

Image Processing Methods for Automatic *in-vitro* Morphology Analysis



Maria Luisa Davila Garcia

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

Department of Automatic Control and Systems Engineering
The University of Sheffield
July, 2018

Image Processing Methods for Automatic *in-vitro* Morphology Analysis



Maria Luisa Davila Garcia

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

Supervisor: Professor Lyudmila L. Mihaylova
2nd Supervisor: Professor Robert Harrison

Department of Automatic Control and Systems Engineering
The University of Sheffield
July, 2018

Abstract

The study of male infertility has become a priority for biologists and researchers in the last decades as a consequence of the declining birth rates. This problem has also become a major public health with economic and psychosocial consequences.

Analysis of human sperm cells, for instance, is widely used in investigations related to male infertility and assisted conception. Sperm samples are usually analysed by health professionals using microscope devices following a manual process to count and describe their morphology. Nevertheless, this practice is prone to errors and time consuming.

This thesis proposes a novel framework based on image processing and machine learning methods to automate the analysis of sperm cells. The proposed method presented an average accuracy performance of 96.4% classify automatically sperm cells in three classes: normal, abnormal and non sperm cell. Performance results have been obtained in challenging conditions: presence of uneven illumination, unwanted noise and blurring caused by the focus drift and occlusion of objects as a result of the overlapping of sperm cells, among others. The object of interest, sperm cells, captured in the images used in this research did not receive any staining or fixation treatment prior to their capture.

A novel and robust methodology based on deep neural learning is developed as part of the automatic feature selection prior to the classification. Also, video and image database of sperm samples was produced at the Andrology laboratory of the University of Sheffield as part of this work. The database was used to validate the proposed framework for the segmentation and classification of *in-vitro* cells.

Acknowledgements

First and foremost, I would like to thank my main supervisor, Professor Lyudmila Mihaylova for her patience, understanding, and incredible support and depth of knowledge and experience along this journey. Without her guidance and encouragement, this thesis would not have been possible.

I would also like to thank Prof. Rob Harrison, my second supervisor for his accurate advice and useful feedback.

Thank you to Professor Allan Pacey and The Academic Unit of Reproductive and Developmental Medicine of The University of Sheffield to allow me access to his facilities and share his clinical knowledge.

Thank you to my sponsor The Mexican National Council for Science and Technology (CONACyT) for funding my research.

I would like to sincerely thank my husband and children, for all of their love and patience. My parents, my parents-in-love whom have supported me my entire life. My brothers, sisters and friends who have believed in me, and encouraged me to be who I am.

Especially thanks to Daniel, the best partner along this time, for his support, making me improve and develop myself. Pablo and Freya for always being my inspiration and bring light and happiness despite countless challenges.

Contents

List of Figures	vii
List of Tables	ix
List of Acronyms	xi
List of Notations	xiii
1 Introduction	1
1.1 Analysis of <i>in-vitro</i> Cells Images	3
1.2 Outline and Key Contributions	8
1.3 Publications	12
2 Literature Review	13
2.1 Introduction	13
2.2 Automatic Segmentation for <i>in-vitro</i> cells	13
2.2.1 Threshold-based	16
2.2.2 Edge-based	17
2.2.3 Region-based	18
2.3 Feature Vector	19
2.3.1 Histogram of Oriented Gradients	20
2.3.2 Speeded-Up Robust Features	20
2.4 Classification	21
2.4.1 Bag of features	22
2.4.2 Support Vector Machine	22
2.4.3 Deep Neural Networks	25
2.5 World Health Organization guidelines	31
2.5.1 Sperm Morphology	33
2.5.2 Sperm Motility	35
2.5.3 Concentration	38
2.6 Computer Assisted Sperm Analysis Systems	39
2.7 Summary	40
3 Analysis and Selection of Features	43
3.1 Introduction	43
3.2 Database	44
3.3 Preprocessing	45

3.3.1	Segmentation with Otsu	46
3.3.2	Edge detection	52
3.3.3	Segmentation based on K-means	52
3.3.4	Principal Component Analysis	54
3.4	Support Vector Machine	58
3.5	Performance Evaluation	61
3.6	Summary	61
4	Speeded-up Robust Features and Bag of Features	65
4.1	Introduction	65
4.2	Histogram of Oriented Gradients	68
4.3	Feature Extraction	69
4.3.1	Bag-of-features (BoF)	69
4.4	Training a Support Vector Machine (SVM)	70
4.5	Performance Evaluation	71
4.6	Summary	74
5	Deep Neural Networks for Morphology Analysis	75
5.1	Introduction	75
5.2	Dataset	77
5.3	Training a new model	79
5.4	Evaluation of pre-trained model	81
5.5	Performance Evaluation	84
5.6	Summary	86
6	Classification of Motile <i>in-vitro</i> Cells	87
6.1	Introduction	87
6.2	Movement Detection	88
6.2.1	Optical Flow	88
6.3	Performance Evaluation	90
6.4	Summary	92
7	Conclusions	93
7.1	Contributions	93
7.2	Future Work	94
7.2.1	Bayesian Convolutional Neural Networks	95
7.2.2	Collaborative analysis with Cloud Computing	95
	Bibliography	96

List of Figures

1.1	Fixed and stained sperm cells	6
1.2	Non-fixed and non-stained sperm cells	6
1.3	Frame extracted from UK NEQAS database	7
1.4	Frame extracted from database created from this thesis	8
1.5	Framework developed in this thesis	11
2.1	Taxonomy of image segmentation methods for mammalian cells . . .	15
2.2	Use of segmentation parameter optimization	16
2.3	First-order derivative representation of a 1-D smoothed step edge: (a) Edge profile (b) First-order derivative.	18
2.4	Approximation gaussians derivatives with box filters	21
2.5	Hyperplane through two linearly separable classes	23
2.6	Hyperplane through two non-linearly separable classes	24
2.7	Basic model of artificial neuron	26
2.8	Activations functions	28
2.9	Convolution	29
2.10	Pooling	29
2.11	General structure of Deep Neural Networks	30
2.12	Frame extracted from video sequence showing sample's components. .	31
2.13	Example of elements found in samples: epithelial cells and leukocytes.	32
2.14	Spermatozoa anatomy.	33
2.15	Abnormal sperm head morphology examples.	34
2.16	Characterization of standard terms used in CASA systems [1].	37
2.17	Improved Neubauer counting chamber.	39
3.1	Morphology classification with SVM	44
3.2	Histogram equalization applied in original image.	46
3.3	Grayscale image from the WHO manual database.	47
3.4	Histogram distribution and threshold for image from the WHO manual.	47
3.5	Sperm cells and background correctly separated for images from WHO.	48
3.6	Original image from the database created, converted to grayscale. . .	48
3.7	Histogram distribution and threshold for image from the database created.	49
3.8	Background extracted, image from the database created.	49
3.9	Distribution of the ratings data	51
3.10	Edge detection of sperm cells using methods based on derivatives and noise reduction: (a) Original image; (b) Prewitt; and, (c) Canny. . .	53

3.11	Results varying the K-means distance and seeds location	54
3.12	PCA applied to dataset, two dimensional projection	55
3.13	Percent of variability of each principal component	57
3.14	First three principal components and features distribution	57
3.15	Prediction using linear kernel.	59
3.16	Prediction using SVM polynomial kernel.	59
3.17	Prediction using SVM gaussian kernel.	60
3.18	Prediction using SVM with sigmoid kernel.	60
3.19	Precision (Press), Recall and accuracy (Acc) obtained analysing 10 features	62
3.20	Precision (Press), Recall and accuracy (Acc) considering Area and Eccentricity	62
3.21	Precision (Press), Recall and accuracy (Acc) analysing Area and En- tropy	63
4.1	Apparent change in shape, size and intensity of two sperm cells . . .	66
4.2	The classification method based on SURF, BoF and SVMs.	67
4.3	Original image, sperm cell with normal morphology	68
4.5	HOG features with cell 8X8 Prediction using linear function	68
4.4	HOG features with size cell 4X4	69
4.6	Template for normal cell and the distribution of the ratings data . . .	70
4.7	Selection of multiples templates for normal class	72
4.8	Performance of the proposed classification method based on SURF and BoF.	73
5.1	Appearance variation of two sperm cells (upper and lower row) along consecutive video frames $t \in [0, 1, \dots, 4]$	76
5.2	First Convolutional layer weights from sperm cell image	77
5.3	Original images used for training	78
5.4	The proposed architecture for DNN	79
5.5	Monitoring the learning process and validation accuracy of the model	80
5.6	Structure for one new model of DNN	80
5.7	Transfer learning process	82
5.8	Structure of CNN AlexNet	83
5.9	Accuracy of classification with Transfer Learning	84
5.10	Classification with SVM based on the deep Neural Network trained .	85
6.1	Movement detection and classification	89
6.2	Detection of ROI with movement	90
6.3	Detection of motile sperm and classification	91
6.4	Confusion matrix for classification of cells with progressive movement	92

List of Tables

2.1	Lower references limits (5th centiles and their 95% confidence intervals) for semen characteristics, WHO 2010	33
5.1	Details of data augmentation	77
5.2	Parameters and Accuracy with DNN created from scratch	81

List of Acronyms

ACC	Accuracy
ALH	Amplitude of the Lateral Head displacement
(C)ASMA	Computer Assisted Sperm Morphometry Analysis
AUC	Area Under The Curve
AVI	Audio Video Interleave
BHAT	Bhattacharyya Distance
CAD	Computer Aided Diagnosis
CASA	Computer Assisted/Aided Sperm Analysis
CNN	Convolutional Neural Network
DNN	Deep Neural Network
ESHRE	European Society of Human Reproduction and Embriology
FC	Fully Connected
FPS	Frames Per Second
GPU	Graphics Processing Unit
HSI	Hue Saturation Intensity colour space
ICSI	Intra-Cytoplasmic Sperm Injection
IM	Immobility
IVF	<i>In-Vitro</i> Fertilization
LoG	Laplacian of Gaussian
LIN	Linearity
LVQ	Learning Vector Quantization
MCMC	Markov Chain Monte Carlo
NP	Non-Progressive Motility
PC	Principal Components

PCA	Principal Component Analysis
PDF	Probability Distribution Function
PF	Particle Filter
PR	Progressive Motility
PRESS	Precision
RBF	Radial Basic Function
ReLU	Rectified Linear Unit
RGB	Red-Green-Blue Colour Model
ROI	Region Of Interest
ROC	Receiver Operating Characteristic
SGD	Stochastic Gradient Descent
SSIM	Structural Similarity
STR	Straightness
SVC	Support Vector Classifiers
SVM	Support Vector Machine
VAP	Average path velocity
VCL	Curvilinear velocity
VSL	Straight line velocity
WHO	World Health Organization

List of Notations

I	a digital image
T	a threshold
x	a feature
$G_{x(i,j)}$	a gradient of x on (i, j)
\mathcal{H}	Hessian matrix
$K(x_i, x_j)$	Kernel function
x_i	an entire row or column
x_{ij}	an specific element from row i and column j
\mathbf{A}	a matrix
\mathbf{a}	a vector
a	a scalar

Chapter 1

Introduction

The analysis of morphological variation of microscopic cell images and their accurate interpretation, especially in the biomedical field, is nowadays a demanding and critical requirement. New technologies in optics and electronics and automated microscopy have made specialists able to visualise and record thousands of images, resulting in the need of integration of computational tools to replace inaccurate and usually monotonous manual inspection procedures.

Analysis and interpretation of images from *in-vitro* experiments are usually conducted into laboratories and research groups where there are a limited image processing hardware resources. Furthermore, their observations follow specific procedures or guidelines involving interpretation based on experience, subjective personal criteria and generally without the possibility of evaluation of the information obtained from observations.

Analysis of microscopy images is usually carried out to find specific characteristics within them. This process can be a time-consuming task depending on the nature of the images and the features analysed. It is important to highlight that a visual scoring can remain inconsistent and unverifiable even if the same professional analyses the same image in a different time.

A wide range of research work has been developed in recent years to analyse microscopic images. Features like shape [2]; size and texture [3]; special components [4], [5]; translation movement [6],[7]; motility level [8]; migration behaviour [9]; concentration [10],[11]; among others features have been identified and evaluated as

part of an analysis report.

Most of the diagnosis procedures and treatment for healthcare are supported by analysis of digital images using computers. Researchers from pathology fields recognise the requirement of automatic tools helping to analyse samples, including quantitative measures as a metric of assessment of samples image-based.

Two areas have been widely developed in recent years: digital cytology and histopathological image analysis [12]. The first means isolated cell and cell clusters with three-dimensional nature. This characteristic reduces the ability to focus [13] as the three dimensions object is represented into a plane. As a result, feature extraction is an even more challenging task requiring special approaches to handle this loss of information.

In the case of histopathological images, tissues from a range of organs are scanned and analysed, distinguishing features such a structure, colour, texture and morphological information. Here, the image for analysis is highly adaptable to single-plane digital scanning without compromising the resolution or features observed by an expert [13].

Recently, approaches using spatial analysis to evaluate histopathology and cells images have been widely investigated [12]. For example, in [14] blood cells are analysed, breast cancer tissues are studied in [15],[16] and [5], stem cells [17] and brain tumor tissue is evaluated in [18],[19].

An important example in cytology image analysis is sperm cells for male infertility and assisted reproduction techniques. In the UK, it is estimated that one in seven couples have fertility problems and the male side is responsible for the 30% [20]. Health professionals diagnose infertility in men following the World Health Organization guidelines, which evaluate motility grade, concentration and morphology of sperm cells [1].

The intra-cytoplasmic sperm injection (ICSI) is an example of assisted conception method where sperm morphology plays a crucial role, as the sperm has to be selected manually and then used to fertilize the ovule. The sample of sperm has to be “fresh” and have not to be previously stained or fixed [21]. The fixation procedure essentially locks cellular structures in place increasing the mechanical strength and stability.

Despite multiples commercial computer systems have emerged, combining powerful Graphical Processing Units (GPU) and algorithms for *in-vitro* image analysis they represent a high-cost technology for most of the laboratories. Furthermore, their optimal performance is usually reached under controlled conditions such as illumination and with high-magnification, with pre-stained cells or pre-processed images.

Computer-Aided Diagnosis (CAD) system for breast mass classification, for instance, typically needs to be fed with images where the damaged tissue has been previously segmented with precision [22], [23]. In some cases, the operator is required to select specific features to be analysed based on his experience [16]. Moreover, most of the previous classification approaches are built upon the extraction of hand-crafted features, such as size, shape and texture.

This thesis addresses the limitations mentioned above by developing a method to classify images of *in-vitro* cells without prior processing (stained or fixed). This thesis presents an image processing method for morphology analysis and evaluation of *in-vitro* cell without prior manual selection of features, using deep learning convolutional neural networks.

1.1 Analysis of *in-vitro* Cells Images

Nowadays, improvements in computers to process information have let to automatic computer systems being used to assist in the health diagnosis and medical treatments. Modern imaging systems, for instance, are capable to visualise, record and store thousands of datasets including microscope images.

On the other hand, analysis of organic components for *in-vitro* models has been traditionally carried out by experts worldwide using manual procedures. The reasons for this is the high cost of computer-based technologies and the lack of adaptability the hardware and software to specific conditions. Laboratories can have light, humidity, temperature, space available variations. The object to be analysed can be deformable, rigid, static, three-dimensional, among others variables.

Analysis of sperm cells, for instance, can be defined as a subjective interpretation

which can be influenced by multiple factors. Also, the analysis may be conducted with a poor quality control. It is generated by omission about standards of quality and safety for the donation, procurement, testing, processing, preservation, storage and distribution of sperm samples published by the World Health Organization (WHO) for this specific procedure.

Also, conditions in the area of examination (e.g. illumination), hardware properties of the imaging devices and properties of the material under examination such as the movement of cells in the mean they are contained can affect the visual definition of those wanted cells for their examination. An additional component affecting the imprecision in visual inspection can be the experience of the operator, visual fatigue and subjectiveness when interpreting datasets.

In order to minimise the subjectiveness and reduce the time to analyse datasets, image processing techniques and computer systems have been used. However, previous approaches studied in this work have presented a low performance in accuracy, precision and reliability, mainly because they do not follow established guidelines or criteria [24]. Additionally, in some areas such as analysis of sperm cells, clinical and technical definitions (e.g. sperm concentration, motility and morphology) remain, to some extent, imprecise [25].

In sperm samples, cells can be presented at a range of size and shape. This produces a subjective classification if the analysis of morphology is carried out manually. Thus, sperm cell classification can vary from one observation to another even though if the analysis is carried out by the same observer at different time, also known as intra-observer variation.

There are three common procedures performed on sperm cells before the analysis: fixation, using a specific chemical to keep rigid cellular structure; air-drying, reducing humidity left to air dry; and staining, adding dye to enhance the component of the cells. Although the main intention is to facilitate the observation, all of them can induce human errors and alterations. This has been published by Maree *et al.* [26].

Although the ideal scenario is having a professional analyst properly trained, and quality control performed regularly on laboratories, in the real practice it is not always possible as means higher cost.

Facing the problems detailed in the above paragraphs, there is a need for a computer based solution. The system or framework should satisfy other requirements such as listed below:

- **Automate the classification and reduce the processing time.** The systems should be able to classify normal or abnormal morphology without the intervention of a human. A reliable computer-based system has to be able to distinguish cells from rest of the material found in a sperm sample.
- **Be objective and verifiable.** Based on the WHO (World Health Organization) guidelines, sperm cells classification must meet formal criteria and results should be verifiable.
- **Be robust.** Frequently the *in-vitro* cell images are affected by low resolution, uneven illumination and low contrast. The systems should then, classifies with good performance under this challenging conditions. As an important condition, the systems should be able to analyse non-stained sperm cells.

Sperm cell classification based on morphology is a challenging task, especially because sperm cells move following random trajectories and sperm cells are non-stained. This introduces multiple occlusions when images are captured making their detection more difficult.

In the Figure 1.1 an example of fixed and stained sperm cells from the Manual of World Health Organization [1] is shown. Cells have been stained with the objective of bold the shape and components and enhancing the contrast against the background. Furthermore, the blur presented is minimum as the cells are static.

Definition and contrast are higher than non-stained cells, as the sample used on this research. An example of the image extracted from the database created for this research is shown in Figure 1.2.



Figure 1.1: Fixed and stained sperm cells. Image taken from the WHO [1], Laboratory manual for examination and processing of human semen.

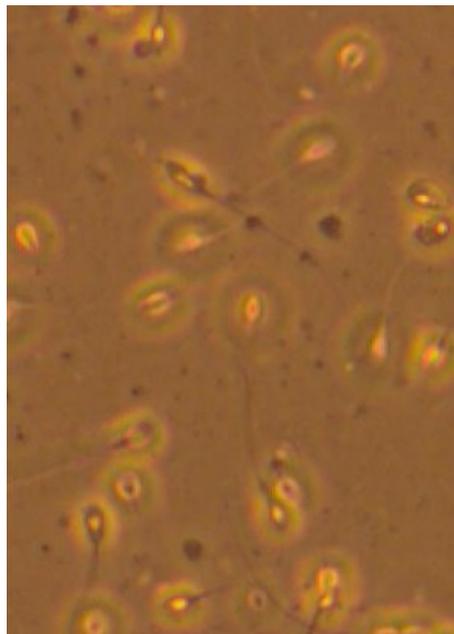


Figure 1.2: Non-fixed and non-stained sperm cells. Example of images analysed on this thesis

This thesis proposes a robust methodology—based on image processing and machine learning methods—addressed to detect and classify, non-fixed and non-stained, sperm cells. There are different and complex scenarios to solve.

As an example in Figure 1.3 and Figure 1.4 there are two samples of frames

extracted from two different sources. Figure 1.3 is part of the database from UK NEQAS (<https://ukneqas.org.uk>) a charitable consortium of external quality assessment laboratories . Figure 1.4 belong to the sequences recorded as part of the database created for this thesis, which is more detailed on Section 3.2.

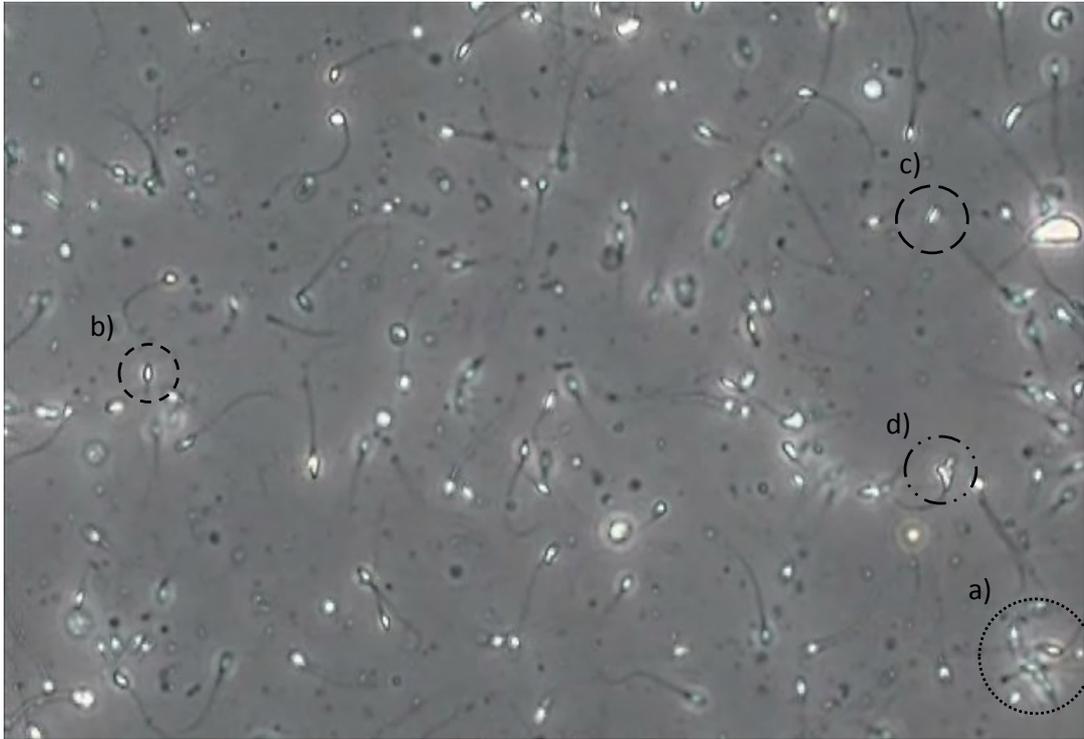


Figure 1.3: Frame extracted from UK NEQAS database. (a) Collision and occlusion between sperm cells. (b) Well defined sperm's head. (c) Blurred sperm's head. (d) Other component joint with sperm cell.

Although both are from the same type of cells, prepared following the same formal guidelines prior to sperm cells analysis, they have significant differences related with the grade of illumination, definition of the sperm head, tails, size, concentration. Definition, resolution, illumination and general conditions are problems to cope with it. The methodology is developed along the next Chapters.

This thesis presents a novel method based on image processing offering an accurate classification of sperm cells. The proposed approach is adaptable and generic, so that it can be used alone or alongside traditional fluorescence single-cell processing to perform objective and accurate morphology analysis for biological applications.

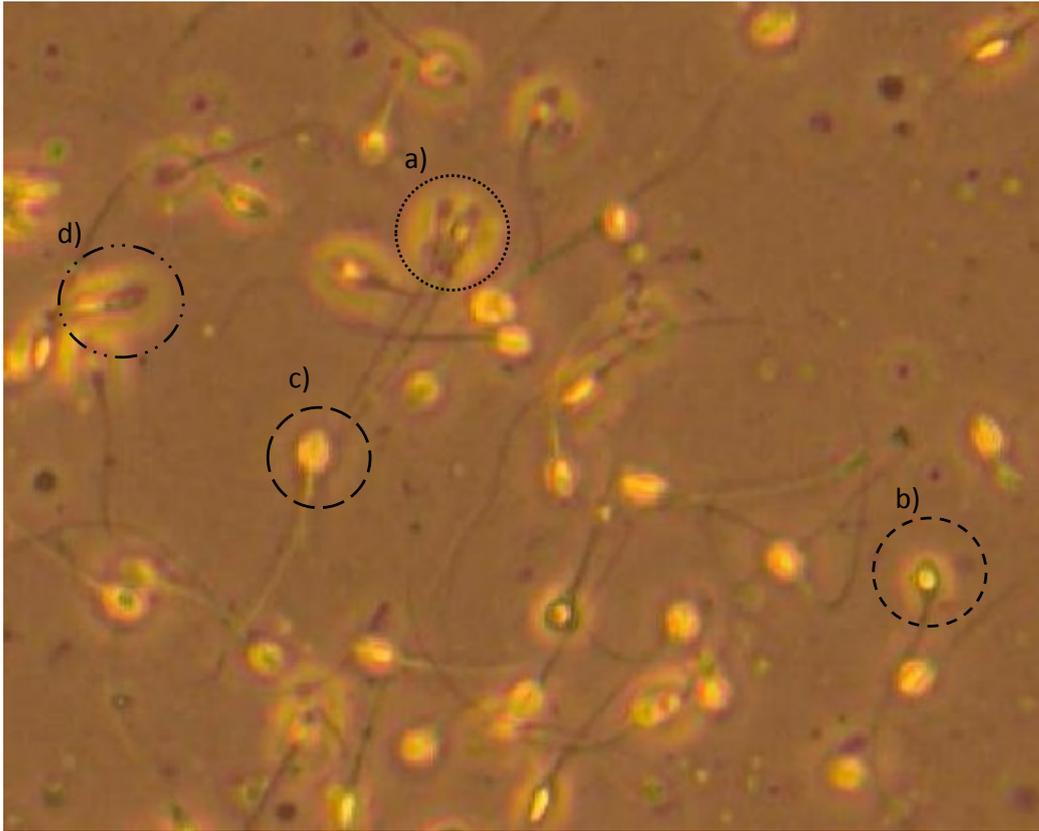


Figure 1.4: Frame extracted from database created for this thesis. (a) Occlusion between sperm cells. (b) Well defined sperm's head. (c) Non-defined-edges sperm's head. (d) Collision between two and more sperm cells.

1.2 Outline and Key Contributions

The structure of this thesis and main contribution is outlined below:

- **Chapter 1.** This Chapter presents an introduction, describing motivation, purpose and challenges of this thesis.
- **Chapter 2.** Literature Review. A review of the current state-of-the-art, newest research done in automatic analysis of *in-vitro* cells and an overview of basics concepts and procedures required to understand next chapters. Description of the guidelines for sperm analysis provided by the WHO.
- **Chapter 3.** Analysis and selection of features. This Chapter presents a methodology for analysing shape, size and colour of non-stained sperm head identifying characteristic features for normal or abnormal morphology labelling.

The main contributions purposed for this chapter is a characteristic feature vector able to represent each class for sperm cells.

Also as a contribution is a database with 1000 images per class, validated with the WHO [1] recommendations for morphology analysis and a professional from the Academic Unit of Reproductive and developmental medicine of The University of Sheffield. This database can be used for validation and test in further research in this field.

- **Chapter 4.** Speeded Up robust features in combination with Bag of Features approach. This Chapter contribution is a new approach for feature selection for our database. As a result of the extraction of SURF features, each class can be represented as a vector.

This vector is invariant to rotation and translation to some extent. Also in this chapter, a bag of feature approach is designed, creating a new representation for each class.

- **Chapter 5.** Deep Neural Network. This Chapter presents an accurate and reliable classification model based on machine learning techniques. Also, transfer learning is analysed and tested, demonstrated better performance.

As the main contribution in this chapter, the use of a pre-trained network, AlexNet, is presented, improving the performance and reducing processing time.

- **Chapter 6.** Classification of motile *in-vitro* cells. This Chapter presents a novel approach with the combination of Convolutional Neural Network and movement detection algorithm. The contribution is a robust method able to classify sperm cells over video sequences.

- **Chapter 7.** Conclusions and future work

This thesis is motivated by the requirements for laboratories to perform sperm morphology analysis following two main sources of validation: WHO guidelines [27] and experience from professionals doing sperm cells analysis. As a result, a novel

framework is presented, using reliable and robust techniques to analyse images, contributing to standardization and automation with regard to sperm morphology classification.

The main contribution of this research is the analysis of motile and non-stained sperm cells based on their morphology using image processing and machine learning techniques, developing a new algorithm for automatic, reliable, repeatable and accurate classification. As far as the author is aware, there is no published work carried out analysing this kind of images.

As a general view of the framework developed in this thesis the diagram in Figure 1.5 shows the main fields studied and their results. There are five main stages, they do not occur in that particular order.

In the framework presented on Figure 1.5 the main topics presented are:

- Capture of images and the creation of the database.
- Several image processing techniques implemented, tested and analyse.
- Analysis and selection of features.
- Analysis, training and test of machine learning models.
- Movement detection.

The contributions from each stage are summarised as results in the bottom of Figure 1.5.

- A validated database with sperm cells images.
- Segmentation of object of interest, sperm cells.
- Characteristic vector for each class.
- Trained model for automatic classification based on morphology.
- Quantification of motile sperm cells.

Analysis of results are presented along this thesis.

