Università degli Studi di Napoli "Federico II"
PhD Program in Industrial Products and Process Engineering,
Cycle XXXVII

THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

CLAUDIA DE CLEMENTE

SINGLE-CELL TRACKING ANALYSIS FOR LABEL-FREE SEMEN DIAGNOSIS BY MEANS OF NEURAL NETWORKS

SCUOLA POLITECNICA DELLE SCIENZE DI BASE
DIPARTIMENTO DI INGEGNERIA CHIMICA, DEI MATERIALI E DELLA PRODUZIONE INDUSTRIALE

Coordinator
Ch.mo Prof. Andrea D'Anna
Supervisor
PhD Ing. Maria Isabella Maremonti
PhD David Dannhauser
Tutor
Ch.mo Prof. Filippo Causa

November 2021 – October 2024

Index

Abstract

Single-cell analysis allows the retrieval of precious information from a pool of cells that would not be appreciated with a bulk analysis. In fact, the characterisation of cell biophysical properties -such as area, perimeter, axes, density, stiffness, and electric potential- has proven to be a powerful mean for phenotyping, sorting, detection, drug testing applications, and to gain more insight into physiological cell functions in responses to external stimuli.

Besides biophysical properties, single-cell tracking analysis is a powerful tool for high-precision diagnostics, gathering information on cell kinematics and morphology. Object tracking, which is the process of locating and following a specific object and its behaviour in sequential images, is a difficult task in biology because of the large size of cell image sequences, which makes the analysis more convenient to automatize. Indeed, computational object tracking can improve reproducibility and computational time significantly. However, the very small cell sizes often require fluorescent labels to ease the location and the tracking process, reducing the sample viability or potentially altering cell properties.

Semen infertility diagnosis is based on the evaluation of a comprehensive set of sperm cells quality indicators, including motion and morphology. For this reason, it is the optimal context for developing innovative single-cell analysis approaches. The World Health Organization (WHO) establishes stringent criteria for assessing semen health. Presently, the gold standard for evaluating semen quality relies on visual examination of motility and morphological parameters, introducing limitations in accuracy and repeatability. However, in 1980, computer-assisted sperm analysis (CASA) systems emerged trying to reduce these analysis biases. In detail, CASA measures the motion of live spermatozoa through sperm head centroid tracking at low magnification, while -for morphology- it requires higher magnification as well as fixing and staining procedures, possibly compromising sample viability and integrity. Additionally, the inability to perform a simultaneous morphological and kinematic assessment on a single cell hinders precise sperm characterization. Addressing these limitations, various label-free techniques have been devised for both motility and morphology assessments, offering promising alternatives in semen quality evaluation.

On this line, the integration of computer science, machine learning (ML) and data science allows the objective analysis of large amounts of data with high accuracy. In neural-network-based ML applications, the aim is the creation of an algorithm to train a machine to solve problems similarly to the human mind, capturing complex relationship between input and output data. For example, convolutional neural networks (CNNs) have been adopted for sperm motility categorization as well as sperm tracking. Deep convolutional neural networks (DCNNs) instead are more commonly used for single-sperm morphology evaluation, DNA fragmentation index retrieval and abnormalities identification. These quality metrics may be further evaluated through advanced ML algorithms, which are able to handle large amounts of data finding possible correlations between sperm parameters at the single-sperm level. This further reduces analysis subjectivity and enhances diagnostic potential.

The objective of the thesis is the development of an automated approach for label-free analysis of sperm cells, based on the automatic tracking of single cells to extract kinematic and biophysical features related to morphology and movement dynamics. Being label-free, the method preserves the viability of the sample, making the process fast and easily reproducible in a clinical context. The procedure involves capturing a brightfield image sequence of swimming sperm in a glass capillary at

40x magnification, followed by single-sperm detection, tracking and segmentation. The detection and segmentation steps are performed with two neural networks, YOLO4-tiny and U-net tiny respectively. Subsequently, kinematic and morphometric parameters are evaluated.

Of note, three parameters are introduced as novel features to better characterize sperm motion. To verify the neural network generalization ability and the new parameters to detect sperm heterogeneity, the analysis was performed under several conditions resembling physio-pathological contexts. In particular, variations in viscosity, osmolarity, absence of calcium, and high doses of caffeine were adopted. This allowed the retrieval of a final dataset of 3900 cells. Principal Component Analysis (PCA) and clustering with the K-means algorithm were adopted to characterize sperm populations in more detail compared to the standard classifications found in the literature. Four clusters were identified, identifying different grades of progressivity related to both motility and head morphology.

The outcome demonstrate that the proposed approach allows for advanced characterization of sperm populations, providing more information on motility compared to traditional CASA systems. Furthermore, the employed cascade of convolutional neural networks provided a more accurate and comprehensive description of sperm movement in heterogeneous conditions.

1. Introduction

1.1 The Male Infertility Issue

Infertility, defined as the inability of a couple to conceive after one year of regular unprotected intercourse, is a global health concern affecting $10-15\%$ of couples worldwide [1]. Male factors account for $20-30\%$ of cases [2]. Male infertility, however, remains underdiagnosed and undertreated in many contexts, partly due to sociocultural stigmas and partly because of the limitations of current diagnostic techniques. As a matter of fact, the global burden of infertility is not distributed equally. Data show that infertility prevalence is particularly high in regions such as South Asia, Sub-Saharan Africa, North Africa, and the Middle East, where healthcare resources may be limited and reproductive health policies are often inadequate [1, 2].
The great advance of technology in this field has led to the development of several Assisted Reproduction Technologies (ARTs). These techniques involve the manipulation of the eggs or embryos outside the female body to achieve successful pregnancies. The most prevalent technique is In-Vitro Fertilization (IVF), where eggs are retrieved from ovaries, combined with a fertile sperm and implanted back in the female body [3]. However common, IVF suffers high failure rates especially in cases where sperm cells from the male counterpart are below the standards dictated by the World Health Organization (WHO). This led to the development of the Intra-Cytoplasmic Sperm Injection (ICSI), where a single spermatozoon is injected mechanically into an oocyte in vitro to achieve fertilisation. This radically increased the success rates of such treatments [4].
In the context of male infertility, several biological, environmental, and lifestyle factors contribute to the issue. Causes of male infertility can range from hormonal imbalances and genetic disorders to exposure to environmental toxins, such as endocrine-disrupting chemicals (EDCs), that negatively affect sperm production [2]. Sperm quality, including motility, morphology, and DNA integrity, plays a crucial role in male fertility, and abnormalities in these parameters can significantly impact the chances of conception. Over the last few decades, there has been growing concern over a global decline in sperm quality. A comprehensive study revealed a significant decrease in sperm concentration and total sperm count in men from Western countries between 1973 and 2011 [5]. These findings suggest that environmental and lifestyle changes, such as increased exposure to pollution, poor diet, and stress, may be contributing to male infertility.
Despite these alarming trends, male infertility remains less studied and less understood than its female counterpart. This disparity in research and clinical attention may partly come from the fact that diagnosing male infertility is often more complex. Traditional diagnostic approaches, such as Semen Analysis (SA), focus on measuring sperm count, motility, and morphology. However, these tests may fail to detect subtle defects in sperm function that contribute to infertility [6]. Furthermore, up to 70% of male infertility cases are considered idiopathic, meaning no clear cause can be identified using current diagnostic methods [2]. This diagnostic gap highlights the need for more advanced tools to assess male reproductive health.

1.2 Sperm journey: from spermatogenesis to the female reproductive tract

Spermatogenesis is the process by which spermatozoa are developed from spermatogonial stem cells in the testes, specifically within the seminiferous tubules (Figure 1A). In mammals, spermatogenesis

occurs in three steps, with different span times among species [7, 8]. The first step involves mitotic cell division, allowing the early cell stage to multiply. The second step requires meiosis, in which the diploid cells form haploid cells, and division occurs until a round spermatid formation occurs. The final stage of spermatogenesis includes spermatozoa production, maturation and motility acquisition through a process called spermiogenesis [9]. During this last phase, spermatids undergo morphological changes and acquire their mature structure (Figure 1B). Such structure comprises head, midpiece, and tail (flagellum), all bound by a continuous plasma membrane, tighter and more compact in the head region than in the flagellum for protection and ease of movement flexibility [10]. The head contains the nucleus, where DNA is stored, and is covered by the acrosome, a membrane-bound vesicle that plays a crucial role in fertilization. As a matter of fact, the acrosome contains hydrolytic enzymes essential for penetrating the egg's protective layer, the *zona pellucida*. During the acrosome reaction (AR), these enzymes are released, allowing the sperm to break down the zona pellucida and access the egg for fusion. Sperm head shape changes among species. For example, a human sperm head shows an oval-like shape of $3-5\ \mu m$ in length and $1.5-3\ \mu m$ in width, while a bovine sperm head has a paddle-like shape of $9-10\ \mu m$ long, $4-5\ \mu m$ wide, and $1\mu m$ thick [11, 12]. The midpiece of the sperm contains a high concentration of mitochondria, which provide the energy required for sperm motility. This energy is critical as the sperm must travel considerable distances through the female reproductive tract to reach the egg. The mitochondria generate ATP, which fuels the movement of the sperm's tail. The tail, or flagellum, is responsible for the motility of the sperm cell. It is composed of a core structure known as the axoneme, consisting of microtubules whose coordinated movement, driven by the motor protein dynein, generates the whip-like motion of the flagellum, propelling the sperm forward [13]. As well as the head, also the midpiece and flagellum are different among species. Human midpiece is $7\ \mu m$ long, while bovine's is significantly longer, approximately $12\ \mu m$. As for the flagellum, human sperm flagellum is approximately $50\ \mu m$; bull sperm flagellum is longer, with a length of approximately $70\ \mu m$ [14].

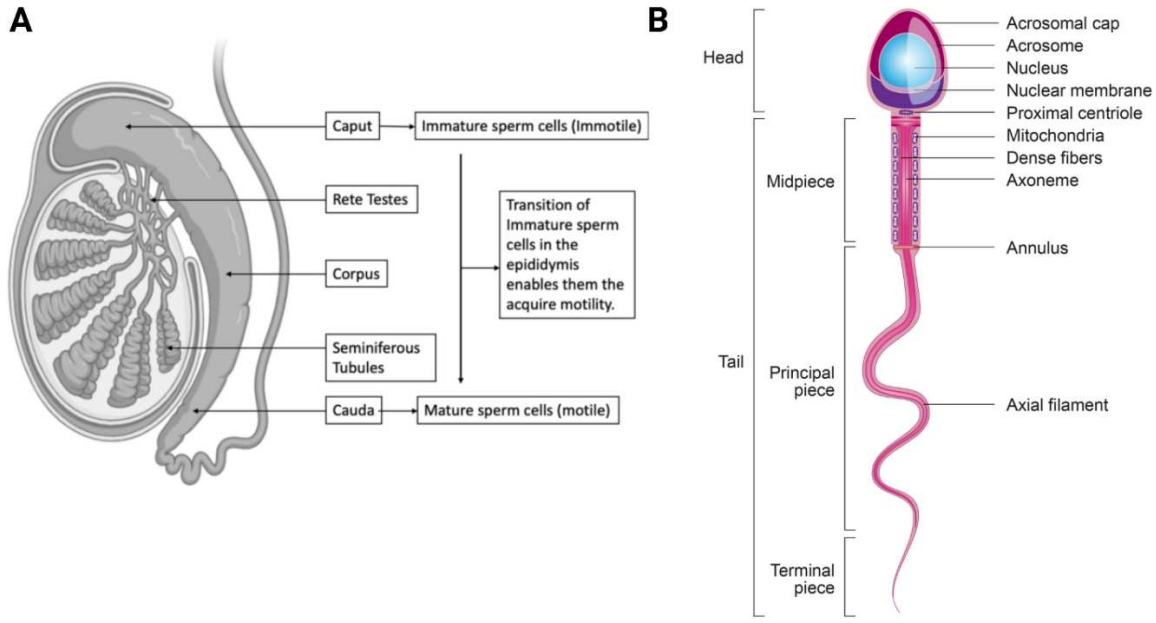

Figure 1: Male epidydimys and sperm structure. A) Diagram of the epididymis. The epididymis is comprised of four anatomical regions: the initial segment, caput, corpus, and cauda. Alterations in membrane composition are driven by concentration gradients of specific enzymes and molecules along the tubule lumen. Sperm maturation occurs during epididymal transit by the interaction of sperm cells with the unique luminal environment of each epididymal region. Reproduced from [15] B) Sperm structure, reproduced from [16].

Sperm are immotile during storage in the cauda epididymis, acquiring motility across their passage along the epididymis, emerging with activated progressive motility after maturation (Figure 1A). This initial pattern of motility, referred to as "progressive" or "activated" motility, consists of relatively symmetrical, low amplitude flagellar bends that produce linear, forwardly progressive sperm movement in aqueous media [13]. The ultimate functional maturation, initiated by secretions from the male accessory glands (prostate and seminal vesicles) and fully completed in the female genital tract, is called capacitation [12]. When in the female genital tract, the exposure to high bicarbonate ions concentration and basic pH trigger biochemical changes both in the sperm head and tail. In the former there is a significant efflux of cholesterol from the plasma membrane, leading to increased membrane fluidity and permeability to ions such as bicarbonate, calcium and chloride [12]. These complex ionic fluxes induce alkalinisation of the cytoplasm, which in turn induce plasma membrane hyperpolarization and flagellar protein phosphorylation [17]. In particular, there is an increased influx of Ca^{2+} due to activation of CatSper channels, which are pH-sensitive channels distributed in the principal piece of the tail [18]. This determines the onset of a more vigorous and asymmetrical flagellar beating termed as hyperactivated motility, which is essential for propagation through the viscous fluid of the oviduct [19]. Hyperactivation and capacitation are thought to be concomitant but distinctive, since hyperactivation can be replicated in vitro exposing sperm to culture media containing calcium but not containing serum or albumin. Both events are partially triggered by extracellular calcium, but hyperactivation manifests as changes in motility, whereas capacitation is principally seen as membrane remodelling [20]. The biochemical changes occurring to the sperm through the capacitation process allow the possibility of a successful interaction with the oocyte, making capacitation a requirement for fertilization [21].
As ejaculated sperm migrate through the female reproductive tract (FRT), they encounter a series of biochemical and biophysical barriers to ensure only those cells with the highest quality reach the site of fertilisation. The first barrier sperm encounter is the vagina, where the pH is acidic. pH constitutes also another relevant barrier as it varies within the FRT. As a matter of fact, it usually increases going from the vaginal tract toward the oviduct, favouring sperm motility and survival [22, 23]. Under normal conditions, only about 200 out of approximately 280 million spermatozoa deposited in the upper vagina upon ejaculation are capable of successfully traversing the cervical canal [24]. Here they encounter the cervical mucus, which is thick and viscous during most part of the menstrual cycle but becomes more fluid during ovulation, allowing sperm to pass through [25]. Osmolarity also plays a crucial role in sperm function regulation and fertilization potential, as both in male and female reproductive tracts sperm experience a gradient in osmolarity [26]. For example, while human sperm cell experience hypo-osmotic conditions going from the seminal plasma ($360-380\ mOsm/Kg$) to the uterus ($280\ mOsm/Kg$), the opposite occurs for bovine sperm ($280-300\ mOsm/Kg$ to $355\ mOsm/Kg$) [27]. The osmotic stress arising from these differences require sperm to adapt and regulate their volume to maintain structural integrity and motility through a process called osmoadaptation. This process enables sperm to regulate water movement across the membrane through aquaporins, ion transporters, and osmoregulatory proteins [28]. Only sperm able to adapt and regulate its behaviour are able to reach the Fallopian Tubes, where chemotaxis and rheotaxis guide the highly motile and progressive sperm towards the oocyte [29].

1.3 Sperm motion

Sperm kinematics is fundamental for overcoming the barriers of the FRT, and is strongly influenced by environmental conditions. For example, it has been seen that sperm with low lateral displacement are not progressive enough to penetrate into or move through the cervical mucus [30]. Hyperactivation is strongly correlated with fertilization potential both in vivo and in vitro, and is characterized by the development of high amplitude flagellar waves but with little net space gain

(Figure 2A) [31]. Hyperactivation can be transitional, where the sperm moves forward with a meandering trajectory, or "star-spin," where the sperm shows vigorous bouncing without forward movement. Similar results were found through 3D holographic tracking, where four distinctive motion patterns have been identified, namely "typical", helical, hyperactivated and hyperhelical (Figure 2B) [32]. The typical trajectory is the most common, exhibited by more than 90% of motile sperm cells. The sperm head moves forward swiftly along a slightly curved axis, with small lateral displacement. The head's direction changes arbitrarily in 3D space. The helical trajectory features sperm moving forward with stable revolutions around a central axis, forming a tight helix. Hyperactivated trajectories are rarer (less than 3%). Hyperhelical trajectory, seen in fewer than 0.5% of sperm, combines hyperactivation and helical movement features. It exhibits enlarged and less stable revolutions around a helix axis while maintaining some forward movement.

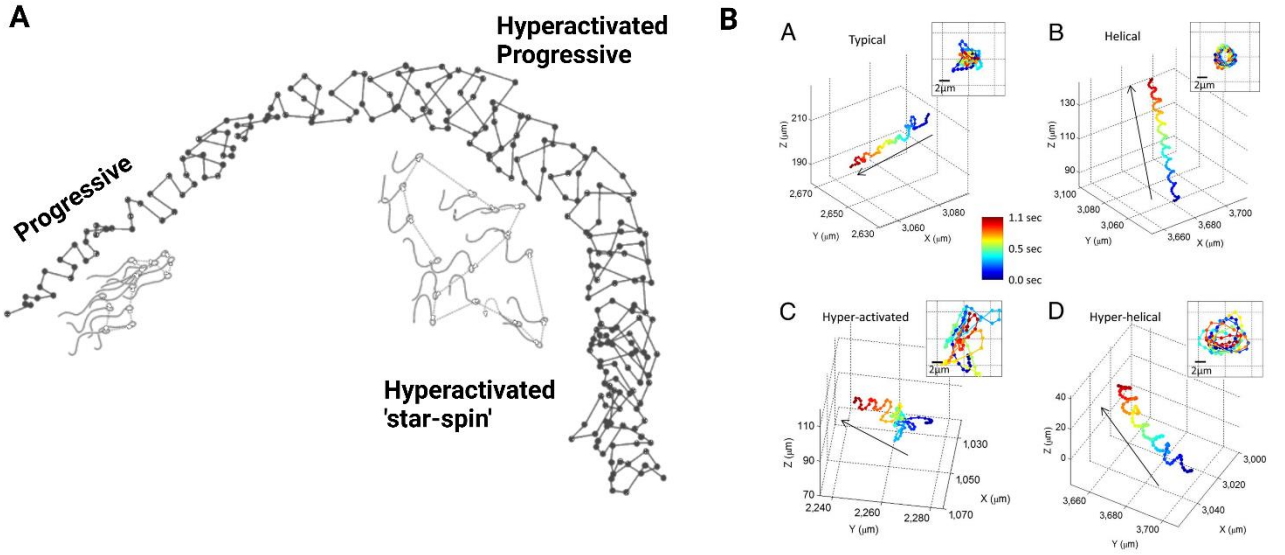

Figure 2: Human sperm motion patterns. A) Spontaneous transitions of motion patterns occurring in washed human sperm, tracked at 60fps. The trajectory should be read from left to right, and is representative of various sperm characteristic motion patterns. Illustrative sperm-tail beating patterns are shown for the non-hyperactivated and hyperactivated progressive regions. Reproduced from [30]. B) Four major categories of human sperm swimming patterns: the typical pattern (upper left), helical pattern (upper right), hyperactivated pattern (lower left), and hyperhelical pattern (lower right). The inset in each panel represents the front-view of the straightened trajectory of the sperm. The arrows indicate the directions of the sperms' forward movement. Reproduced from [33].

