared with control blastocysts. Hum Reprod Oxf Engl 18:399—407
12. Heid PJ, Voss E, Soll DR (2002). 3D-DIASemb: a computer-assisted system for reconstructing and motion analyzing in 4D every cell and nucleus in a developing embryo. Dev. Biol. 245:329–347.
13. Hoffman R (1977) The modulation contrast microscope: principles and performance. J Microsc 110:205—222
14. (2008) Human Fertilisation and Embryology Act http://www.legislation.gov.uk/ukpga/2008/22/contents. Cited 30 Oct 2013
15. (2013) Human Fertilisation and Embryology Authority http://www.hfea.gov.uk/ivf-success-rate.html. Cited 30 Oct 2013
16. Isenberg, C (1992) The Science of Soap Films and Soap Bubbles. Dover Publications, New York
17. Kankaanpää P, Paavolainen L, Tiitta S, Karjalainen M, Päivärinne J, Nieminen J, Marjomäki V, Heino J, White DJ (2012) BioImageXD: an open, general-purpose and high-throughput image-processing platform. Nature Methods 9:683–689.

18. Lamprecht MR, Sabatini DM, Carpenter AE (2007) CellProfiler: free, versatile software for automated biological image analysis. BioTechniques 42,1:71–75.
19. Lichtman JW, Conchello J-A (2005) Fluorescence microscopy. Nat Methods 2:910—919
20. Liu Y-C, Chiang A-S (2003) High-resolution confocal imaging and three-dimensional rendering. Methods 30:86-–93
21. Mölder A, Czanner S, Costen N (2013) Multidimensional visualisation to improve the understanding of biological data sets. Central European Seminar on Computer Graphics 17
22. Nomarski G (1955) Microinterferometrie differentiel a ondes polarisees. J Phys Radium 16:9S—11S
23. Payne D, Flaherty SP, Barry MF, Matthews CD (1997) Preliminary observations on polar body extrusion and pronuclear formation in human oocytes using time-lapse video cinematography. Hum Reprod Oxf Engl 12:532-–541
24. Price JH, Gough DA (1994) Comparison of phase-contrast and fluorescence digital autofocus for scanning microscopy. Cytometry 16:283-–297
25. Rafferty K, Drury S, Hartshorne G, Czanner S (2012) Use of Concave Corners in the Segmentation of Embryological Datasets. Rev Bioinforma Biom 1:1—8
26. Rajaram S, Pavie B, Wu LF, Altschuler SJ (2012) PhenoRipper: software for rapidly profiling microscopy images. Nature Methods 9:635–637.
27. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez J-Y, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A (2012) Fiji: an open-source platform for biological-image analysis. Nature Methods 9:676–682
28. Schoolcraft WB, Gardner DK, Lane M, Schlenker T, Hamilton F, Meldrum DR (1999) Blastocyst culture and transfer: analysis of results and parameters affecting outcome in two in vitro fertilization programs. Fertil Steril 72:604-–609
29. Scott LA, Smith S (1998) The successful use of pronuclear embryo transfers the day following oocyte retrieval. Hum Reprod Oxf Engl 13:1003-–1013
30. Sun Y, Duthaler S (2004) Autofocusing in computer microscopy: selecting the optimal focus algorithm. Microsc Res Tech 65:139–149
31. Tassy O, Daian F, Hudson C, Bertrand V, Lemaire P (2006) A Quantitative Approach to the Study of Cell Shapes and Interactions during Early Chordate Embryogenesis. Current Biology 16:345–358
32. Welch P (1967) The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. IEEE Trans Audio Electroacoustics 15:70–73
33. Zhong Q, Busetto AG, Fededa JP, Buhmann JM, Gerlich DW (2012) Unsupervised modeling of cell morphology dynamics for time-lapse microscopy. Nature Methods 9:711–713