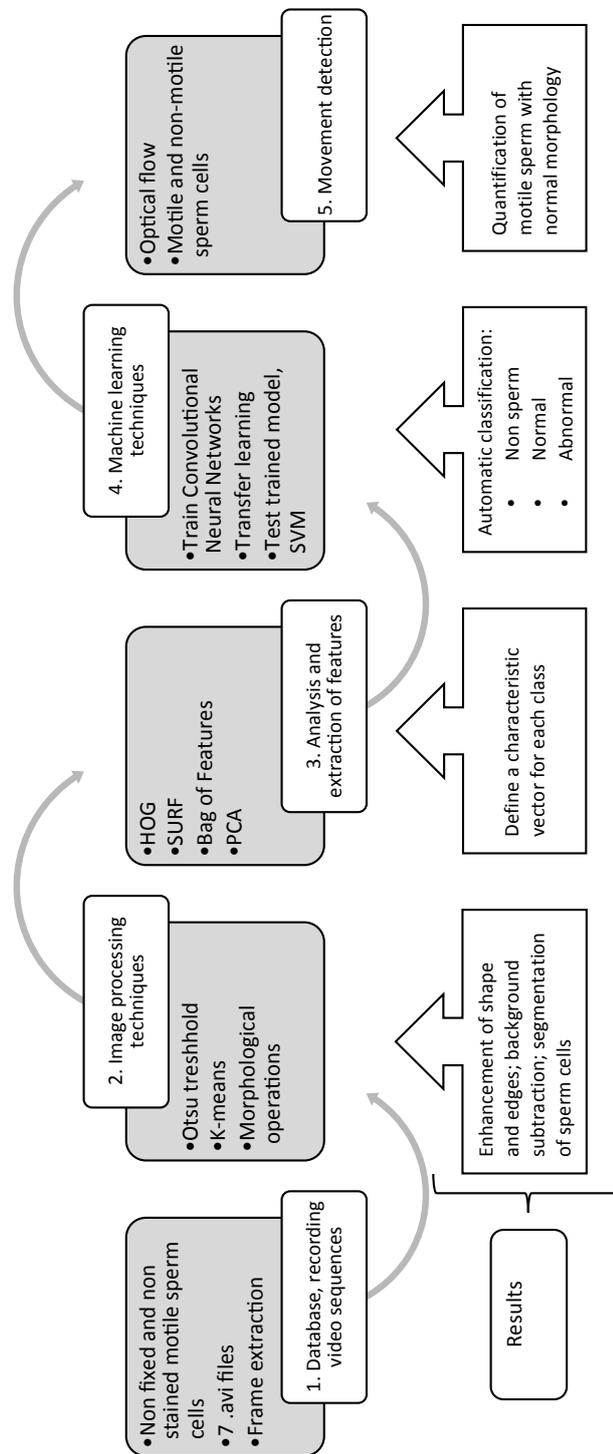


Figure 1.5: Framework developed in this thesis.

1.3 Publications

Some results obtained in this thesis have been disseminated in the following papers:

Peer-reviewed conferences

- LEM Van Raemdonck, A. Ur-Rehman, ML Davila-Garcia, L. Mihaylova, R. Harrison, A. Pacey, "Human Sperm Morphology Analysis for Assisted Reproduction Technique" in *Proceedings of the 10th Sensor Data Fusion: Trends, Solutions, and Applications Conference*, 6-8 October 2015, Bonn, Germany, pp. 1-6.
- ML Davila Garcia; D.A. Paredes Soto; L.S. Mihaylova, "A Bag of Features Based Approach for Classification of Motile Sperm Cells" in *Proceedings of the International Conference on IEEE Physical and Social Computing (CPSCOM) and IEEE Smart Data (SmartData)*, 21-23 June 2017, Exeter, UK, pp. 104 - 109.

Chapter 2

Literature Review

2.1 Introduction

This chapter presents a review of the actual state of the art of the *in-vitro* sperm cells analysis. Also, general definitions and algorithms as a theoretical fundamental for the methodology developed in the following chapters. Then, a section describing formal guidelines established by the World Health Organization about morphology, motility and concentration of sperm cells.

2.2 Automatic Segmentation for *in-vitro* cells

Segmentation is an important part of image processing area. An image is divided into regions or objects according to similar features defined for each class. The objective is to isolate the object of interest.

This process has to be able to distinguish between the objects of interest from the background and other objects.

For this thesis an image is defined as a two-dimensional function

$$I(x, y) \tag{2.1}$$

where x and y are spacial (plane) coordinates, and the amplitude of I is called intensity of the image in the coordinate (x, y) .

Three popular approaches are presented based on the features analysed are presented: threshold-based, edge-based and region-based.

Wide variety of segmentation methods were found in the literature reviewed which was focused on *in-vitro* biological and medical image informatics. Their aim is to improve the efficiency, accuracy, usability, and reliability of diagnostics and treatments.

Segmentation has been one of the fundamental challenges on digital image processing. Segmentation of *in-vitro* cells sperm images require isolate spermatozoa from its background for cellular measurements. At a very low level, segmentation is a partition of an image into connected groups of pixels that have semantic meaning [28], in this case the object of interest is the sperm cell.

The paper published by Lupica *et al.* [29] proposes a multi-resolution-based image fusion pre-processing step to extend the apparent depth of field in cell images for robust automated identification of unstained cells using bright-field microscopy and high magnification lenses. Nuclear features are detected with binary Support Vector Classifiers (SVCs), other features detected include the edge and the intensity.

Bacjy *et al.* [30] present an overview of an automated segmentation methodology applied to optical imaging of mammalian cells. They presented a taxonomy of image segmentation methods which is reproduced in Figure 2.1, presents a general view of a wide range of techniques used for cell segmentation in images.

Most of the cases the noise affecting digital images makes difficult the segmentation process. In order to improve segmentation performance, specific parameters can be optimized, improving robustness to other variables.

This parameters are able to enhance regions or objects with similar features. Intensity threshold, intensity distribution and geometric characteristics of the segmented object are the most popular.

In Figure 2.2 a table with segmentation parameter optimization is shown, the information presented is based on Bacjy *et al* [30]. On this table specified the parameter, the optimization approach for the specific parameters, the segmentation workflow or procedure and the references to some related works published about Otsu approach [31], [32]; intensity threshold [33], [34]; and intensity distribution [6].

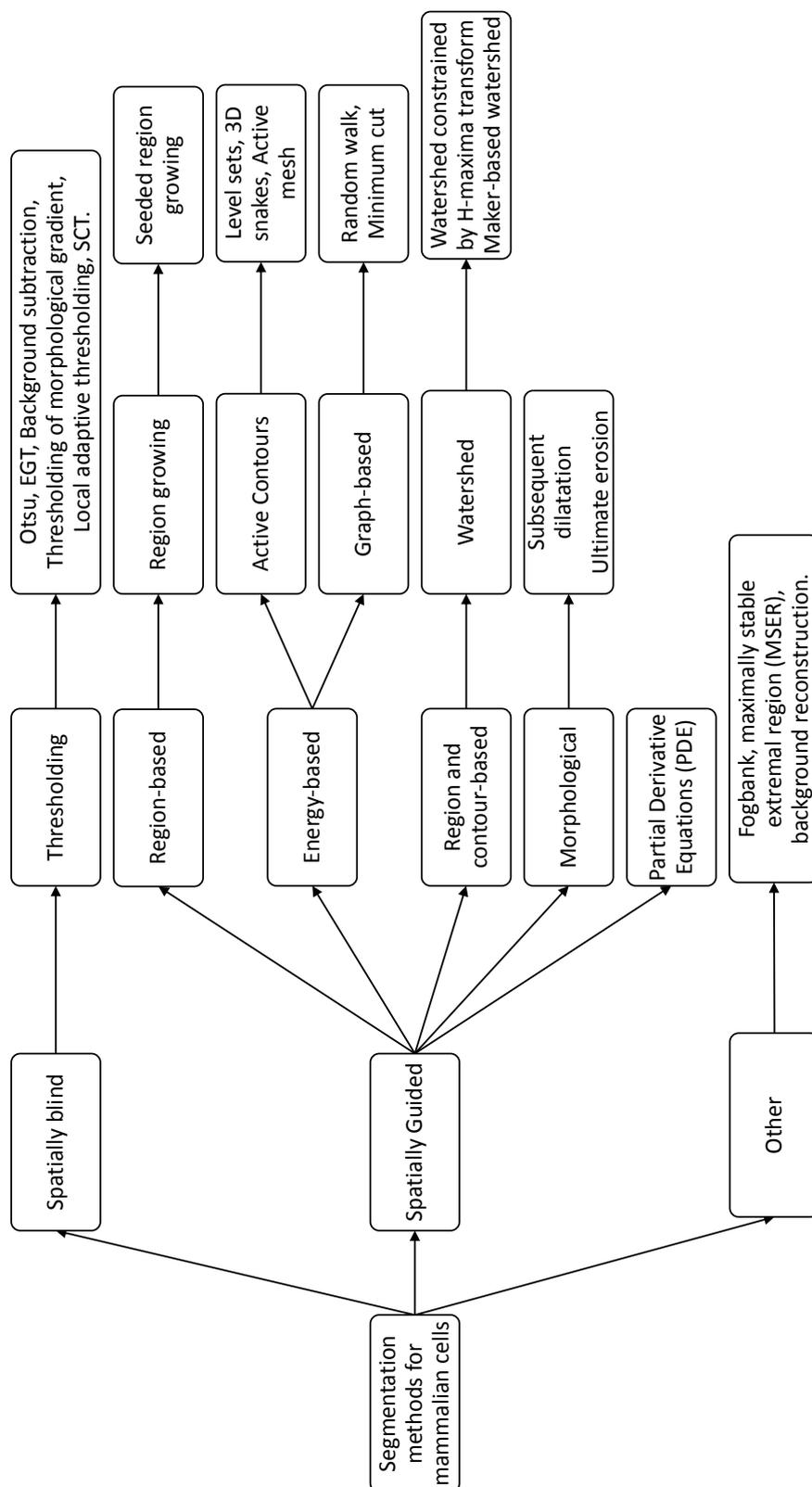


Figure 2.1: Taxonomy of image segmentation methods for mammalian cells presented by Bajcsy *et al.*[30].

Parameter	Optimization approach	Procedure	Ref
Intensity threshold, intensity distribution	Otsu technique	Thresholding, morphological seeded, watershed.	[31]
		DIC-based nonnegative-constrained convex objective function minimization, thresholding	[32]
Intensity threshold, intensity distribution, geometric characteristic of segmented object	Threshold able to fit expected size and geometric characteristics	Gaussian filtering, exponential fit to intensity histogram, thresholding, morphological refinements.	[33]
		Thresholding, morphological refinements	[34]
Intensity distribution, geometric characteristics of segmented objects	Hessian based filtering and medial axis transform for enhanced intensity-based centroid detection	Iterative non-uniform correction, Hessian based filtering, weighted medial axis transform, intensity based centroid detection.	[6]

Figure 2.2: Use of segmentation parameter optimization based on Bajcsy *et al.*[30].

Using image segmentation for medical applications can be a useful tool, especially when huge data is generated making impossible analysis for human. A classification system could have some false positives but this factor can be measured and modelled.

A false positive is when the object as been identified has an object of interest but this detection is wrong because the fact is the object does not belong to the class. This also know as false alarm and is part of the performance measure defined as specificity.

Sperm morphology analysis has been conducted for decades, using wide range of approaches for specifics image conditions. However, only few publications have been found where the images have been capturing with motile and non-stained sperm cells.

The work developed by Ghasemian *et.al.* 2015 presents a methodology for recognition and analysis of each part of non-stained sperm cell. Nonetheless, the measurements of accuracy is no presented [35].

2.2.1 Threshold-based

As a second step, after detection, segmentation is also widely investigated. Some segmentation techniques can be considering as threshold-based if they are based on

local pixel intensity levels.

Each pixel value is compared with the background, and with a specific threshold value, then is labelled. This technique is based on a parameter T or threshold applied to the image $I(i, j)$ as follows:

if $I(i, j) \geq T$ then $I(i, j) = 1$ is object,

but if $I(i, j) \leq T$ then $I(i, j) = 0$ is background.

Threshold-based segmentation is one of the simplest segmentation methods where basically the pixels are classified depending on their intensity value.

Is possible define a variable threshold T along the image. The value of T may depend of the values around (x, y) or neighbours or the value of T is a function of (x, y) which is called adaptive threshold.

2.2.2 Edge-based

An edge can be defined as a set of connected pixels that are located in the limit between two regions. The edge-based segmentation technique is done making a distinction between the values of connected pixels in one region from pixels belonging to other region. Making this differentiation segmentation can be done as the area in between is labelled as an edge.

One basic technique for enhancement of the boundaries is highlighting the contrast of the image. Essentially, the edge is a step-change in the intensity levels of pixels. Thus, edge can be defined as changes in the intensity of pixels. Edge is where the first derivative of the intensity is larger in magnitude than some threshold or where the second derivative of the intensity has a zero crossing [28].

The main problem using edge detection is the definition of what an edge might be. Usually it is a discontinuity in the image signal. The localisation of this is the same where is local maxima in the first derivative of the signal, or zero-crossings in the second derivative of the signal [36].

Then, assuming that an edge occurs where there is a discontinuity in the intensity function or a very steep intensity gradient in the image, see Figure 2.3, the derivative is taken of the intensity values across the image and finding points where the derivative is a maximum, is where an edges.

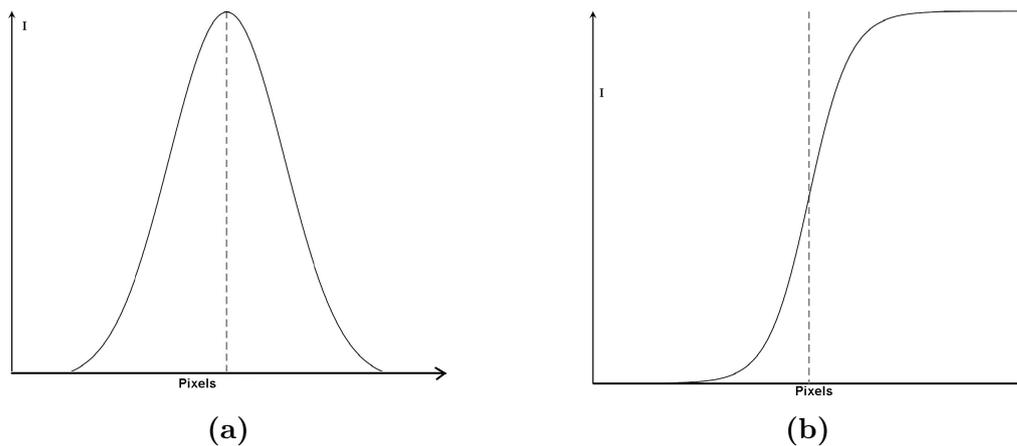


Figure 2.3: First-order derivative representation of a 1-D smoothed step edge: (a) Edge profile (b) First-order derivative.

The *gradient* of the image function I is given by the vector

$$\Delta I = \begin{pmatrix} \frac{\partial I}{\partial x} \\ \frac{\partial I}{\partial y} \end{pmatrix} \quad (2.2)$$

The *magnitude* of this gradient is given by

$$\sqrt{\left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2} \quad (2.3)$$

and its *direction* can be estimated as

$$\theta = \tan^{-1} \left(\frac{\frac{\partial I}{\partial y}}{\frac{\partial I}{\partial x}} \right) \quad (2.4)$$

The simplest gradient operator is the Robert's Cross operator [37]. A 3x3 approximation to $\partial I/\partial x$ is given by the Prewitt [38] and Sobel [39] operator using a mask in order to strengthen of the edge along the horizontal and vertical axes.

2.2.3 Region-based

Region-based techniques try to segment specific area called region formed by a closed path [28].

A classical example for region based segmentation is K-means. K-means clustering or Lloyd logarithm is an iterative, data-partitioning algorithm that assigns

n observations to exactly one of k clusters defined by centroids, where k is chosen before the algorithm starts [40].

Given a set of observations (x_1, x_2, \dots, x_n) , where each observation is a d -dimensional real vector, K-means minimizes the distance between each point to the K centre of the cluster that it belongs to. The objective of K-means is to find to what cluster belongs each observation, in a iterative procedure until cluster assignments remain unchanged.

For all the observations (x_1, x_2, \dots, x_n) , and all the centroids $c_k, k = 1, \dots, K$ compute the distance d_{ik} between the observation vector x_i and the centroid c_k .

Assign each observation vector x_i to the closest centroid c_k (with minimal d_{ik}) record the cluster membership and finally set each cluster centroid to the mean of it is assigned cluster.

One disadvantage of this approach is the iterations can easily get stop in suboptimal cluster formations.

2.3 Feature Vector

The measurement used for the automatic object classification are defined as a features. Feature selection is crucial for object recognition problems. The general definition is presented as a feature vector that represents uniquely a single object. A feature vector is given by:

$$x = [x_1, x_2, \dots, x_l] \quad (2.5)$$

where l is the number of features x_i .

Features extracted are classified into two main groups: low level and high level features. Typically features such as shape, size, colour are considered as high level.

Features extracted with computers systems and algorithms are considered as low level features. For instance, pixel intensity, pixel gradient and orientation. The following sections present a review of the feature selection and extraction techniques for segmentation and later, classification.

2.3.1 Histogram of Oriented Gradients

Initially used for human silhouette recognition from Dalal *et al.* [41] the histogram of Oriented Gradients (HOG) is now widely used in text and face recognition [42], [43]. Both areas, have defined structure for their object of interest, characters and face.

HOG is based on the distribution of the intensities of gradients G_x and G_y of a given image. Gradient estimates at a pixel (i, j) are given by Equations (2.6) and (2.7).

$$G_{x(i,j)} = f(i + 1, j) - f(i - 1, j), \quad (2.6)$$

$$G_{y(i,j)} = f(i, j + 1) - f(i, j - 1) \quad (2.7)$$

where $f(i, j)$ is the intensity value at pixel location (i, j) . Gradients can be used to estimate the local orientation θ and magnitude H of the gradient:

$$\theta(i, j) = \arctan(G_{x(i,j)}/G_{y(i,j)}), \quad (2.8)$$

$$H(i, j) = \sqrt{G_{x(i,j)}^2 + G_{y(i,j)}^2} \quad (2.9)$$

An image is segmented into cells or sections and the orientation of gradients are calculated. Different kernels can be used to calculate gradient, for example a Sobel-based kernel $D_x = [-1 \ 0 \ 1]$ and $D_y = [-1 \ 0 \ 1]^T$.

2.3.2 Speeded-Up Robust Features

Speeded-Up Robust Features (SURF)—based on the Scale-Invariant Feature Transform (SIFT) [44]—was introduced by Bay *et al.* [45] and has been shown to be successful in object recognition approaches [46]. Mehrotra *et al.* [47], for instance, use an adaptive SURF descriptor for human iris recognition; Feulner *et al.* [48] employ SURF for human-body region detection in Computed Tomography (CT) data; and, Han *et al.* [49] use SURF for traffic sign recognition in colour images.

Similarly to SIFT [50], SURF analyses the spatial distribution of gradients. In addition, SURF divides the image in sub-regions that make the method faster and

less noise-sensitive [45].

SURF is based on the Hessian matrix and rely on its determinant to select the best response across a range of scales. Hence, it integrates the scale-space theory introduced by Lindeberg [51]. The Hessian matrix $\mathcal{H}(x, \sigma)$ at a location $\mathbf{x} = (x, y)$ and scale σ in an image I is given by Equation (2.10).

$$\mathcal{H}(x, \sigma) = \begin{bmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_{yy}(x, \sigma) \end{bmatrix} \quad (2.10)$$

where $L_{xx}(x, \sigma)$ is the second-order partial derivative of the image I at point \mathbf{x} and can be estimated by the convolution of I with a second-order derivative of a Gaussian kernel, also known as Laplacian of Gaussian (LoG); $L_{yy}(x, \sigma)$ and $L_{xy}(x, \sigma)$ can be estimated similarly.

Unlike SIFT, SURF estimates the second-order Gaussian derivatives using box filters based on integral images, as is shown in Figure 2.4.

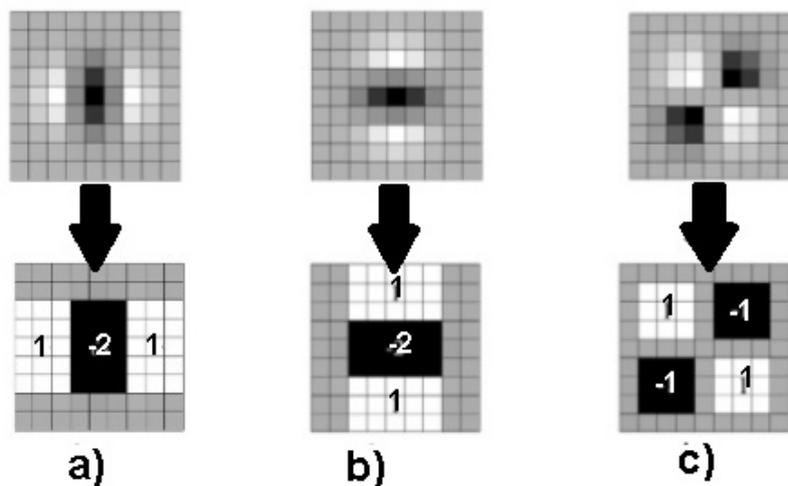


Figure 2.4: Approximation of Gaussian derivatives with box filters: a) L_{xx} b) L_{yy} c) L_{xy}

2.4 Classification

The main goal of object recognition is classifying the object into a specific category, assigning a label.

Object recognition techniques can be divided in two: decision-theoretic, objects described using quantitative descriptors (area, size, length); and structural, objects represented by symbolic information, and described by the properties and relations between this symbolic information.

Next section describes basic concepts from techniques used in the next chapters: Bag of Features, Support Vector Machines and Convolutional Neural Networks.

2.4.1 Bag of features

Bag of Features (BoF) has been used typically for natural language processing and text-mining. Initially named as Bag-of-Words (BoW) [52], this technique uses a histogram to represent the extracted features from images. It has been used previously for cell classification [53] in biomedical images and for scene classification [54], [55], with good results.

The BoF model is summarised in the following steps:

1. *Selection of features.* For each task it is important the selection of the characteristic features. For example: SURF, HOG. A small region (patch) surrounding every key point is used to extract the features of the image.
2. *Learning vocabulary—also termed visual codebook.* In this step, the extracted features are divided into groups (clusters). The clustering process can be carried out by using the K -means approach, where the centroid of a cluster represents a visual word of the codebook [56].
3. *Feature quantisation.* The final step in the BoF model is the mapping process of every feature (patch) into a specific codeword by using a distance metric (*e.g.* Manhattan or Euclidean). Eventually, quantisation of features yields to the histogram representation of the codewords.

2.4.2 Support Vector Machine

This classification method involves separating data into two sets: one set for training, one set for testing. Each element for training is related to one class-label and several

features. The aim is get a model after a training process able to predict target values of the test data using only features.

Given a training set of instances-label pairs (x_i, y_i) , $i = 1, \dots, l$ where $x_i \in R^n$ and $y \in (1, -1)^l$ being x features and y labels with value of l 1 or -1.

Let's assume that data is linearly separable then, it is possible to draw a line separating the two classes when $n = 2$ and a hyperplane on graphs of x_1, x_2, \dots, x_n when $n > 2$. This hyperplane can be described by $w \cdot x + b = 0$ where

- w is normal to the hyperplane
- $\frac{b}{\|w\|}$ is the perpendicular distance from the hyperplane to the origin

The aim of the SVM is to orient the hyperplane maximising the distance between the nearest elements of each class [57], as shown in Figure 2.5 where SVM are circled.

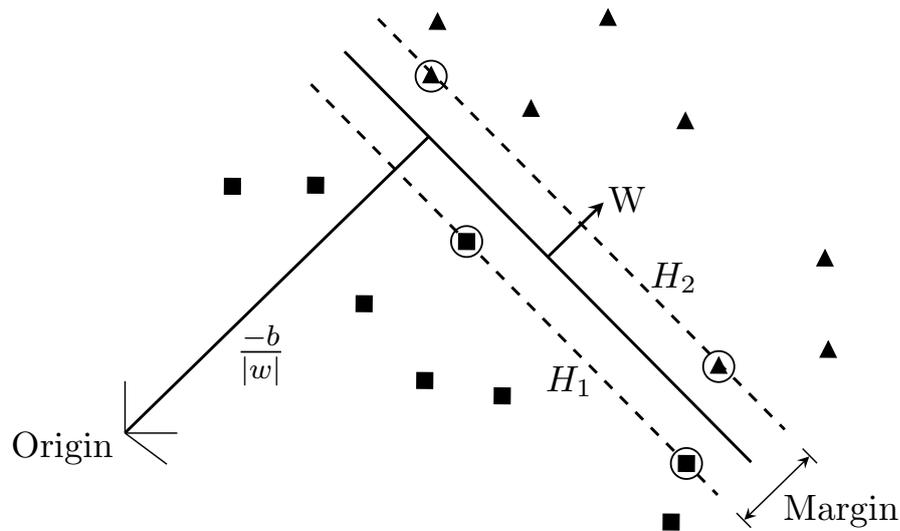


Figure 2.5: Hyperplane through two linearly separable classes

Support vector machines [58], [59] require the solution of the following optimization problem:

$$\min_{w, b, \varepsilon} \frac{1}{2} w^T w + C \sum_{i=1}^l \varepsilon_i \quad (2.11)$$

where w and b allows us to explicitly treat the intercept term b separately from the other parameters. w is the width of the margin.

This equation is subject to $y_i(w^T \phi(x_i) + b) \geq 1 - \xi_i$,
 where $\xi_i \geq 0$.

Using Equation (2.11) cases where data are not fully linearly separable are covered with the positive slack variable $\xi_i, i = 1, \dots, l$.

Using kernel functions it is possible to transform the data into a higher dimensional feature space to make it possible to perform the linear separation.

Figure 2.6 shows how data points on the wrong side have a penalty that increases with the distance from the margin and where ξ_i is.

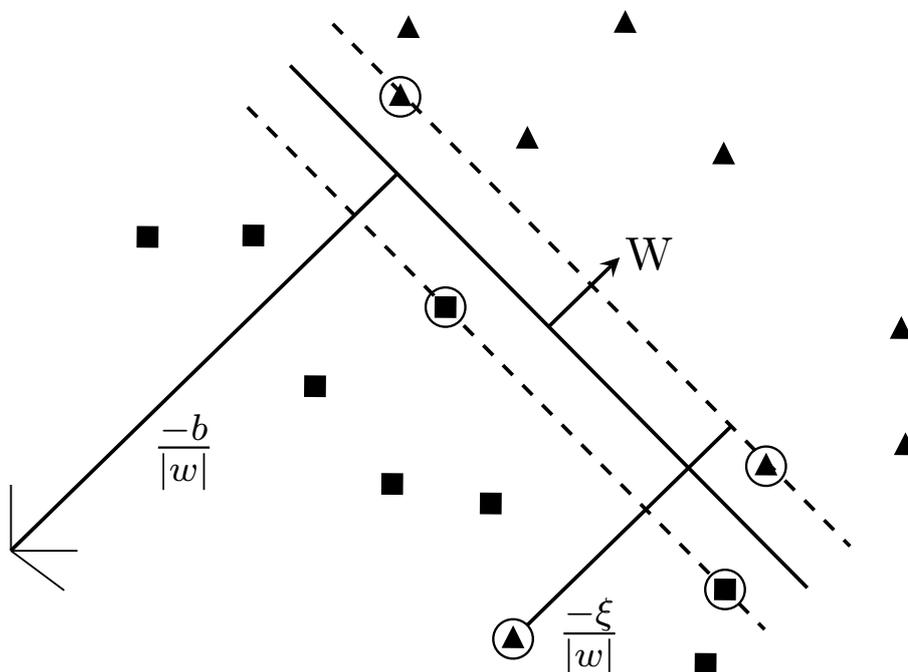


Figure 2.6: Hyperplane through two non-linearly separable classes

Where the training vector x_i are mapped into a higher dimensional space by the function ϕ .

The constant C is the penalty parameter of the error term. This parameter tells the SVM optimization how much wants avoid misclassifying each training example. For large values of C , the optimization will choose a smaller-margin hyperplane, if this yields into training points classified correctly.

When the value of C is small the optimizer will find for a larger-margin separating hyperplane, although this could mean higher misclassification rates.

The generic Kernel function, is defined by Equation 2.12 as follow.

$$K(x_i, x_j) = \phi(x_i^T) \phi(x_j) \quad (2.12)$$

There are multiples kernels, the most used are defined as:

- Linear: $K(x_i, x_j) = x_i^T x_j$
- Polinomial: $K(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0$
- Radial Basic Function (RBF): $K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0$
- Sigmoid: $K(x_i, x_j) = \tanh(\gamma x_i^T x_j + r)$

Where γ is a parameter specified by the user, called the functional margin. A large functional margin represents a high confident and correct prediction.

2.4.3 Deep Neural Networks

Convolutional Neural Networks (CNN) have been used for medical image analysis since 1990s [60], [61]. From the last decade, with the development of graphics processing unit (GPU), the number of areas where CNN is applied has been impacted positively as the training and classification of CNNs can be done in shortest time.