The 3D patterns of sperm motility are crucial in ensuring efficient navigation through the female reproductive tract, especially as sperm approach the egg for fertilization.

1.4 Semen analysis

The assessment of male fertility is performed through semen analysis (SA) according to strict guidelines provided by the WHO. The standard protocol recommends a period of abstinence of 7 days prior to sperm examination. It includes the evaluation of both macroscopic (sperm volume, pH, and viscosity) and microscopic parameters (concentration, vitality, motility and morphology) [6].
Semen volume helps assess the accessory glands' functional capacity, which contribute fluids to the ejaculate. Abnormal volumes can indicate potential issues such as obstructions or glandular dysfunction**.** Normal values fall in the range between 1.5 and $5ml$. The evaluation of pH indicates the balance between the acidic prostatic secretions and the seminal fluid. An abnormal pH could suggest infections or blockages in the reproductive tract. Normal values are $7.2 - 8.0$. Viscosity is

assessed by observing how semen flows through a pipette or by drawing the semen up with a pipette and allowing it to drop. The sample is considered hyper-viscous if the drop forms a long thread ($>2\ cm$). High viscosity could be an indicator of infection, dehydration, or dysfunction of the accessory glands, with a consequent reduction of fertility potential [6].
Regarding microscopic parameters, the evaluation is usually performed under a microscope at $20x$ or $40x$ magnification. Sperm concentration is assessed by diluting the sample in a haemocytometer or Makler chamber, and counting for at least 200 cells across multiple fields. The results are averaged to provide the total number of sperm cells per ml. Low sperm concentration, known as oligozoospermia, can indicate reduced fertility, while a lack of sperm in the ejaculate is termed azoospermia. Vitality assessment provides precious information regarding the sperm sample and it is measured usually through eosin-nigrosin staining procedure. In cases where sperm should not be stained as in ICSI procedure, the Hypo-osmotic Swelling Test (HOS) is preferred. In this case, sperm is placed in a hypo-osmotic solution; live sperm swell in response to the solution, whereas dead sperm remain unchanged. This test helps determine the functional integrity of the sperm membrane, which is crucial for sperm to fertilize an egg.
For motility assessment, a small volume (usually $5 - 10\ \mu L$) of liquefied semen is placed on a clean microscope slide with a coverslip, and is kept at $37°C$. Cells are observed with phase-contrast optics at $20x$ or $40x$. The WHO recommends the classification of sperm into four motility categories. For humans the definitions are the following:

- Rapid progressive: Sperm that move actively in a straight line or large circles. The velocity from the starting point to the endpoint of the movement should be greater than $25\mu m/s$ (equivalent to approximately half of tail's length).
- Slow progressive: similar to the previous, but with a velocity between $5\mu m/s$ and $25\mu m/s$.
- Non-progressive: Sperm that exhibit movement but do not make forward progression, or small circular movements that result in a velocity below $5\mu m/s$.
- Immotile: Sperm that do not move at all.

The WHO guidelines use these categories to assess the percentage of motile sperm as it is considered a strong indicator of fertility potential [34]. A sample is considered normal if at least 40% of sperm are motile, whether progressive or non-progressive.
Sperm morphology refers to the size and shape of sperm cells and is an important parameter in semen analysis as abnormal sperm shape can negatively impact fertility [35]. Its evaluation differs from motility, as sperm is air dried, stained and observed at $100x$ magnification. Strict criteria define standard dimensions of good quality for head, midpiece and tail. A sperm sample is considered normal when there are more than 4% normal sperm. This low percentage is accepted because most men produce a high number of sperms with minor abnormalities, which do not significantly reduce fertility potential. This classification acts as a surrogate tool for selecting the most suitable ART procedures for infertile couples [19]. If the percentage of sperm with normal morphology is $\geq 4\%$, the likelihood of successful fertilization with ART is high. In cases where sperm count and motility are normal but morphology is poor, this factor alone may influence the decision to proceed with ICSI. When the percentage of sperm with normal morphology is $< 4\%$, the chances of successful fertilization through IVF are low, and ICSI is typically recommended [36]. Semen analysis outcome *per se* is influenced by a great variety of factors, related to both sample collection and handling and evaluation process itself). For example, results can be influenced by the duration of sexual abstinence before semen collection. The WHO recommends a period of $2 - 7$ days of abstinence. Shorter periods may result in lower sperm concentration, while longer periods can lead to increased sperm concentration but reduced motility and vitality [6]. The time between sample collection and analysis constitutes another factor to account for, as delays in analysis can lead to sperm degradation, reduced motility, and

changes in pH, potentially leading to altered results. The environmental temperature plays a role as well [37]. The major criticality for this analysis is subjectivity. As a matter of fact, especially for motility and morphology, the analysis outcome depends on the experience and precision of the laboratory technician. Moreover, the lack of standardization within and between laboratories contributes to lack of reproducibility and comparability [38].

1.5 Computer-Assisted Sperm Analysis (CASA)

Computer-aided sperm analysis (CASA) systems revolutionized the field of reproductive biology by providing an automated, objective method to evaluate sperm motility, concentration, and morphology. Initially developed in the late 1980s, CASA systems were designed to overcome the subjectivity and variability associated with manual sperm assessments. Such systems combine video microscopy, digital image processing, and sophisticated software algorithms to track and analyse sperm. There are several CASA systems that are currently available on the market. The Sperm Class Analyzer (SCA), developed by Microptics SL (Barcelona, Spain), evaluates semen concentration and motility by analyzing images captured through phase-contrast microscopy. On the other hand, the SQA-V GOLD, produced by Medical Electronic Systems (Los Angeles, California, USA), utilizes electro-optic technology to assess sperm concentration and motility parameters. IVOS and CEROS systems by Hamilton-Thorne (Beverly, Massachusetts, USA) incorporate advanced image processing capabilities, delivering detailed sperm analysis [39].

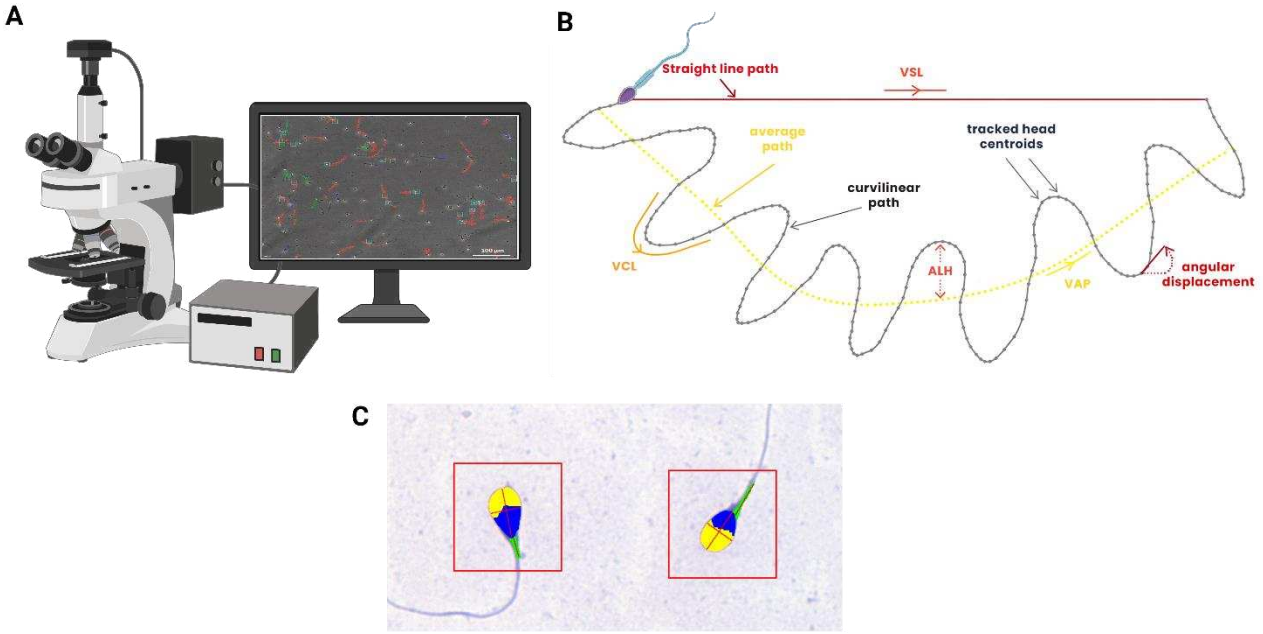

Figure 3: Computer-Assisted Sperm analysis (CASA) for motility and morphology. A) Typical CASA acquisition setup for motility analysis. It usually comprises of an inverted microscope equipped with a camera, and is connected to a computer. An example of output from CASA motility module is reported. Sperm tracks are color-coded according to the motility class as described by WHO. B) Sketch of CASA motility parameters. Reproduced from [6] C) Example of output from CASA morphology module for sperm head segmentation. Different parts of the sperm are coded with a different colour. Reproduced from [40].

These systems capture high-resolution video images of spermatozoa and convert them into a series of digitized frames (Figure 3A). Usually a phase contrast microscope is adopted since it increases the sperm head contrast with respect to the surroundings, making it easier for the algorithm to locate it. It is equipped with a heated plate to fix temperature. Images are acquired at different magnifications

depending on the application and digitized. The computer determines the number of picture elements (pixels) the sperm head covers, and if this number is within a certain range of pixel numbers that is acceptable for a sperm head, given the expected minimum and maximum size for the species, the computer will recognize it as a sperm head [41].
Sperm motility is evaluated starting from the trajectory of the single sperm cell (Figure 3B). The acquisition is usually done in phase contrast, at 20x magnification, and the sample is loaded onto slides of $10 - 20\ \mu m$ in depth. The following parameters are computed [31]:

- Straight-line velocity (VSL): The velocity of a sperm cell measured along the straight line between its starting and ending points during its movement.
- Curvilinear velocity (VCL): The actual speed of the sperm cell measured along its curved trajectory.
- Average-path velocity (VAP): The average velocity of a sperm cell along a smoothed trajectory that approximates the path.
- Linearity (LIN): The ratio of VSL to VCL, expressed as a percentage (VSL/VCL × 100). Indicates how straight a sperm's path is. A higher LIN suggests more linear, progressive movement, while a lower value indicates a more circular trajectory.
- Wobble (WOB): The ratio of VSL to VAP (VSL/VAP × 100). WOB measures the consistency of the sperm's path in relation to its curvilinear velocity. It provides an indication of how much the sperm deviates from a straight line as it moves
- Mean Angular Displacement (MAD): average change in direction (angle) per unit of time as the sperm moves. MAD is used to evaluate the change in movement direction and is relevant for assessing sperm that move in non-linear paths, such as during hyperactivation.
- Fractal Dimension (FD): A measure of the complexity of the sperm's trajectory, quantifying how irregular or chaotic the sperm's movement is.
- Amplitude of Lateral Head displacement (ALH): The magnitude of the side-to-side displacement of the sperm head during its movement.
- Beat-Cross Frequency (BCF): The frequency at which the sperm's head crosses the average path. BCF reflects the rate at which the sperm's flagellum beats and is related to the energy efficiency of sperm motility. Higher BCF can indicate more vigorous sperm motility.

The CASA systems output these parameters as mean values to qualitatively describe the sample and identifies motility groups based on static thresholds that change among species [42].
Sperm morphology is evaluated acquiring snapshots of fixed and stained cells at higher magnifications, typically $100x$ (Figure 3C). The system simultaneously evaluates the dimensions of the acrosome, head, midpiece and tail, and classifies sperm morphology according to standard of custom criteria [6].
Commercial CASA systems have revolutionised the field of semen evaluation, allowing the analysis of high number of cells in a small amount of time and reducing the subjectivity of standard semen analysis. However, several limitations hinder the use of such systems. As a matter of fact, CASA kinematic parameters related to the computation of the average path suffer lack of reproducibility among different instruments due to the different available computations [43-45]. It is usually evaluated smoothing the curvilinear trajectory with a fixed-length window, distorting the average path of irregularly swimming sperm. Moreover, they are very sensitive to the sample preparation and technical conditions. Factors such as the type of counting chamber used, the depth of the chamber, and the dilution of the sample can significantly influence the kinematic parameters measured by CASA [46] For example, the choice of the observation chamber can affect the space in which the

sperm move, which in turn alters the measurements of motility and concentration [47]. Frame rate and image acquisition settings are also critical. CASA systems typically rely on high frame rates between 30 and 60 Hz to capture sperm movement accurately. If the frame rate is too low, it may miss rapid movements, particularly in hyperactivated sperm, leading to inaccurate results [46, 48]. Expensiveness and inflexibility are other factors to be considered. The restrict access to raw data and the limited acquisitions settings led researchers to propose several open-source solutions to meet experimental needs.

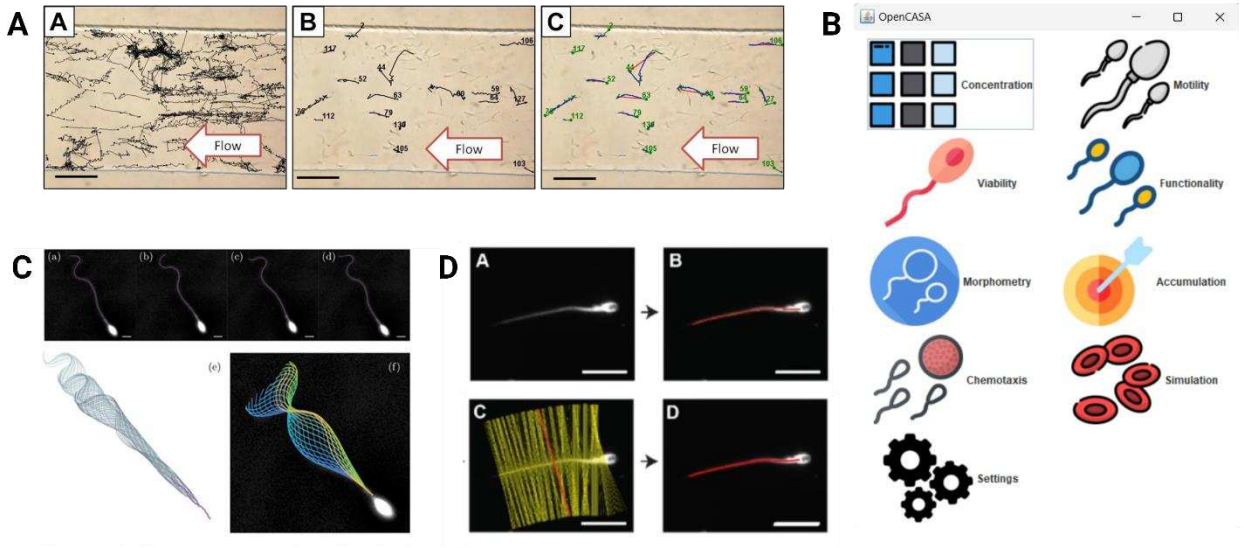

Figure 4: Alternatives to commercial CASA systems. A) Image-j plugin to track sperm motility into a microfluidic device, in the presence of a counterflow. Reproduced from [49]. B) OpenCASA multi-modular software for sperm analysis [50]. C) FAST software to reconstruct flagellar dynamics [51] D) SpermQ algorithm for flagellar tracking [52].

For these reasons, some authors have proposed free open-source alternatives to perform semen analysis.

In 2007, Wilson-Leedy and Ingermann proposed the first open-access ImageJ plugin to analyse zebrafish spermatozoa. The algorithm is based on the MTrack2 plugin for ImageJ, which is an open-source software for tracking particles. The system calculates the sperm motility parameters by following the trajectory of sperm heads across multiple frames of video. The sperm cells were acquired with a $10x$ phase-contrast objective at a frame rate of $\sim 100\ fps$ in $12\mu m$ deep slides. To isolate sperm from the background thresholding is applied, and motion is tracked by assigning x and y coordinates to each sperm head and linking them across frames according to proximity. Points distant more than a certain threshold are not linked. [53]. This algorithm was applied to mammalian sperm, obtaining results in good agreement with commercial sperm analysis systems [54].

As mentioned before, commercial CASA systems are also limited in the acquisition setup. Based on Wilson-Leedy plugin, an improvement was proposed to account also for sperm motion in a microfluidic device (Figure 4A) [49]. The microfluidic device used in the study consists of a straight microchannel designed using soft lithography techniques. The channel dimensions are 200 μm width, 20 μm height, and 3.3 cm in length. Human, chicken, and bull sperm were analysed separately at $40x$ in brightfield at a frame rate of $40fps$. Improvements of the previous algorithm regard the possibility to use brightfield images as input and the possibility to account for cells temporarily going out of focus. A sliding window method is applied to search for sperm cells over multiple following frames if they are not located in the next immediate frame. The search area expands with each subsequent frame to accommodate the possible distance the sperm may have travelled while out of focus. If a sperm cell is found in one of these later frames, it is linked to the ongoing track, and the

search process proceeds as before. The number of frames for this search was optimised to 6 to not increase risks of collisions. Besides traditional CASA parameters, three novel parameters were proposed:

- Orientation: Measures the angle between the sperm's VSL and the flow direction.
- Proximity to wall: Determines how close the sperm swims to the microchannel walls.
- SwimVAP: The sperm's swimming velocity relative to the liquid flow.

This approach allowed the study of sperm rheotactic and wall-tracking behaviour, which would have not been feasible through commercial CASA systems (Figure 4A).

In order to expand the kinds of analysis of open-source CASA software, OpenCasa was proposed (Figure 4B) [50]. This software further expands the Wilson-Leedy and Ingermann plugin to incorporate morphometry, viability and chemotaxis analysis. The motility module works with video sequences of sperm captured under phase-contrast microscopy. The settings of the detection and tracking algorithm can be tuned to accommodate for different acquisitions. The implementation of the fractal dimension was incorporated as it is a good descriptor of hyperactivated behaviour [55]. The chemotaxis module allows the evaluation of sperm instantaneous directionality angles ψ with respect to a user-defined gradient direction. In this way, it is able to account for sperm behaviour in presence of a chemoattractant. The morphometry module works on stained sperm images and computes morphometrical parameters such as area, perimeter, mean intensity, length, width, ellipticity, roughness, elongation and regularity. As for vitality, it is evaluated identifying stained cells with respect to non-stained ones from RGB snapshots. OpenCASA offers a comprehensive suite for sperm analysis, making it a flexible tool for fertility studies in multiple contexts.