A new approach for CNN is nowadays widely applied: Deep Neural Network. Although the principle is the same as CNN, the deep concept means two o more hidden layers.

Also, the single input deep Convolutional Neural Networks is an image itself. Meanwhile, for the traditional neural network, the input is a set of manually or automatically selected features computed from the images.

The lower layer of the deep neural networks learn and extract the image features. The higher layer classifies images into the different classes using the extracted feature patterns.

New publications have been using CNN as a tool for diagnosis, detection, validation and treatment for health issues. Some of applications are: detection polyp

in colonoscopy videos [62]; detection of mitotic cells in histopathology images [63]; segmentation of pancreas on CT images [64]; brain tumor on magnetic resonance imaging (MRI) scans [65] and nuclei segmentation [5], among others.

Machine learning has been changed radically with neural networks. Neural Networks are algorithms successfully applied to problems about learning.

Neural Networks is one kind of implementation of machine learning, tries to simulate the natural mechanism of communication between neurons in the human brain.

Machine learning techniques are used to learn a model from the data. There are three types of machine learning: supervised, unsupervised and reinforcement learning.

The general representation of a neuron is illustrated on Figure 2.7.

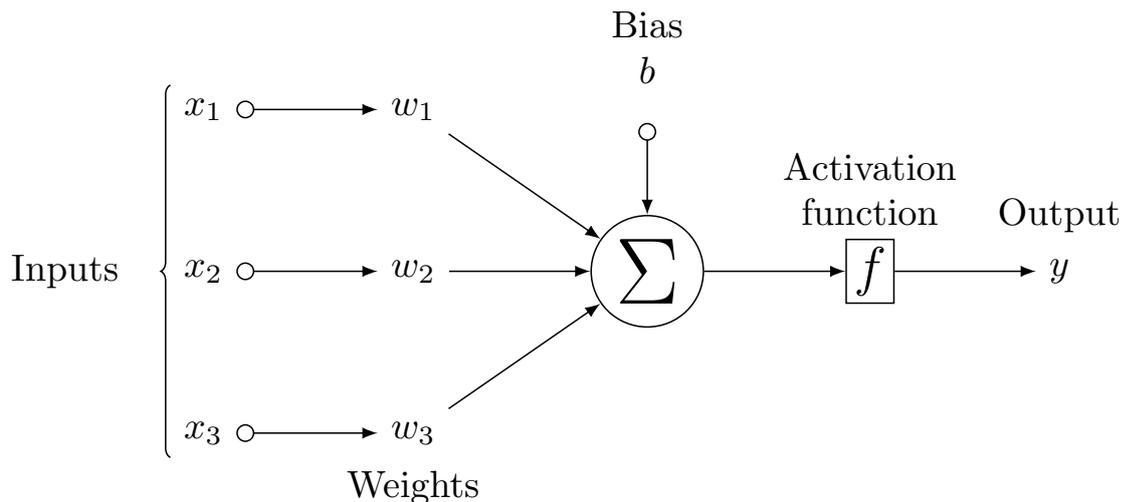


Figure 2.7: Basic model of artificial neuron

Circle and arrow denote node and signal flow, respectively, x_1 , x_2 , and x_3 are input signals and their weights w_1 , w_2 , w_3 . And b or bias is a factor for the storage information. Input signal is multiplied by the weights before reaches the node.

The weighted sum of this example is expressed by Equation (2.13) :

$$v = (w_1 \times x_1) + (w_2 \times x_2) + (w_3 \times x_3) + b \quad (2.13)$$

Finally the node enters the weighted sum into the activation function f which

determines the behaviour of the node. This model usually have more than one node, have layers of multiples nodes, called hidden layers. This kind of model are used for learning. Key concepts are given below [66]:

- **Learning process** means change the model in order to reduce the differences between correct output and model's output.
- **Learning rule** is the systematic approach to update the weights according to the given information.
- **Training** of the network is precisely based on how after input it new informations and adjusts weight.
- **Delta rule** adjusts the weight following the next statement: if an input node contributes to the error of the output node, the weight between the two nodes is adjusted proportionally to the input value, x_1 and the output error, e_i . As is indicated in Equation (2.14)

$$w_{ij} \leftarrow w_{ij} + \alpha e_i x_j \quad (2.14)$$

where

x_j = output from the input node j , ($j = 1, 2, 3$)

e_i = error of the output node i

w_{ij} = weight between the output node i and input node j

α = Learning rate ($0 < \alpha \leq 1$)

- **Learning rate** α determines how much the weight is changed per unit of time.
- **Epoch** the number of times the neural network goes through the repeated training processing with the same dataset.
- **Stochastic gradient descent (SGD)** calculates the error for each training data and adjust the weights immediately. Calculates the weights updates as

is expressed in the Equation (2.15):

$$\Delta w_{ij} = \alpha \delta_i x_j \quad (2.15)$$

- **Batch** is a method where each updated weight is calculated for all errors of the training data. This method uses all of the training data and updates only once.
- **Minibatch** select a part of the training dataset and then the same is used for training in the batch method. The number of training cycles of the neural network equals an epoch, this is because this method uses all the dataset for one training process.
- **Momentum m** term that is added to the delta rule (Equation 2.14) for adjusting the weight to a certain direction in some extent rather than produce an instant change.
- **Non-linear activation function** is a function applied after each Convolution (CONV) or Fully Connected (FC) layer to introduce non-linearity and reduce the *vanishing gradient*. The representative solution is the Rectified Linear Unit (ReLU) but can be sigmoid or hyperbolic tangent. The Figure 2.8 shown how this function transform the input value.

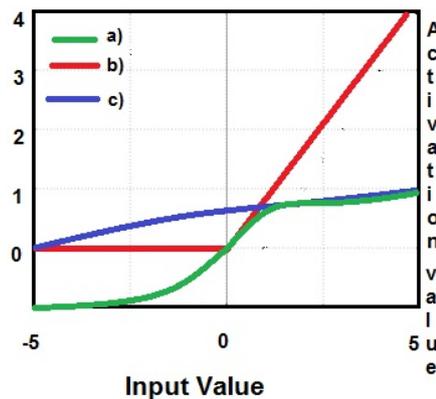


Figure 2.8: Activations functions: a) Tanh, b) ReLU, c) Sigmoid

A basic operation is related to Convolutional neural networks, the convolution. Convolution is a mathematical operation, combines the input and a pre-defined

filter or kernel. The objective is keep the relationship between pixels by learning image features. The Figure 2.9 shown how convolution transforms the input image.

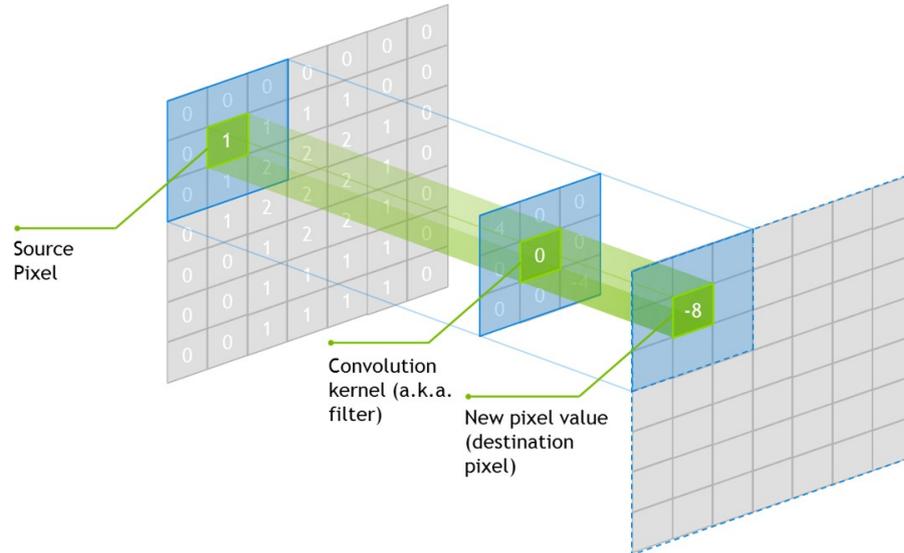


Figure 2.9: Convolution

- **Stride** is the number of pixels shifts over the input matrix. This means if stride is 1 then the filters moves 1 pixel at a time.
- **Pooling** are computations to reduce the dimensionality of a feature map. Typically is done overlapping blocks, when stride is equal to the pooling size. Figure 2.10 shown how pooling is executed.

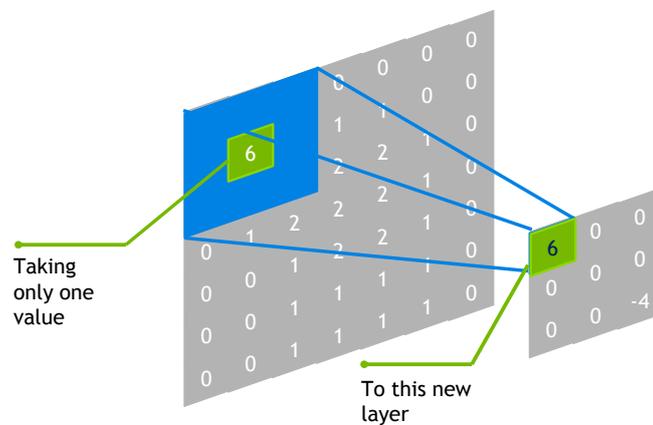


Figure 2.10: Pooling

The backpropagation algorithm solves the problem of training a multilayer neural network. But then, training a network with multiple hidden layers is difficult and slow because the error at the output needs to be propagated back to update the weights in all the preceding layers, and there is interference when there are many parameters.

This was solved in the 1980 when deep learning was introduced. The use of deeper neural networks improve the performance significantly. In the Figure 2.11 a general structure of the Deep Neural Network is show.

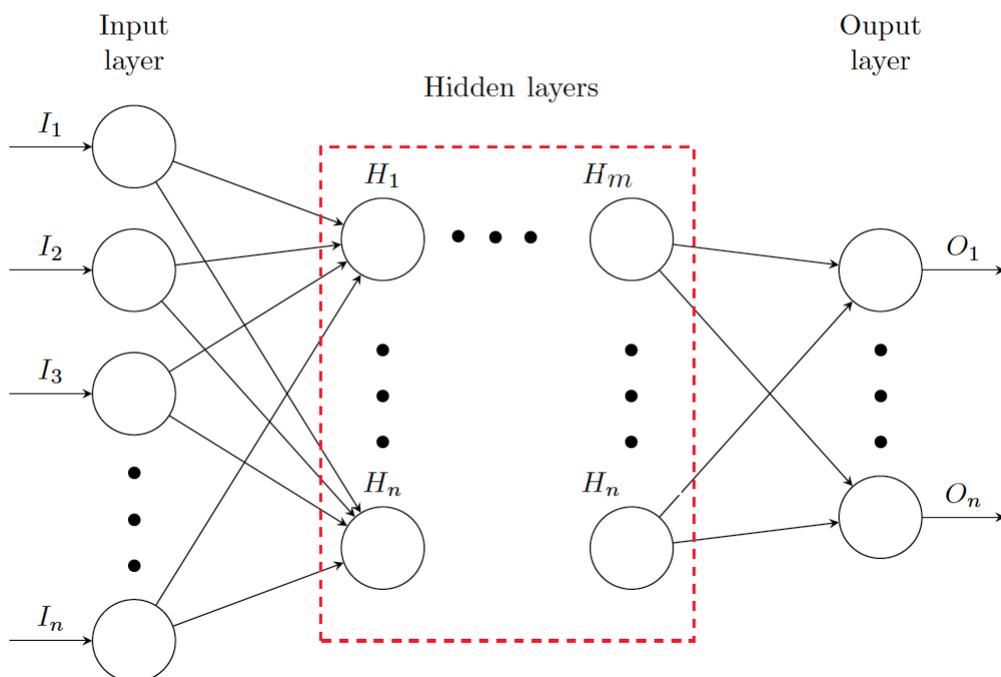


Figure 2.11: General structure of a Deep Neural Network

Deep Neural Networks (DNN) are able to extract high-level features from raw data and capable to manage large amount of data to obtain an accurate representation of an input space. However, this superior performance and accuracy means the cost of high computational complexity.

2.5 World Health Organization guidelines

This section presents an overview of the formal guidelines for sperm assessments. Is important highlight that despite the WHO [1], [67] and ESHRE (European Society of Human Reproduction and Embriology)[68] guidelines, nowadays there is not a world wide accepted protocol for this procedure performance, specially those for sperm morphology due to the use of different techniques and interpretations [69].

According to surveys conducted by the World Health Organization (WHO), near 45% of couples cannot get pregnant because of male infertility. Specific assessment in morphology, sperm concentration and motility values suggests that probability of pregnancy is affected [1].

Sperm analysis has been an essential test in identifying male infertility for more than 50 years. The achievement of pregnancy is directly related with the proportion of sperm with normal morphology [70] and motility grade [25].

The biological nature of sperm makes analysis liable to produce wrong assessments. Other cells like bacteria, leukocytes or epithelial cells can be present in the sample, making difficult the analysis. Figure 2.12 shown an image of the whole sample where this others components are rounded in red.

Also, Figure 2.13 shown image of this kind of elements segmented and scaled, they have a round shape or are smallest than sperm cell.

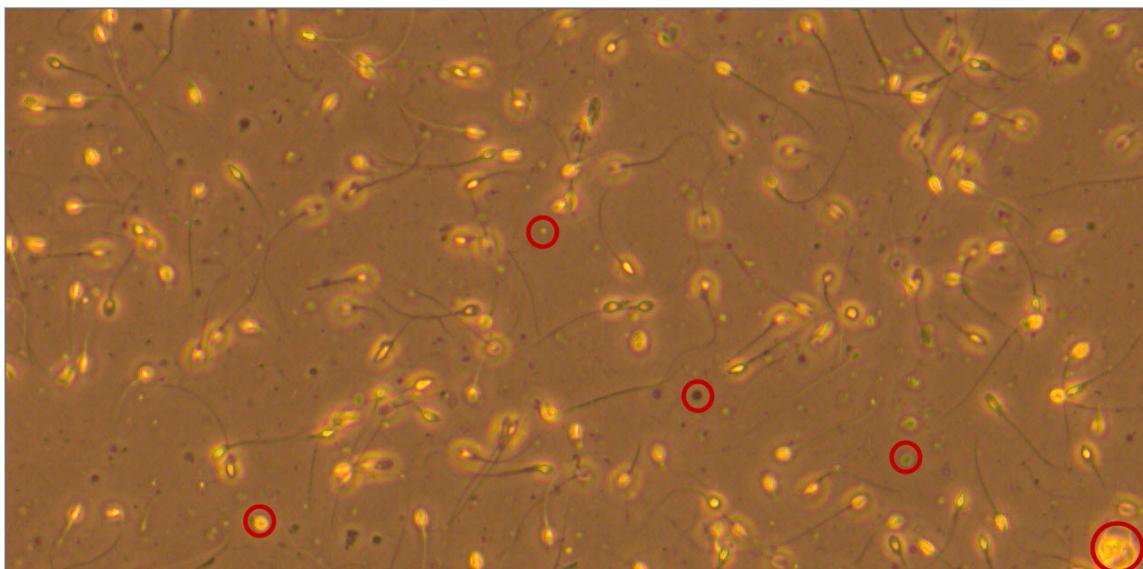


Figure 2.12: Frame extracted from video sequence showing sample's components.



Figure 2.13: Example of elements found in samples: epithelial cells and leukocytes.

The particular characteristic of sperm such as spermatozoa with filamentous structure, dynamic behaviour, microscopic size, visual occlusion as a consequence of constant movement, short period of life and grade of motility, increases complexity. Also, different material or method to prepare the sample before observation could influence in results [71].

According to the World Health Organization (WHO) there are specific recommendations for preparing the samples when semen analysis is performed using a computer system [1].

However, conditions like sperm concentration due dilution techniques, counting chambers, time since sample collection or room temperature can generate different behaviour in sperm [72] and consequently wrong classification.

The analysis of sperm is traditionally carried out using a microscope as follows:

- First a sperm droplet smear is spread on a slide and left to dry.
- The sperm is stained adding one or more dyes to enhance sperm structure.
- Magnification and focus are adjusted, to get a close look at the fixed stained sperm [1].
- Spermatozoa is analysed and using a manual counter, a number of normal and progressive motile sperm is registered.

Usually, this procedure is done twice. Manual method results can be influenced mainly by observer skills. Furthermore, different staining techniques change the morphology dimensions of the human sperm head [26]. It is one of the main reasons why this research is over fresh sperm, non-stained and non-fixed.

The WHO published specific reference limits in order to define ranges for normal or reference value sperm, see Table 2.1. However, these values are generated by expert consensus rather from clinical trials.

Table 2.1: Lower references limits (5th centiles and their 95% confidence intervals) for semen characteristics, WHO 2010

Semen Volume (ml)	1.5(1.4-1.7)
Concentration (10^6 sperm/ml)	15(12-16)
Total sperm number ($10^6/$ ejaculated)	39(33-46)
Total Motility (%)	40(38-42)
Forward progression (%)	32(31-34)
Sperm Morphology (% normal)	4(3-4)
Viability/Vitality (% alive)	58(55-63)
White blood cells (10^6 sperm/ml)	<1.0

2.5.1 Sperm Morphology

A spermatozoa has basic form, consisting of head and flagellum or tail. The head is usually represented by an ellipse, this diagram is shown in Figure 2.14.

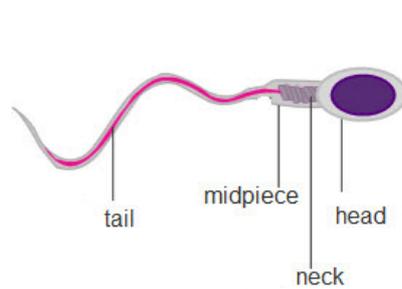


Figure 2.14: Spermatozoa anatomy.

Based on WHO's manual using computational techniques allows classifies spermatozoa in normal or abnormal [1]. At least 200 individual cells should be assessed and counted to get an accurate estimate of the percent normal morphology in the whole sample.

The form of the sperm has major importance than its dimensions. Figure 2.15 shown schematic drawings of head defects of human spermatozoa.

Morphology classification as a normal or abnormal specified by WHO [1] is based on Kruger Strict analysis and is valid using the specific Papanicolaou stained procedure [1]. According to this, the head length for normal spermatozoa may vary



Figure 2.15: Abnormal sperm head morphology examples.

between $4.0 \pm 4.5 \mu m$, with a mean length of $4.07 \pm 0.19 \mu m$ and a mean width of $2.98 \pm 0.14 \mu m$.

Usually, the sperm head can be measured helped by an eye-piece micrometer to be placed in the microscope tube, however doing analysis by eye is time-consuming and tiring process.

Different approaches to solve sperm morphology classification have been published, not only for human sperm but for different species [73],[74].

Morphology evaluation can be developed analysing shape and measurement of sperm head dimensions. Severa *et al.* [75] assessed and evaluated sperm head length to width ratio, position of the centre of gravity, curvature, and degree of roundness using elliptic Fourier descriptors and inverse Fourier transformation. They concluded that for sperm head the predominant parameter is length to width ratio.

Morphological filters are usually applied over sperm images in order to segment from others particles usually present in semen. Example of this is Ravanfar and Moradi [76] by using proper structure element the image were filtered based on Top-Hat operator to separate sperm cells from other debris. An adaptive temporal median filter was used to remove background.

Segmentation of sperm cells into their components is an approach also studied. Khachane *et al.*[77], implemented fuzzy logic for classification, head area and eccentricity are taken. In this work they removed noise using median filtering and segmented head, mid-piece and tail using morphological opening-close operator.

Pattern recognition is another technique used to identify spermatozoa with tools like Principal component analysis (PCA) and K-nearest neighbour algorithm [78].

Few laboratories have the capacity to acquire microscope with high power of magnification technology in order to get better resolution images. Berkovitz *et al.* [79] developed a method for the detailed morphological evaluation of motile sperm called MSOME (motile sperm organelle morphology examination) which is performed with an inverted light microscope equipped with high-power Normarski optic enhanced by digital imaging to achieve a magnification of 6400x able to observe the sperm nucleus.

The development of new techniques to evaluate spermatozoa normality grade in variety of species using image processing techniques is an actual trend. Sanchez *et al.* [80] analyses images of boar spermatozoa heads to describe and classify them by means of their intracellular distribution.

An expert veterinary identified normal heads and extract a model distribution of this training group. For each head image they defined a measure of the deviation from the model intensity distribution.

Biehl *et al.* [81] classified boar sperm heads onto normal and non-normal based on grey-scale microscopic images only and assess the quality of semen using Learning Vector Quantization (LVQ).

Information obtained by visual inspection and manual morphology analysis is not reliable and hardly verifiable. Analysis of sperm morphology is a challenging task, particularly when images are low resolution and noising conditions.

Morphology analysis, overcoming noise, low resolution and high computational are published nowadays looking for an efficient solution. This goal is still being a challenge with image processing methods.

2.5.2 Sperm Motility

Sperm motility assessment is one of the most important parameters for semen quality evaluation, according to WHO [1], for natural or assisted conception. Classified sperm based only on their morphology is not enough evidence to determine a problem related with sperm low quality. Sperm quality assessment can be more accurate and

objective if an estimation of sperm motility is developed [82].

In laboratories worldwide, sperm motility is evaluated by visualisation via a microscope with phase contrast optics [74]. Along the video sequence the movement of the sperm cells is recorded marking on an acetate film stick to a monitor.

Yang *et al.* [83] purposed a quantitative assessment of sperm motility providing head trajectories and flagellum beat patterns.

Several algorithms have been reported for tracking and movement analysis. Template matching is published by Nafisi *et al.* [84] using images acquired by normal light microscope instead of phase contrast configuration and doing a manual selection of the initial point of the sperm trajectory.

Lui *et al.* [85] proposed motion history image to estimate the position of each head along frames, also use motion template method to detect and track sperm head and tail.

Block matching method was published by Yang *et al.* [83] used for detect head sperm and register position. Also, purpose a flagellum tracking algorithm, based on a Markov chain Monte Carlo (MCMC) sampling method, to obtain the flagellum beat pattern.

Sperm cells are difficult to recognize when are alive due to the fast movement and partial or full occlusion in some moments. Furthermore, using hardware to capture images with different configuration and resolutions yield different results. In order to get an accurate description about motility of sperm cells Hu *et al.* [86] recommend a frame rate higher of 30-60 Hz. Severe agglutination can affect the assessment of sperm motility and concentration [1].

The newest version of WHO for assessment of sperm motility recommend categorizing spermatozoa movement as progressively motile, non-progressively motile and immotile [1].

- *Progressive motility (PR)* Actively movement, could be linearly or in a large circle, despite the speed.
- *Non-progressive motility (NP)* Absence of progression, may be moving in small circles or having only flagellum beat.

- *Immotility (IM)* No movement, neither head or tail.

There is some standard terminology for motility variables measured (Figure 2.16) (figure based on [1]).

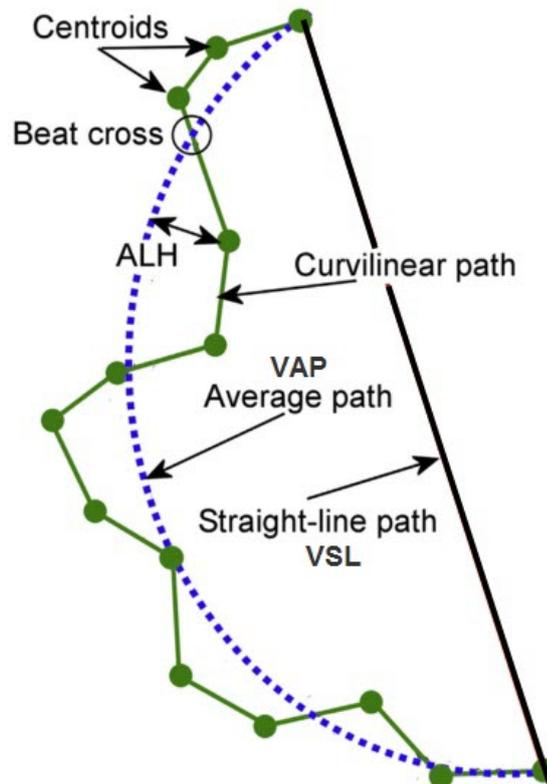


Figure 2.16: Characterization of standard terms used in CASA systems [1].

According to WHO, parameters are defined as follow:

1. VCL ($\mu\text{m/s}$) Curvilinear Velocity of a sperm head along the actual curvilinear path, as perceived in two dimensions in the microscope.
2. VSL ($\mu\text{m/s}$) Rectilinear Velocity velocity of a sperm head following the trajectory from the first point detected to the last one.
3. VAP ($\mu\text{m/s}$) Velocity of a sperm head along the path computed by smoothing the curvilinear trajectory. Different algorithms can be used for estimation.
4. ALH (μm) Amplitude of lateral head displacement compared with its average path.

5. LIN Linearity of a curvilinear path.

$$LIN = \frac{VSL}{VCL} \quad (2.16)$$

6. STR Linearity of an average path.

$$STR = \frac{VSL}{VAP} \quad (2.17)$$

All the measurements are calculated based in a two dimensional plane, as perceived by the microscope. Authors like Kathiravan *et al.* [74] and Castellini [87] studied sperm motility in bull,rabbit and ram. Sorensen *et al.* [88] published they work related to multi-object tracking of human spermatozoa using particle filter and Kalman filters. Also, they purposed The Hungarian algorithm on a rectangular cost matrix to solve the problem of linear assignment and hidden Markov models to calculate the cost. It reports difficulties distinguishing between round cells and non-motile spermatozoa.

2.5.3 Concentration

There is a direct relation between the total number of spermatozoa per ejaculate and the sperm concentration with time to pregnancy [89] and pregnancy rates [90]. Manually, cells can be counted using Haemocytometer. It is an instrument with gridded area in a thick glass slide, has a counter chamber with specific depth created by a rectangular indentation. WHO [1] recommend the improved Neubauer haemocytometer (Paul Marienfeld GmbH and Co.KG, Lauda-Knigshofen, Germany) shown in Figure 2.17.

Sperm concentration ..“refers to the number of spermatozoa per unit volume of semen and is a function of the number of spermatozoa emitted and the volume of fluid diluting them. Total sperm number refers to the total number of spermatozoa in the entire ejaculate and is obtained by multiplying the sperm concentration by the semen volume” according to WHO [1].

Concentration of motile sperm using automatic systems can be estimated with



Figure 2.17: Improved Neubauer counting chamber.

stained head using fluorescent DNA stains, however must follow the technique strictly [91]. Preparation of the sample and calculate procedures to determinate sperm concentration are established in WHO manual [1].

2.6 Computer Assisted Sperm Analysis Systems

Computer assisted sperm analysis (CASA) systems have been developed for the beginning of the 1980s reducing time due process frames automatically from video recording. CASA system is usually a microscope attached to a video camera, sperm images are captured and digitalized. Different techniques of image processing, segmentation, classification have been used by a computer.

Using computer algorithms each component of the image (pixel) is analysed. However, there is a lack of reliability to CASA systems because is not possible to make a validation since algorithms applied are not fully exposed [74].

CASA systems are usually expensive and needs pre-configured parameters for quality control before the analysis. It is recommended to have *a priori* knowledge of standard operating procedure, considering: temperature, dilution factor, mixing method, chamber depth, and technicians with experience using CASA systems[92].

Using CASA systems can reduces human error due to automation of observation and enable the repetition of analysis.

Analysis of sperm using computer systems allows measures like stain content,

length, width, perimeter, area, and arithmetically derived combinations had been considered in order to classified spermatozoa.

Some common deficiencies detected in current CASA systems:

1. Low reliability due to different results even if is the same CASA systems with same sample and identical settings.
2. CASA systems, in some cases, are unable to discriminate between immobility spermatozoa and another kind of components or cells part of the sample.
3. Number frames per second (fps) makes a difference because when rate fps is low, sperm concentration is more accurate; when fps is high, motility features are more realistic. However, in this case losing information about quantitative features is common [93].
4. Is highly recommendable have a priori specification of standard operating procedure and technician's training with computer assisted semen analysis [92].
5. Most of the commercial CASA systems are expensive.
6. The test cannot be repeated to verifies results.

Some common commercial CASA systems are: IVOS Integrated Visual Optic System (Hamilton-Thorne Biosciences, Beverly, MA, USA) able to measures head area and elongation; SCA Sperm Class Analyzer (Microoptics, Barcelona, Spain) reports information from head, mid-piece and tail anomalies and MedeaLAB (Medical Technology MTG, Altdorf, Germany) offer length of the sperm tail but with very specific requirements of microscope and camera [94].

2.7 Summary

This Chapter presented a background review of image processing techniques used in *in-vitro* cell analysis. Three main tasks are involved in this process: segmentation, feature selection and classification, a brief introduction of this is given in this chapter.

Also, the official guidelines from WHO for sperm morphology evaluation are summarised. After this review the necessity of an automatic and efficient computer-based technique is justified. Based on evidence and recommendations from the WHO this thesis is focused on the head's sperm morphology.

After the background review, this research is taking image processing techniques as a fundamental principle. One of the objectives of this thesis is to find the best feature vector for each class, which needs to be invariant to translation and rotation.