Several closed-source software implement motility analysis through different acquisitions and with more complex tracking algorithms. For example, the Joint Probabilistic Data Association Filter (JPDAF) was adapted from radar tracking algorithms [56]. Each sperm's position and velocity are estimated with independent Kalman filters. The JPDAF addresses occlusions and tracks sperm in undiluted samples even during cell collisions. Sperm paths are reconstructed, and motility parameters are continuously calculated to account for temporal variation of kinematic parameters. The same tracking logic was adopted in [57], where the Interacting Multiple Models (IMM) framework is incorporated to address the sperm's nonlinear, manoeuvring motion. IMM is based on the idea that an object can move in different ways depending on its state. For example, sperm may alternate between Constant Velocity (CV), moving in a straight line, and Constant Turn (CT), moving in a circular fashion. Through Kalman filters, the algorithm predicts the sperm head position accounting for both models of motion, being able to handle sperm that exhibit either straight or circular swimming patterns. In another work, the algorithm uses frame difference for background subtraction, followed by non-linear diffusion filtering to remove noise [58]. A Hungarian search method is used for optimal matching of sperm cells across frames, handling occlusion and complex tracking scenarios. A novel approach based on smartphone image acquisition was proposed in [59]. In this work, a smartphone was mounted on the ocular part of the microscope via a designed tool and videos were recorded at 30 frames per second. The tracking software is based on the Recursive Kalman Filter (RKF) to estimate the position and velocity of human sperm cells over time by minimizing tracking errors caused by noise. The RKF works by predicting the next position of the sperm based on its current velocity and trajectory. This prediction is then updated based on the actual observed position from the next video frame. By recursively updating the position and velocity estimates, the filter smooths out any sudden noise or inaccuracies in the tracking data.

CASA systems dedicated to the analysis of the tail were proposed to gain more insight into sperm behaviour. For example, FAST (Flagellar Analysis and Sperm Tracking) calculates simultaneously both human sperm head and tail kinematics (Figure 4C) [51]. This approach utilizes negative phase-

contrast microscopy for image capture, ensuring that both the head and flagellum are clearly visible, and a high frame rate for accurately capturing the rapid oscillations of the flagellum. The software analyses flagellar waveforms to measure the detailed movement of the sperm's tail. It tracks the flagellar beat pattern and extracts parameters such as the flagellar beat frequency, curvature, arc-wavelength, and amplitude of the waveform. Furthermore, FAST estimates the power dissipation or the energy lost during the flagellar movement. The study examined the connections between flagellar parameters and traditional CASA metrics, uncovering more distinct variations in sperm swimming patterns under different viscosity conditions. An analogous software is SpermQ, which is able to track both sperm head and flagellum [52]. A key feature is the estimation of the flagellum z position starting from a 2D acquisition in a custom-made observation chamber $150\ \mu m$ deep (Figure 4D). This allows the study of the kinetics of the flagellar beat in the direction vertical to the focal plane, providing more insight into sperm 3D motion.
While many research efforts have focused on improving tracking algorithms, there has been less emphasis on the characterization of sperm motility, resulting in a gap in clinically relevant insights. Some studies in fact highlight the issue of poor statistical data and the absence of detailed descriptive parameters, which complicates comprehensive sperm analysis [60]. As a matter of fact, morphology and motility are still evaluated following two different protocols. Thus, these pieces of information remain uncorrelated at the single-cell level.

1.6 Single-cell tracking

Unravelling cell complexity is still a major challenge in biomedical research. While on one hand many advantages were done in recent years, on the other there are still many things to be discovered regarding, for example, signalling pathways and molecular responses to stimuli [61]. One element contributing to this difficulty is the fact that is very hard to account for the variety of cell characteristics even among the same population, making an average response insufficient to provide such in-dept information. Compared to the average analysis of millions of cells together in bulk, single-cell analysis can provide more detailed information on molecular contents of cells related to cell state and type, spatial and temporal transformations, and micro-environment, being a valuable tool in fields like stem cells characterization, tumor progression, sequencing (genomics, proteomics, transcriptomics, metabolomics), drug screening and diagnosis [61, 62]. In this context, single-cell tracking analysis aims at providing information on moving cells, thus being able to capture cells heterogeneity in the interaction and response to different signals. Moreover, this approach holds significant potential in diagnostics and personalized medicine [63]. For example, following single-cells migrations paths allows the identification of cancerous cells with altered motion patterns or biophysical properties [64-66] , cells lineage monitoring [67, 68], and phenotyping [69, 70]. Due to the large amount of data typically produced by a live-cell imaging experiment, downstream extraction of biologically relevant information becomes the rate-limiting step. With the development of image and vision computing technology, computer-vision-based cell characterization has been applied in label-free cell analysis to overcome these difficulties, often adopting machine learning and deep-learning solutions for high throughput.

1.7 Artificial intelligence for sperm analysis

Artificial Intelligence (AI) is revolutionizing the field of semen analysis by offering precise, automated methods that mitigate human error and variability. AI technologies, such as machine learning (ML), neural networks (NNs) and deep learning (DL), enhance the accuracy of sperm assessment. As AI continues to evolve, its integration in clinical settings promises to offer new insights into reproductive health [71].

1.7.1 Neural networks

In the brain, neurons communicate by passing signals to one another through synapses. Artificial Neural Networks (ANNs) are built to mimic neuronal communication. As a matter of fact, they are composed of artificial neurons (or nodes) organized into layers, with connections between them. Information is processed as it passes through these layers, where the output from one layer serves as the input to the next. Eventually, the data passes through all the layers, until it reaches the final output layer, where the network delivers the final result. Over time, a wide variety of neural network architectures have been developed, each designed to address specific types of tasks, ranging from image recognition to natural language processing. The basic structure of ANNs comprises of layers made of a number of interconnected neurons. Each neuron in a given layer is connected to all neurons in the next layer, forming a fully connected structure. In mathematical terms, considering a neuron i having m inputs (Figure 5A), the output is a function of the weighted sum of its inputs x_i, expressed as:

$$z_i = f\left(\sum_{j=1\ldots m} w_{ij}x_j + b_i\right) \quad (1)$$

where w_i are the weights associated with each input x_i, b is the bias associated to the neuron, and f is the activation function. There are three kinds of layers: input, hidden and output layers (Figure 5B). Networks that have a high number of hidden layers, from dozens to hundreds, are called deep neural networks (DNNs). The learning process of a neural network involves adjusting its weights and biases to minimize the error in predictions. The output z is the neuron's prediction, and is propagated forward to the following layers until the output layer. The predicted value of the output layer is compared to the ground truth input label to estimate the loss. This minimization process occurs through backpropagation. The network calculates the gradient of the loss with respect to each weight and bias, propagating the error back to the input layer and updating the parameters accordingly. The activation function f is necessary to introduce non-linearity in the network, enabling it to learn from non-linear complex data. Among the most used activation functions there are the sigmoidal function, the rectified linear unit (ReLu) and the Leaky ReLu (FIGURA 5C) [72]. The sigmoidal function transforms any number in input into an output comprised between 0 and 1, being suitable for binary classification tasks. This function, however, is subject to the vanishing gradient problem, that is when components of the gradients of the loss function get close to 0. Since the derivative of the sigmoid function is a bell curve centred on 0, when the gradient is very small a stall in the parameters update arises and the network cannot learn further [73]. When the ReLU (Rectified Linear Unit) activation function is used instead of the sigmoid function, the partial derivative of the loss function takes on values of either 0 or 1 avoiding the vanishing gradient problem. However, ReLU can lead to deactivated nodes when the gradient is 0, meaning the weights stop updating and the neuron ceases to contribute to learning. To prevent this issue, the Leaky ReLU function is often used. When the input is negative, the weights keep updating and nodes deactivation is prevented by introducing a small, non-zero gradient.

The choice of the loss function depends on the task [74]. The minimization of the loss function and the subsequent updates of the weights and biases follows an algorithm called gradient descent (Figure 5D) [75]. The gradient descent algorithm computes the gradient of the loss function with respect to all the training data at once. After calculating this gradient, the model updates its parameters. the algorithm performs the partial derivative of the loss function with respect to the network parameters, retrieving information about the direction and rate of the fastest increase. The rule for the parameters update is the following:

$$\theta_{i+1} = \theta_i - \alpha \nabla J(\theta_i) \quad (2)$$

Where θ is the network parameters at the end of the *ith* iteration, ∇J the gradient of the loss function and α the learning rate. This hyperparameter dictates how large a step is taken in the direction of the gradient. If the learning rate is too high, the algorithm may overshoot the minimum and fail to converge. If it is too low, the convergence will be slow, and the algorithm might get stuck in local minima. The gradient descent, despite being easy to comprehend and implement, can be computationally intensive and require much time to converge to an optimal minimum. For this reason, other gradient optimization algorithms have been proposed [76]. The Stochastic Gradient Descent (SGD), for example, updates the model parameters after loss computation on each training sample, alike the canonical gradient descent which updates parameters after calculating the gradient on the whole training samples. It is therefore usually much faster with great loss fluctuations. If on one hand this could be good to jump to new possibly optimal minima, on the other it may cause overshooting. This issue is solved by the Mini-Batch SDG, which introduces a new hyperparameter, namely the mini-batch. This algorithm computes the loss for a subset of the training sample whose size is dictated by the batch. The advantages are reduction of variance of the parameter updates and less memory requirement. Momentum was introduced for further reducing high variance in SGD and soften convergence [77]. It accelerates the convergence towards the relevant direction and reduces the fluctuation to the irrelevant direction, by adding a fraction γ of the update vector of the past time step to the current update vector v_i:

$$v_i = \gamma v_{i-1} + \alpha \nabla J(\theta_i) \quad (3)$$

$$\theta_{i+1} = \theta_i - v_i \quad (4)$$

The momentum term γ is usually set to 0.9 or a similar value. A variation of Eq.3 is proposed in the RMSprop algorithm, one of the most used optimization strategies for neural networks [76]. Instead of keeping the learning rate α constant, it is scaled down by an exponentially decaying average of squared gradients, ensuring smoother and more consistent updates. The concept of decaying learning rate was further developed in the Adaptive Moment Estimation (Adam) algorithm [78]. It combines the ideas from RMSProp and momentum, making it one of the most widely used optimization algorithms in deep learning. As a matter of fact, it adapts the learning rate individually for each parameter using both first-order moments (mean) and second-order moments (variance) of the gradients.

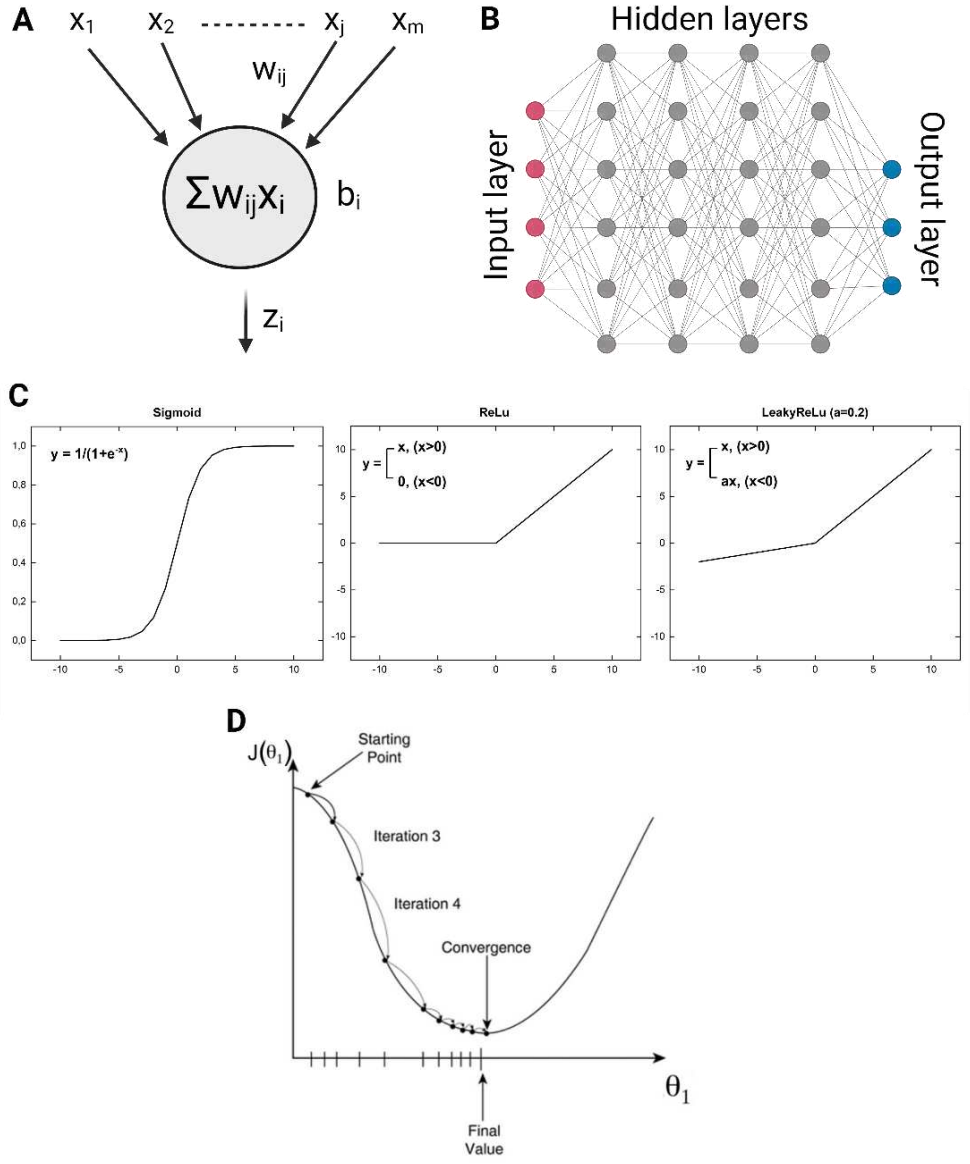

Figure 5: Elements of neural networks. A) Scheme of a neuron. The inputs x are weighted with the set of weights w and biases b. The activation function determines the value of the output z. B) Layers composing a neural network. C) Activation functions that introduce non-linearities in the models, favouring the learning process. D) Gradient descent logic for the minimization of the loss function. At each iteration, the algorithm performs the partial derivative of the loss function with respect to the network parameters. Parameters are updated in the opposite direction with respect to the gradient, with a step depending on the learning rate. The learning process is optimum when the minimum of the loss function gradient is reached. Reproduced from [79].

The final goal of every neural network application is generalisation, namely the ability to perform well on unseen data. To achieve this, besides hyperparameters tuning, the training dataset should be managed properly. Usually the dataset is divided in three parts:

- Training set: the portion of dataset used to train the model. It should be large enough to yield meaningful results and representative of the dataset.
- Validation set: helps monitor the model's performance during training and tune hyperparameters accordingly.
- Test set: Provides an unbiased evaluation of the model's performance on unseen data after training.

A common approach is a 70-80% training set, 10-15% validation set, and 10-15% test set. This split ratio ensures that the model has enough data to learn from, while maintaining sufficient data for unbiased evaluation. The dataset dimension is fundamental to avoid errors in the training process. As a matter of fact, it may happen that if the training set is too small, the model learns the training data too well and performs poorly on unseen data. Overfitting is occurring when the validation loss starts increasing while the training loss keeps decreasing. To manage overfitting, techniques such as regularization, dropout, and early stopping can be used. Regularization techniques add a penalty to the loss function, discouraging the model from becoming overly complex or fitting noise in the training data. Dropout involves randomly ignoring neurons during training to prevent over-reliance on specific neurons. Early stopping monitors the model's performance on a validation set during training and terminates training when validation reaches a plateau or starts increasing.[80] On the other hand, underfitting happens when the model is too simple and fails to capture the underlying patterns in the data. This results in poor performance on both training and validation sets. The solution often involves increasing the model's complexity or gathering more data. To improve the generalizing ability of a network in cases of limited labelled data availability or imbalances in the dataset, data augmentation has been proposed, especially for computer vision applications [81]. Classical data augmentation techniques are affine transformation like rotation and flipping, scaling, intensity and contrast transformations, cropping and padding. It has been shown that data augmentation not only improved accuracy but also enhanced neural networks' robustness to unseen data variations, proving to be of great value for medical applications where the limited data availability is critical [82, 83].
Network performances are evaluated through different metrics according to the application [84]. For example, in a binary classification problem, data instances are classified as either positive or negative. The threshold that marks the distinction between positive and negative, called Intersection over union (IoU), is usually set by default as 0.5. A positive label typically indicates the presence of a condition, anomaly, or deviation, while a negative label suggests no deviation from the norm. There are four possible outcomes for each predicted label: a true positive (TP) occurs when a positive instance is correctly identified, a true negative (TN) is when a negative instance is accurately predicted, a false positive (FP) happens when a negative instance is incorrectly classified as positive, and a false negative (FN) occurs when a positive instance is mistakenly labelled as negative. Combination of these labels are used as performance metrics, defined accordingly to the application.

1.7.2 Convolutional neural networks

In computer vision applications the key tasks are detection, classification and localization [85]. Classification involves determining the category of an object within an image or video by assigning a class label either to the entire image or to a specific region of interest. Localization focuses on finding the object's position in an image or video by generating a bounding box around it, which provides the coordinates that define where the object is located. Segmentation, on the other hand, is the process of identifying the individual pixels that make up an object in an image or video. This enables the creation of a pixel-level mask that outlines the object's shape. Unlike bounding-box localization, segmentation provides more accuracy as it captures precise object boundaries, making it particularly useful in applications like biomedical research [86].
CNNs are specifically designed for processing structured grid-like data such as images. Alike fully connected neural networks, where every neuron in the network is connected to every neuron in adjacent layers and so spatial features among pixels are lost, CNNs are able to take spatial correlation into account. To do this, they use convolutional layers to automatically detect local patterns such as edges or textures in images. The traditional CNN structure is shown in Figure 6A.

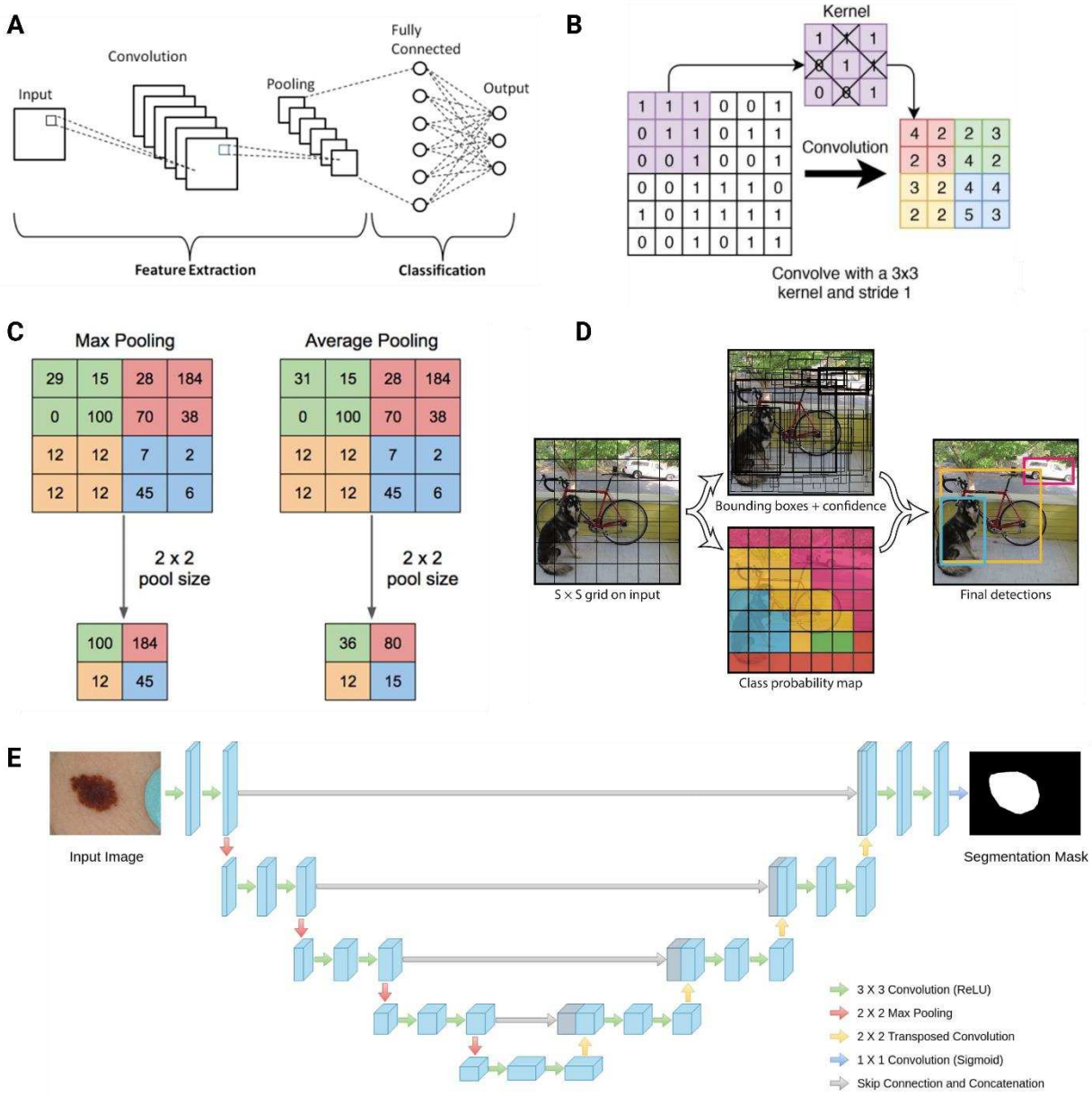

Figure 6: Convolutional neural networks for computer vision tasks. A) Structure of a convolutional neural network. Reproduced from [87]. B) Example of convolution operation. The kernel slides on the input with a fixed stride, and computing the dot product between the values in the kernel and the input at each position. The output is a feature map of reduced dimensions. Reproduced from [88]. C) Max and Average Pooling operations. At max pooling, each filter is taken the maximum value, then arranged into a new output with. While the average pooling value taken is the average value of the filter size. Reproduced from [89]. D) YOLO bounding box generation logic. The network divides the input in grids, predicting a set of predefined bounding boxes called anchors, with confidence scores and class probabilities for each box. Those having highest scores are selected for the final output. Reproduced from [90]. C) Segmentation of a skin lesion with the U-net. The model has an encoder part which downsizes the input, retrieving low-level spatial information. The decoder part upsizes the input, and information from the initial layers are correlated through skip connections (grey arrows). This structure allows the retrieval of multi-scale features. Reproduced from [91].

The basic structure of CNN consists of two blocks, the feature extraction layer and the feature mapping layer [87]. The feature extraction block is mainly composed of a convolutional layer and a pooling layer. The convolutional layer consists of several convolution units, and the weights and biases of each convolution unit are optimized by the backpropagation algorithm. The set of weights of a convolutional unit are called kernel or filter [92]. In the convolutional layer, the same kernel

slides on the input image with fixed stride forming the feature map (Figure 6B). In this way, each pixel in the feature map is only connected to the part of the previous feature map. This strategy is called weight sharing and reduces significantly the number of network parameters, thus the computational cost [93]. The filtering process is computed with more than one kernel at a time, scanning the image to search for multiple features and producing multiple feature maps. The max pooling layer reduces the size of the feature map, aiding in retaining the crucial information or key features of the input image while also improving computational efficiency. Pooling generates a lower-resolution representation of the input that still preserves the primary or important aspects of the original image [93]. There are two types of pooling techniques, Max and Average Pooling (Figure 6C). While Average performs down-sampling by dividing the input into rectangular pooling regions and computing the average values of each region, max pooling selects the higher value pixels discarding less relevant information [94]. After the feature extraction phase, the next stage is classification. In this step, the input image is assigned a label using this layer. The information from both the convolutional and pooling layers is passed through this layer, which connects them to the output layer, where the input is ultimately categorized based on the classification criteria.

Object detection models typically aim to perform detection, localization and segmentation simultaneously to comprehensively understand the objects in an image or video. Deep learning-based object detection methods generally fall into two categories, Two-stage and one-stage object detection. In Two-stage approaches, the first stage involves generating object region proposals analysing image features with a CNN backbone, finding possible candidate regions from the feature map. Then, sliding window operations are performed on the candidate regions for classification and bounding box refinement[95]. While this method is slower compared to others, it achieves the highest accuracy. Algorithms such as RCNN, Faster-RCNN, and Mask RCNN are examples of two-stage detectors. One-stage detectors on the other hand skip the region proposal phase and directly predict the bounding boxes and class labels from the input images. Although faster than the two-stage approach, it may struggle to detect smaller objects. One-stage detectors are highly suitable for real-time applications due to their speed of inference [85]. In this context the YOLO network was proposed [90], rapidly gaining fame for its valuable performances in a wide variety of fields [85]. As a matter of fact, it marked the first instance of a real-time, fully integrated method for object detection. The acronym "YOLO" stands for "You Only Look Once", emphasizing that the network performs the detection task in a single pass. This approach contrasts with older methods that relied on sliding windows combined with classifiers, which required running multiple times per image to detect objects. The first step of YOLO is to divide the picture into grids of equal size (Figure 6D). Each grid cell is responsible for detecting objects whose centre falls within that cell. For each grid cell, YOLO predicts several bounding boxes, confidence scores for those boxes, and class probabilities. The confidence score represents the model's certainty that an object exists within the box and the accuracy of the predicted bounding box, while the class probability score indicates what class the detected object belongs to. A threshold can be set for the minimum confidence score that the network will consider the prediction to be a true prediction. Each cell grid predicts several bounding boxes. The less relevant are discarded according to the Non-Maximal Suppression logic [96]. Thanks to the outstanding performances, YOLO is constantly evolving, with each updated version introducing enhancements to boost both accuracy and speed for real-time object detection and face new challenges related to real-time implementations [85, 97].

Image segmentation is a fundamental task in computer vision. However, the requirement for low-level spatial information poses great challenges. Deep learning techniques, particularly convolutional neural networks (CNNs), have revolutionized the field of image segmentation by automating feature extraction [98]. There are mainly two kinds of segmentation, semantic and instance. Semantic segmentation principle is to connect each and every pixel of an image to a class name, assigning items belonging to the same class to the same colour class. Instance segmentation instead allows to distinguish individual objects belonging to the same class constructing a mask around each object in the image [99]. Models like Fully Convolutional Networks (FCNs) enable pixel-wise predictions,

improving performance over older machine learning methods [100]. The most widely used CNN for medical field is the U-net, thanks to its encoder-decoder structure (Figure 6E) [101]. The encoder part is contractive, meaning that the input is down-sampled through a series of convolutional and pooling operations while feature information are retrieved. After passing through the encoder, the input data is reduced in dimension but the depth, namely the number of feature channels, is increased. This allows the bottleneck to capture abstract, high-level features of the input image that are crucial for understanding global structures within the image. The decoder performs up-sampling operations, which increase the spatial resolution of the feature maps, restoring the details lost during down-sampling. One of the major innovations of U-Net is the use of skip connections between the encoder and decoder paths (grey arrows in Figure 6E). These connections allow concatenation of the decoder feature map with the corresponding feature map from the encoder, using information from earlier layers to detect fine details in the input image [98, 101].
Neural networks have brought significant advancements in computer vision, but they come with both benefits and limitations. Automatic feature extraction allows to remove manual subjectivity from the analysis, especially in biomedical applications, where images (e.g., from MRI or CT scans) contain intricate patterns that manual methods might miss [86]. Furthermore, deep learning models have demonstrated superior performance in tasks like segmentation, detection, and classification. In medical imaging, this translates to more accurate diagnoses and detection of diseases such as tumors and lesions [100]. Finally, the ability to handle big datasets enhances generalisation potential and results reliability. However, the learning ability of a neural network depends very much on the dataset dimension and quality. This is especially difficult in cases where data are complex and hard to retrieve as the medical field. Furthermore, training neural networks has a high computational cost, requiring high-end hardware (like GPUs) and significant memory [102]. Neural networks, particularly deep models, are often seen as "black boxes." This lack of interpretability can be a concern in the medical field, where understanding why a model made a specific decision is crucial for clinical adoption [86].

1.7.3 Sperm analysis with neural networks

Neural networks have made significant strides in sperm analysis, enhancing the accuracy and efficiency of assisted reproductive technologies (ART). AI has proven to be an invaluable tool for evaluating key sperm parameters such as motility, morphology, and DNA integrity across various species, including humans and animals like bulls and rodents [103, 104]. There are several studies in literature where the classification task was applied to the male fertility research. Sperm head and acrosome were classified as normal or abnormal based on shape and vacuole presence in [105] with the use of a sequential CNN (Figure 7C). The dataset comprised 1540 brightfield low magnification snapshots (40x and 60x) of unlabelled human sperm cells. Similarly in [106] human sperm head was classified according to WHO guidelines using a convolutional neural network called VGG16 (Figure 7A). An approach based on holography was proposed to investigate stress-induced abnormalities on sperm conditions [107]. The authors used a custom-built partially spatially coherent digital holographic microscope (PSC-DHM) to capture interferometric images of sperm cells (Figure 7D). These images are then processed to generate quantitative phase maps, which contain detailed information about the cells' refractive index and local thickness. The phase maps serve as input to a deep neural network (DNN), which classifies the sperm cells into four categories: normal, cryopreserved, oxidative stress-affected, and ethanol-affected. Seven different DNN architectures were tested, but ResNet-101 provided the highest accuracy, achieving 85.6%, with a sensitivity of 85.5% and a specificity of 94.7%. Neural networks have been used to predict sperm motility. In [108] the authors explored the application of a deep CNN, specifically the ResNet-50 architecture, to predict sperm motility categories according to the WHO guidelines. The study utilizes ResNet-50 to classify sperm motility into different WHO categories: progressive motility, non-progressive motility, and immotile spermatozoa. The DCNN model showed significant potential in automating sperm motility classification, particularly for progressive and immotile sperm (Figure 7B). A similar task was implemented in [109]. The authors sought to improve existing methods by incorporating temporal

data through the use of 3D CNNs, rather than relying solely on frame-by-frame analysis. To achieve this, they trained three different 3D Residual Network (ResNet) architectures, which were designed to capture both the spatial and temporal features of sperm motility by processing the video data as 3D tensors. The best-performing model achieved an accuracy of 88.89% on the test data. There are several studies in literature implementing CNN for DNA integrity evaluation to improve cell selection for ICSI, offering non-invasive methods for live sperm DNA assessment [110-112]. With this purpose, McCallum *et al*. aimed to predict sperm DNA integrity using bright-field images of human sperm cells and trained VGG16 with DNA Fragmentation Index (DFI) values labelled via acridine orange staining (Figure 7E) [111]. A similar application has been implemented with phase-contrast snapshots of stained human sperm (Figure 7F) [112]. While the aforementioned methods evaluate DFI from fixed sperm, Noy L. *et al.* proposed an approach that works on live cells acquired through quantitative phase imaging (Figure 7E) [110].

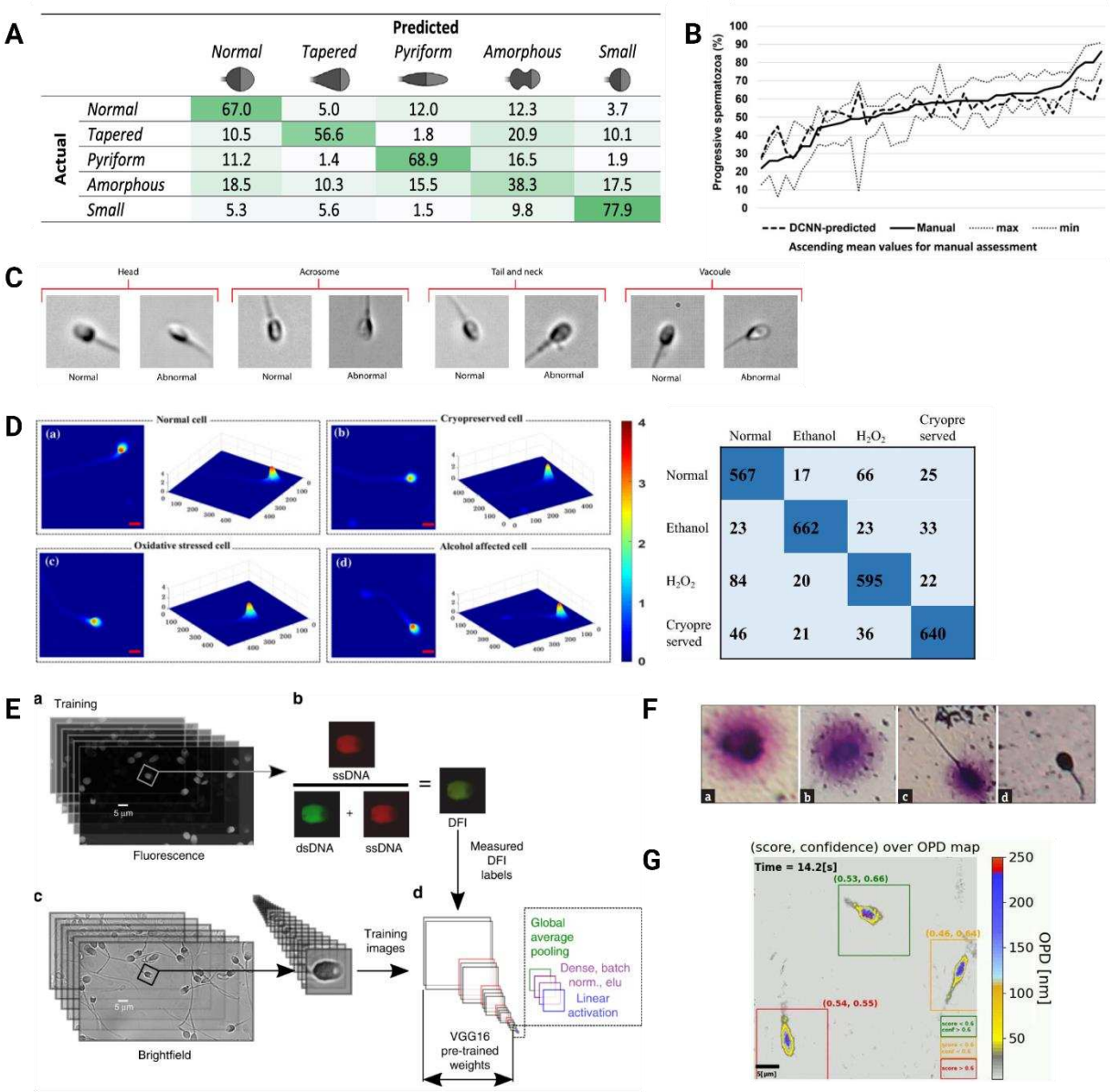

Figure 7: Classification tasks in fertility. A) Sperm head classification according to head shape. Reproduced from [106] B) Prediction of progressive motility based on optical flow. Reproduced from [108]. C) Classification of human sperm according no normality or abnormality of head, acrosome, tail and neck and based on presence of vacuoles. Reproduced from [105]. D) Phase maps generated with spatially coherent digital holographic microscope. Sperm cells are classified

in normal, cryopreserved, oxidative stress-affected, and ethanol-affected. Reproduced from [107]. E) Scheme of the CNN used to predict DNA integrity. Human sperm was acquired in brightfield and fluorescence to estimate the DNA fragmentation index, used to train the network. Reproduced from [111]. F) Phase-contrast snapshots of stained human sperm, used to predict DNA fragmentation index. Reproduced from [113]. G) Quantitative phase image for live cells' DNA fragmentation index prediction. Different colors encode for different DNA score predictions. Reproduced from [110].