The image of the sperm in movement is affected for shrinking, collision with other cells, illumination and have to be correctly classified even when is the image is captured in a different place along the video sequence.

Chapter 3

Analysis and Selection of Features

3.1 Introduction

This Chapter presents a first evaluation and analysis of *in-vitro* sperm cells. In this chapter an automatic method able to achieve an accurate classification extracting features from multiple template of non-stained sperm cells is presented.

Features extracted are used to train a multi-class model for Support Vector Machine (SVM) and finally used to classify new cell images. The main purpose of this chapter is finding a suitable feature vector able to represent each class of sperm cells. Features are extracted from pre-selected templates from 3 classes: normal cell, abnormal cell and others. Also, is presented the dataset created for this thesis is presented.

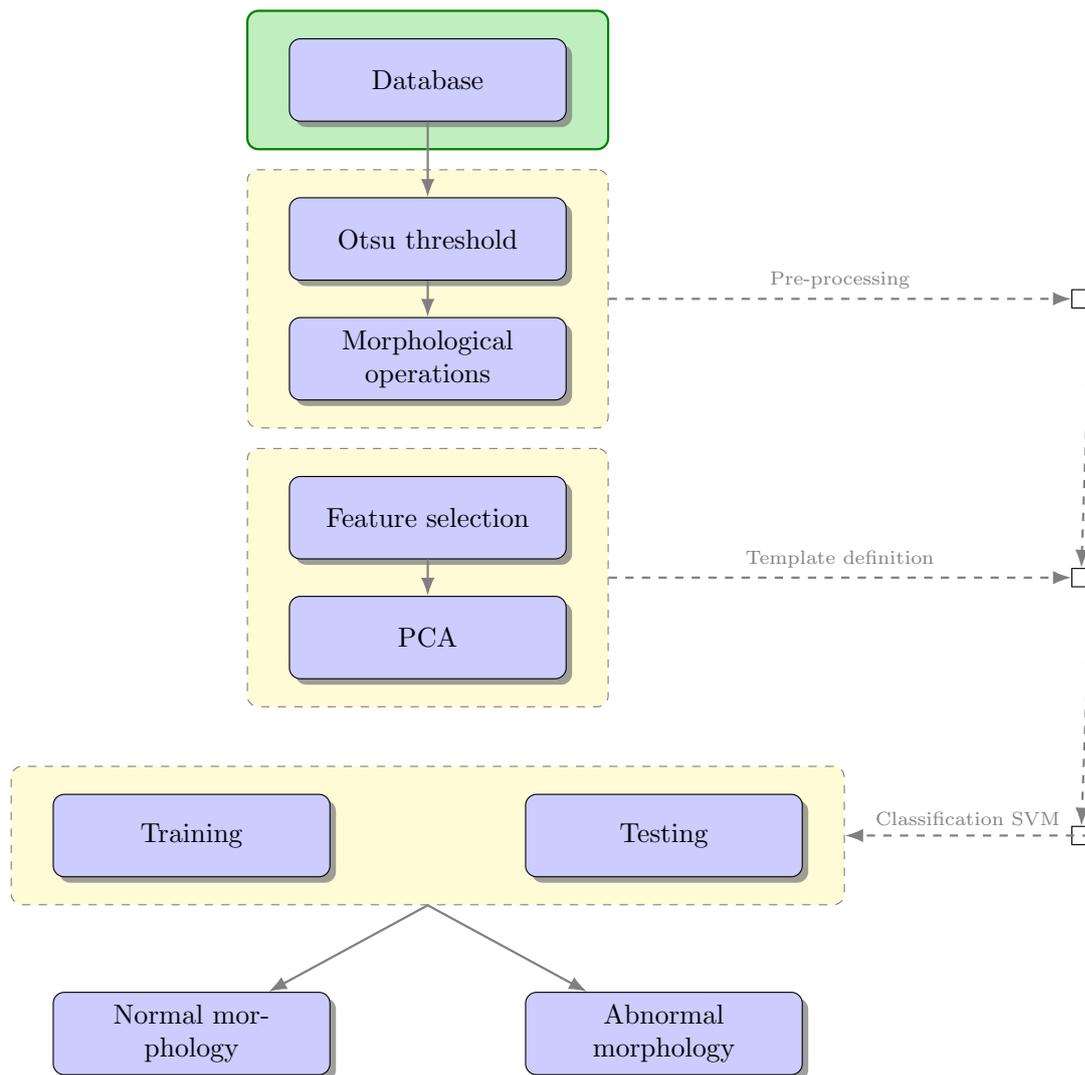


Figure 3.1: Morphology classification with SVM

3.2 Database

Images from sperm cells non-stained and non-fixed are difficult to access, for confidentiality issues and because the laboratories do not have technologies for recording and storing this kind of images.

In order to get images from sperm cells, 7 video sequences of fresh and unstained sperm were recorded. The videos were 30 seconds in length, rate of 10 frames per second (fps) and 2040x1086 resolution.

The recording process was taken at the Academic Unit of Reproductive and De-

velopmental Medicine (University of Sheffield) using a camera Basler model acA2000-50gc (Gholographic Company, Rajasthan, India) mounted over microscope Zeiss Axioskop 40 (Carl Zeiss Microscopy, LLC, United States). Sperm samples were obtained from six volunteers, with previous consent. All of the procedures had been approved by the University Research Ethics Committee (Ref: SMBRER293).

Since the obtained videos are compressed in *.avi* format, they can be decomposed into a sequence of images for their analysis. The images are coded into de RGB colour space, so they are formed by three individual monochromatic images representing each of the colour space components. This allowed pixels being parametrized by their (x,y) location, intensity and time in the video sequences.

3.3 Preprocessing

The objective is dividing the content of the image into wanted and unwanted objects where sperm cells as regarded an wanted objects and the rest of the pixels are taken as background or unwanted objects. Three main approaches were implemented: segmentation based on threshold, based on edges and region-based. The theoretical background of this techniques have been presented on Section 2.2.1, 2.2.2 and 2.2.3.

Based on the morphology of sperm cells, they can be classified according to WHO (World Health Organization), defects in the sperm's head mean -to some extent- a degree of abnormality producing less ability to fertilise. Also, after literature review, it concluded that the shape of sperm's head is a determinate variable (see Section 2.7)

Thus, the initial stage of the process is detecting and isolating sperm's head. In Digital Image processing, segmentation can be obtained by analysing pixels by their intensities, to detect shapes, textures, discontinues. In the next sections, some of the work completed approaching segmentation of sperm head is detailed.

As a previous step to segmentation, the technique of histogram equalization [28] has been applied to images from database created. The objective is adjusts image intensities to enhance contrast and enhance differences between object of interest and background.

Figure 3.2 shows the result of this processing over original image extracted.

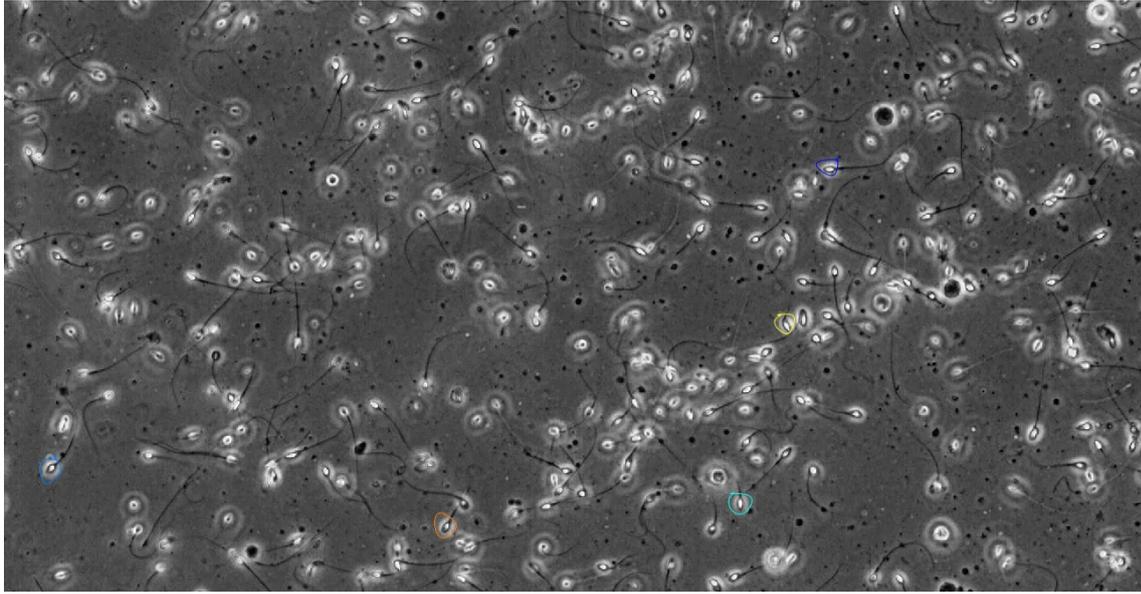


Figure 3.2: Histogram equalization applied in original image.

3.3.1 Segmentation with Otsu

Digital images are usually corrupted by noise. Changes in illumination of the environment and hardware are some of the most common responsible of introducing noise to captured images. As part of the preprocessing stage, noise reduction with the least loss of essential data has been carried out.

The Otsu thresholding [95] method aims to find a deep valley between two peaks in a histogram distribution, in a two class segmentation one peak represents background, and other represents the distribution of the pixels of interest.

The segmentation method using automatic Otsu threshold was applied, using images from two sources: first, the database created from non-fixed and non-stained sperm cells, and the second, images from WHO's manual for sperm cells analysis [1], with fixed and stained sperm cells.

First, the image was changed from RGB colour space to grayscale, as is shown in Figure 3.3. Red crosses are showing centroids of cells detected.

The histogram distribution with the position of the peak calculated by Otsu automatic threshold applied to the image presented on Figure Figure 3.3 is shown in Figure 3.4.

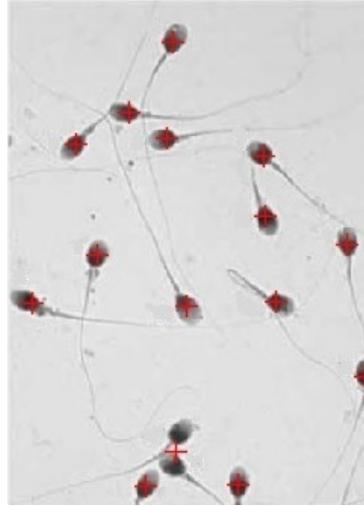


Figure 3.3: Grayscale image from the WHO's manual database, red crosses are centroids detected.

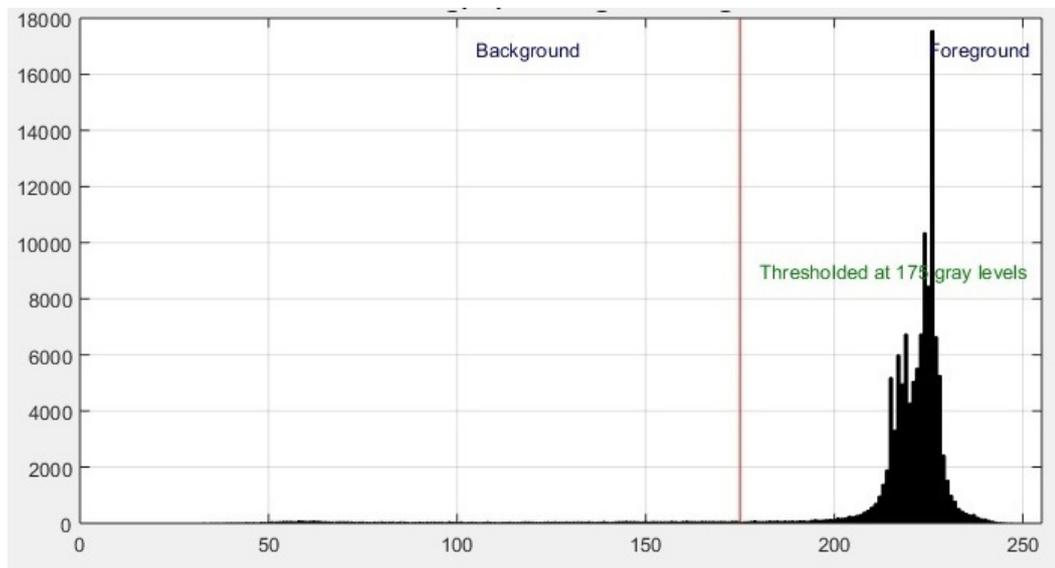


Figure 3.4: Histogram distribution and threshold for image from the WHO manual.

Background and sperm are separated with high accuracy, as shown in Figure 3.5.

Using an image from the database created, where sperm cells are non-fixed and non-stained, the grayscale image is shown in Figure 3.6. Also, red crosses are showing centroids. Centroids have been calculated based on the extraction of features from blob detected, explained in the next paragraphs.

Figure 3.7 shows the histogram distribution for the image in Figure 3.6.

Applying the same method, automatic Otsu threshold, onto images from the database created, there are sperm cells correctly segmented. However, the segmentation cuts some area belonging to the sperm's head. Furthermore, there are many



Figure 3.5: Sperm cells and background separated, the image is from WHO.

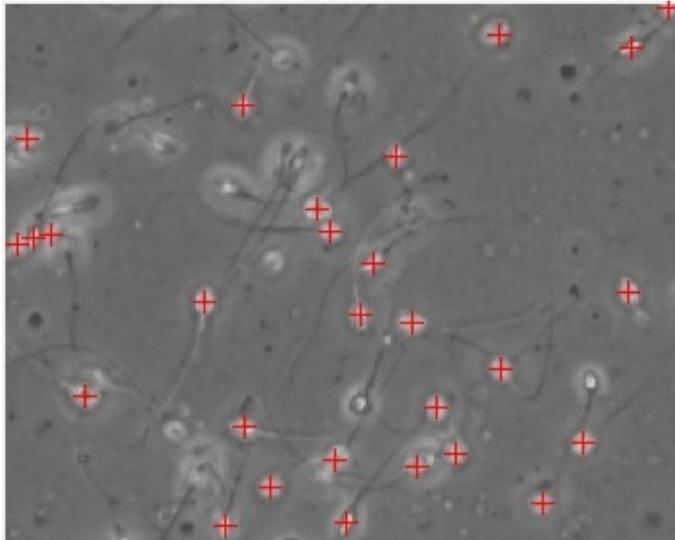


Figure 3.6: Original image from the database created, converted to grayscale.

cells missed. This is shown in Figure 3.8.

Afterwards, a binary image can be obtained by applying the Equation (3.1) to every single pixel to evaluate its grey level.

$$g(x, y) = \begin{cases} 1 & f(x, y) > T \\ 0 & f(x, y) \leq T \end{cases} \quad (3.1)$$

where $f(x, y)$ is the grey-level at the point (x, y) of the processed image and T is the threshold. An object is defined as 1 and the background as 0.

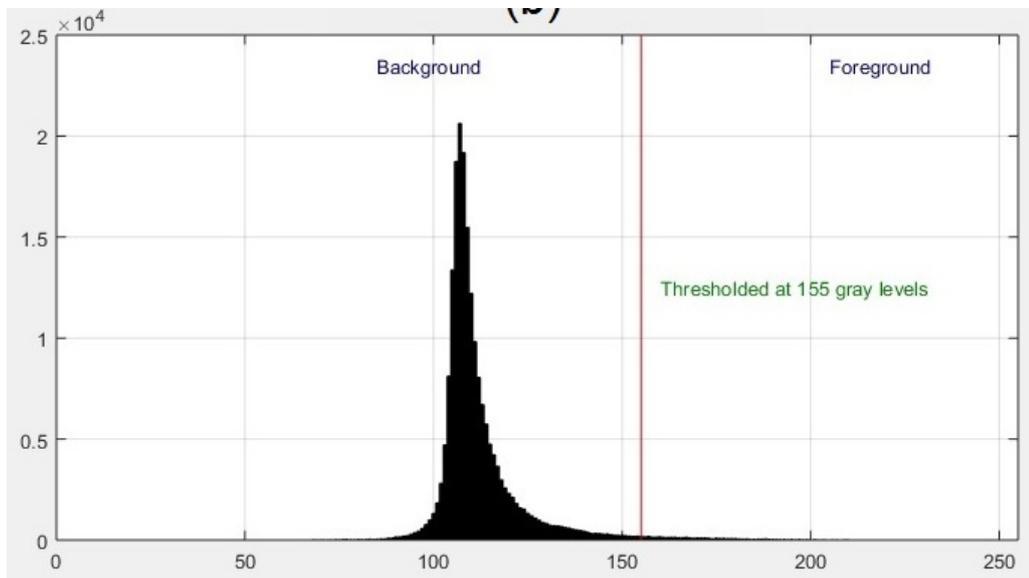


Figure 3.7: Histogram distribution and threshold for image from the database created.



Figure 3.8: Background extracted, image from the database created.

The binary image can have “holes” (a set of background pixels that are out of reach by filling in the background from the edge of the image). Some morphological operations are applied to get a better definition of the sperm cells shape extracted from the background.

After subtracting the background and filling holes, connected pixels are segmented as *blobs*. A set of properties for each detected blob in the image was estimated:

- *Area*: (A) defined as number of pixels forming the blob.
- *Major Axis Length*: (MaAx) length of the major axis of the ellipse around the blob.
- *Minor Axis Length*: (MinAx) length of the minor axis of the ellipse around the blob.
- *Eccentricity*: (E) ratio of the distance between the focal point of the ellipse and its major axis length. The value is between 0 and 1, 0 is actually a circle.
- *Convex Area*: (CArea) numbers of pixels in the smallest convex polygon that can contain the blob.
- *Filled area*: (FArea) number of pixels correspond to the blob, with all holes filled.
- *Extent*: (Ext) scalar that specifies the ratio of pixels in the region to pixels in the total bounding box. Computed as the Area divided by the area of the bounding box.
- *Solidity*: (Sol) scalar specifying the proportion of the pixels in the convex hull that are also in the region. Computed as $\text{Area}/\text{ConvexArea}$.
- *Entropy*: (Entr) statistical measure of randomness that can be used to characterize the texture of the blob segmented.

The list of intensity for each pixel belonging to each blob and their mean of this list (Mean Intensity, MeanInt), are also extracted as part of the values to analyse.

The Otsu thresholding and blobs segmentations methods were applied to 30 frames of a video sequence from the video captured. This process was repeated using 5 different video sequences.

The distribution of the values extracted from the properties of blobs detailed above is summarized in the Figure 3.9.

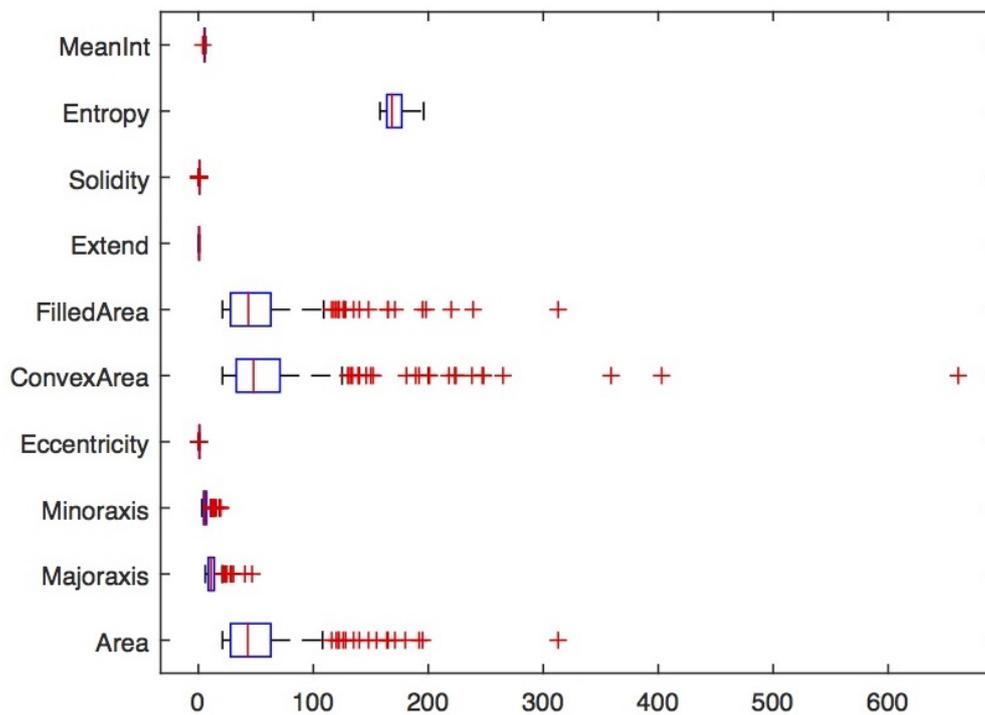


Figure 3.9: Correlation between values of features

Along the x axis is the numerical value after being normalized. The y axis is the description of each feature calculated.

The objective of this graph is identified how the variables are correlated, and how the values for each feature is distributed. Area, convex area and filled area have similar distribution.

The rest of the features are related between their as the values are in a similar range. Thus, any of these features can be used to distinguish sperm' heads from other components based on the properties of extracted blobs.

3.3.2 Edge detection

In Figure 3.10 the original image RGB and the edge detection using Prewitt [38] and Canny [96] are shown. Although Roberts [37] and Sobel [28] were also tested, the results does not improve the segmentation. Within the image, head's sperm can be defined as an elliptical shape with lighter colour inside.

Edge detection alone does not provide enough information to enhancement and further detection of the head of sperm cells.

3.3.3 Segmentation based on K-means

The K-means algorithm was used to segment using clustering approach to detect sperm's heads. Theoretical background about this technique is detailed on Section 2.2.3.

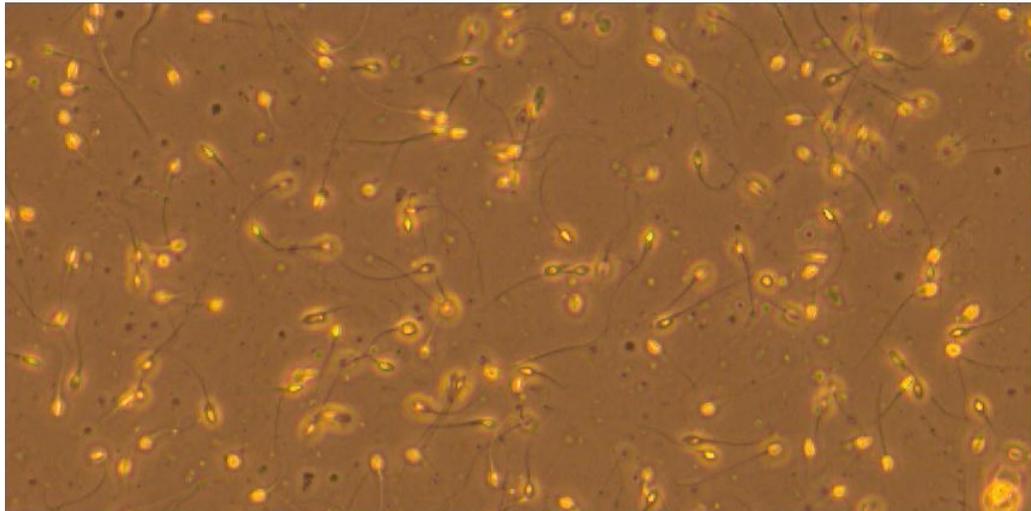
In this test images were converted from RGB to to the HSI (hue saturation intensity) colour spaces. HSI decouples the intensity component from the colour carrying information (hue and saturation) in to a colour image [28]. Sperm head has highest intensity in sperm sample image.

Cityblock -defined as the shortest distance in unit steps along each axis between two points- and square euclidean distance were evaluated. The initial position of cluster centroid were set using three different methods:

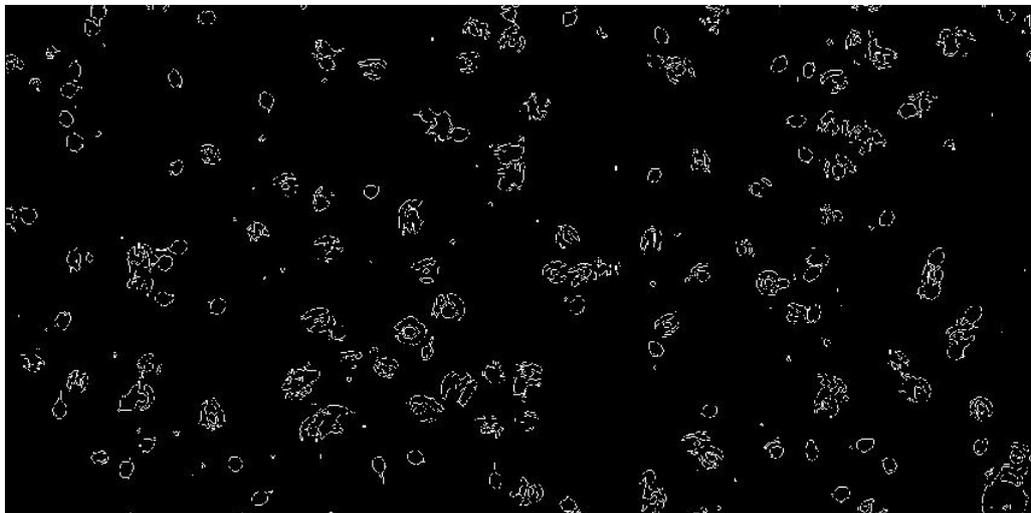
- Sample: randomly selection of K observations from data
- Uniform: selecting K points uniformly from the range of data
- Cluster: performing a preliminary clustering phase using a random sub-sample, usually 10% of data.

The K-means clustering algorithm was used to separate of the sperm's head cells from the background based on the intensity of pixels at the HSI colour space. Some results obtained using K=3 are shown in Figure 3.11.

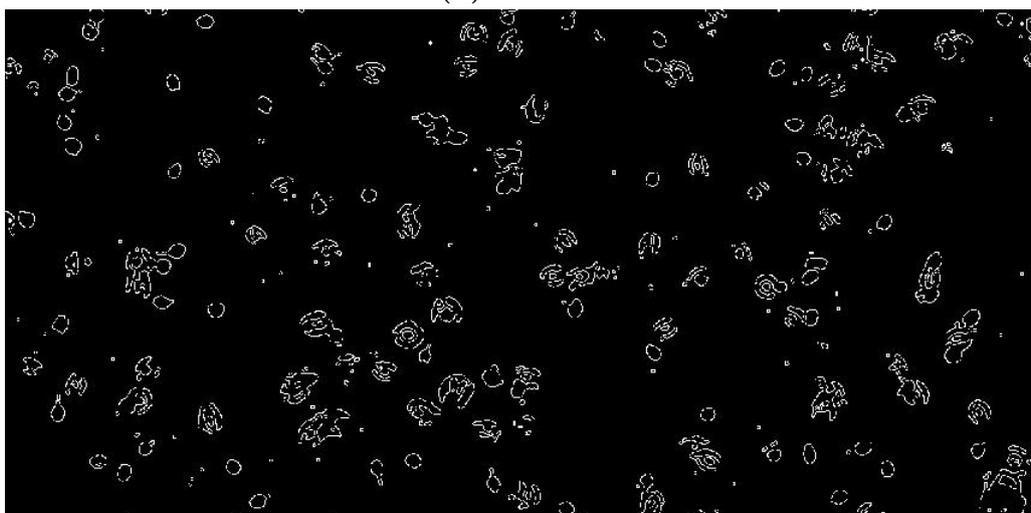
K-means for clustering segmentation give us good information from sperm's head, however this is part of pre-processing process aims to get accurate measurements.



(a)



(b)



(c)

Figure 3.10: Edge detection of sperm cells using methods based on derivatives and noise reduction: (a) Original image; (b) Prewitt; and, (c) Canny.