Sperm detection is the location of cells in the field of view and is the base of motility analysis. For this reason, there is a growing interest into the implementation of this task with neural networks to develop fast and effective workflows. In [114] a solution was proposed to detect bull sperm cells in videos with high sperm density (Figure 8A). The main challenge the paper addresses is the difficulty in detecting small, fast-moving objects in low-resolution videos with partial occlusion and artefacts. As a matter of fact, the dataset comprises phase-contrast videos of bovine sperm cells acquired at 10x magnification, with an average number of 370 cells in each frame. The proposed network called DeepSperm is based on YOLOv3, although the number of layers is reduced for speed and the detection heads of the network is only one. In fact, with respect to the first version of YOLO, YOLOv3 introduces a multi-scale detection mechanism by using three detection heads, each working at a different scale of feature maps [115]. DeepSperm has only one head because the proposed method is meant to detect only small objects, hence a multi-scale prediction is not required. Compared to the performances of YOLOv3 and YOLOv4, this approach obtained comparable results in terms of accuracy, slightly better results in terms of F1score and much better speed and training time. The DeepSperm network was further modified in [116] to detect bull sperm acquired at higher magnifications. Moreover, following the detection of sperm heads, a tracking logic was implemented to calculate VSL, VCL, and LIN. These parameters were used to train a Support Vector Machine (SVM) to predict the percentage of progressive and non-progressive sperm according to WHO standards. The resulting accuracy with respect to the ground truth labelled by expert operators was 92%. Along these lines is TOD-CNN [117], which incorporates YOLO concepts but integrates features from ResNet and SPPNet for multi-scale feature extraction (Figure 8B). Following detection, human sperm cells are tracked using a k-nearest neighbours (kNN) algorithm, which matches detected sperm cells across consecutive video frames to generate movement trajectories. TOD-CNN's tracking accuracy was high, with lower error rates in calculating sperm motility parameters like VSL**,** VCL**,** and VAP**.** The network achieved an AP (Average Precision) of 85.60%, an F1 score of 90.00, and a detection speed of 35.7 FPS. It was compared against YOLOv3, YOLOv4, SSD, RetinaNet, and Faster R-CNN, outperforming these models in sperm detection tasks. An approach based on sperm detection and tracking was proposed specifically for ICSI applications, in order to address the subjective nature of manual sperm evaluation [118]. The study utilized a newly created dataset, the Jikei sperm dataset (JSD), consisting of over 4,000 unstained sperm images from ICSI patients acquired at 40x with an inverted microscope (Figure 8C). Yolov3 was trained to classify normal, abnormal, unclassifiable sperm and vacuoles. The bounding boxes generated by the network were tracked with the simple online and real-time tracking (SORT) algorithm. For abnormal sperm detection, the model achieved a high sensitivity of 0.881, while for normal sperm the sensitivity was 0.794. A similar solution was proposed in [119], with the aim to identify whether a semen sample is suitable for an artificial insemination procedure. The Faster R-CNN architecture was trained to detect sperm heads on the VISEM dataset [83], which includes sperm images captured under a 40x magnification microscope. These videos had a resolution of 640x480 pixels and a frame rate of 50 fps. Each frame in the dataset had annotations for sperm heads, while other parts of the sperm (like the tail) were not included. The network was tested with three different learning rates and four different numbers of iterations (Figure 8D). Once the sperm heads are tracked, their velocity is evaluated calculating the path covered in one second and the percentage of motile and non-motile sperm is evaluated. In [120], a comparison was made among several CNNs to detect human sperm and non-sperm object in video frames. The study concluded that YOLOv5 was the most effective model for detecting sperm cells due to its balance between accuracy and speed. A modified version

of this network was proposed in [121], where a strategy called Shuffle Attention (SA) was introduced to improve the detection of small objects like sperm. To improve the training process, mosaic data augmentation was used. The model achieved an average precision (AP) of 87.5% on the test set, with a detection speed of 57.3 frames per second (fps). The YOLOv4 network was implemented to assess boar sperm vitality in a microfluidic device (Figure 8E). The sperm images were acquired with a portable micro-imaging system from samples diluted with saline and stained with eosin-aniline black, allowing differentiation between live and dead sperm based on membrane integrity. The network achieved an accuracy of 94.0% in detecting sperm survival rates. [122]. In order to improve sperm morphological analysis, several approaches adopting CNNs have been proposed. In [123] a workflow comprising two CNNs, k-means and SVM was used to segment head, nucleus, acrosome, midpiece and tail of human sperm semen smears. One CNN was deputed to the generation of a probability map of the sperm head region, while a second one segmented midpiece and tail. K-means clustering was applied to segment the internal parts of the head such as the acrosome and nucleus, while a SVM classifier was utilized to classify each pixel of the axial filament segments to extract the tail and mid-piece regions (Figure 8F). Several solutions adopting U-net have been proposed [124-126]. In [124], the dataset used for the study was created from 1207 microscopic images of human sperm cells, captured at 100x magnification. The aim of the study was to process sperm images containing multiple cells more effectively compared to previous methods, which typically handle only single sperm cells (Figure 8G). Dice Similarity Coefficient (DSC) of 95.14% and an F1-score of 77.23% resulted for sperm head segmentation. A similar application on RGB images achieved an Intersection over Union (IoU) score of 75% and a Dice similarity score of 86% on the test set [125]. To improve the quality of sperm images enhancing the segmentation process, some applications were proposed using more elaborated image acquisitions [126, 127]. In [126] the authors proposed a deep learning approach to improve sperm analysis through Image-Based Flow Cytometry (IBFC). The pipeline involves two steps. The first is the classification of sperm snapshots as in-focus or out-of-focus with ResNet50. U-Net was employed for the segmentation of focused images, specifically targeting the sperm head and flagellum. Another work introduced a method to predict reproductive outcomes by analysing sperm ultrastructure using label-free phase imaging combined with deep learning [127]. The authors employed a technique called Phase Imaging with Computational Specificity (PICS), which integrates Spatial Light Interference Microscopy (SLIM) with deep learning to analyse bull sperm cells (Figure 8H). This method offers non-destructive imaging of the sperm cell's ultrastructure, including its head, midpiece, and tail, and quantifies their dry mass. Sperm samples from five bulls were processed, fixed on microscope slides, and imaged using the SLIM system, which was attached to a phase-contrast microscope. The literature findings mentioned above are summarized in Table 1.

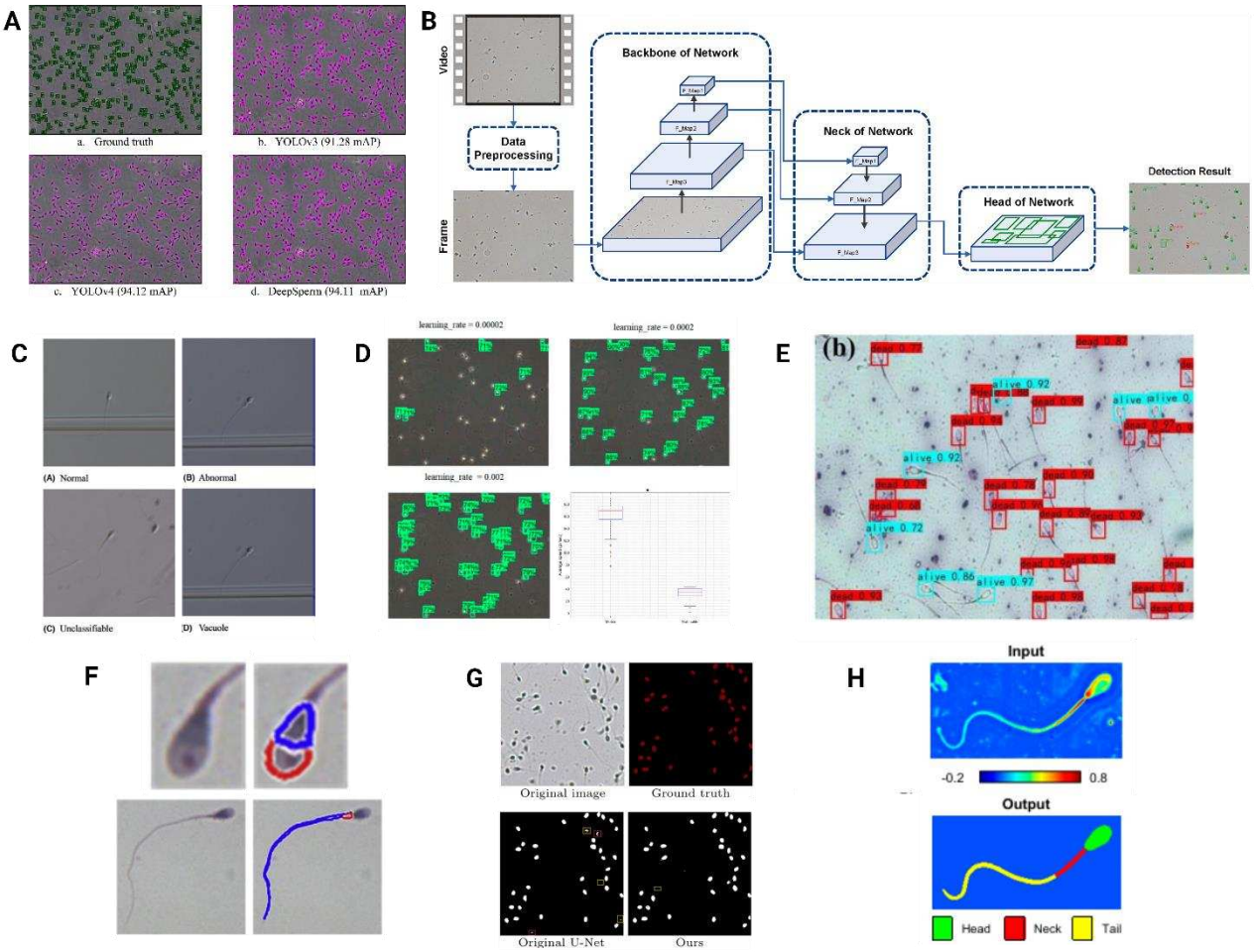

Figure 8: Sperm detection and segmentation applications with CNNs. A) DeepSperm for bovine sperm detection in crowded field of view [114]. B) TODCNN architecture for human sperm head detection. Reproduced from [117]. C) Detection of normality or abnormality of human sperm in IVF applications using YOLOv3. Reproduced from [118]. D) Faster R-CNN training with different learning rates to detect human sperm head in phase-contrast snapshots. Once detected, sperm are tracked and the percentage of motile and non-motile sperm is estimated. Reproduced from [119]. E) Detection of live and dead stained boar sperm in a microfluidic device with YOLOv4. Reproduced from [122]. F) Segmentation of head, acrosome, midpiece tail from human semen smears [123]. G) Comparison of original U-net and improved U-net model for sperm head segmentation. Yellow boxes indicate missing detections with respect to the ground truth. Reproduced from [124]. H) Segmentation of bovine sperm from high specificity phase imaging. Reproduced from.[127].

Table 1: Summary of the mentioned literature works regarding application of CNNs for sperm detection and segmentation.

NEURAL NETWORK MODEL	SPECIE AND IMAGE ACQUISITION	DETECTION	PERFORMANCE METRICS	REF.
Modified YOLOv4	Bull sperm 100x Phase contrast	Sperm head	AP = 91%	[114]
TOD-CNN	Human sperm, 40x Brightfield	Sperm heads and impurities	$AP50_{sperm}$ = 86% $AP50_{impurities}$ = 57%	[117]
YOLOv3	Human sperm, 40x Brightfield	Normal, abnormal and unclassifiable sperm	Positive Predictive Value (PPV) < 90%	[118]
Faster R-CNN	Human sperm, 40x brightfield	Sperm heads	Accuracy = 91.77%	[119]
YOLOv5x	Human sperm, 100x Brightfield	Sperm and non-sperm objects	AP_{sperm} = 84% $AP_{nonsperm}$ = 93%	[120]
Light-Weighted and Improved YOLOv5	Human sperm, 100x Phase contrast	Sperm head	AP = 88 %	[121]
YOLOv4	Human sperm, 40x RGB	Live and dead sperm	Accuracy = 94%	[122]
		SEGMENTATION		
CNN, K-means clustering for sperm segmentation	Human sperm 100x, RGB	Head, Axial filament Acrosome Nucleus Tail Mid-piece	Dice score = 0.90, 0.77 0.77 0.78 0.75 0.64	[123]
U-net	Human Sperm 100x Brightfield	Head	Dice = 95.14% F1 = 77.23%	[124]
	Human Sperm 63x RGB	Sperm	Dice = 86%	[125]
	Mouse Sperm, Flow cytometry snapshots	Head and tail	Dice = 84%	[126]
	Bull Sperm, Label free phase imaging	Head, midpiece and tail segmentation	F1 = 0.87	[127]

To merge multiple tasks for a more precise SA, there are several pipelines involving the use of more than one CNN to do sperm detection, tracking and segmentation. The cascade of CNNs proposed in [128] focused on human sperm head detection and segmentation in real-time using bright-field microscopy (Figure 9A). A three-stage method was proposed where the first step is a YOLOv3-tiny network deputed to sperm head detection. Then, a second convolutional neural network classifies sperm heads detected in the previous stage as in-focus or out-of-focus. Only in-focus snapshots are sent to a U-Net-tiny segmentation model to segment the head and vacuoles, from which morphological features are computed. This work adopts a microscope equipped with a Low-Magnification Wide-Field (LMWF) unit and high magnification small-field (HMSF) bright-field microscopy unit. The LMWF unit is used for simultaneous tracking of multiple freely swimming human sperms. Images of single sperms were collected using the HMSF unit with an objective lens of 60× or higher magnification at 200 frames per second. No kinematic parameters were evaluated in

this work. A Dice score of 0.948 for segmentation was obtained. A very similar approach was implemented in [129], where four neural networks operate multi-target tracking, instance and semantic segmentation, and classification (Figure 9B). At the end of the pipeline, human sperm acquired with 100x magnification in brightfield are categorized as progressive or non-progressive and as morphologically normal or abnormal according to the WHO guidelines. This work aimed at giving information about progressivity and morphology, without computing any descriptive parameters of sperm motion. None of the methods mentioned above correlate in any way to sperm kinematics at the single-cell level. In [130] the authors developed a method to assess morphology, motility, and DNA fragmentation of live sperm cells simultaneously (Figure 9C). The approach combines quantitative phase imaging (QPI), obtained with an interferometric module mounted on an inverted microscope, with deep learning to virtually stain sperm cells without chemicals. The dataset consists of 51809 dynamic human sperm cells from eight donors. The study utilized 4 CNNs of which one performs morphological virtual staining, the second classifies virtually stained images according to WHO morphological categorisation, the third computes the DNA fragmentation scoring and the fourth creates the virtual DNA staining on the basis of the fragmentation score computed in the preceding step. Motility according to WHO definition is evaluated through least-squares linear approach, which correlates the CASA parameters computed on virtually stained images to the WHO motility classes. The approach achieved high accuracy: 93.1% for morphology, 88% for motility, and 90% for DNA fragmentation [130]. To compute this simultaneous analysis on lower magnifications SFCnet was proposed [131]. This approach employs advanced image processing techniques to ensure reliable results at 20x magnification, computing both motility and morphology analysis as shown in Figure 9D. The Dice coefficient however was 64.14% for segmentation. The literature findings mentioned above are summarized in Table 2.

Table 2:Summary of literature findings regarding the application of cascades of neural networks for single-cell motility and morphology analysis.

NEURAL NETWORK MODEL	SPECIE AND IMAGE ACQUISITION	APPLICATION	PERFORMANCE METRICS	REF.
YOLOv3-tiny	Human sperm, 60× Brightfield	Head detection	F1 = 0.951	[128]
U-Net-tiny		Head segmentation	Dice = 94.8 %	
SegNet	Human 100x Brightfield	Head, midpiece and principal piece segmentation	Dice = 0.99, 0.976 and 0.998	[129]
EfficientNet		Motility classification	Accuracy = 90.82%	
Custom DNNs	Human sperm, Quantitative phase imaging	Detection	Accuracy = 93.1%	[130]
		Segmentation	Accuracy = 88%	
		Evaluation of DNA fragmentation index	Accuracy = 90%	
SFCnet	Human Sperm, 20x brightfield	Sperm segmentation	Dice = 64.14%	[131]

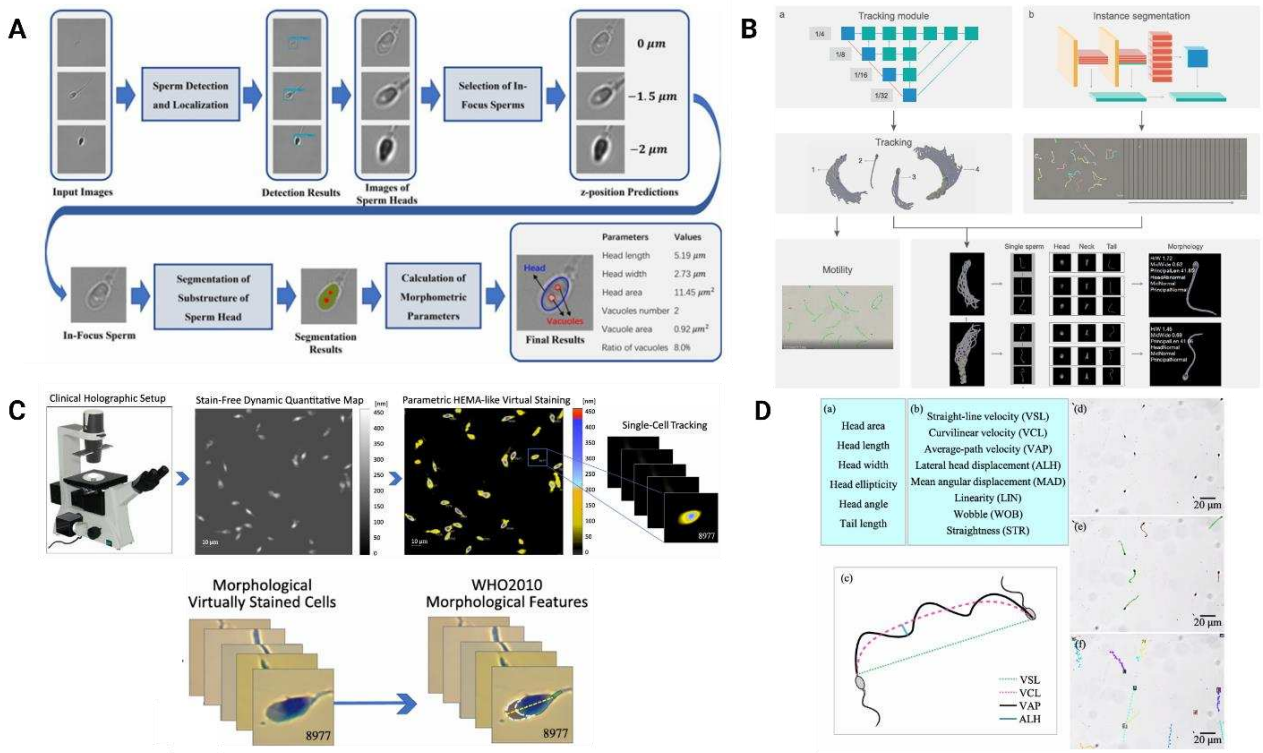

Figure 9: Pipelines of different studies employing cascade of CNNs for sperm analysis. A) Label-free morphological analysis of human. The pipeline detects sperm heads and classifies them as in or out of focus, segmenting only in focus sperm to do morphology analysis. Reproduced from [128]. B) Tracking and instance segmentation approach that outputs the percentage of motile sperm in the sample and the classification of morphology according to the WGO standards [129]. C) Workflow for the simultaneous assessment of kinematics, morphology and DNA integrity of human sperm. Reproduced from [130]. D) Kinematic and morphological analysis of sperm cells at low magnification [131].

AI has the potential to transform semen analysis by delivering consistent, efficient, and highly accurate results. They also process large datasets quickly, providing faster results and efficient clinical workflows. Additionally, many AI-based approaches are non-invasive, which helps preserve sperm viability for further use in clinical treatments. However, the effectiveness of neural networks systems relies heavily on access to large, well-annotated datasets. Inadequate or poorly labelled data can lead to inaccuracies. Furthermore, implementing solutions using complex deep learning models can be costly due to the need for advanced hardware like GPUs [71, 103].