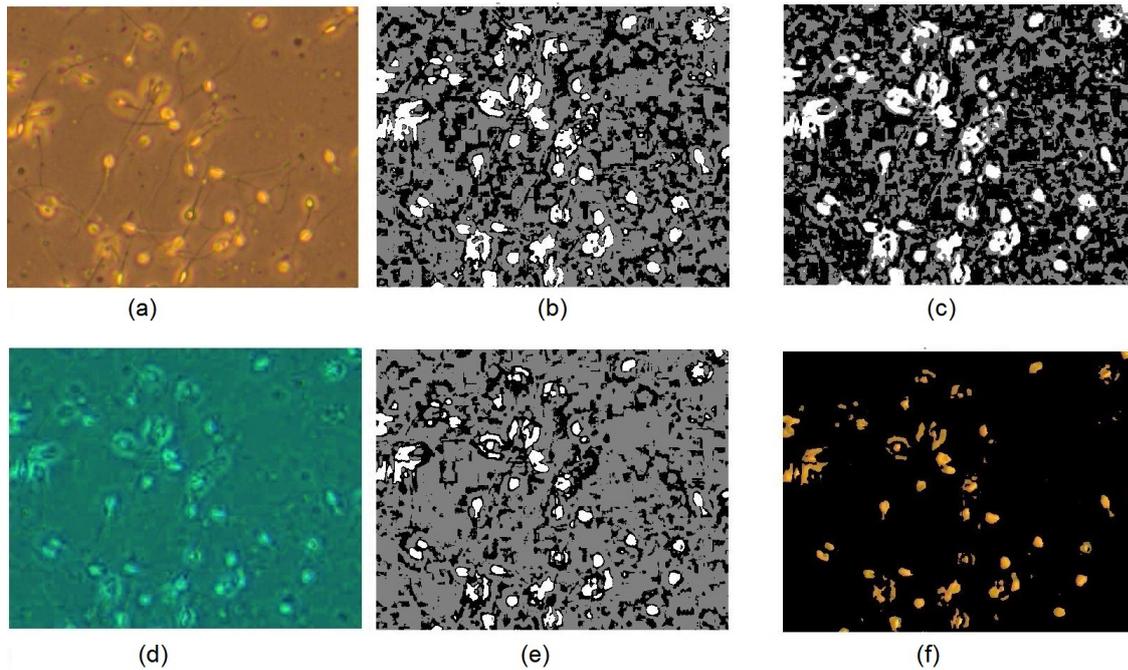


Figure 3.11: Results varying the K-means distance and seeds location, (a) Original Frame; (b) City block distance, sample; (c) City block distance, uniform; (d) Frame using HSI colour space; (e) Square Euclidean distance, plus and (f) Sperm head segmented

3.3.4 Principal Component Analysis

The use of PCA in this research can contribute to reduce computational cost as can reduce dimensionality and remove redundant data.

A prior feature selection and feature extraction are crucial to get better classification. In feature selection, a set of the original measurements is selected making the *Feature Space*.

On the other way, feature extraction is just the combination of measurements to have a new reduced dimensionality of a measurement space[97]. Principal Components Analysis is a common example of feature extraction method.

The Principal Components Analysis (PCA) aims to reduce the excessive dimensionality of data by combining features. This approach seeks a projection that best represents the data in a least-square sense [98]. The information is no redundant because, spatially, each principal component is orthogonal to each other in the space of the data.

Given a set of data on n dimensions, PCA aims to find a linear subspace of dimension d lower than n such that the data points lie mainly on this linear subspace. In Figure 3.12 an example of a two-dimensional projection found by PCA is shown. Such a reduced subspace attempts to maintain most of the variability of the data.

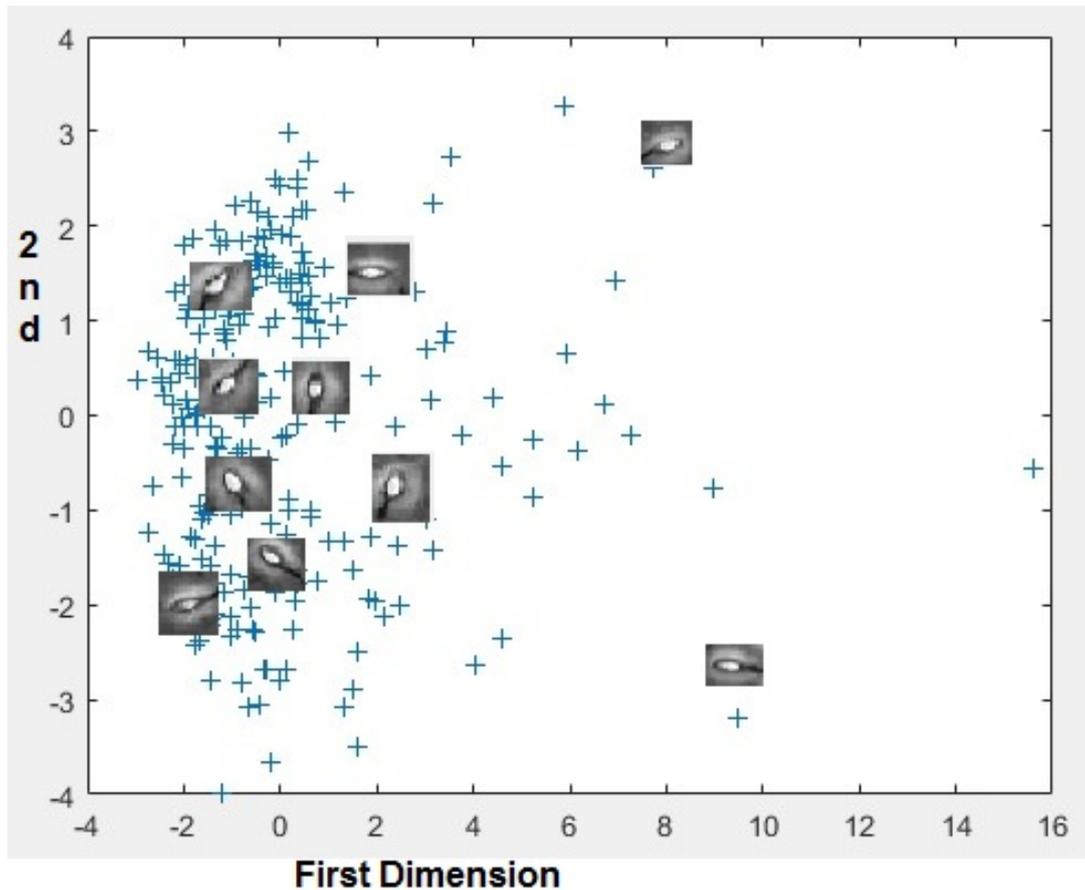


Figure 3.12: PCA applied to dataset, two dimensional projection

The linear subspace can be specified by d orthogonal vectors that form a new coordinate system, called the ‘principal components’. The principal components are orthogonal, linear transformations of the original data points, so there can be no more than n of them.

The objective is to find a vector x_0 such that the sum of the square distance between x_0 and the various x_k is as small as possible. Data obtained from blobs’ measurements (from a single frame) are processed using PCA. Their concept was defined in section 3.3.1

The vector of features extracted from each blob is defined by the value of:

1. Area (A)
2. Majoraxis (MaAx)
3. Minoraxis (MinAx)
4. Eccentricity (E)
5. ConvexArea (CArea)
6. FilledArea (FArea)
7. Extend (E)
8. Solidity (Sol)
9. Entropy (Entr)
10. Mean Intensity (MeanInt)

The initial step, prior to PCA is standardizing the data, subtracting the sample mean from each observation, then dividing by the sample standard deviation. This centers and scales the data.

Then, finding the eigenfunctions of the sample covariance matrix calculating the coefficients of the PCA and their respective variances [99]. The proportion of each Principal Component represents of the total variance can be observed in Figure 3.13.

This figure shows the first four components forming 98% of the total variance (area , majoraxis, minoraxis and eccentricity). There is a clear difference between the first and the second component. The first component by itself covers more than 50% of the variance, so that might be a reasonable way to reduce number of components used for the classification process.

Figure 3.14 shows a view of the data and the original variables (blobs properties) in the space of the first three principal components.

Component 1 (Area) has a wide distribution than the others components.

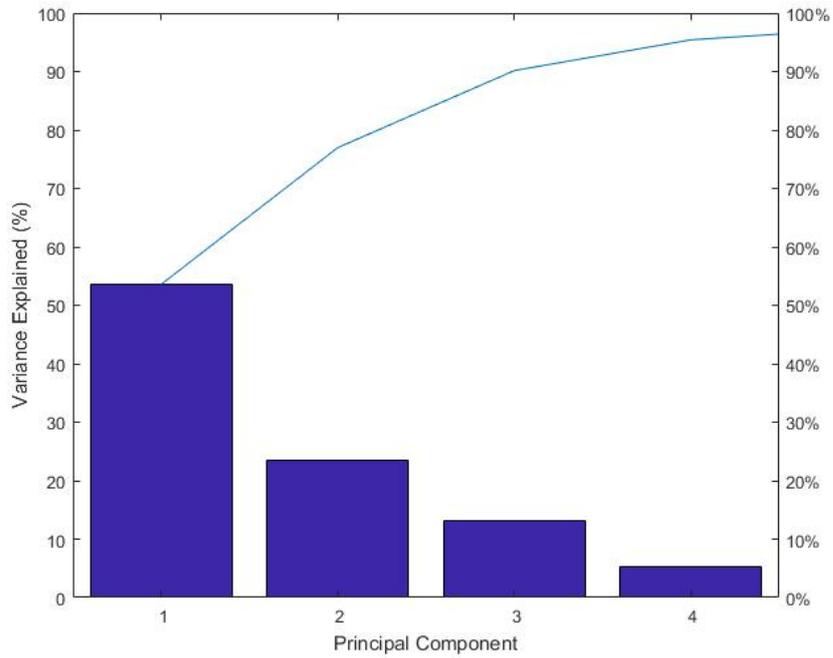


Figure 3.13: Percent of variability of the first four principal component. Line light blue is variance

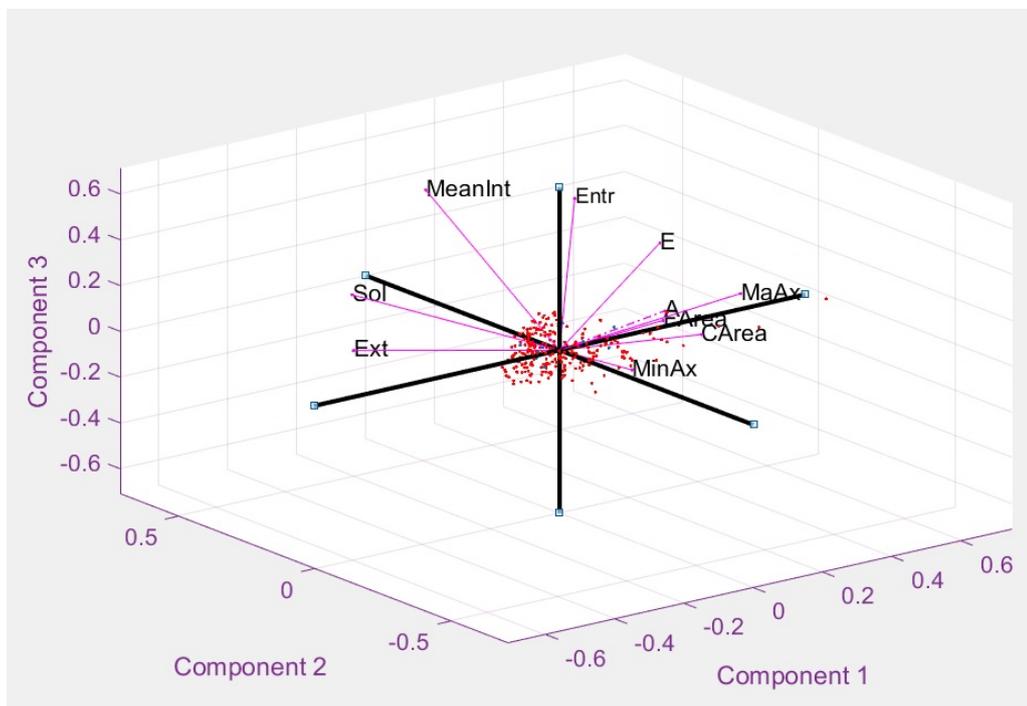


Figure 3.14: First three principal components and features distribution

3.4 Support Vector Machine

The SVM has been identified as a useful tool for sperm morphology. Based on the theory presented on 2.4.2 is presented as solution for classification.

Images taken from the WHO [1] manual were used as a ground truth. The reason for this is having images validated as normal and abnormal morphology.

The process for the SVM classification was carried out as follows:

- Re-scale and centre data, training and testing.
- Select a different kernel function: linear, polynomial, gaussian (radial basis function), RBF sigma.
- Experiment with different method to find the separating hyperplane: quadratic programming and least square.
- Train and predict with 10 features per instance.
- Train and predict using only Area and Eccentricity (based on result obtained from PCA analysis).
- Train and predict using only Area and Entropy.

The training process was performed using a sample of 400 detected blobs, to be classified into two classes: normal and abnormal. The “normal” class set 1 and “abnormal”, 2. For the prediction process a new data set of 100 samples was used.

The proportion was 80% for training and 20% test. The sample has 500 images of sperm cells.

Results of the classification using Area and Eccentricity shown better performance. The samples wrongly classified are circled in red. Classification using a linear kernel can be shown in Figure 3.15.

Results obtained using a polynomial kernel, gaussian (RBF) kernel and sigmoid can be shown in Figure 3.16, Figure 3.17 and Figure 3.18 respectively.

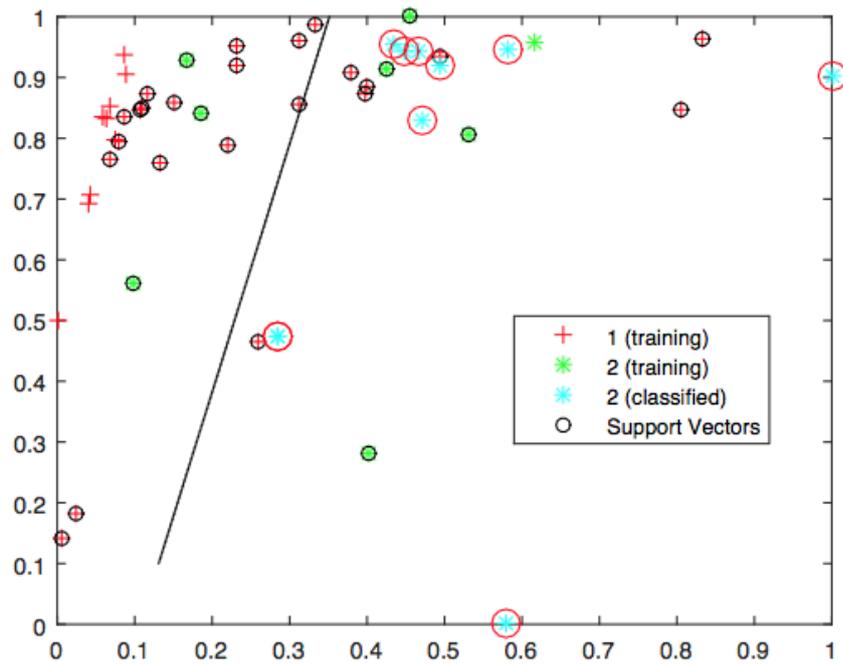


Figure 3.15: Prediction using linear kernel.

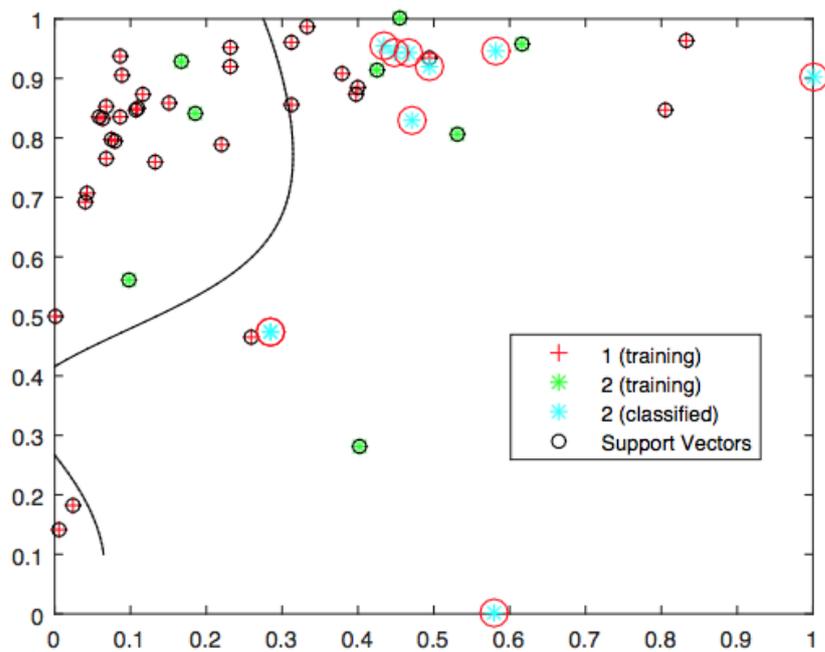


Figure 3.16: Prediction using SVM polynomial kernel.

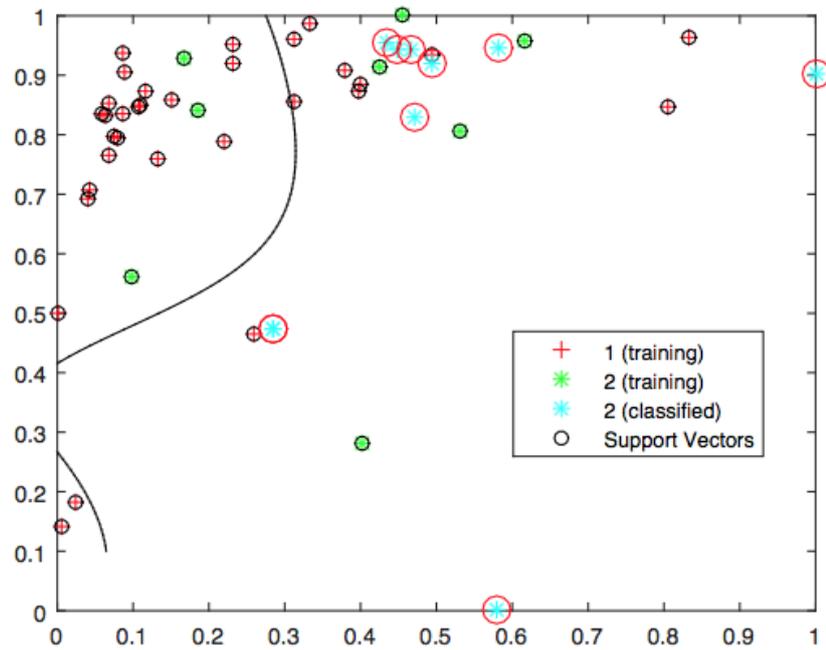


Figure 3.17: Prediction using SVM gaussian kernel.

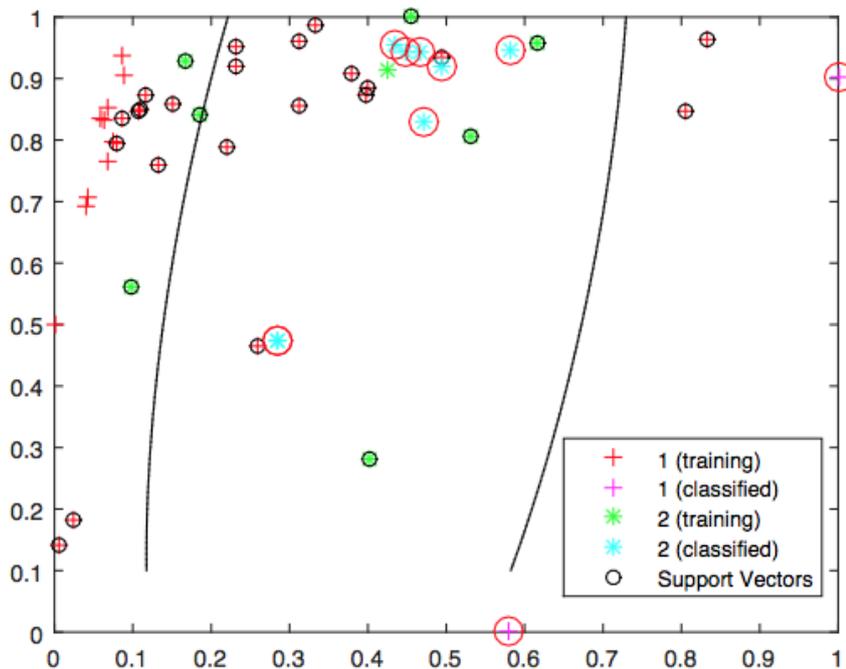


Figure 3.18: Prediction using SVM with sigmoid kernel.

3.5 Performance Evaluation

After training SVM with 4 different kernels, 100 new sperm cells were classified. Performance obtained is based on predicted labels and true labels. Precision (Press) is the proportion of cases correctly identified as belonging to class 1 among all cases of which the classifier claims that they belong to class 1. Recall is the proportion of cases correctly identified as belonging to class 1 among all cases that truly belong to class 1. And Accuracy is a measure to know overall, how often is the classifier correct.

Base on the number of blobs classified as True Positives (TP), True Negative (TN), False Positive (FP) and False Negative (FN), Press, Recall and Accuracy are obtained as follow:

- Precision = $TP / (TP+FP)$
- Recall = $TP / (TP+FN)$
- Accuracy = $(TP + TN) / (TP + TN + FP + FN)$

Results analysing 10 features are shown in Figure 3.19. Performance analysing just Area and Eccentricity are shown in Figure 3.20. And performance analysing Area and Entropy are shown at Figure 3.21. Results obtained have good performance, however number of elements used for the training process is limited, thus prediction have low accuracy.

3.6 Summary

This Chapter presents an initial approach to the problem of analysis and classification of sperm cells. The image processing technique allows separate background from the sperm cells (equalization, Otsu thresholding, filling holes, extraction of blobs). However, there is a high grade of missed information as some sperm cells cannot be segmented under this approach.

Then, the analysis of features with PCA allows identified the best combination of features. Finally, the SVM model was trained and tested with new samples given



Figure 3.19: Precision (Press), Recall and accuracy (Acc) obtained analysing 10 features

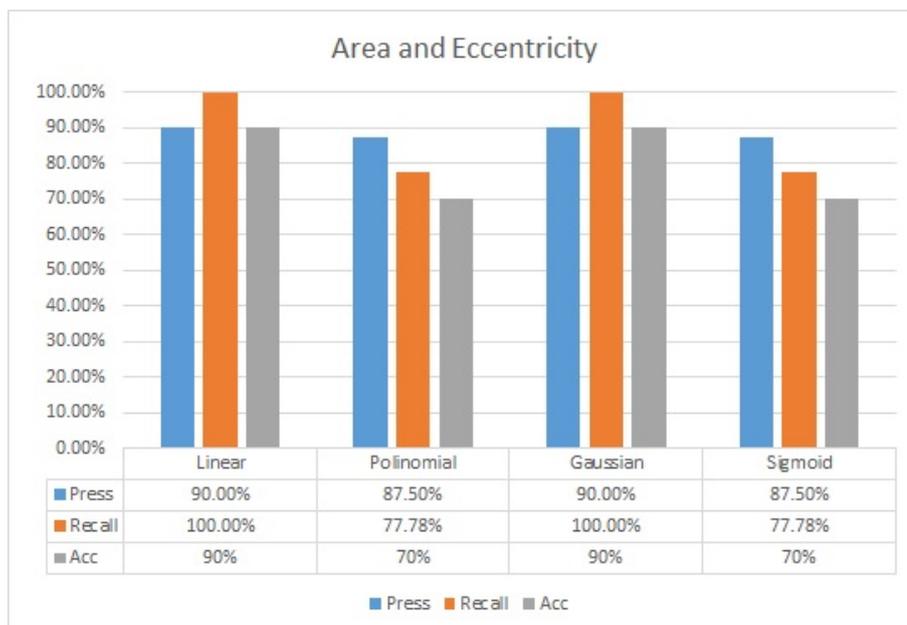


Figure 3.20: Precision (Press), Recall and accuracy (Acc) considering Area and Eccentricity

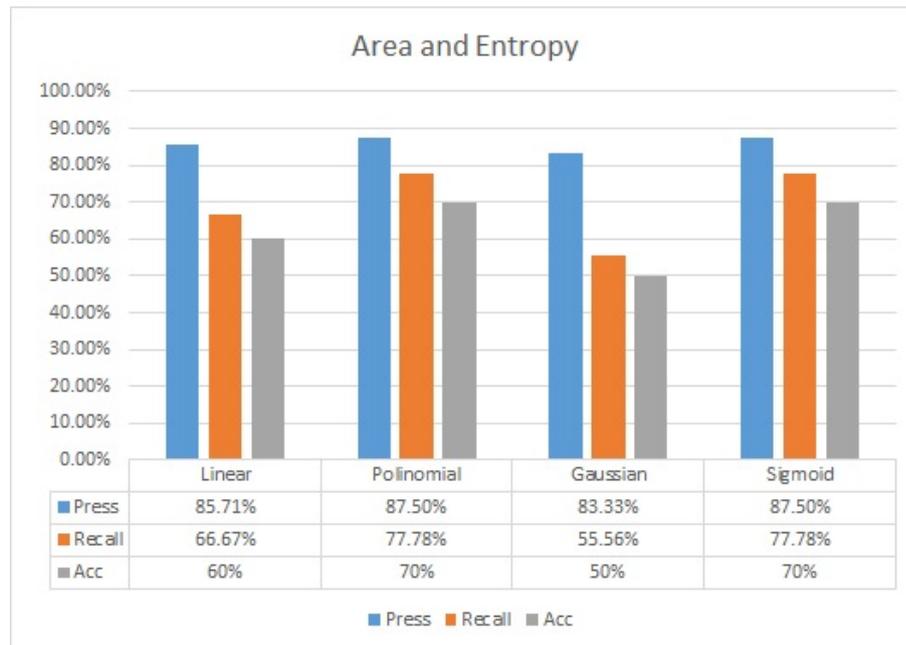


Figure 3.21: Precision (Press), Recall and accuracy (Acc) analysing Area and Entropy

an acceptable accuracy of %90 classification for normal or abnormal morphology. This performance is obtained using the features: area and eccentricity with linear kernel and Gaussian kernel.

However the parameters and techniques can be optimized to get better results. Particularly, the use of SVM for classification is a bit uncertain as there are several kernels, threshold and function customizable depending of the nature of the input data.

In this Chapter the features selected using PCA demonstrated to be robust and stables, however the cost factor C when SVM is trained impact directly into the accuracy of classification.

Chapter 4

Speeded-up Robust Features and Bag of Features

4.1 Introduction

This chapter presents a new approach for classification of in-vitro sperm cells. The framework proposed is based on Speeded-Up Robust Features (SURF) combined with Bag-of-Features (BoF) models to quantise and simplify the large amount of low-level features computed by SURF.

The objective of the framework presented is first, identify non-sperm cells (see Figures 2.12 and 2.13 in Section 2.5 for examples), sperm cells with normal morphology and sperm cells with abnormal morphology.

Support Vector Machines (SVMs) are used to classify the simplified feature vectors, extracted from sperm cells images, into normal, abnormal and non-cell categories. The performance of this framework is compared to a similar model where Histogram of Oriented Gradients (HOG) is used to extract features and SVMs is used for their classification.

After the analysis of classification results from Chapter 3, has been detected an apparent change on the shape of the same sperm cell along the video, as a consequence the same sperm cell has been wrongly classified with different label.

Figure 4.1 shown how two sperm cells, one in the circle other into the square, are changing their size, shape and intensity along 5 consecutive frames. This variations

are consequence of movement, as the sperm cells are alive, swimming in a water-based medium, their visual characteristics are affected.

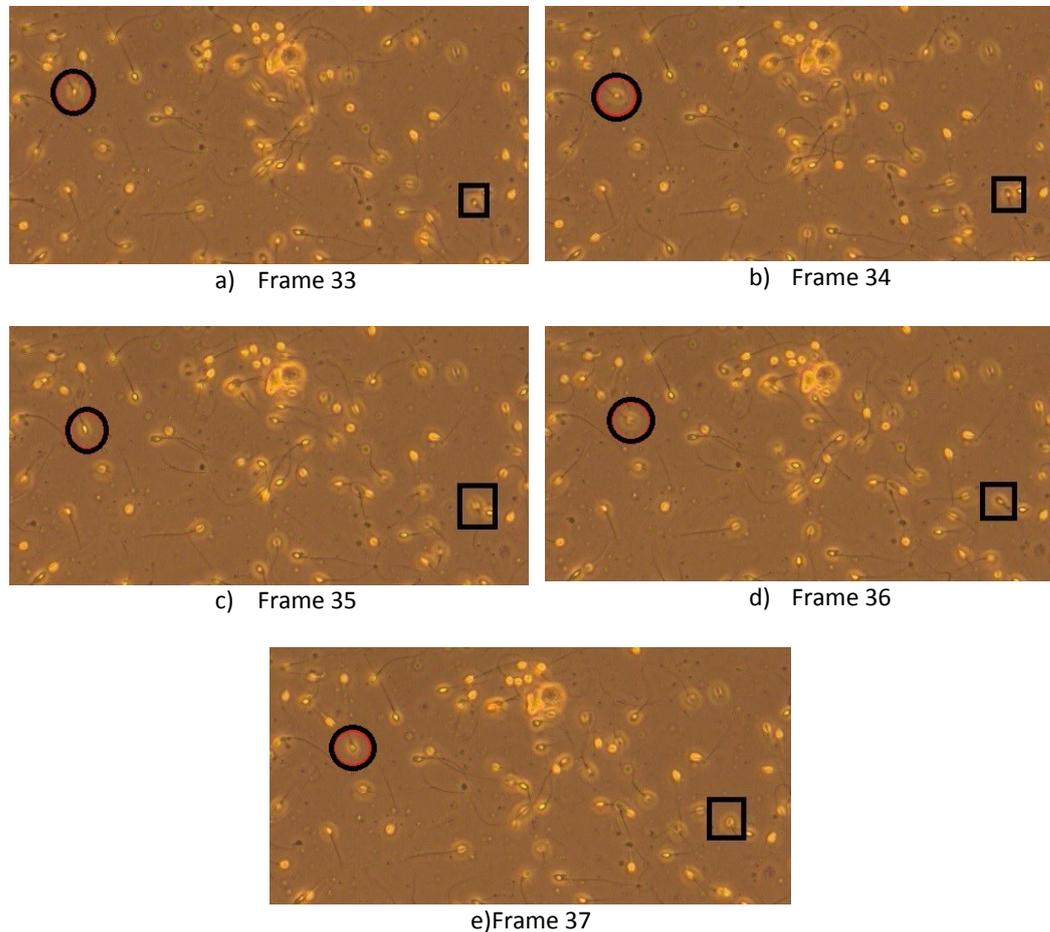


Figure 4.1: Apparent change in shape, size and intensity of two sperm cells along 5 consecutive frames

After the analysis of previous results, has been possible detect that the appearance of sperm cells in video sequences can change due to the conditions at every scene captured. Variation of illumination resulting from the flow of the seminal fluid, for instance, can produce shades in the background.