1.8 Clustering methods for sperm classification

Good semen quality -based on a traditional subjective analysis of sperm number, motility, and morphological normality- does not guarantee acceptable fertility because other factors are involved, many of which interest the female partner. As a matter of fact, several trials used to define the significance of semen analysis results on predicting fertility success after ARTs have produced different conclusions [132]. Motility average values computed by CASA do not provide cell distributions nor adequate capture of spermatozoa heterogeneous populations. Therefore, there is a growing interest in the use of multivariate statistics to identify motility sub-populations as possible markers of fertilizing potential [133]. The discriminant and the cluster analyses are the two methods used for multivariate analysis classification. The first one is based on an a priori classification, while the second one defines a conjunct of classes required for a post hoc description of their meaning by

estimating the mathematical distances among the variables under consideration. Since there is no a priori knowledge about the presence of sub-classes, cluster analysis is the most suited technique as it allows unsupervised grouping of observations into subsets, called clusters, so that observations in the same cluster are similar depending on a given criterion [132]. The challenges of clustering analysis in sperm motility data also regard the dataset dimensionality and redundancy. As a matter of fact, CASA data include numerous kinematic parameters that are often correlated, complicating the clustering process and necessitating dimensionality reduction. Principal Component Analysis (PCA) reduces the dimensionality of sperm motility data while retaining most of the relevant information. PCA creates new variables (principal components) that are linear combinations of the original variables, aiming to simplify the analysis and improve clustering by reducing the correlation among variables. The richness of data allows PCA to construct components that summarize the key characteristics of the dataset effectively, even if individual variables do not contribute directly to the main findings. By including more variables, even if not all are highly relevant, PCA can extract more robust patterns, especially in exploratory contexts where the relevance of individual parameters might not be evident from the start. PCA is widely adopted for dimensionality reduction on CASA parameters of different species [134-141]. For cluster retrieval, cluster analysis allows for unsupervised grouping of observations into subsets (clusters) based on a similarity criterion. The goal is to assign similar sperm cells to clusters, thereby resolving the heterogeneity in sperm motility data. Clustering methods include hierarchical and non-hierarchical (or partitional) methods. Hierarchical clustering can be agglomerative (bottom-up) or divisive (top-down) [142]. Agglomerative clustering starts with each observation as its own cluster, merging them step by step until a desired number of clusters is obtained. Divisive clustering, on the other hand, starts with all observations in one cluster, splitting them iteratively until each observation is isolated. K-means is popular partitional clustering algorithm used for grouping data based on similarity, and it operates through the following steps:

1. Initialization: The user specifies the number of clusters, denoted by k, and the initial cluster centres (centroids) are selected randomly from the dataset.
2. Assignment: Each data point in the dataset is assigned to the closest cluster centroid based on a similarity measure, typically the Euclidean distance. This step groups each data point being grouped into one of the k clusters.
3. Centroid Update: After all points are assigned to clusters, the centroids are recalculated as the mean of the data points in each cluster.

Iteration: Steps 2 and 3 are repeated until convergence is reached, i.e., when cluster assignments no longer change or the sum of squared distances between data points and their respective cluster centroids is minimised [143]. K-means is widely used for sperm clustering applications, although it has limitations, especially sensitivity to outliers and noise [132, 144].

Many studies focusing on clustering sperm based on kinematic parameters have reported the presence of three to four clusters. In general, the “fast and linear” subpopulation is considered a good indicator of sample quality, while a predominant “slow and non-linear” one would be a marker of poor quality [145]. Besides identifying of sub-groups based solely on kinematic parameters [37, 139], some studies have correlated these parameters with other sperm quality indicators. For example, two different clustering approaches were investigated to identify sperm subpopulations (SPs) in bull semen [146]. In particular, sperm quality parameters evaluated through flow cytometry were associated with fertility outcomes and kinematic parameters. A positive correlation was found between spermatozoa with fast and nonlinear movements before freezing and bull non-return to estrous rates [146]. In another work, blue fox sperm was clustered based on both kinematic and morphometry parameters such as head area, perimeter, length, and width, ellipticity, rugosity,

elongation, and regularity. It revealed four sperm subpopulations each characterized by specific kinematic and morphometric attributes, indicating that motility and morphology can be combined to provide a new perspective to assess the ejaculate quality [141]. A similar approach was used to seek for a correlation among semen parameters and DFI with successful outcome of IVF procedures [147]. K-means clustering classified participants into four clusters and multivariable logistic regression assessed the association of clusters with low DFI and good semen parameters with good IVF outcomes. Another study based only on sperm head morphometric parameters identified subpopulations in four domestic species [148]. The resulting clusters showed high variability among species, differing mainly for size and elongation. To provide a more comprehensive assessment of human sperm quality, kinematic and morphometric parameters were combined to establish potential sperm subpopulations [149]. PCA was used to reduce the dimensionality of the dataset, resulting in three PCs for both kinematic and morphometric data. Then, a two-step clustering procedure was applied to identify distinct sperm subpopulations based on the principal components. Four distinct sperm subpopulations were identified. However, motility and morphology were evaluated with two different procedures, one requiring 10x magnification and acquisition in a shallow chamber, and the other requiring observation at 40x of stained smears on glass slides. Table 3 reports a summary of the research works reported above.

Table 3: Summary of the aforementioned literature findings regarding sperm cluster analysis. SPs stands for sub-populations.

SPECIE	MAIN GOAL	USED PARAMETERS	CLUSTERING METHOD	OUTPUT	REF.
Ram	Retrospective study on ram cryoresistance	Motility percentages, morphology, viability (plasma membrane integrity, HOST viability, mitochondrial activity	K-means and PCA	3 motility SPs; 2 morphometric sperm SPs; Correlations of parameters with cryoresistance	[137]
Boar	Multivariate analysis of boar sperm morphology	Morphometric parameters	PCA, K-means	4 SPs	[138]
Bull	Comparison of classical and Bayesian clustering in bull sperm	CASA motility parameters	PCA, K-means	4 motile sperm SPs	[139]
	Investigate the association of sperm sub-population with flow cytometric characteristics	Total sperm motility, progressive sperm motility, CASA parameters	PCA, K-means, Twostep clustering	4 SPs; Fast/linear and fast/nonlinear SPs correlated with good sperm quality	[146]
	Comparative study of sperm morphometry in artiodactyls	Morphometrical parameters (area, perimeter, length, width, ellipticity, rugosity, elongation, regularity)		4 SPs	[148]
Blue Fox	Integration of motility and morphology in blue fox sperm	Motility and morphology (head area, perimeter, length, and width, ellipticity, rugosity, elongation, and regularity)	PCA, Hierarchical clustering	3 morphometric and 3 kinematic SPs	[141]
Human	Correlation of DNA fragmentation index and CASA parameters with IVF outcomes	DFI, Volume, concentration, Slow and rapid motility percentages	K-means	4 clusters; Cluster with Low DFI and high motility % has good correlation with IVF outcomes	[147]
	Sperm classification based on a complete set of seminal variables	CASA and morphometrical parameters,	PCA, K-means	4 SPs based on kinematic parameters (84% of variance) and 4 SPs based on morphometrical parameters (74%)	[149]

1.9 Proposed approach

Current semen diagnostic techniques pose several limitations, above all the subjectivity of the gold standard evaluation. CASA systems on the other hand are very costly and inflexible due to limited access to acquired images and a lack of standardization in the algorithms, which hampers comparability and consistency of results. Kinematic parameters provided by the standard analysis are insufficient to account for the complexity and variability of sperm motion patterns. Assessing sperm morphology also presents challenges for CASA systems, as accuracy can be affected by staining techniques that alter sperm head dimensions, distorting morphometric analysis. Moreover, morphology and motility cannot be assessed on the same sperm cell due to different acquisition protocols, leaving these data uncorrelated at the single-cell level. In this context, Neural Networks have been widely explored in fertility research thanks to the computational power. Tasks like

classification, detection and segmentation have been widely used for multiple applications related to sperm analysis, ranging from motility and morphology evaluation to classification of healthy and unhealthy sperm based on WHO guidelines.

The aim of the thesis is the development of a computer-aided approach for a label-free analysis of sperm cells. The analysis is conceived as an automatic tracking of single cells for the retrieval of kinematic and biophysical features related to morphology and motion dynamics, in order to characterize sub-populations in the biological samples. Being label-free, it preserves the sample viability, making the process fast and easily reproducible in a clinical context. Bovine cells were employed as a representative model for mammalian sperm. Such approach involves the use of a cascade of two neural networks. The first one, YOLOv4-tiny, performs sperm head detection among the image sequence. The information regarding the sperm location is used to track single-cell heads, which are then segmented by the second neural network, U-net-tiny. Once the sperm head shape has been acquired across the image sequence, both kinematic parameters and morphometrical parameters can be acquired. Moreover, three additional parameters were introduced to describe sperm motion at the single-cell level, overcoming the limitations of standard sperm analysis parameters. To enrich the dataset heterogeneity and test the generalization ability of the approach, bovine sperm cells are analysed in different environmental conditions to replicate motility alterations that can occur in non-physiological states. In particular, we treated sperm cells in conditions of altered viscosity, hyper-osmolarity, absence of Ca^{2+} and in medium containing high caffeine concentration. The resulting dataset comprises a vast number of cells, and is then elaborated with Principal Component Analysis for dimensionality reduction and k-means for clustering. The result is a characterization of sperm populations that goes beyond the standard classifications adopted in literature, providing a parametrization based on parameters that link both sperm motility and morphology. The identified motility alterations among different conditions confirms the generalisation ability of the employed neural networks and underlies the diagnostic potential of the proposed approach. Alongside with the European regulations for the use of AI in healthcare -which emphasize the importance of transparency, safety, and accountability- the MATLAB code employed in this study is provided in Appendix A. This step aims to address concerns regarding the opacity of AI systems, particularly in diagnostic applications, where the interpretability of results remains a critical challenge.

As a future perspective, a graphical user interface (GUI) will be proposed to enable intuitive navigation through the data generated by the employed neural networks and the clustering analysis. This interface could be added to CASA software to achieve a more comprehensive characterization of sperm cells. Furthermore, such a tool would enhance the interpretability and accessibility of results for future users, building trust and effective integration in clinical workflows.

2. Materials and methods

2.1 Sample preparation and video acquisitions

Cryopreserved bovine sperm were purchased in frozen straws from ABC Love Genetix (CONSWORK srl, Lodi, Italy) and stored in liquid nitrogen until the experiment. The straws contained $250\ \mu l$ of volume and a concentration of $30\ x\ 10\text{^}6\ sperm/ml$. The procedures for sperm preparation were performed in sterile conditions under a biological hood. To start, the straws were retrieved from the liquid nitrogen container and left in the oven in a water bath for $1\ min$ to defreeze at $37°C$. Then, $100\ \mu l$ of pure sample were diluted in $500\ \mu l$ of Sydney IVF Gamete Buffer (Cook medical) to achieve a final concentration of $3 * 10\text{^}6\ cells/ml$.
As regards the acquisition of image sequences, the optical setup includes an inverted microscope (IX81, OLYMPUS) and a CMOS camera (ORCA flash 4.0, HAMAMATSU) with a 40X brightfield objective. This magnification was chosen because it provides a sufficiently large field of view to allow for the simultaneous analysis of an adequate number of cells, while also enabling precise morphological evaluations. Bovine sperm samples were loaded in a $0.05\ x\ 2\ mm$ glass capillary with rectangular section (CM Scientific) by means of capillarity, and the two extremities were sealed by means of a capillary wax (Vitrex Medical A/S, Denmark) to avoid drifting of the liquid. The videos, $3\ sec$ long, were acquired at each timestep at a frame rate of $50\ fps$, with an exposure time of $3\ msec$.

2.1.1 Cypermethrin drug treatment

Cryopreserved bovine sperm were obtained in frozen straws from ABC Love Genetix (CONSWORK srl, Lodi, Italy) and stored in liquid nitrogen until use. To minimize data biases caused by potential differences in individual animal characteristics and health status, all experiments utilized sperm from the same bull. A single experimental replicate was conducted for each test. Two straws were used, one for each condition. Each straw contained $250\ \mu l$ of sperm at a concentration of $30\ \times\ 10^6\ sperm/ml$. For each experiment, a straw was removed from the liquid nitrogen container and thawed in a water bath at $37°C$ for 1 minute. Following this, the sperm were washed three times with phosphate-buffered saline (PBS) and separated from seminal plasma via centrifugation (500g for 5 minutes at 21°C). After the supernatant was discarded following the final centrifugation, the pellet was resuspended in Sperm Preparation Medium (Origio, Copenhagen, Denmark) to a final volume of 100 µl at a concentration of $30\ \times\ 10^6\ sperm/ml$. Cypermethrin (CYP) (ML-PR100-0050) was sourced from Enzo Life Sciences Inc. (Farmingdale, New York, USA) and prepared according to the manufacturer's instructions by dissolving it in 100% ethanol. Two distinct samples were prepared for the experimental conditions: a control (CTRL), without any inhibitor, and a CYP-treated condition ($0.1\ nM$). Both samples were incubated at $37°C$ in a 5% CO_2 environment. For the CTRL condition, $10\ \mu l$ of the sample was diluted in $190\ \mu l$ of sperm preparation medium, while for the CYP condition, 10 µl of the sample was diluted in $190\ \mu l$ of the CYP solution. Sperm motility was evaluated at three time points: 10, 30, and 60 minutes.

2.1.2 Human sample preparation

As a proof of concept, the approach was applied on human sperm. Fresh samples were donated by the Azienda Ospedaliera Universitaria (AOU) “Federico II” (Naples). The samples were diluted into

PBS to a final concentration of $3 \cdot 10^6\ sperm/ml$. Sperm were loaded in a $0.05\ x\ 2\ mm$ glass capillary with rectangular section (CM Scientific) by means of capillarity, and the two extremities were sealed by means of a capillary wax (Vitrex Medical A/S, Denmark) to avoid liquid drift. The image sequence was acquired with an inverted microscope (IX81, OLYMPUS) and a CMOS camera (ORCA flash 4.0, HAMAMATSU) with a 40X brightfield objective (Figure 54A).

2.2 Tracking

The tracking algorithm was first optimised for tracking apparent sperm centroids acquired at $20x$ magnification [150]. The acquisitions were analogous to those seen for the neural network analysis, with the exception of the lower magnification, larger loading chamber (capillary with rectangular cross section of $0.2\ x\ 2\ mm$) and frame rate, which was set to $100\ fps$. To test the approach, sperm were treated with Cypermethrin, a pesticide known to induce motility alteration [151].
Single sperm heads were tracked using a custom MATLAB routine (version 2022b). Initially, a sequence of image filtering techniques, including Gaussian filtering, morphological opening, mean subtraction, and Wiener filtering, was applied to emphasize the sperm heads compared to their surroundings [56]. The MATLAB function 'imfindcircles' was used to identify the coordinates of the apparent head centroid (AC) — the brightest point in the sperm head region after filtering. The sensitivity parameter (ranging between 0 and 1) directly influenced the likelihood of false detections. To link centroids across frames, a Hungarian algorithm was implemented, which establishes connections between ACs found within a maximum distance, D, thereby minimizing the total distance among pairs of ACs across consecutive frames [152]. The algorithm also includes a mechanism to handle gaps in tracking when a centroid is not detected in the subsequent frame, which helps minimize breaks or incorrect connections. After the initial linking, a second iteration is performed for gap-closing, which checks for track interruptions by examining the endpoints of tracks. If a new track's starting point is located within a distance, $D_{gap\ closing}$, of an existing track endpoint, a link can be created over a maximum of G frames to restore continuity. This gap-closing process uses a nearest-neighbour approach that optimizes local distances by identifying the closest pairs of ACs between tracks, ensuring that the linking is not influenced by the order of the centroids. For tracking, the maximum distance allowed for linking was approximately $5\ \mu m$. To mitigate errors arising from the gap-closing process, a validation step was introduced. This involved calculating the mean distance between consecutive points along the track and checking whether any two points were separated by more than five times this mean value—selected through trial and error. In such cases, the trajectory was split into two separate segments for evaluating motility. A threshold was set on track length and VSL to filter out noise. VSL, VCL and LIN were computed at 0, 30 and 60 min. They were converted from pixels to μm using the known pixel size, which was determined based on the camera used during acquisition. All parameters could be adjusted to suit specific experimental requirements, such as noisy environments or variations in frame rate. The MATLAB code for tracking and CASA parameters computation is reported in Appendix A.
To translate the analysis at higher magnifications, parameters were adapted as can be seen in Table 1. In this case, the centerpoints of the bounding boxes detected by YOLOv4-tiny are tracked. All trajectories are cut at 75 points to be at the same length. Once the bounding boxes have been tracked, an 80x80 crop is performed, centred on each bounding box, to focus the image on the sperm head. The cropped image is then provided to the U-net for segmentation.

Table 4: Set of acquisition parameters used for the analysis.

ALGORITHM PARAMETER SETTINGS		
	20x	**40x**
Frame Rate	100 fps	50 fps
Pixel Size	$\frac{6.5}{20}$ $\mu m/pixel$	$\frac{6.5}{40}$ $\mu m/pixel$
Centroid location function sensitivity S	0.6	-
Gap Closing frame interval G	5 $frames$	3 $frames$
Gap Closing Distance $D_{gap\ closing}$	8.5 μm	10.7 μm
Max Linking Distance D	4.2 μm	7.4 μm

2.3 Sperm detection

2.3.1 Dataset

As mentioned before, sperm exhibit a complex 3D motion. Bovine sperm have a paddle-like shaped head, so during their motion two features emerge, which we termed the 'Face-on' and 'Rim-on' (Figure 10A). Face-on refers to exposure of the bigger portion of the head, ideally parallel to the plane of acquisition, while Rim-on refers to the opposite case. These two configurations were used as classes for the detection network training. Images were manually labelled using MATLAB Labeler. The training dataset comprised 1317 images, of which 1137 randomly extracted from 25 different videos. To balance Face and Rim class numerosity, 180 mosaic images containing only Rim configurations were created and labelled, as can be seen in Figure 10B [121]. 20% of the training data was set aside for validation and 200 more images were labelled for testing. The rest was augmented through several augmentation techniques reported in Figure 10C [83].

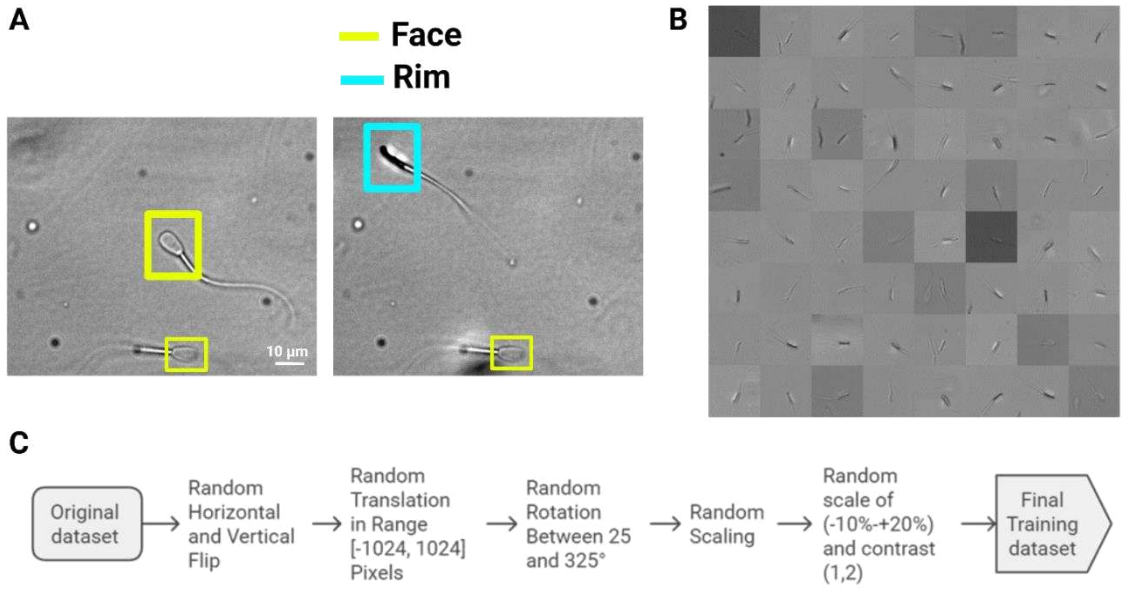

Figure 10: Training set and Image augmentation pipeline. A) The detection network was trained on two classes, namely Face and Rim. B) Mosaic images created only with Rim configurations to implement the class numerosity. C) Image augmentation techniques employed on the training set.

2.3.2 Yolov4 Network Architecture

For sperm head detection YOLOv4-tiny was chosen for its balance between accuracy and efficiency, making it suitable for low-cost embedded systems requiring real-time processing capabilities [153]. The architecture is shown in Figure 11.

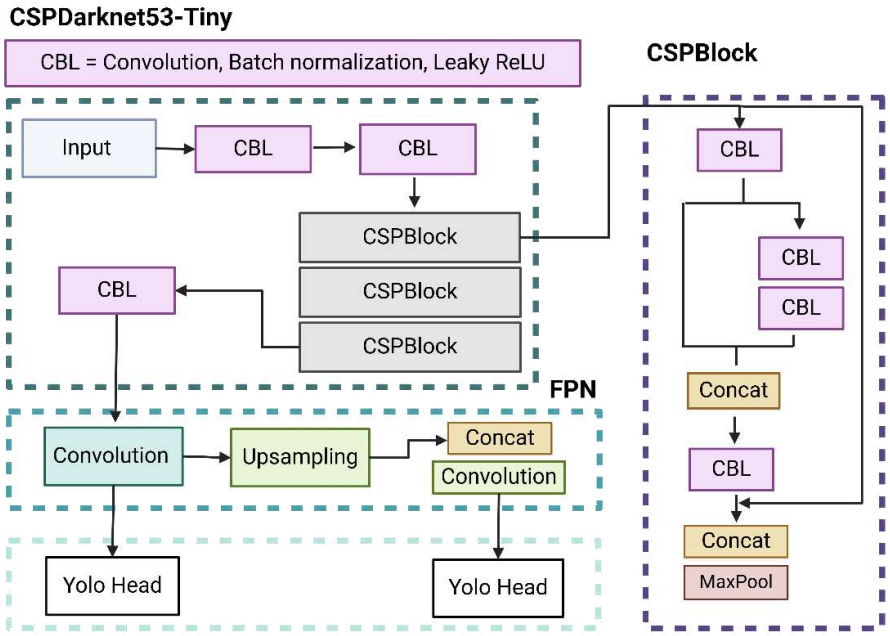

Figure 11: Yolov4-tiny architecture. The model processes an input image of size [640x640x3] through several convolutional layers, each followed by Batch Normalization and Leaky ReLU activation. The backbone consists of multiple CSPBlocks that enhance feature extraction efficiency while reducing computation. The Feature Pyramid Network (FPN) is used to combine multi-scale features, followed by two YOLO heads for object detection. The CSPBlock structure is shown on the right, including its base layer, concatenation, and max-pooling operations, highlighting the use of cross-stage partial connections to improve gradient flow and model performance.

It comprises three primary modules: CSPDarknet53-Tiny, Feature Pyramid Network (FPN), and YOLO Head. CSPDarknet53-Tiny serves as the backbone of the network, responsible for initial feature extraction from the input image. CSPDarknet53-Tiny is a simplified version of CSPDarknet53, adapted to reduce computational load. The CSP (Cross Stage Partial) block divides the feature maps into two parts, with one passing through several convolutional layers while the other remains as a residual connection. This structure helps enhancing gradient flow and reducing redundancy, leading to more efficient feature extraction. Convolutional layers include batch normalization and leaky ReLU activation functions. Batch normalization is used to standardize inputs, thereby improving stability and convergence, while leaky ReLU helps prevent the vanishing gradient problem by allowing a small gradient when the input is negative. The FPN is employed to combine features from different network layers, effectively merging high-level information from deeper layers with the spatial details found in shallower layers. This combination allows the network to detect objects of different sizes with greater accuracy. The FPN structure helps preserve critical geometric information while enhancing the overall robustness of the feature maps, making it suitable for detecting objects even when they appear at different scales or positions within the video. The YOLO Head is the final component of the YOLOv4-tiny network, responsible for generating the output predictions. It performs dense predictions, producing bounding boxes, class labels, and confidence scores for each detected object. Specifically, it outputs a vector containing the coordinates of the predicted bounding box (center coordinates, height, width), the object class, and a confidence score. The bounding box prediction includes key parameters such as the center location (x, y) and the dimensions (height and width). The confidence score indicates the network's confidence in whether the detected box contains the object of interest. In YOLOv4-tiny, the number of YOLO heads are reduced to two (compared to three in YOLOv4), which helps maintain accuracy while improving processing speed and reducing model size. The total number of layers of YOLOv4-tiny is 37.

2.3.3 Training

The network was trained on MATLAB using the '`yolov4ObjectDetector`' function from the Computer Vision Toolbox. Adam was chosen as the optimization approach for the gradient of the loss function [117]. As mentioned in paragraph 1.4.1, Adam keeps an exponential moving average of both the gradients and the squared gradients, called respectively the first moment m_t and the second moment v_t. The formulations are reported below.

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \tag{5}$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2)\, {g_t}^2 \tag{6}$$

g_t is the gradient of the loss function at time *t*, β_1 is the gradient decay factor, usually set to 0.9 and β_2 is the squared gradient decay factor, usually set to 0.99 [76]. The network parameters are updated according to the following update rule:

$$\theta_{t+1} = \theta_t - \left(\frac{\alpha}{\sqrt{\widehat{v_t}} + \epsilon}\right) * \widehat{m_t} \tag{7}$$

where θ_t is the network parameter at time t, α is the learning rate, $\widehat{m_t}$ and $\widehat{v_t}$ are bias-corrected first and second moment estimates. While the first moment m_t helps accelerate convergence in relevant directions while dampening oscillation, the second moment v_t helps to adapt the learning rate for each parameter individually.

Two values for batch size and input size were chosen for training in neural network sperm recognition, namely 16 and 32 for batch size and 512x512 and 640x640 pixels for input size. The learning rate was scaled after a certain number of epochs, where an epoch is the number of iterations required for the network to train on the whole training set in one cycle. After 20 epochs the learning rate was halved for 5 additional epochs, then scaled by 5 for 5 more epochs. This strategy was adopted to ensure the reach of the optimal minimum for gradient loss function. The networks performances on the validation set were assessed every 1200 iterations. The best validation loss is used to select the final model for evaluation, as it represents the state where the model performed the best on data it wasn't trained on. To avoid overfitting, the inputs are shuffled every epoch. The GPU employed for the training was a NVIDIA RTX A5000. The training time was ~9 hours for batch size 32 and ~7 hours for batch size 64. Table 5 reports the fixed training hyperparameters for YOLO training. The code for the network training can be found in Appendix A.

Table 5: Fixed training options.

HYPERPARAMETER	VALUE
Learning rate	0.001 (first 20 epochs) 0.0005 (5 epochs) 0.0001 (5 epochs)
Number of anchors	18
Gradient decay factor	0.9
Squared gradient decay factor	0.99
Validation frequency	1200 iterations
Shuffle	Every epoch
Output	Best validation loss

2.3.4 Testing

The different configurations of the detection network were tested on the same dataset, comprising 200 manually labelled images randomly extracted from 10 videos. The metrics for the performance evaluation were chosen accordingly to the task, and are reported in Table 6. The IoU threshold for assigning a detection to a ground truth box was set as 0.5 as commonly done [154]. Moreover, the AP^{50} was used as it is widely used in literature for object detection evaluation [155]. It is computed as the mean average precision for an IoU of 0.5. The confidence score threshold was fixed to 0.5 for testing. The code for the network testing can be found in Appendix A.

Table 6: Performance evaluation metrics.

METRICS	COMPUTATION	DEFINITION [84]
Precision	$Pr = \frac{TP}{TP + FP}$	Proportion of true positive predictions among all positive predictions. It is useful when the cost of false positives is high.
Recall	$Rec = \frac{TP}{TP + FN}$	Measures the proportion of actual positives that were correctly identified. Crucial in scenarios where missing positive cases is more costly than false positives.
F1 score	$F1 = \frac{2 * Precision * Recall}{Precision + Recall}$	Harmonic mean of Precision and Recall, balancing the two metrics. It is particularly useful when there is an uneven class distribution.

2.4 Sperm segmentation

2.4.1 Dataset

To build the training dataset for the segmentation task bovine sperm were stained with Hoechst, a DNA-staining compound. For instance, starting from a stock concentration at $1{:}\,100\ v/v$, $10\ \mu l$ Hoechst were put in $1\ ml$ of diluted sample, which was then left at $37°C$ for $30\ min$. After staining, $100\ \mu l$ of samples were put in $1000\ \mu l$ of Paraformaldehyde 4%, and left at room temperature for $15\ min$. Then, after a centrifuge at $1100\ rpm$ for $5min$, the supernatant was removed and the pellet was re-suspended with $500\ \mu l$ of PBS. The acquisition was done with the same settings and loading capillary used for the live sperm experiments. Each field of observation was acquired in brightfield and fluorescence. A post-processing step was performed to remove noise, binarize the fluorescent images and crop each one to have a single sperm head centred in every image, with a final dimension of 80x80 pixel (Figure 12). The choice to crop the images to this size was based on the fact that it allows the entire sperm head to be captured at 40x magnification within the cropping. Larger sizes would have been unnecessarily large, while smaller sizes risked excluding parts of the cell's head. Two classes were defined, "Head" and "Background", encoded respectively with 1 and 0 in the binary masks. Since when cells sediment they tend to lay on the flatter side that is the Face configuration, snapshots of Rim and Intermediate configuration were cropped and labelled manually from videos of live sperm cells (Figure 12). This approach allowed to train the network on all the possible configurations that a sperm can acquire during its rolling motion. The final dataset before data augmentation comprised 2540 images. The training set was split in 75% training – 25% validation, and the training fold was augmented through the following transformations: random horizontal and vertical flip, random translation in a range of [-20, 20] pixels, random scaling, random shift and scale of brightness (-10%-+20%) and contrast (1,2), addition of random Gaussian blur and random rotation in an interval between 25° and 335°.

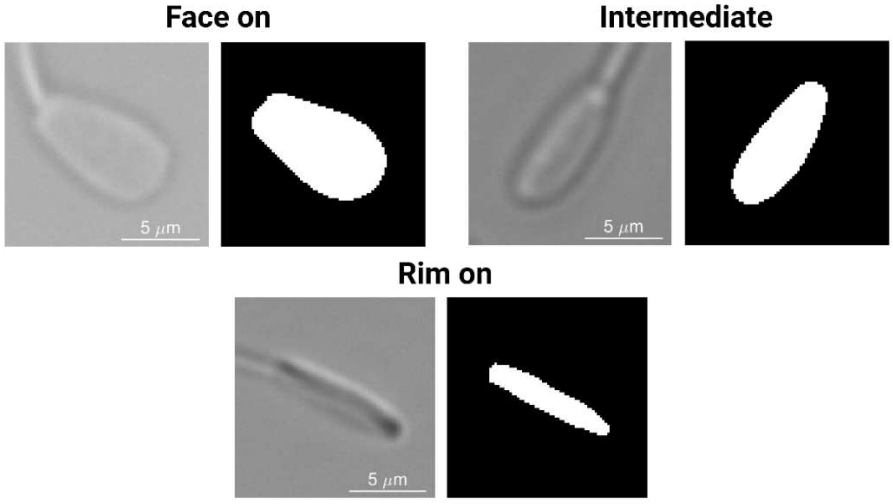

Figure 12: Training set for the U-net segmentation, comprised of Face on, Rim on and Intermediate. Three head configurations were considered, namely Face on, Rim on, and a middle configuration between the two which we identified as intermediate. The latter is of interest as our method aims at segmenting moving sperm. For all three cases, the Background pixels was labelled as '0' while Head pixels with '1'.

2.4.2 U-net-tiny architecture

To identify the pixels composing the sperm head, a modified version of the U-net was adopted. As a matter of fact, this network is widely used for evaluation of medical images, including sperm cells

[91], [101], [124], [125]. A light-weight version of U-net called U-net-tiny was used for computational speed (Figure 13) [128].

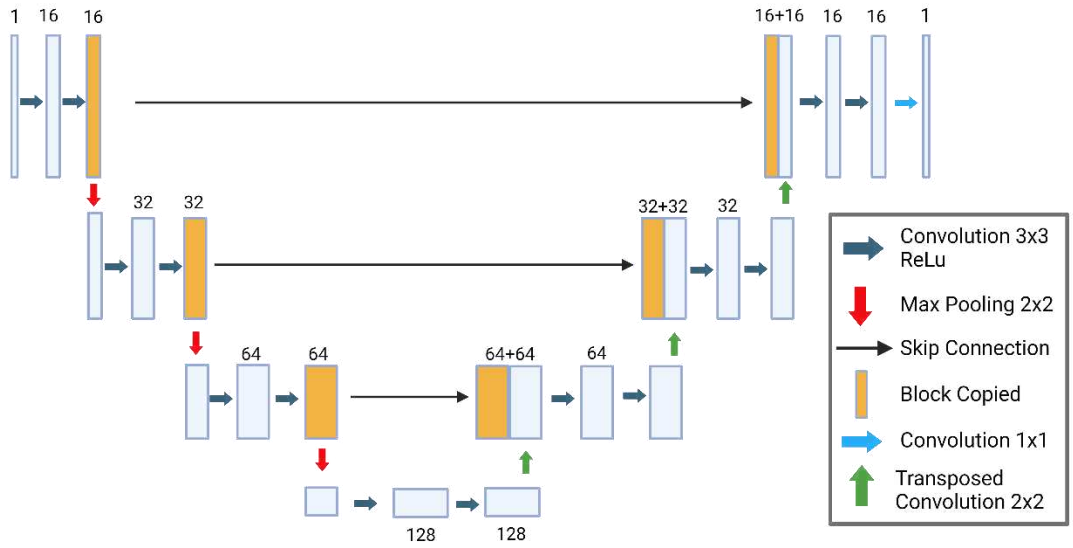

Figure 13: U-net-tiny architecture. The encoder part reduces input size through convolution, batch normalization and ReLU and max pooling blocks. The lowest point of the architecture is the backbone, were finest details are extracted from the feature maps. Then, the decoder part up-samples the input, correlating the information coming from the decoder part through skip connections to recover the spatial context lost during down-sampling. The sketch was created with Biorender©.

Specifically, from the original U-Net, the order of down-sampling is reduced from 4 to 3, which reduce its number of channels to a quarter. The number of kernels of the first layer is set to 16. The model size of U-Net-tiny is only 1.8 MB, approximately 1.6% of the model size of the U-Net. Each level of the encoder consists of two 3x3 convolutions, followed by a ReLU activation (blue arrows in Figure 13), and a 2x2 max pooling operation (Red arrows). The max-pooling operation progressively reduces the spatial dimensions of the feature maps, compressing the information into higher-level features. Between the encoder and decoder there is the bottleneck that is the lowest point of the architecture. In the bottleneck features are highly compressed, capturing critical information about the input image. In the decoder, the feature maps are up-sampled using transposed convolutions, and they are combined with the corresponding layers from the encoder through skip connections (black arrows in Figure 13). This helps recover the spatial context that was lost during down-sampling.

2.4.3 Training

To classify sperm head with respect to the background, two classes were defined, ‘Head’ and ‘Background’, encoded by 1 and 0 respectively. The Generalized Dice Loss (GDL) function was chosen as it is recommended for unbalanced class segmentations, particularly common in medical imaging where the region of interest is often very small. The GDL mitigates the bias towards larger classes by weighting each class differently based on its size, ensuring that small classes (often underrepresented) are properly considered during training [156]. In GDL, each class is assigned a weight that is inversely proportional to its size. This ensures that smaller classes contribute more significantly to the loss, thereby counteracting the dominance of larger classes that can lead to poor segmentation performance for underrepresented regions. The weight for each class is calculated as:

$$w_l = 1/(\sum_n r_{ln})^2 \tag{8}$$

r_{ln} represents the reference values for a specific class l, while n refers to the element in the image and goes from 1 to N. In this case, there are 2 classes, namely sperm head and background. The function loss becomes:

$$GDL = 1 - (2\sum_{l=1}^{2} w_l \sum_{n}^{N} r_{ln}p_{ln}) / (\sum_{l=1}^{2} w_l \sum_{n}^{N} (r_{ln+}p_{ln})) \tag{9}$$

where r_{ln} are the reference labels and p_{ln} the predicted labels for class l. This formula emphasizes contributions from underrepresented classes by using class weights w_l.
RMSprop was chosen as the optimiser for the loss function. The rule for the parameter update is the following:

$$\theta_{t+1} = \theta_t - \left(\frac{\alpha}{\sqrt{E[g^2]_t + \epsilon}}\right) * g_t \tag{10}$$

Where:

- α is the learning rate;
- g_t is the gradient of the loss function at time *t;*
- $E[g^2]_t$ is the moving average of squared gradient, whose decay is regulated by the squared gradient decay factor;
- ϵ is a smoothing factor to prevent division by zero.

The learning rate was fixed to 0.0001, and the network was trained for 50 epochs. The performance of the network was validated every 50 iterations, and every epoch the mini-batch of images was shuffle to improve the learning process. To avoid overfitting, early stopping was introduced if the validation loss does not improve for a specified number of validation steps, determined by the validation patience. The best validation loss is used to select the final model for evaluation. There settings are summarised in Table 7. The MATLAB code is reported in Appendix A.

Table 7: Settings for U-net training.