A specific visual phenomenon knowing as ‘halo’ surrounding sperms head can be observed in some images making heads edge more or less distinguishable. For the class normal-h, the letter -h means the presence of ‘halo’. In order to overcome this visual effect, the classes defined to classify sperm cells based on their morphology are:

- Normal. Normal morphology criteria and with to some extent sharp heads edge.
- Normal-h. Normal sperm cell with a blurred heads edge.
- Abnormal. Does not meet the normal morphology definition.
- Other. Non sperm cells.

The proposed solution is show in Figure 4.2

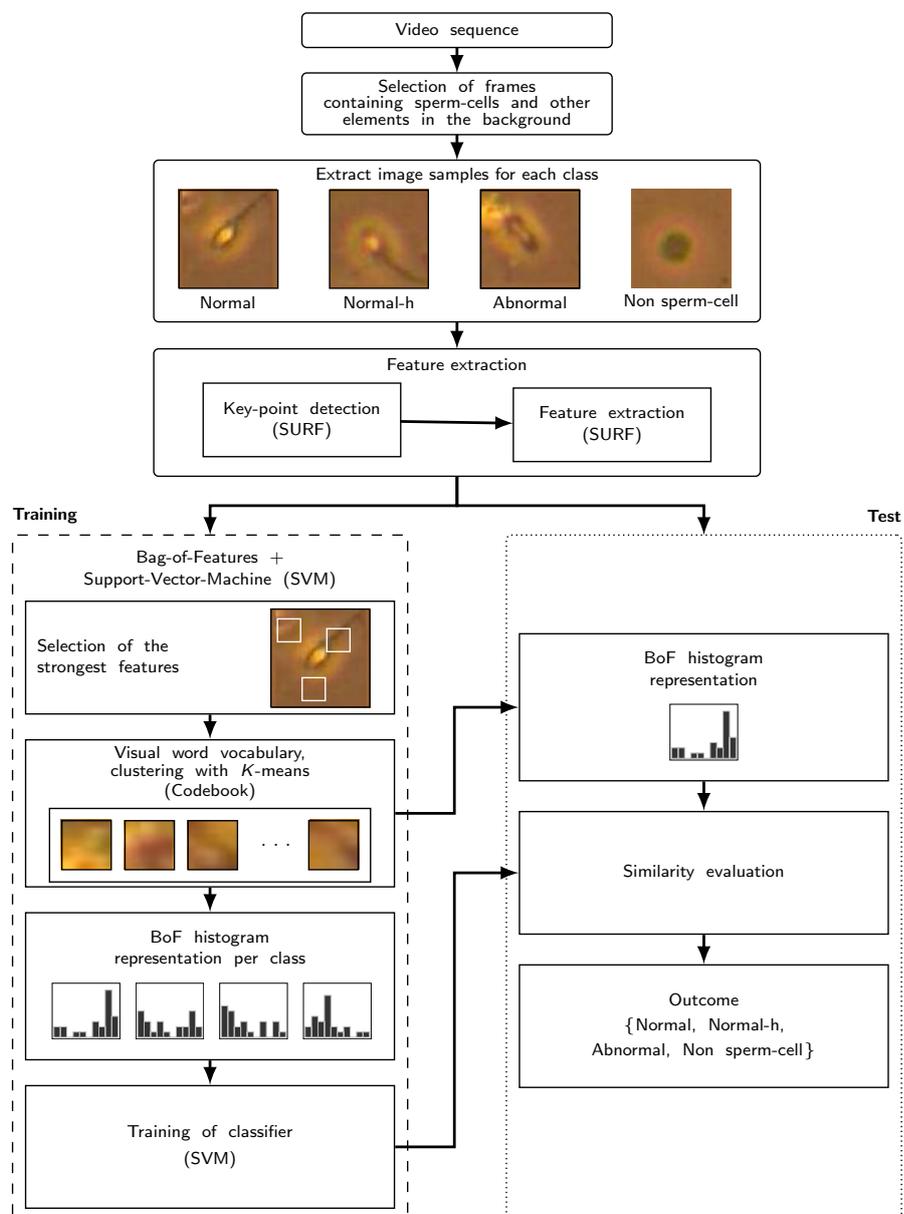


Figure 4.2: The classification method based on SURF, BoF and SVMs.

4.2 Histogram of Oriented Gradients

HOG can solve the apparently variation of motile sperm cells because are image descriptors invariant to 2D rotation as take localized portions of the binary image and counts occurrences of gradient orientation. This method is computed on a dense grid of uniformly spaced cells and uses overlapping local contrast normalization for improved accuracy.

To compute the HOG of a given image, it is divided into small regions termed cells. Then the orientation of gradients are computed and the histogram of the gradients is calculated. Gradient can be performed by filtering the image using a Sobel-based kernel $D_x = [-1 \ 0 \ 1]$ and $D_y = [-1 \ 0 \ 1]^T$. The concatenation of all histograms produces the feature vector of the image.

A visualization of HOG features extracted from a normal sperm cells, using size cells 4x4 and 8x8 are shown in Figures 4.4 and 4.5 respectively. The original image is shown in Figure 4.3.

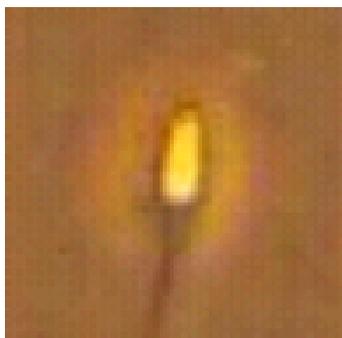


Figure 4.3: Original image, sperm cell with normal morphology

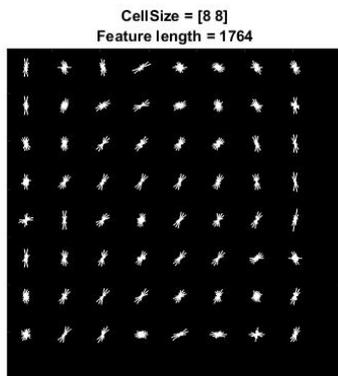


Figure 4.5: HOG features with cell 8X8 Prediction using linear function

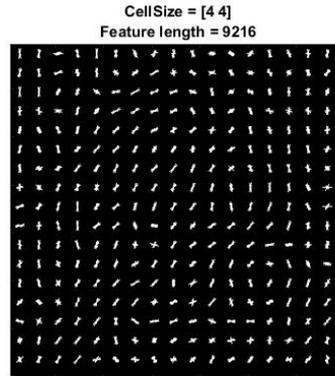


Figure 4.4: HOG features with size cell 4X4

4.3 Feature Extraction

The SURF [45] method is used to detect key points from the images. SURF selects the most representative pixels based in low-level features—using the Hessian matrix. Even though SURF is invariant to rotation, sperm cell samples at different orientations are used to consider the pixellation effect when capturing the images.

For every key point, the SURF features are computed using a region of $m \times n$ pixels encompassing the key point location yielding to the initial feature vector. More details about the theoretical fundament of SURF has been presented in Section 2.3.2

The number of features is reduced by selecting the minimal number of features n found across the four classes. The SURF method is used to choose the n —strongest features from each image. Thus, the same number of features for each image across all classes can be transferred to the training process of a classifier.

4.3.1 Bag-of-features (BoF)

Processing large SURF vectors can represents a high computational cost. Therefore, the BoF model is employed to reduce the feature representation of the images.

More about Bag of features approach has been presented on 2.4.1. The aim of BoF is creating a codebook—also termed Visual Word Vocabulary—by transforming feature vectors into visual words. In order to do so, the SURF feature vectors are clustered by using K -means. The number of clusters K defines the codebook size

and the centroids of clusters define the words.

Therefore, each feature of the SURF vectors is assigned to one word. This can be represented by a histogram where the x -axis represents the words and y -axis the number of features mapped to the words (see Figure 4.6). The BoF method is repeated to produce the histogram of BoF for each of the four classes defined. The resulting BoF model is used to train the classifier.

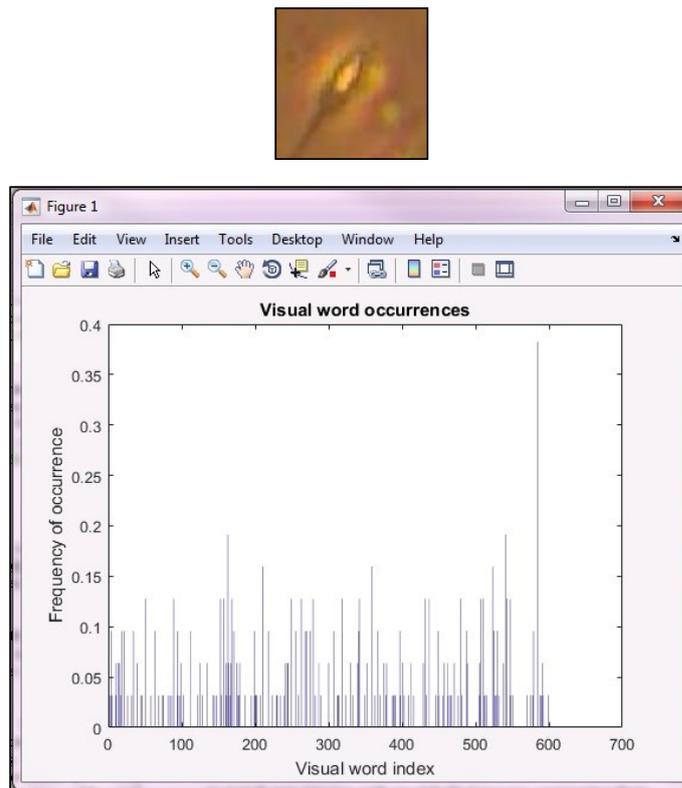


Figure 4.6: Template for normal cell and the distribution of the ratings data. The x -axis represents the words and y -axis the number of features mapped to the words

4.4 Training a Support Vector Machine (SVM)

It is important to highlight our objective: classifying non-stained, non-fixed sperm cells extracted from a video sequence suffering from occlusion, collision, uneven illumination and presence of debris. Aiming to fulfill this requirement, a dictionary

of templates has been defined as the source of the ground truth for this framework.

The dictionary of images is generated after the selection and extraction of multiples templates for normal, abnormal and other components wide enough to include all the variations of each class.

Selection of templates is done manually from original frames. The bank of images defined as template are for the training stage in the form of Bag of Features (BoF).

Each element for training is related to one class-label and several features. The aim is get a model after a training process able to predict target values of the test data using only features.

A multi-class Support Vector Machine (SVM) classifier is trained with the BoF model obtained. SVMs are inherently binary classifiers. Thus, an approach is set-up to define the class for a new entry. A 'one-to-one' model is used to define the outcome based on the most voted class by the binary classifiers. Eventually, the trained SVM model is used to process BoF histograms instances from the training dataset.

4.5 Performance Evaluation

A collection of video frames showing sperm cells and other objects were selected to create two datasets: training using 70% and testing with the 30% of the samples.

The dictionary of templates is generated after the selection and extraction of multiples templates from image patches.

Image patches contain a single sperm cell with one unique label. The patches were manually selected to include the possible variates that can be found in practice (*e.g.* orientation and morphology of sperm cells and other objects in the background). Figure 4.7 shown how templates have been selected from extracted frames.

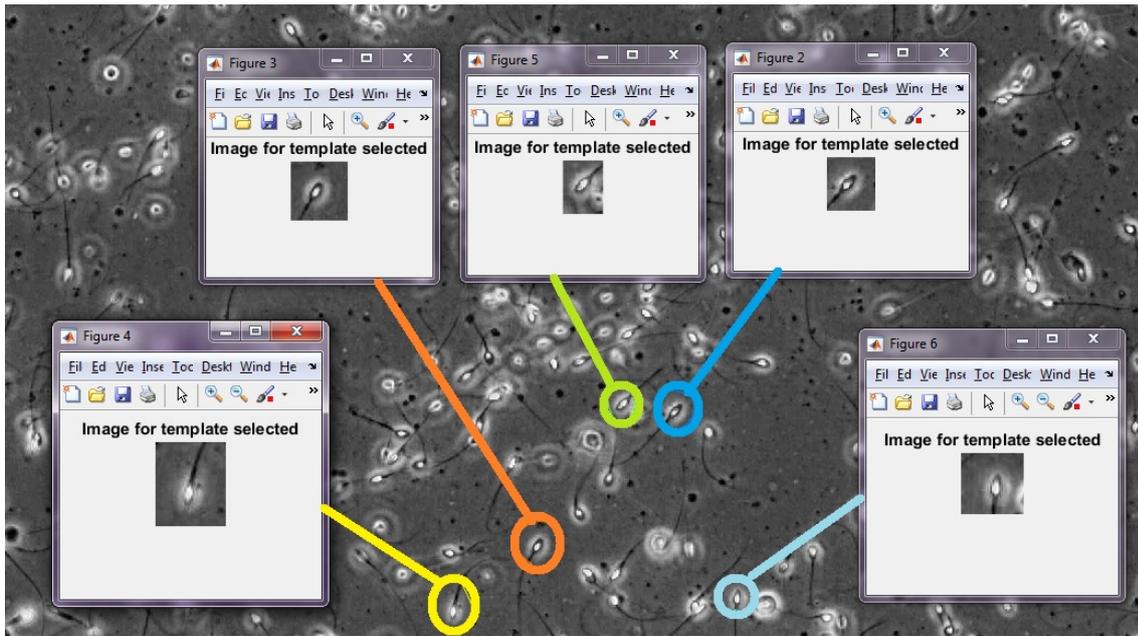


Figure 4.7: Selection of multiples templates for normal class

The testing dataset was used to validate and measure the performance of the method proposed. For each image in the testing dataset, the BoF histogram is computed as follows: Using the BoF model—codebook, a histogram is computed by mapping every SURF feature to a word in the codebook. The distance $l_2 - norm$ is used to estimate the closest distance between a feature and the words in the codebook. The resulting histogram is normalised and classified by the SVM model which eventually measures the similarity to the class models obtained in the training process.

The database is the presents before on Section 3.2. Sperm cell samples for both, training and testing, were extracted from frames which were selected by using a random function. A full dataset—images of all classes—was formed by: 235 images labelled as *Normal*; 145 *Normal-h*; 105 *Abnormal*; and, 172 images containing other objects considered as *Non-cell* objects. The labelling process was based on the WHO guidelines and the visual appearance. The images ranged from 60×60 to 70×85 pixels in size and were compressed into JPEG (Joint Photographic Experts Group) format.

105 images were used for each class since it is the smallest number of samples for one class in the full dataset. This process aims to make a fair comparison between

the number of features for each class in the training process. The full dataset was divided into two datasets: 70% of the images were used for training and 30% for testing.

Parameters of the implemented methods were hand-crafted to obtain the performance described in this section. 95% of the strongest SURF features were used to choose a reduced number features which eventually were used by the BoF method. A $K=2325$ was used in the K -means clustering process to create a BoF model which was passed to an SVM classifier—based on a $K(K+1)/2$ kernel—for training. This K was calculated automatically based on the general K -means algorithm presented on Section 2.2.3.

The performance obtained with the framework proposed reached a 90% in average accuracy. Other performance measures are summarised in Figure 4.8.

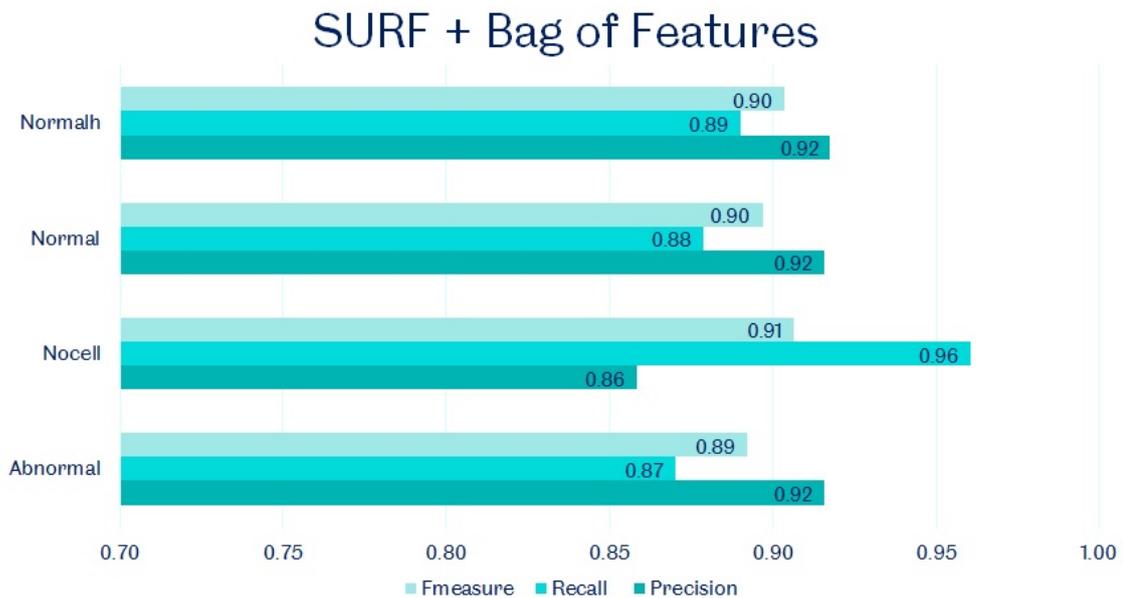


Figure 4.8: Performance of the proposed classification method based on SURF and BoF.

The performance of the framework proposed was compared to an approach based on HOG combined with SVM, where HOG was used to extract feature vectors directly from images with an average accuracy 78%.

4.6 Summary

In this Chapter a fast and efficient method for sperm cells classification combining Bag of Features model where, using K-means, SURF features have been clustered into visual words. The best performance was obtained when we defined a vocabulary size equal to 2300 keeping 95 percent of strongest features for codebook. We aimed at achieving a good classification rate with an average of 90% percent in accuracy. After some experimentation, it is possible conclude that size of vocabulary and width of the patch have a direct impact on the classification results. The method proposed showed better performance and accuracy against classification using HOG features.

Chapter 5

Deep Neural Networks for Morphology Analysis

5.1 Introduction

This Chapter presents a novel approach for analysis of non-stained and non-fixed sperm cells. As far as the author is aware, there is no published work carried out analysing of this kind of images using deep convolutional neural networks.

There are important challenges to solve :the resolution offers by images of sperm cells are affected by noise and uneven illumination. Moreover, the layer of pooling in convolutional neural networks reduce resolution even more, as has been explained before in Section 2.4.3. In the other hand, CNN is a technique could be overfitting easily.

The challenge of accuracy on classification is incremented by the nature of living cells affected by movement causing effects like cell elongation, expansion and shrinkage [100].

In Figure 5.1, for instance, a set of two *normal* sperm cells along a sequence of five consecutive frames is shown. Variation in size and brightness can be observed as sperm cell swims.

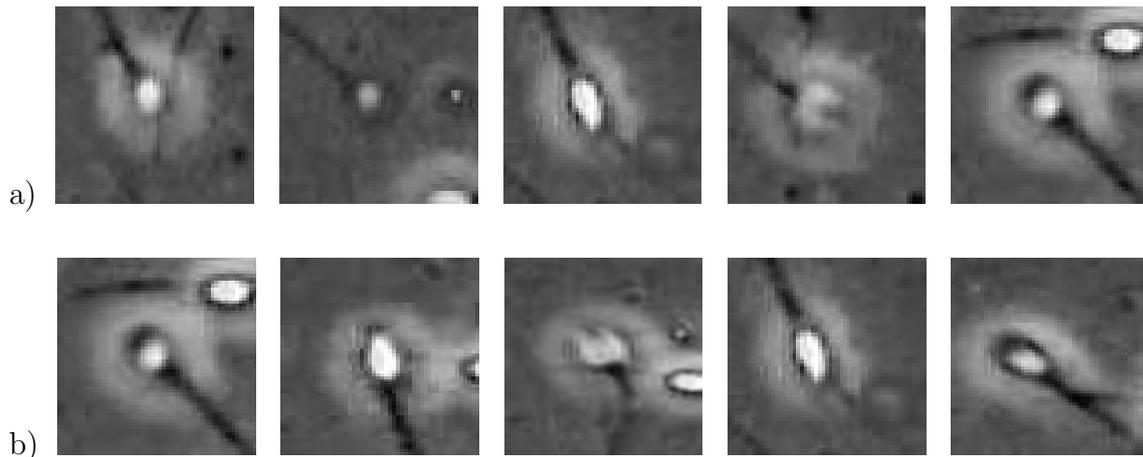


Figure 5.1: Appearance variation of two sperm cells (upper and lower row) along consecutive video frames $t \in [0, 1, \dots, 4]$.

The hand-crafted features representations such the presented in Chapter 4 and Chapter 5 (HOG, SURF, SIFT, Bag of Features) represent lower accuracy rates for classification; the reason: deep learned features are automatically learned from the images or videos. Furthermore, CNN can be invariant to translation, viewpoint, size, illumination or a combination of all of them.

Each layer in the architecture of CNN extract features at a different level. An image of sperm cells, for example, is initially represented like an array of pixel values, and the first layer of representation give information about edges their orientation and location.

In Figure 5.2 a visualization of the weights networks filters from first convolutional layers, is shown. This visualization gives a clue about how the CNN extract features and how this technique “sees” hidden characteristics.

The second layer can detect changes on the edges positions if they have particular distributions. Usually the existence of edges, blobs and basic image features, with location and orientation is information extracted from the first layer. The second layer also, can identifies a particular arrangement from the edges.

The third, could combine repeated patterns to form larger combinations. The layers of features are learned from data using a learning procedure without intervention of human [101].

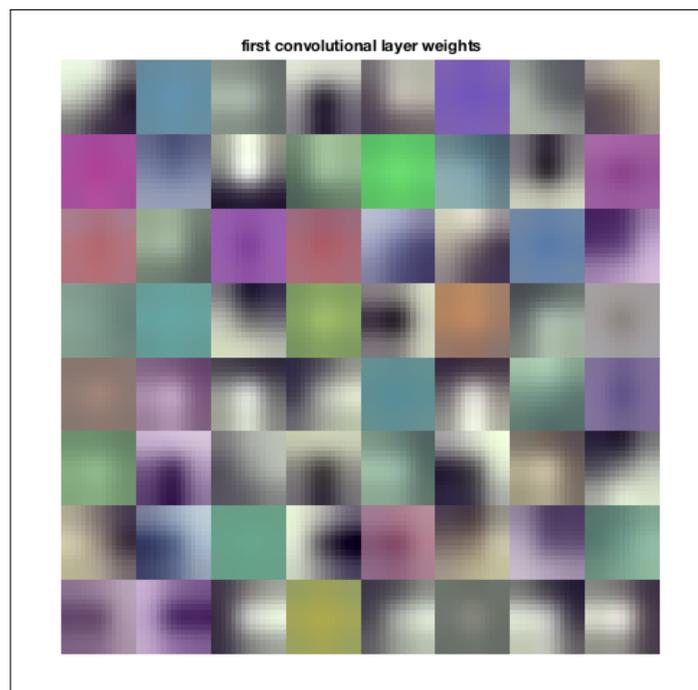


Figure 5.2: First Convolutional layer weights from sperm cell image

Table 5.1: Details of data augmentation

Class	Original Size	Operations	Final Size
1. Normal morphology	470	Mirror vertical =470 Rotation 90° =470	1410
2. Abnormal morphology	173	Rotate 90° =173 Rotate -45° =173 Gaussian blur 3x3 =173 Mirror Hrztl. =173 Rotate 180° =173	1038
3. Non-sperm	235	Mirror Hrztl = 235 Deskew image = 235 Mirror vertical = 235 Rotate 180° = 235	1175

5.2 Dataset

Training a deep CNN requires a large amount of pre-labelled data. Also, computational and memory resources. Microscopy video sequences were recorded at a rate of 10 frames per second producing 8-bit colour images of 2040 x 1086 pixels in size. Patches with a single sperm cell have been extracted manually from a video sequence where debris, complex background and other components are present.

The original dataset created for Chapter 4 has been augmented in order to reduce over-fitting. Details are shown in Table 5.1.

Sperm cell samples for both, training and testing, were extracted from frames which were selected by using a random function. A full dataset images of all classes was formed by: 1410 images labelled as Normal; 1038 as abnormal cell; and, 1175 images containing other objects considered as Non-cell objects.

The labelling process was based on the WHO guidelines and the visual appearance. The images ranged from 60×60 to 70×85 pixels in size and were compressed into JPEG (Joint Photographic Experts Group) format. In this chapter, 1038 images were used for each class since it is the smallest number of samples for one class in the full dataset, making a fair comparison between the number of features for each class in the training process. A sample of the cropped images used for training is shown in 5.3.

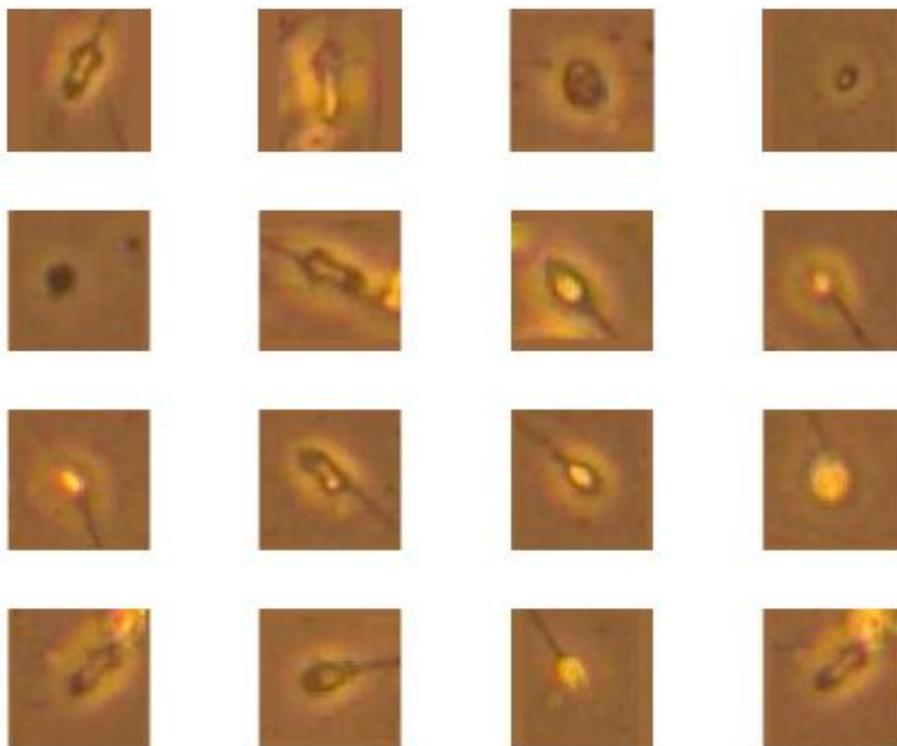


Figure 5.3: Original images used for training

The full dataset was divided into two datasets: 70 % of the images were used for training and 30% for testing. The same dataset is the input for the two approaches presented: a new deep convolutional neural network and a pre-trained deep neural network.

5.3 Training a new model

The first approach presented is the construction of an own DNN and training of a new model, from the beginning. The architecture implies multiples linear and non-linear transformations changing the data representation into something more useful and abstract. The objective is creating a learning network able to extract features, train the networks and then use it for new data prediction.

For training and classification using DNN the hardware used in this thesis is: iMac, processor 2.7.Ghz, Intel Core i5, Memory 8GB 1600 Mhz, Graphics NVIDIA GeForce GT 640M, 512 MB.

Training a deep CNN is often complicated by over-fitting and convergence issues, whose resolution frequently requires repetitive adjustments in the architecture or learning parameters of the network to ensure that all layers are learning with comparable speed. The architecture of the DNN is shown in Figure 5.4.

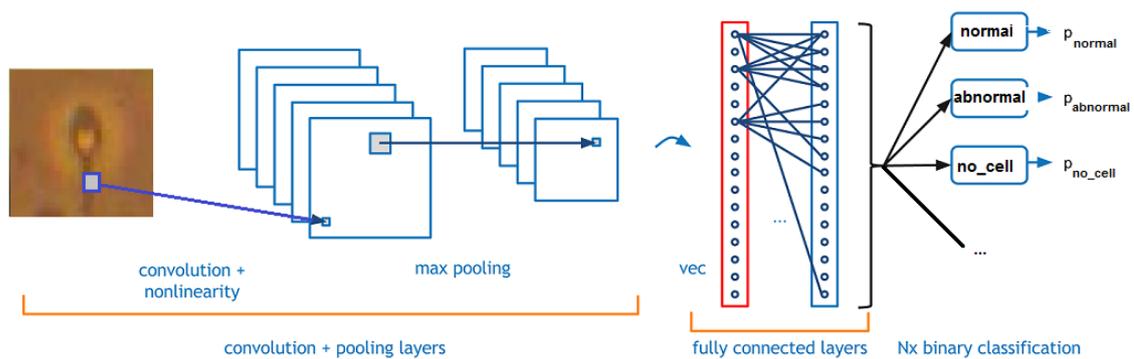


Figure 5.4: The proposed architecture for DNN

The training and has been monitored to stop early the learning by the model and evaluate how accurate and quick is the model in each batch. Figure 5.5 shown how the training status per epoch. For fine-tuning, the network has to be changed step by step. As part of the experimentations different learning rates have been tested. The learning rate concept has been already presents in Section 2.4.3.

The architecture found with better performance, with training and learning from scratch is shown is Figure 5.6.

Different combinations of convolutional layer, filter, pooling layer and finally classification has been tested. The best performance reached are summarised on

```

Training on single CPU.
Initializing image normalization.

```

Epoch	Iteration	Time Elapsed (seconds)	Mini-batch Loss	Validation Loss	Mini-batch Accuracy	Validation Accuracy	Base Learning Rate
1	1	10.09	1.2223	2.2361	33.33%	38.95%	0.0010
1	50	758.01	0.0053		100.00%		0.0010
1	100	1028.77	0.5618		66.67%		0.0010
1	150	1302.99	1.0856		66.67%		0.0010
1	200	1575.08	0.3047		66.67%		0.0010
1	250	1846.46	0.0363		100.00%		0.0010
1	300	2126.70	0.1623		100.00%		0.0010
1	350	2403.43	0.0319		100.00%		0.0010
1	400	2679.23	0.0506		100.00%		0.0010
1	450	2959.78	0.0117		100.00%		0.0010
1	500	3243.70	0.0007		100.00%		0.0010
1	550	3540.41	0.1322		100.00%		0.0010
1	600	3819.85	0.1139		100.00%		0.0010
1	650	4104.07	0.0814		100.00%		0.0010
1	700	4383.69	0.1611		100.00%		0.0010
1	750	4662.17	0.0284		100.00%		0.0010
1	800	4940.49	0.2597		100.00%		0.0010
1	845	5183.04	0.6802	0.2202	66.67%	91.90%	0.0010

```

Elapsed time is 6144.395937 seconds.

```

Figure 5.5: Monitoring the learning process and validation accuracy of the model

1	1 × 1 ImageInputLayer
2	1 × 1 Convolution2DLayer
3	1 × 1 ReLULayer
4	1 × 1 MaxPooling2DLayer
5	1 × 1 Convolution2DLayer
6	1 × 1 BatchNormalizationLayer
7	1 × 1 ReLULayer
8	1 × 1 MaxPooling2DLayer
9	1 × 1 Convolution2DLayer
10	1 × 1 BatchNormalizationLayer
11	1 × 1 ReLULayer
12	1 × 1 MaxPooling2DLayer
13	1 × 1 Convolution2DLayer
14	1 × 1 BatchNormalizationLayer
15	1 × 1 ReLULayer
16	1 × 1 Convolution2DLayer
17	1 × 1 BatchNormalizationLayer
18	1 × 1 ReLULayer
19	1 × 1 FullyConnectedLayer
20	1 × 1 SoftmaxLayer
21	1 × 1 ClassificationOutputLayer

Figure 5.6: Structure for one new model of DNN

Table 5.2: Parameters and Accuracy with DNN created from scratch

Layers	Epochs	m	Initial l.r.	MBatch	Accuracy		
					N	A	O
3(conv+norm+relu+pool) +conv+norm+relu+fully+ s ofsoftmax+class	4	0.9	1e-4	10	90.03	97.11	94.86
2(conv+norm+relu)+fully+ softmax+class	30	0.9	1e-2	128	76.02	84.02	92.01
transfer+fully+softmax+class	6	0.9	1e-4	40	91	91.20	88.3
4(conv+norm+relu+pool)+fully+ softmax+class	7	0.9	1e-4	10	82.86	94.29	87.14

Table 5.2 where parameters (epochs, momentum, initial learning rate, miniBatch size) defined in Section 2.4.3 have been adjusted by test and error.

The mean accuracy for each class is showed (N=normal morphology; A=abnormal morphology; O= others). The combination of the layers specified in the table in the first row trained using 10 epochs with a mini-batch size of 10 reach the average accuracy of 94%.

5.4 Evaluation of pre-trained model

A different approach is presented in this section, known as transfer learning. Transfer learning is defined as exporting knowledge from previously learned source to a target task [34].

The solution presented is shown in the Figure 5.7. The CNN model is not going to be used for the original classification task. It is going to be re-purposed to solve a different classification task on the sperm cells dataset.

As is shown in the Figure 5.7 the steps followed for training and testing using transfer learning are:

1. Pre-process input images, have to be RGB images with size 227-by-227 pixels.
2. Divide data into training 70% and testing sets 30%
3. Load the pre-training network, in this case AlexNet. The last three layers must be fine-tuned for the new classification problem, then all layers have been extracted, except the last three from AlexNet.

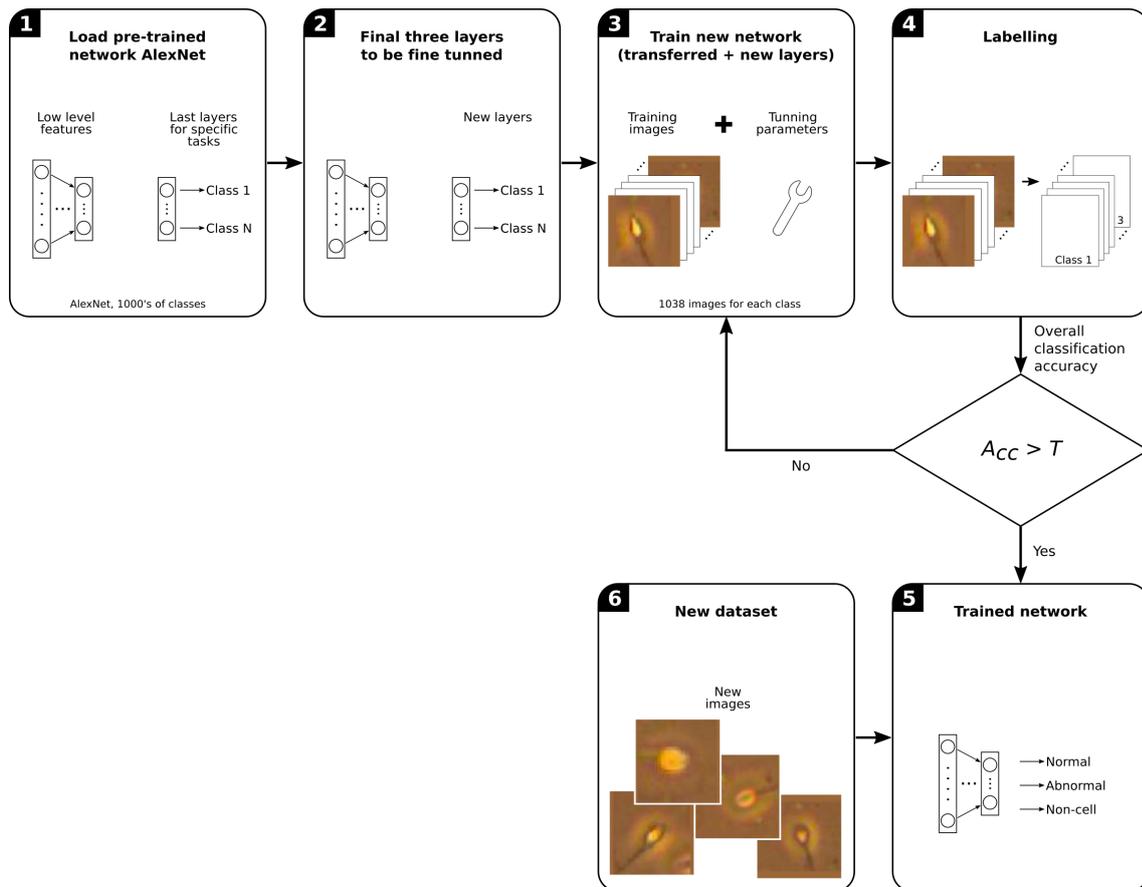


Figure 5.7: Transfer learning process

- Transfer the layers to the new classification task by replacing the last three layers with a fully connected layer, a softmax layer, and a classification output layer.
- Specify the options of the new fully connected layer according to the new data.
- Set the fully connected layer to have the same size as the number of classes in the new data. In this case, there are 3 classes: normal cells, abnormal and non cell.
- To learn faster in the new layers than in the transferred layers, increase the learn rate factor and the bias of the fully connected layer.
- Specifies the layer to extract features, in this case experimentation which the layers: Conv3, Conv5, Fc6, Fc7 and Fc8. Usually the layer right before the classification layer is recommended.

Learning from sperm cells images from scratch is often not the most practical strategy due to its computational cost and the convergence problem [32]. The technique success of knowledge transfer embedded in the pre-training CNN, is affected directly for the measure of dissimilarity, between the database selected for train the CNN and the database to which the knowledge is to be transferred.

For this thesis the pre-trained AlexNet has been selected. AlexNet [102] is a public pre-trained deep network. Has been trained using 1.2 million images and is able to assign till 1000 labels. In the Figure 5.8 the structure of AlexNet is showed, begins with L1 and L2 two pairs of convolutional and pooling layers with 6×6 kernels, mapping the 227×227 input images to 13×13 feature maps. Then there are 3 convolutional layers (L3,L4,L5) that efficiently implement a convolutional layer with 9×9 kernels, yet with a larger degree of non-linearity. After this, there is a pooling layer and 3 fully connected layers (L6 and L7).

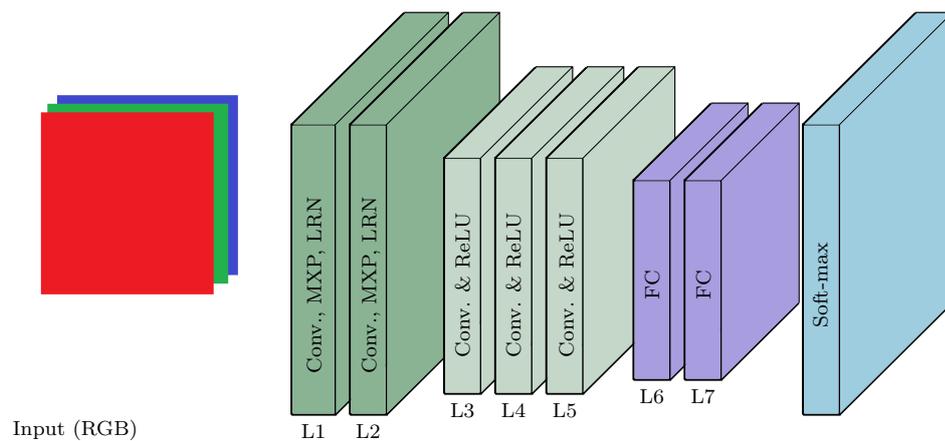


Figure 5.8: Structure of CNN AlexNet

Although in this case a model already trained has been used, some fine-tuning process have been adopted. The learning rates of the original layers are actually small so are keep in the same value. But, the learning rates on the new layers added has been boosted, so they can change faster than the rest of the network and learn is shortest time the weights of the newer layer.

In this case the knowledge of a rich set of features, provided by a pre-trained network AlexNet, is used to learn new patterns in images of sperm cells.

Is important to evaluate the model trained calculating the classification accuracy on the validation set. This value is giving the accuracy, the fraction of labels that the network predicts correctly.

5.5 Performance Evaluation

We trained our models using stochastic gradient descent with a batch size of 128 examples, momentum of 0.9, and weight decay of 0.0005. We found that this small amount of weight decay was important for the model to learn. In other words, weight decay here is not merely a regularize: it reduces the models training error. Multiples combination of parameters have been experimented, the results are showed on Figure 5.9.

Mini-batch size	Layer	Learning rate	Classification (%)			
			Abnormal	No-cell	Normal	Accuracy
6	Conv3	0.01	77.580	95.600	93.670	88.950
10	Conv3	0.01	79.230	96.420	95.870	90.507
60	Conv3	0.01	78.400	97.250	93.810	89.820
10	Conv5	0.01	88.310	94.910	94.770	92.663
10	Conv5	0.01	90.920	92.160	94.360	92.480
10	Conv5	0.01	51.310	92.570	99.040	80.973
40	Conv5	0.01	83.080	93.540	96.700	91.107
60	Conv5	0.01	92.600	97.440	99.290	96.443
16	Fc6	0.01	81.020	92.710	97.110	90.280
32	Fc6	0.01	76.200	96.010	92.300	88.170
40	Fc6	0.01	86.240	91.610	95.320	91.057
20	Fc7	0.01	95.050	82.670	89.680	89.133
20	Fc7	0.01	93.540	86.660	86.380	88.860
25	Fc7	0.01	83.610	88.550	96.510	89.557
32	Fc7	0.01	95.460	88.860	85.690	90.003
40	Fc7	0.01	89.680	84.320	93.950	89.317
50	Fc7	0.01	80.610	92.430	95.460	89.500
20	Fc8	0.01	87.900	94.770	70.700	84.457

Figure 5.9: Accuracy of classification with Transfer Learning

To evaluate the performance of the deep neural network designed , the features extracted from the training images are used as predictor variables and fit a multiclass support vector machine (SVM), using a fast linear solver, as a classifier. The results of classification using SVM are shown in Figure 5.10.

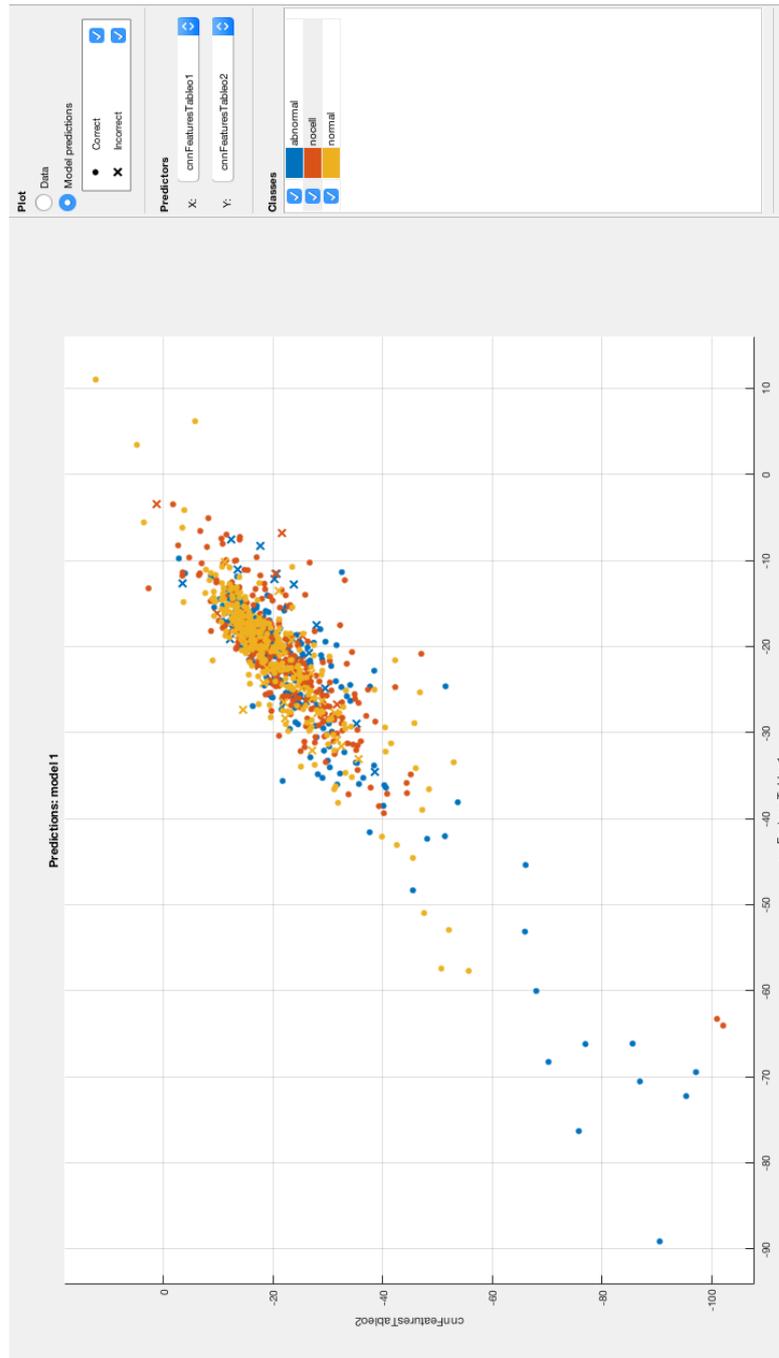


Figure 5.10: Classification with SVM based on the deep Neural Network trained

The SVM classifier reaches an accuracy of 96% with a quadratic kernel using a multiclass method one-vs-one, which is a good performance.

5.6 Summary

A solution for accurate and reliable classification of sperm cells based on deep neural network is presented in this Chapter. Design, training and evaluation of multiple architectures for Deep Neural Networks made from scratch are presented. After fine-tuning, the best performance with an average accuracy of 94% was achieved. However, the time for training to reach this percent was almost 50 hours.

Using transfer learning has demonstrated better average accuracy with 96.44%. Using a pre-trained network (AlexNet) improving performance and reducing processing time to 30 hours. The trained deep convolutional neural network is able to classify with high accuracy images of sperm cells without previous stained or fixed procedures.

Chapter 6

Classification of Motile *in-vitro* Cells

6.1 Introduction

This Chapter presents an initial approach for a robust and complete methodology for motile grade quantification for future works. The present approach is a combination of the classification proposed on Chapter 5 and a classic approach for movement detection.

There are several algorithms related with cell motility detection published with different focus: human sperm [88], neural stem cells [103], [32], stem cells [9]. They usually applied a specific procedure to enhance the edges and definition of cells, for sperm cells the use of fluorescent DNA stains is recommended by the WHO.

However, for this research non-stained sperm cells have been analysed. The challenges about identifying and classifying sperm cells in these conditions have been exposed before in Section 1.1 and, for the movement detection the complexity is bigger.

The WHO recommend the use of CASA systems to obtain movement parameters. Also, specified 200 as the minimum number of sperm cells to be analysed from the same specimen [1]. There is no formal agreement about for how long the sperm has to be tracked in order to calculate their motility grade, but the recommendation is at least 1 second [104].

This Chapter presents a combination between two frameworks that have been successfully used in the Chapter 5 and Chapter 3: deep neural network for classification; and features analysis for movement detection.

This approach has also the advantage of being computationally efficient because the deep convolutions neural network has been trained beforehand. Secondly, the model to detect first the area with movement and then classify the element found in this specific area. The framework developed is shown in Figure 6.1.

6.2 Movement Detection

Detection of live sperm cells is performed in this thesis based on detection, localization and evaluation of movement in the sample. After while the live sperm is classified based on morphology. This Chapter is not aiming to calculate the motility level as has been formally defined by WHO in section 2.5.2, the objective of this work is classify live cells to determinate their morphology.

Movement detection is still a challenging task because of situations already explained in Section 1.1: illumination, hardware properties, movement of cells and the viscosity of the smear among others can generate false positives.

6.2.1 Optical Flow

Background subtraction and optical flow methods have been tested. Optical flow is a popular technique in image processing, particularly on analysis of motion. There is a relative motion between the observed and the scene, this motion generates some patterns along the scenario. Optical flow is the distribution of observed velocities of object in a image.

Optical flow represents describes a sparse vector field, where a particular pixel from the image I have a position (x, y) , in this position, there is a displacement vector where that pixel can be found in another image.

Some pre-processing steps have been applied to the original image before using optical flow estimation on frames:

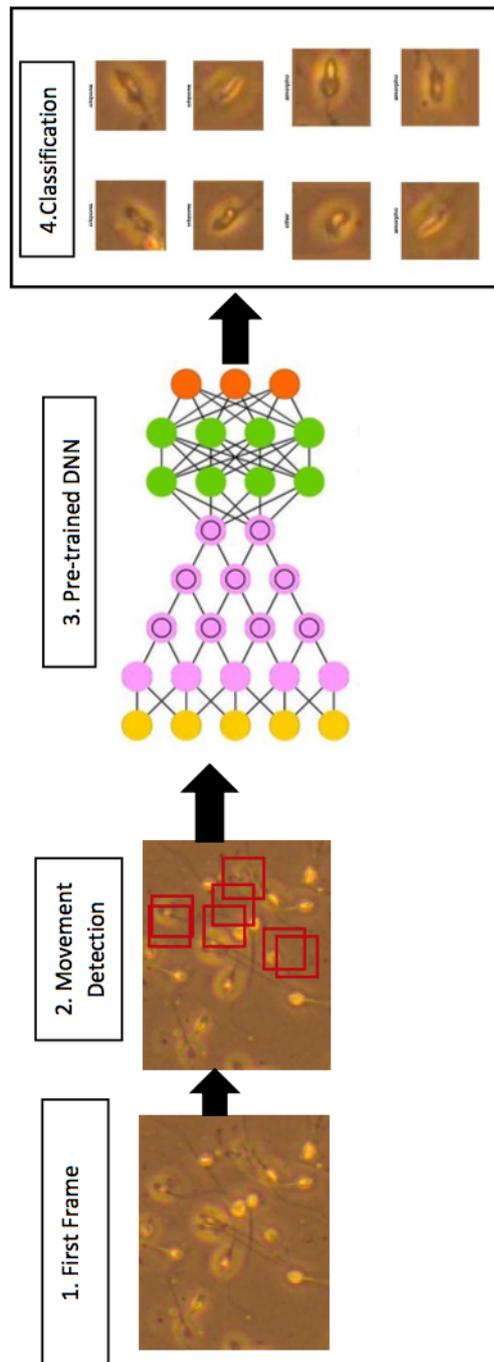


Figure 6.1: Movement detection and classification

- The image format is converted from RGB to grayscale because intensity have a good performance on grayscale frames.
- The median filter is performed to reduce noise corruptions.

- Optical flow is estimated.

Optical flow estimation regardless of mobility, results in high contrast spots which are focused on, over this patches the detection and classification of sperm cells is executed.

The algorithm developed is based on the presented by Farneback on [105]. A sample of a frame where motile sperm has been successfully detected is show in Figure 6.2 the area where there is movement are showed in blue. The analysis have been done in a 10 seconds sequence.

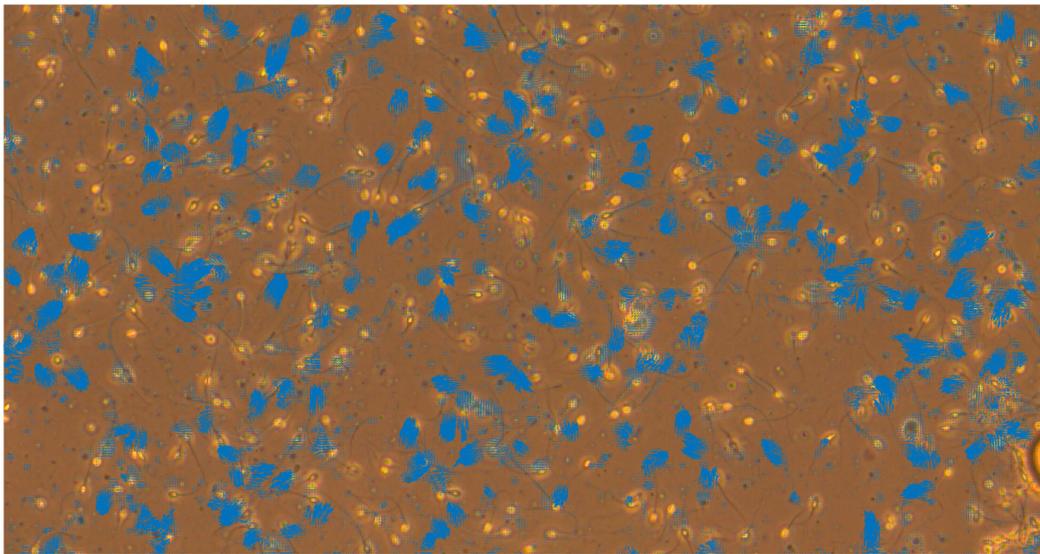


Figure 6.2: Detection of ROI with movement

The results of optical flow for movement detection is usually affected by the image brightness and velocity of objects detected. Sometimes a motion smoothness constraint is included to improve the detection accuracy and stability of the estimation. However issues like uneven illumination, noise and occlusion are still unsolved with this approach.

6.3 Performance Evaluation

After the implementation using Matlab_R2015b [106] the detection of movement allows to find where exactly the live cells are along the video sequence, this is the

ROI (region of interest). After this, the algorithm is analysing each ROI and using the image as input for the pre-trained deep convolutional neural network, presented in Section 5.4.

The algorithm proposed has been tested along 5 short sequences. Figure 6.3 shown how the results of classification is displayed along the video. Experiments evaluating until 3 ROI simultaneously were executed.

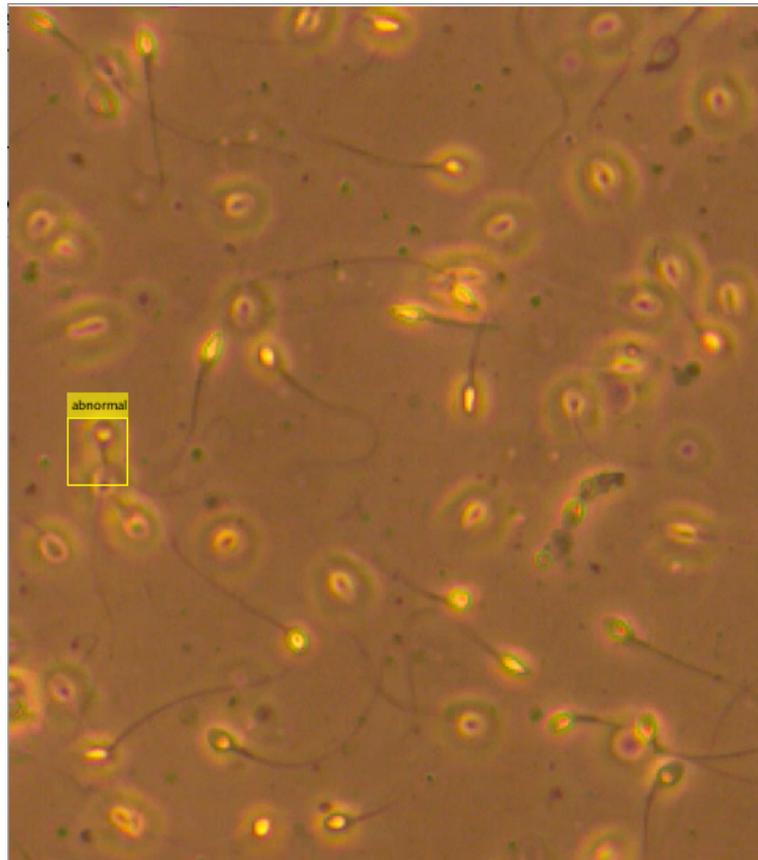


Figure 6.3: Detection of motile sperm and classification

Along a the video sequence with 5 frames and 2 seconds long has been detected objects belongs to the three pre-defined classes: normal, abnormal and no sperm cell. After the classification of cells with movement, along 5 video sequences, detecting 700 ROI, results obtained are show in Figure 6.4.

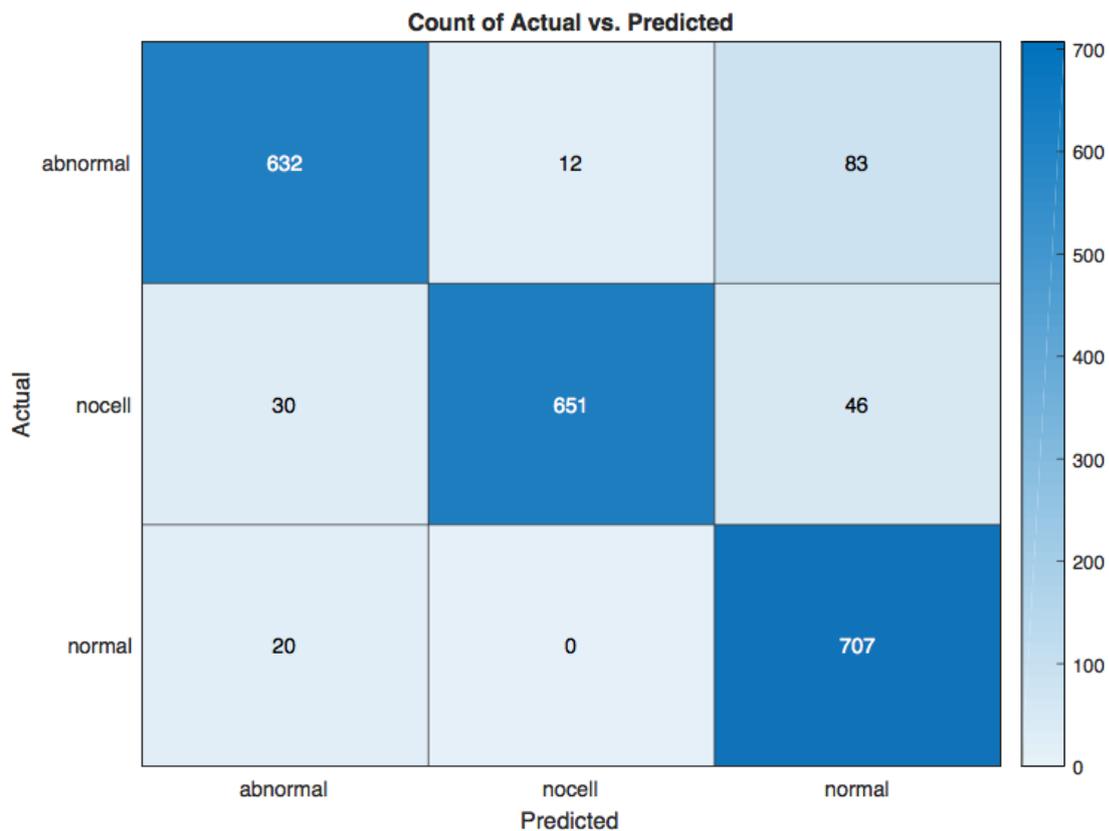


Figure 6.4: Confusion matrix for classification of cells with progressive movement

6.4 Summary

This Chapter presents a combination of optical flow for motion detection and deep neural network technique based as a solution for the need of identify morphology on motile sperm. The classification of motile sperm was done successfully, first detection of ROI and then the analysis of the object in the ROI. The results are obtained after counting analysis along 5 video sequences and summarizing the classification rates for each video.

Chapter 7

Conclusions

Analysis of microscopic images is common in medical diagnosis and treatment. However, in the majority of laboratories this analysis is done manually, meaning time consuming and with a low level of accuracy. This thesis presents a novel methodology based on image processing for automatic *in-vitro* cells analysis. Unlike many others works, this approach is able to classify non-stained and non-fixed sperm cells, with an average accuracy of 96%.

7.1 Contributions

The key contributions of this thesis are detailed below:

- In Chapter 3 is presented the contribution of the files recorded at the Andrology Lab of The University of Sheffield, following high standard procedures to manually analyse sperm samples using a microscope. Therefore, the image sequences show real challenges, e.g. illumination is not controlled and sperm-cells remain in movement during the recording for their analysis.

As an important contribution also, is the construction of a dataset with 1038 samples for normal morphology, abnormal morphology and others component. This dataset can be used for further analysis and research.

- Identification of common problems when analysing *in-vitro* images. Phenomenons found in this analysis have common characteristic in other studios

(occlusion, collusion, shades, lack of focus). The procedures presented in this thesis can be applied for other areas, especially related with analysis of in-vitro biological images.

- The analysis of features, their selection and combination with Support Vector Machine, presented on Chapter 4 offers a new approach of two popular techniques.
- A novel and efficient method is proposed to deal with the illumination variability effect to a certain extent. The use of transfer learning is presented in Chapter 5 as a solution for extracting the right features when there is a variation in them along the time.
- As a contribution, the ability to use the proposed framework to reduce the processing time, since currently, manual analysis is a highly time consuming task and health professionals can be overloaded of work. A framework like the proposed can be used to verify the estimated measures done by a health professional.
- The proposed framework provides an objective analysis of a dataset. Thus, the analysis of a sperm sample can be repeated to verify the same output.
- Also, the proposed framework is written using a programming language providing the ability to extend the current range of methods to make the framework more robust.

7.2 Future Work

The analysis of *in-vitro* images still faces many challenges and there are wide opportunities to expand. Two particular trends are mentioned as a future works: Bayesian convolutional neural networks and Cloud computing.

7.2.1 Bayesian Convolutional Neural Networks

Bayesian Convolutional Neural Network has been presented as a trusted classification approach as a future work for this research.

Bayesian networks and neural networks are not mutually exclusive. In fact, it is recommended to make predictions when the level of uncertainty is high. For this case, sperm cells have a random and unpredictable movement being the tracking process an adequate area of application. Therefore, the morphology of the cells seems change along the video sequences.

7.2.2 Collaborative analysis with Cloud Computing

As a general limitation for analysis of in-vitro images in medical area is the lack of direct communications with other professionals. The objective of sharing experience with the use of Cloud computing is a great technological opportunity to improve results and formal guidelines combining with different levels experience of professional doing this kind of assessment.

Cloud computing is a model to enable the convenient access to the same images and resources in multiples location. Then the diagnosis or evaluation of in-vitro cells could be then a collaborative experience.

References

- [1] W. H. Organization, “Laboratory manual for the examination and processing of human semen,” *World Health Organization*, vol. 30, 2009.
- [2] T. Ahmad, K. Aggarwal, B. Pattnaik, S. Mukherjee, T. Sethi, B. K. Tiwari, M. Kumar, A. Micheal, U. Mabalirajan, B. Ghosh, S. S. Roy, and A. Agrawal, “Computational classification of mitochondrial shapes reflects stress and redox state,” *Cell Death & Disease*, vol. 4, no. 1, pp. e461–e461, 2013.
- [3] D. J. Tozer, E. Zeestraten, A. J. Lawrence, T. R. Barrick, and H. S. Markus, “Texture analysis of t1-weighted and fluid-attenuated inversion recovery images detects abnormalities that correlate with cognitive decline in small vessel disease,” *Stroke*, 2018.
- [4] F. Shaker, S. A. Monadjemi, J. Alirezaie, and A. R. Naghsh-Nilchi, “A dictionary learning approach for human sperm heads classification,” *Computers in Biology and Medicine*, vol. 91, pp. 181–190, 2017.
- [5] P. Naylor, M. Lae, F. Reyal, and T. Walter, “Nuclei segmentation in histopathology images using deep neural networks,” in *14th International Symposium on Biomedical Imaging 2017*. IEEE, 2017.
- [6] M. K. Hadjidemetriou S, Gabrielli B, “Detection and tracking of cell divisions in phase contrast video microscopy,” in *Third Workshop on Microscopic Image Analysis with Applications in Biology*, 2008.
- [7] T. He, H. Mao, J. Guo, and Z. Yi, “Cell tracking using deep neural networks with multi-task learning,” *Image and Vision Computing*, vol. 60, pp. 142–153, 2017.
- [8] A. Zaritsky, S. Natan, J. Horev, I. Hecht, L. Wolf, E. Ben-Jacob, and I. Tsarfaty, “Cell motility dynamics: A novel segmentation algorithm to quantify multi-cellular bright field microscopy images,” *PLoS One*, vol. 6, no. 11, 2011.

- [9] F. P. Cordelieres, V. Petit, M. Kumasaka, O. Debeir, V. Letort, S. J. Gallagher, and L. Larue, “Automated cell tracking and analysis in phase-contrast videos (itrack4u): development of java software based on combined mean-shift processes,” *PLoS ONE*, vol. 8, no. 11, 2013.
- [10] E. Suryani, W. Wiharto, and N. Polvonov, “Identification and counting white blood cells and red blood cells using image processing case study of leukemia,” *CoRR*, vol. abs/1511.04934, 2015.
- [11] F. Piccinini, A. Tesei, G. Paganelli, W. Zoli, and A. Bevilacqua, “Improving reliability of live/dead cell counting through automated image mosaicing,” *Computer Methods and Programs in Biomedicine*, vol. 117, no. 3, pp. 448 – 463, 2014.
- [12] M. Gurcan, L. Boucheron, A. Can, A. Madabhushi, N. Rajpoot, and B. Yener, “Histopathological image analysis: A review,” *IEEE Reviews in Biomedical Engineering*, vol. 2, pp. 147–171, 2009.
- [13] D. C. Wilbur, “Digital cytology: Current state of the art and prospects for the future,” *Acta Cytologica*, vol. 55, no. 3, pp. 227–238, 2011.
- [14] S. N. Mazhir, A. H. Ali, N. H. Harb, and F. W. Hadi, “The effect of dielectric barrier discharge plasma on smear of leukemia blood cells by texture analysis images,” *Journal of Applied Sciences Research*, vol. 13, pp. 35–42, 2017.
- [15] Y. Qiu, S. Yan, R. R. Gundreddy, Y. Wang, S. Cheng, H. Liu, and B. Zheng, “A new approach to develop computer-aided diagnosis scheme of breast mass classification using deep learning technology,” *Journal of X-Ray Science and Technology*, vol. 25, no. 5, pp. 751–763, Oct 2017.
- [16] Y. Qiu, S. Yan, R. Gundreddy, Y. Wang, S. Cheng, H. Liu, and B. Zheng, “A new approach to develop computer-aided diagnosis scheme of breast mass classification using deep learning technology,” *Journal of X-Ray Science and Technology*, vol. 25, no. 5, pp. 751–763, 2017.
- [17] S. Christof, G. Amiq, R. Mauricio, B. L. M., G. Thomas, S. K. A., and G. Benjamin, “Timelapse microscopy and classification of 2d human mesenchymal stem cells based on cell shape picks up myogenic from osteogenic and adipogenic differentiation,” *Journal of Tissue Engineering and Regenerative Medicine*, vol. 8, no. 9, pp. 737–746, 2012.

-
- [18] S. Pereira, A. Pinto, V. Alves, and C. A. Silva, “Brain tumor segmentation using convolutional neural networks in MRI images,” *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1240–1251, 2016.
- [19] H. Su, F. Xing, and L. Yang, “Robust cell detection of histopathological brain tumor images using sparse reconstruction and adaptive dictionary selection,” *IEEE Transactions on Medical Imaging*, vol. 35, no. 6, pp. 1575–1586, 2016.
- [20] R. B. Leaver, “Male infertility: an overview of causes and treatment options,” *British Journal of Nursing*, vol. 25, no. 18, pp. S35–S40, 2016.
- [21] G. Palermo, H. Joris, P. Devroey, and A. Van Steirteghem, “Pregnancies after intracytoplasmic injection of single spermatozoon into an oocyte,” *The Lancet*, vol. 340, no. 8810, pp. 17–18, 1992.
- [22] M. Tan, W. Qian, J. Pu, H. Liu, and B. Zheng, “A new approach to develop computer-aided detection schemes of digital mammograms,” *Physics in Medicine and Biology*, vol. 60, no. 11, p. 4413, 2015.
- [23] B. Zheng, J. H. Sumkin, M. L. Zuley, D. Lederman, X. Wang, and D. Gur, “Computer-aided detection of breast masses depicted on full-field digital mammograms: a performance assessment,” *The British Journal of Radiology*, vol. 85, no. 1014, pp. e153–e161, 2012.
- [24] D. Riddell, A. Pacey, and K. Whittington, “Lack of compliance by uk andrology laboratories with world health organization recommendations for sperm morphology assessment.” *Human reproduction (Oxford, England)*, vol. 20, no. 12, pp. 3441–5, Dec. 2005.
- [25] D. S. Guzick, J. W. Overstreet, P. Factor-Litvak, C. K. Brazil, S. T. Nakajima, C. Coutifaris, S. A. Carson, P. Cisneros, M. P. Steinkampf, J. A. Hill, D. Xu, and D. L. Vogel, “Sperm morphology, motility, and concentration in fertile and infertile men,” *The New England Journal of Medicine*, vol. 345, no. 19, pp. 1388–1393, 2001.
- [26] L. Maree, S. S. du Plessis, R. Menkveld, and G. van der Horst, “Morphometric dimensions of the human sperm head depend on the staining method used,” *Human Reproduction*, vol. 25, no. 6, pp. 1369–1382, 2010.
- [27] A. Pacey, “Is quality assurance in semen analysis still really necessary? a view from the andrology laboratory.” *Human reproduction (Oxford, England)*, vol. 21, no. 5, pp. 1105–9, 2006.

- [28] R. C. Gonzalez, *Digital image processing*, 3rd ed. Upper Saddle River, N.J.: Upper Saddle River, N.J. : Pearson/Prentice Hall, c2010, 2010, (Richard Eugene) Includes bibliographical references (p. 937-964) and index.
- [29] G. Lupica, N. Allinson, and S. W. Botchway, "Hybrid image processing technique for the robust identification of unstained cells in bright-field microscope images," in *International Conference on Computational Intelligence for Modelling Control and Automation*, 2008.
- [30] P. Bajcsy, A. Cardone, J. Chalfoun, D. Juba, M. Majurski, A. Peskin, and Simon, "Survey statistics of automated segmentations applied to optical imaging of mammalian cells," *BMC Bioinformatics*, 2015.
- [31] O. Al-Kofahi, R. J. Radke, S. K. Goderie, Q. Shen, S. Temple, and B. Roysam, "Automated cell lineage construction: a rapid method to analyze clonal development established with murine neural progenitor cells," *Cell Cycle*, vol. 5, no. 3, p. 327, 2006.
- [32] R. Bise, K. Li, S. Eom, and T. Kanade, "Reliably tracking partially overlapping neural stem cells in dic microscopy image sequences," in *Workshop on Optical Tissue Image Analysis in Microscopy, Histopathology and Endoscopy*, 2009, pp. 67–77.
- [33] M. Kerschnitzki, P. Kollmannsberger, M. Burghammer, G. N. Duda, R. Weinkamer, W. Wagermaier, and P. Fratzl, "Architecture of the osteocyte network correlates with bone material quality," *Journal of Bone and Mineral Research*, vol. 28, no. 8, pp. 1837–1845, 2013.
- [34] G. Andrew, C.-L. Alejandro, E. C. Tina, R. B. Kirsten, C. Y. Richard, and B. Roger, "Single-cell quantification of molecules and rates using open-source microscope-based cytometry," *Nature Methods*, vol. 4, no. 2, p. 175, 2007.
- [35] F. Ghasemian, S. A. Mirroshandel, S. Monji-Azad, M. Azarnia, and Z. Zahiri, "An efficient method for automatic morphological abnormality detection from human sperm images," *Computer Methods and Programs in Biomedicine*, vol. 122, no. 3, pp. 409–420, 2015.
- [36] D. Marr, "*Vision*". San Francisco: W.H.Freeman and Company, 1982.
- [37] L. Roberts, *Machine Perception of Three-Dimensional Solids*, ser. Massachusetts Institute of Technology. Lincoln Laboratory. Technical Reports. U.S. Government Printing Office, 1969.

-
- [38] J. M. S. Prewitt, “Object enhancement and extraction,” *Picture Processing and Psychopictorics*, pp. 75–149, 1970.
- [39] I. Sobel, “An isotropic 3x3 image gradient operator,” *Presentation at Stanford A.I. Project 1968*, 1968.
- [40] S. Lloyd, “Least squares quantization in pcm,” *Information Theory, IEEE Transactions on*, vol. 28, no. 2, pp. 129–137, 1982.
- [41] N. Dalal, B. Triggs, and C. Schmid, “Human detection using oriented histograms of flow and appearance,” in *Proc of the European Conference on Computer Vision*. Springer Berlin Heidelberg, 2006, pp. 428–441.
- [42] O. Déniz, G. Bueno, J. Salido, and F. D. la Torre, “Face recognition using histograms of oriented gradients,” *Pattern Recognition Letters*, vol. 32, no. 12, pp. 1598–1603, sep 2011.
- [43] S. Tian, S. Lu, B. Su, and C. L. Tan, “Scene text recognition using co-occurrence of histogram of oriented gradients,” in *Proc of the 12th International Conference on Document Analysis and Recognition*, Aug 2013, pp. 912–916.
- [44] D. G. Lowe, “Object recognition from local scale-invariant features,” in *Proc of the IEEE International Conference on Computer Vision*, vol. 2, 1999, pp. 1150–1157.
- [45] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, “Speeded-up robust features (SURF),” *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346–359, jun 2008.
- [46] F. Lecron, M. Benjelloun, and S. Mahmoudi, *Descriptive Image Feature for Object Detection in Medical Images*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 331–338.
- [47] H. Mehrotra, P. K. Sa, and B. Majhi, “Fast segmentation and adaptive surf descriptor for iris recognition,” *Mathematical and Computer Modelling*, vol. 58, no. 12, pp. 132–146, 2013.
- [48] J. Feulner, S. Zhou, E. Angelopoulou, and S. Seifert, “Comparing axial CT slices in quantized n-dimensional SURF descriptor space to estimate the visible body region,” *Computerized Medical Imaging and Graphics*, vol. 35, no. 3, pp. 227–236, Apr 2011.

- [49] Y. Han, K. Virupakshappa, and E. Oruklu, “Robust traffic sign recognition with feature extraction and k-nn classification methods,” in *Proc of the IEEE International Conference on Electro/Information Technology*, May 2015, pp. 484–488.
- [50] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [51] T. Lindeberg, “Feature detection with automatic scale selection,” *International Journal of Computer Vision*, vol. 30, no. 2, pp. 79–116, 1998.
- [52] Y. Liu, H. Zha, and H. Qin, “Shape topics: A compact representation and new algorithms for 3d partial shape retrieval,” in *Proc of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, 2006, pp. 2025–2032.
- [53] L. Shen, J. Lin, S. Wu, and S. Yu, “Hep-2 image classification using intensity order pooling based features and bag of words,” *Pattern Recognition*, vol. 47, no. 7, pp. 2419–2427, 2014.
- [54] L. Zhou, Z. Zhou, and D. Hu, “Scene classification using a multi-resolution bag-of-features model,” *Pattern Recognition*, vol. 46, no. 1, pp. 424–433, 2013.
- [55] L. Nanni and A. Lumini, “Heterogeneous bag-of-features for object/scene recognition,” *Applied Soft Computing*, vol. 13, no. 4, pp. 2171–2178, 2013.
- [56] S. O’Hara and B. A. Draper, “Introduction to the bag of features paradigm for image classification and retrieval,” *CoRR*, vol. abs/1101.3354, 2011.
- [57] N. Cristianini, *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press, 2000, 2000.
- [58] B. E. Boser, I. M. Guyon, and V. N. Vapnik, “A training algorithm for optimal margin classifiers,” pp. 144–152, 1992.
- [59] C. Cortes and V. Vapnik, “Support- vector networks,” *Mach Learn*, vol. 20, no. 3, pp. 273–297, 1995.
- [60] R. M. Nishikawa, M. L. Giger, K. Doi, C. J. Vyborny, R. A. Schmidt, C. E. Metz, C. Y. Wu, F.-F. Yin, and Jiang, “Computer-aided detection and diagnosis of masses and clustered microcalcifications from digital mammograms,” *Proceedings of SPIE*, vol. 1905, no. 1, pp. 422–432, 1993.

-
- [61] H.-P. Chan, K. Doi, C. J. Vyborny, C. E. Metz, H. Macmahon, P. M. Jokich, and S. Galhotra, “Digital mammography: Development of a computer-aided system for detection of microcalcifications,” *Proceeding of SPIE*, vol. 767, pp. 367–370, 1987.
- [62] N. Tajbakhsh, S. Gurudu, and J. Liang, “A comprehensive computer-aided polyp detection system for colonoscopy videos,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9123, pp. 327–338, 2015.
- [63] D. Ciresan, A. Giusti, L. Gambardella, and J. Schmidhuber, “Mitosis detection in breast cancer histology images with deep neural networks,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 8150, no. 2, pp. 411–418, 2013.
- [64] H. Roth, A. Farag, L. Lu, E. Turkbey, and R. Summers, “Deep convolutional networks for pancreas segmentation in ct imaging,” *Progress in Biomedical Optics and Imaging in proceedings of SPIE*, vol. 9413, 2015.
- [65] M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P.-M. Jodoin, and H. Larochelle, “Brain tumor segmentation with deep neural networks,” *Medical Image Analysis*, vol. 35, no. C, pp. 18–31, 2017.
- [66] P. Kim, *MATLAB Deep Learning*. Apress, 2017. [Online]. Available: <https://doi.org/10.1007/978-1-4842-2845-6>
- [67] W. H. Organization, *WHO Laboratory Manual for the Examination and Processing of Human Semen*, ser. Nonserial Publications Series. World Health Organization, 2010.
- [68] U. Kvist and L. Björndahl, *Manual on basic semen analysis: 2002*. Oxford University Press, 2002.
- [69] R. Menkveld, C. A. Holleboom, and J. P. Rhemrev, “Measurement and significance of sperm morphology,” *Asian Journal of Andrology*, vol. 13, no. 1, pp. 59–68, 2010. [Online]. Available: <https://doi.org/10.1038/aja.2010.67>
- [70] J. P. E. Bonde, E. Ernst, T. K. Jensen, N. H. I. Hjollund, H. Kolstad, and Scheike, “Relation between semen quality and fertility: a population- based study of 430 first- pregnancy planners,” *The Lancet*, vol. 352, no. 9135, pp. 1172–1177, 1998.

- [71] G. Di Caprio, A. El Mallahi, P. Ferraro, R. Dale, and Coppola, "4d tracking of clinical seminal samples for quantitative characterization of motility parameters," *Biomedical optics express*, vol. 5, no. 3, pp. 690–700, 2014.
- [72] E. Bailey, N. Fenning, S. Chamberlain, L. Devlin, J. Hopkisson, and M. Tomlinson, "Validation of sperm counting methods using limits of agreement." *Journal of andrology*, vol. 28, no. 3, pp. 364–73, 2007.
- [73] O. Simonik, J. Sichtar, A. Krejcarkova, R. Rajmon, L. Stadnik, and J. Beran, "Computer assisted sperm analysis, the relationship to bull field fertility, possible errors and their impact on outputs : A review," *Indian Journal of Animal Sciences*, vol. 85, no. January, pp. 3–11, 2015.
- [74] P. Kathiravan, J. Kalatharan, G. Karthikeya, K. Rengarajan, and G. Kadirvel, "Objective sperm motion analysis to assess dairy bull fertility using computer-aided system - a review," *Reproduction in Domestic Animals*, vol. 46, no. 1, pp. 165–172, Feb. 2011.
- [75] L. Severa, L. Machal, L. vbov, and O. Mamica, "Evaluation of shape variability of stallion sperm heads by means of image analysis and fourier descriptors," *Animal Reproduction Science*, vol. 119, no. 1, pp. 50–55, 2010.
- [76] M. H. Moradi and M. R. Ravanfar, "Low contrast sperm detection and tracking by watershed algorithm and particle filter," *Biomedical Engineering (ICBME), 2011 18th Iranian Conference of Biomedical Engineering*, pp. 260–263, 2011.
- [77] M. Y. Khachane, R. R. Manza, and R. J. Ramteke, "Fuzzy rule based classification of human spermatozoa," *2015 International Conference on Electrical, Electronics, Signals, Communication and Optimization (EESCO)*, pp. 1–5, 2015.
- [78] J. Ehlers, M. Behr, H. Bollwein, M. Beyerbach, and D. Waberski, "Standardization of computer-assisted semen analysis using an e-learning application." *Theriogenology*, vol. 76, no. 3, pp. 448–54, 2011.
- [79] A. Berkovitz, F. Eltes, Y. Soffer, N. Zabludovsky, Y. Beyth, J. Farhi, D. Levran, and B. Bartoov, "Art success and in vivo sperm cell selection depend on the ultramorphological status of spermatozoa," *Andrologia*, vol. 31, no. 1, pp. 1–8, 1999.
- [80] L. Sanchez, N. Petkov, and E. Alegre, "Statistical approach to boar semen head classification based on intracellular intensity distribution," *Computer Analysis Of Images And Patterns, Proceedings*, vol. 3691, pp. 88–95, 2005.

-
- [81] S. L. Bielh W.M., Pijl M., “Classification of board sperm head images using learning vector quantization,” *European Symposium on Artificial Neural Networks, Proceedings*, pp. 545–550, 2006.
- [82] D. E. Kime, K. J. W. V. Look, B. G. Mcallister, and G. Huyskens, “Computer-assisted sperm analysis (CASA) as a tool for monitoring sperm quality in fish,” *Comparative Biochemistry and Physiology*, 2001.
- [83] S. P. Huei-Fang Yang, Xavier Descombes, “Head tracking and flagellum tracking for sperm motility analysis,” *Biomedical Engineering (ICBME), 2011 18th Iranian Conference of Biomedical Engineering*, 2014.
- [84] V. R. Nafisi and M. H. Moradi, “A template matching algorithm for sperm tracking and classification.” *Physiological measurement*, vol. 26, no. 5, pp. 639–51, Oct. 2005.
- [85] C. Jun Liu, C. Leung, and C. Zhe Lu, “Quantitative analysis of locomotive behavior of human sperm head and tail,” *Biomedical Engineering, IEEE Transactions on*, vol. 60, no. 2, pp. 390–396, 2013.
- [86] Y. a. Hu, J. C. Lu, Y. Shao, Y. F. Huang, and L, “Comparison of the semen analysis results obtained from two branded computer-aided sperm analysis systems.” *Andrologia*, vol. 45, no. 5, pp. 315–8, 2013.
- [87] C. Castellini, A. Dal Bosco, S. Ruggeri, and G. Collodel, “What is the best frame rate for evaluation of sperm motility in different species by computer-assisted sperm analysis?” *Fertility and sterility*, vol. 96, no. 1, pp. 24–7, 2011.
- [88] L. Sorensen, J. Ostergaard, P. Johansen, and M. de Bruijne, “Multibject tracking of human spermatozoa,” *Proceedings SPIE, Medical Imaging 2008: Image Processing*, vol. 6914, pp. 69 142C–69 142C–12, 2008.
- [89] R. Slama, F. Eustache, B. Ducot, T. K. Jensen, N. Jrgensen, A. Horte, and S. Irvine, “Time to pregnancy and semen parameters: a cross-sectional study among fertile couples from four european cities,” *Human reproduction (Oxford, England)*, vol. 17, no. 2, p. 503, 2002.
- [90] M. J. Zinaman, C. C. Brown, S. G. Selevan, and E. D. Clegg, “Semen quality and human fertility: A prospective study with healthy couples,” *Journal of Andrology*, vol. 21, no. 1, pp. 145–153, 2000.
- [91] C. Garrett, D. Y. Liu, and G. N. Rushford, “Automated semen analysis: Zona pellucida preferred sperm morphometry and straight-line velocity are related

- to pregnancy rate in subfertile couples,” *Human Reproduction*, vol. 18, no. 8, pp. 1643–1649, 2003.
- [92] G. Corkidi and B. Taboada, “Three-dimensional image acquisition system for multi-sperm tracking,” *Biomedical Imaging: From Nano to macro, Proceedings*, no. mM, pp. 145–148, 2008.
- [93] A. Vested, C. H. Ramlau-Hansen, J. P. Bonde, A. M. Thulstrup, S. L. Kristensen, and G. Toft, “A comparison of conventional and computer-assisted semen analysis (crismas software) using samples from 166 young danish men,” *Asian journal of andrology*, vol. 13, no. 3, p. 453, 2011.
- [94] V. Chang, J. M. Saavedra, and V. Castaeda, “Gold-standard and improved framework for sperm head segmentation,” *Computer Methods and Programs in Biomedicine*, vol. 117, no. 2, pp. 225 – 237, 2014.
- [95] N. Otsu, “A Threshold Selection Method from Gray-Level Histograms,” *IEEE Transactions On Systems And Cybernetics*, vol. 20, pp. 62–66, January 1979.
- [96] J. Canny, “A computational approach to edge detection,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. PAMI-8, no. 6, pp. 679–698, 1986.
- [97] C. C. Reyes-Aldasoro and A. Bhalerao, “The bhattacharyya space for feature selection and its application to texture segmentation,” *Pattern Recognition*, vol. 39, no. 5, pp. 812–826, 2006.
- [98] R. O. Duda, *Pattern classification*, 2nd ed. New York: New York : Wiley, 2001, 2001.
- [99] W. Zheng and Y. Zhang, “A novel improvement to pca for image classification,” in *Computer Science and Service System (CSSS), 2011 International Conference on*, June 2011, pp. 1964–1967.
- [100] F. Yang, M. Mackey, F. Ianzini, G. Gallardo, and M. Sonka, “Cell segmentation, tracking, and mitosis detection using temporal context,” *Medical Image Computing And Computer-Assisted Intervention - Miccai 2005*, vol. 3749, pp. 302–309, 2005.
- [101] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, may 2015.

- [102] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” *Commun. ACM*, vol. 60, no. 6, pp. 84–90, 2017.
- [103] K. Althoff, J. Degerman, and T. Gustavsson, “Combined segmentation and tracking of neural stem-cells,” *Proc of the 14th Scandinavian Conference Image Analysis*, pp. 282–291, 2005.
- [104] D. Mortimer, “Laboratory standards in routine clinical andrology,” *Reproductive Medicine Review*, vol. 3, no. 2, pp. 97–111, 1994.
- [105] G. Farneback, “Two-frame motion estimation based on polynomial expansion,” *Image Analysis, Proceedings*, vol. 2749, pp. 363–370, 2003.
- [106] MATLAB, *version 8.5.0 (R2015a)*. Natick, Massachusetts: The MathWorks Inc., 2015.