HYPERPARAMETER	VALUE
Learning rate	0.0001
Max Epochs	50
Gradient decay factor	0.9
Squared gradient decay factor	0.99
Validation frequency	50 iterations
Validation patience	40 iterations
Shuffle	Every epoch
Output	Best validation loss
Input size	80x80

Three combinations of datasets were studied and compared. Since the sperm head acquires different intensity and contrast passing from Rim to Face, one dataset was made up of all the labelled images

including Face, Rim and Intermediate configurations (Net FR). Another one comprised only Face and Intermediate and included 1482 images (Net F). The third comprised 1058 images of only Rim configurations (Net R). All the training sets were augmented before training with random scaling, translation, intensity shift, flipping, blurring and rotation transformations. The learning rate was fixed and equal to 0.0001, while three batch size were tested, namely 16, 32 and 64. The GPU employed for the training was a NVIDIA RTX A5000.

2.4.4 Testing

The resulting networks were tested on 100 Face, 100 Intermediate and 100 Rim images. The chosen metrics for performance evaluation were Dice score and IoU, whose computation and definition are reported in the table below.

Table 8: Segmentation performance evaluation metrics.

METRICS	COMPUTATION	DEFINITION [84]
Dice score (Sørensen–Dice)	$D = \frac{2 * TP}{2 * TP + FP + FN}$	Measure of overlap between two samples
Intersection over union (Jaccard)	$IoU = \frac{TP}{TP + FP + FN}$	Measures the similarity between the predicted and actual segmentation.

2.5 Sperm treatments

To induce motility alterations, sperm where treated with different drugs. All the experiments were performed using frozen bovine semen from the same bull and were replicated at least two times. The thawing procedure for sperm preparation is reported in 2.1. Multiple fields of acquisition were acquired for each condition.

2.5.1 Osmolarity

The conditions of altered osmolarity were chosen to mimic the bovine oviduct osmolarity and an hyperosmotic condition, namely $355\ mOsm/Kg$ (HYP1) and $420\ mOsm/Kg$ (HYP2) [27]. These values ensure hyperosmolar conditions with respect to the Gamete buffer, whose osmolarity is $280\ mOsm/Kg$ (ISO), comparable to bovine seminal plasma [27]. To replicate these conditions, both ionic and organic compounds were used. For ease of dilution, a stock was made at $500\ mOsm/Kg$ and, adding respectively $730\ mg$ of NaCl (MW: $58.4\ g/mol$) to $50\ ml$ of Gamete Buffer. From such stock, $0.8\ ml$ were withdrawn and put in $1.7\ ml$ of Gamete Buffer to achieve a final volume of $2.5\ ml$ for HYP1. For HYP2, $1.55\ ml$ were put in $0.95\ ml$ of Gamete buffer.
The second hyperosmolar condition was reproduced using pure powdered Fructose (purchased from Zen store, MW:$180.16\ g/mol$) instead of NaCl. Similarly to the previous case, a stock was made adding $9\ g$ of fructose to $50\ ml$ of Gamete to reach an osmolarity of $1000\ mOsm/Kg$. Then, $0.46\ ml$ from this stock were added to $2.04\ ml$ of Gamete buffer to reach a final osmolarity equal to $420\ mOsm/Kg$ (FR). Two semen straws were thawed for each osmolarity experiment, and each experiment was replicated three times. Thawing and dilution were performed as reported in 2.1. $100\ \mu l$ of sample were diluted in an Eppendorf containing $500\ \mu l$ of pure Gamete buffer for iso-osmotic control (ISO). The same amount was diluted into three different Eppendorf containing the hyperosmolar media. The samples were analysed at 5, 30 and 60 min upon contact with the media, and was kept at $37°C$ in between the time-points. The experiment was replicated three times.

2.5.2 Absence of Ca^{2+}

To create this condition, $100\ \mu l$ of pure sample was diluted in $500\ \mu l$ of PBS, which does not contain calcium. Similarly, to the osmolarity experiment, sperm were analysed at 5, 30 and 60 min upon contact with the media and kept at $37°C$ in the meantime. The experiment was replicated three times.

2.5.3 Viscosity

To create a condition of altered viscosity a highly biocompatible polyethylene-oxide (4MDa – Merck KGaA, Darmstadt, Germany) solution was used. PEO was prepared at two concentrations, 0.5%wt (PEO0.5) and 0.9%wt (PEO0.9) in PBS, having a viscosity of $0.0295\ Pa * s$ and $0.0798\ Pa * s$. Before use, both PEO preparations were filtered with a disposable filter having pores of $0.45\ \mu m$ for debris removal. The sperm from one straw were thawed and suspended in $1.25\ ml$ of Gamete buffer and centrifuged at $200\ rpm$ for $5\ min$. The suspending fluid was removed and the procedure was repeated again. The pellet resulting from the double centrifuge was suspended with an amount of Gamete buffer to bring the sample back to the initial straw volume, $250\ \mu l$. Then $50\ \mu l$ of sample were put in $250\ \mu l$ of PEO0.5 and PEO0.9 solution as well as in a solution of pure Gamete buffer for control (ISO). The experiment was replicated twice.

2.5.4 Caffeine

For caffeine experiments, sperm were exposed to a concentration of $10\ mM$ of pure powdered caffeine (purchased from Zen store) as this value is known from literature to induce motility alteration [61]. To achieve this condition (CAFF), a stock was made adding $2.5\ g$ of pure caffeine powder (MW: $194.19\ g/mol$) to $50\ ml$ of distilled water, reaching a concentration of $260\ mmol/l$. Then, $0.46\ ml$ of such stock were put in $500\ ml$ of sperm sample diluted in Gamete buffer at a concentration of $3 * 10^6 cells/ml$. Sperm were analysed at 5, 30 and 60 min upon contact with the media and kept at $37°C$ in the meantime. The experiment was replicated twice.

2.5.5 Confocal acquisitions

For osmolarity experiments, at each time step $50\ \mu l$ of both ISO and treatment were withdrawn from the samples and fixed in Paraformaldehyde 4% (PAF, Sigma-Aldrich). Following a $1:10$ proportion, $500\ \mu l$ of PAF were used to fix each sample. After $10\ min$ at room temperature, the samples were centrifuged at $1100\ rpm$ for $5\ min$. After supernatant removal, the pellet was resuspended in $500\ \mu l$ of PBS. $200\ \mu l$ were put in an Eppendorf and coloured with Wheat Germ Agglutinin (WGA) with a proportion of $1:400$. After $30\ min$ of contact at room temperature, samples were centrifuged at $1100\ rpm$ for $5\ min$ and resuspended in PBS to adjust the volume back to $200\ \mu l$.

For confocal analysis, $150\ \mu l$ of stained and fixed samples were loaded in rectangular wells on a chambered coverslip (Ibidi GmbH, $\mu -$ Slide with dimensions $9.4\ x\ 10.7\ x\ 6.8\ mm^3$). Prior to sample loading, the wells were treated with polylysine $1:10$ to reduce sperm movement during confocal acquisitions. Specifically, $200\ \mu l$ of polylysine were put in each well and left at room temperature for $15\ min$. Then, washing with PBS was performed to remove residues. To investigate whether the hyperosmolar conditions had effects on sperm volume, ISO, HYP1, HYP2 and FR were investigated using a confocal microscope (LSM 71, Zeiss) equipped with a $63x$ oil-immersion objective. To reconstruct sperm height, z-stacks were acquired using a distance of $0.32\ \mu m$ between slice. An extract is reported in Figure 14A. Images were elaborated with ImageJ. Sperm head contour

was acquired for at least 20 cells per condition, computing Area, Major and Minor axes with ImageJ "Measure" option (Figure 14B). The height was evaluated using the "Z project" option, as it allows to visualize the fluorescent signal profile inside the selected area along the z-stack (Figure 14C). The height was estimated as the width of the resulting profile of the fluorescent signal.

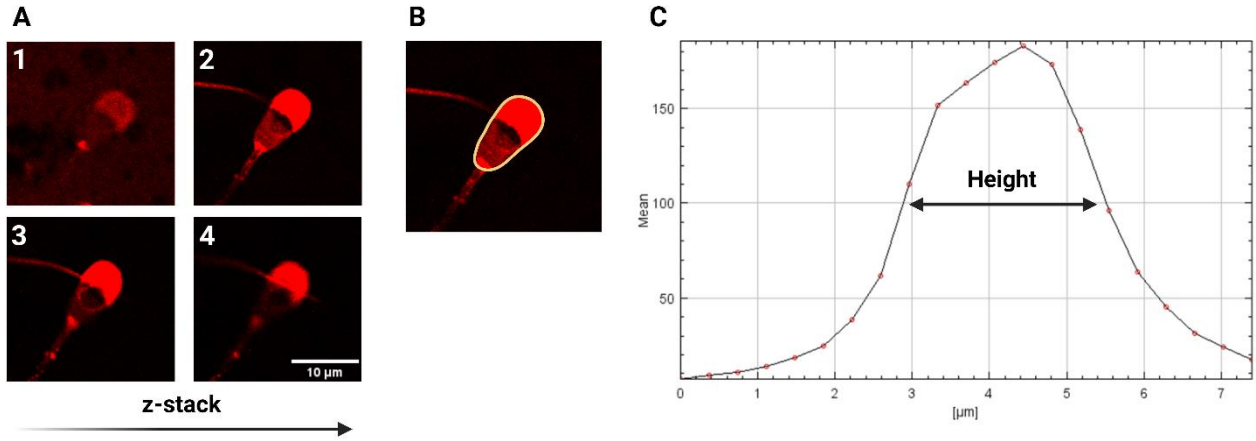

Figure 14: Scheme of the sperm height reconstruction with z-stack acquisition. A) Extract of 4 slices from a z-stack acquisition, acquired with a distance of 0.32 μm between each slice. Each sperm is captured varying the z dimension in a way that captures the whole volume. B) The middle slice is taken as a reference to draw the head contour. C) Z-projection of the mean intensity of the defined head contour. The amplitude of the resulting curve represents the sperm head height.

2.6 Motility and morphometric analysis

Once the sperm head contour is identified along the whole trajectory, morphometrical parameters such as Area can be computed with the MATLAB function 'regionprops'. Furthermore, knowing the precise head centroid, kinematic parameters can be computed as well with high precision on the reconstructed trajectory. Table 9 reports the computed kinematic parameters and their definitions [50], whose representation can be seen in Figure 15A and B. The motility classes categorization of bovine sperm according to CASA is reported in Table 10.

Table 9: Mathematical definition and meaning of the implemented kinematic parameters. Pi is the i-th point of a trajectory of length N; d(P,Q) is the Euclidean distance between point P and Q; FR is the frame rate; γ is the conversion factor from pixel to microns; n is the number of track intervals (N-1); d is the planar extent of the curve, evaluated as the maximum distance between the starting point and any point of the track; L is the curvilinear path length; $n1$ is a positive integer lower than N; $\breve{v}_t$ is the angle of the vector connecting point P_t and P_{t+n1}.

PARAMETER	DEFINITION	COMPUTATION
Straight-Line Velocity (VSL)	Straight-line distance between first and last points of the trajectory and correcting for time.	$d(P_1, P_N) * \frac{FR}{N-1} * \gamma$
Curvilinear Velocity (VCL)	Distance travelled by the sperm along its curvilinear path, corrected for time.	$\left[\sum_{i=1}^{N-1} d(P_i, P_{i+1})\right] * \frac{FR}{N-1} * \gamma$
Linearity (LIN)	Comparison of the straight-line and curvilinear path. It expresses the relationship between the 3D path and the net space gain of the cell.	$\frac{VSL}{VCL} * 100$
Fractal Dimension (FD)	The fractal dimension reflects the extent to which a line occupies a plane. A curve with a low fractal dimension would be smooth and predictable, while a curve with a high fractal dimension would exhibit irregular changes in direction, seemingly at random.	$\frac{\log(n)}{\log(n) + \log\left(\frac{d}{L}\right)}$
Mean Angular displacement (MAD)	MAD is a measure of the trajectory curvature, defined as 'the time average of absolute values of the instantaneous turning angle of the head along its curvilinear	$\frac{\sum_{t=1}^{N-n1} \breve{v}_t}{N-n1}$

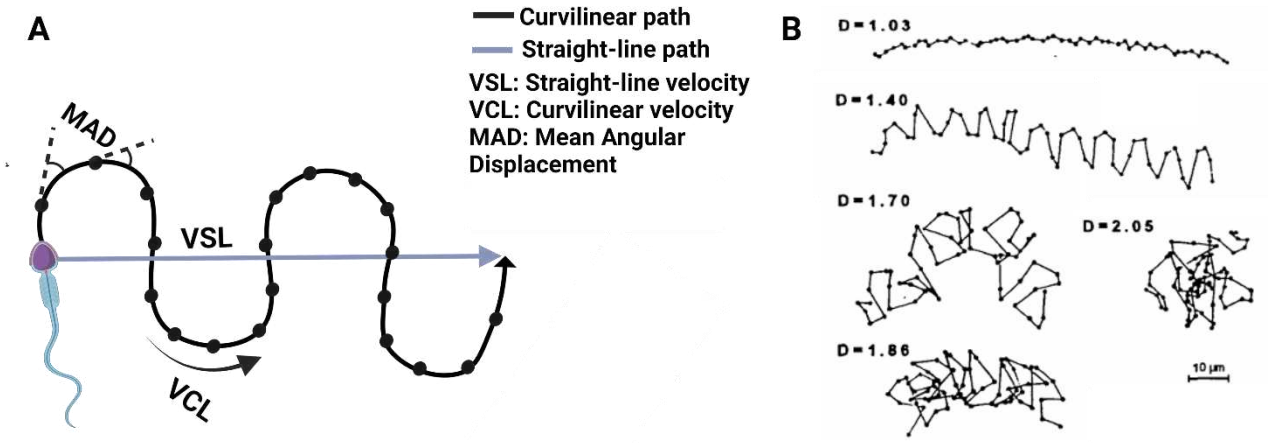

Figure 15: Sketch of CASA motility parameters and fractal dimension. A) Following the sperm head centroid, the trajectory is reconstructed and several kinematic parameters can be computed. B) Representation of sperm trajectories with different fractal dimensions (D), captured at 60Hz. The more irregular the path is, the higher the D. Reproduced from [55].

Table 10: Bovine sperm motility class definitions according to literature [42].

CLASS	VELOCITY RANGES
Slow progressive	$25\ \mu m/sec < VCL < 80\ \mu m/sec$
Medium progressive	$80\ \mu m/sec < VCL < 150\ \mu m/sec$
Rapid progressive	$VCL > 150\ \mu m/sec$

Head Area, VSL and VCL were validated by manually tracking 24 cells, both with normal and abnormal motility, for at least 1 sec of acquisition. The percentage error was evaluated as:

$$Err_{\%} = \left| \frac{X_{manual} - X_{Network}}{X_{manual}} \right| \tag{11}$$

Since the area fluctuates due to sperm motion, the mean area was evaluated considering only Face on areas with values bigger than $30\mu m^2$. The obtained mean value was further compared with literature findings of bovine and bull sperm head.
The MATLAB code for the analysis is reported in Appendix A.

2.7 Statistical analysis

The computed parameters are reported as mean value with standard deviation. Due to non-normality of data distribution the statistical analysis was performed with a pairwise Kruskal-wallis analysis. The MATLAB function 'kruskalwallis' was employed. The comparison was done for each parameter at the same time-step between experimental conditions.

2.8 Clustering analysis

Clustering procedures were performed to identify sperm subpopulations from the complete set of kinematic and morphometric parameters. The data from all treatments were joined in one dataset to account, reaching a final number of cells equal to 3902. The procedure was performed on OriginLab 2024 (Academic version). The first step was to perform a principal component analysis (PCA). Before applying the PCA, the Spearman correlation coefficient was evaluated to remove uncorrelated variables ($\rho < 0.3$). The fractal dimension FD was chosen as a pre-screening parameter, dividing the dataset in two groups based on a threshold of FD equal to 1.5. The PCA and k-means procedures were performed on the two datasets separately. The number of principal components was chosen in order to capture at least 80% of the dataset variance. The second step was to perform a clustering procedure with the sperm-derived indices obtained after the PCA. Measurements were clustered using a non-hierarchical clustering procedure, namely k-means algorithm and Euclidean distance. This analysis allowed the identification of sperm subpopulations. The results are presented as mean values with standard deviation. Statistical significance among clusters' parameters was considered as $P < 0.05$ according to the Kruskal-Wallis test, performed using the MATLAB "kruskalwallis" function.

3. Results and discussion

The proposed approach aims at developing a workflow based on NNs for label-free tracking and analysis of sperm cells. Bovine sperm cells were acquired in brightfield, at 40x magnification, in a glass capillary (Figure 16, Sample preparation). Image sequences were acquired at 50 fps for 3 sec. Such acquisitions were used to train two neural networks, namely YOLOv4-tiny for sperm head detection and U-net-tiny for sperm head segmentation. The workflow was developed using MATLAB2023b. YOLOv4-tiny is a lightweight version of YOLOv4, and was chosen for the speed and reduced computational load [153]. As mentioned before, sperm exhibit a complex 3D motion. Due to bovine sperm shape, during their 3D motion they either show the bigger portion of the head, which we termed 'Face on' or the opposite, which we called 'Rim on'. The intensity and contrast of the sperm head is very different among these two configurations, so the network was trained to detect and classify sperm head as 'Face' or 'Rim' so to gather an optimal detection (Figure 16, 'Neural network training'). Intermediate configurations were labelled as Face as well. A lightweight version of U-net called U-net-tiny was employed to identify the head region [128]. For the training dataset, bovine sperm were fixed and strain with WGA to identify the head region. The images were binarized to obtain segmentation masks and the value '1' was associated to the "Head" label, while '0' to "Background" (Figure 16, 'Neural network training'). While CASA analyses usually perform sperm head segmentation on fixed cells [50, 123-126], our approach proposed a live cell morphometric measurement to correlate morphology features to motion. For this reason, the training dataset was enriched with manually labelled masks of sperm head acquired in Intermediate and Rim configurations. To prove the method, sperm were treated with different conditions to induce motility alterations, namely caffeine, hyperosmolarity, increased viscosity and absence of Ca^{2+} [27, 157, 158] (Figure 16, 'Sample treatment). The analysis workflow comprises using a cascade of the two networks. The first detects the object in the field of observation and classifies the head configuration. Then the bounding boxes detected by the first network are tracked with an optimized tracking algorithm that identifies the single trajectories [150]. Once the single cell tracks are known, each cell head is cropped using the center points of the bounding boxes detected before and is segmented by the second neural network. The segmentation step allows to compute the sperm head contour along the whole trajectory, allowing the precise computation of both kinematic and morphometric parameters. Furthermore, three novel parameters have been proposed to better characterize sperm motion and enrich the pool of existing sperm parameters (Figure 16, "Analysis"). To gather more insight into the obtained dataset, PCA and k-means were adopted to identify clusters present in the dataset. This allowed to account for the motion variability of the samples, identifying motility sub-groups that would have been otherwise overlooked using standard classification criteria (Figure 16, "Clustering").

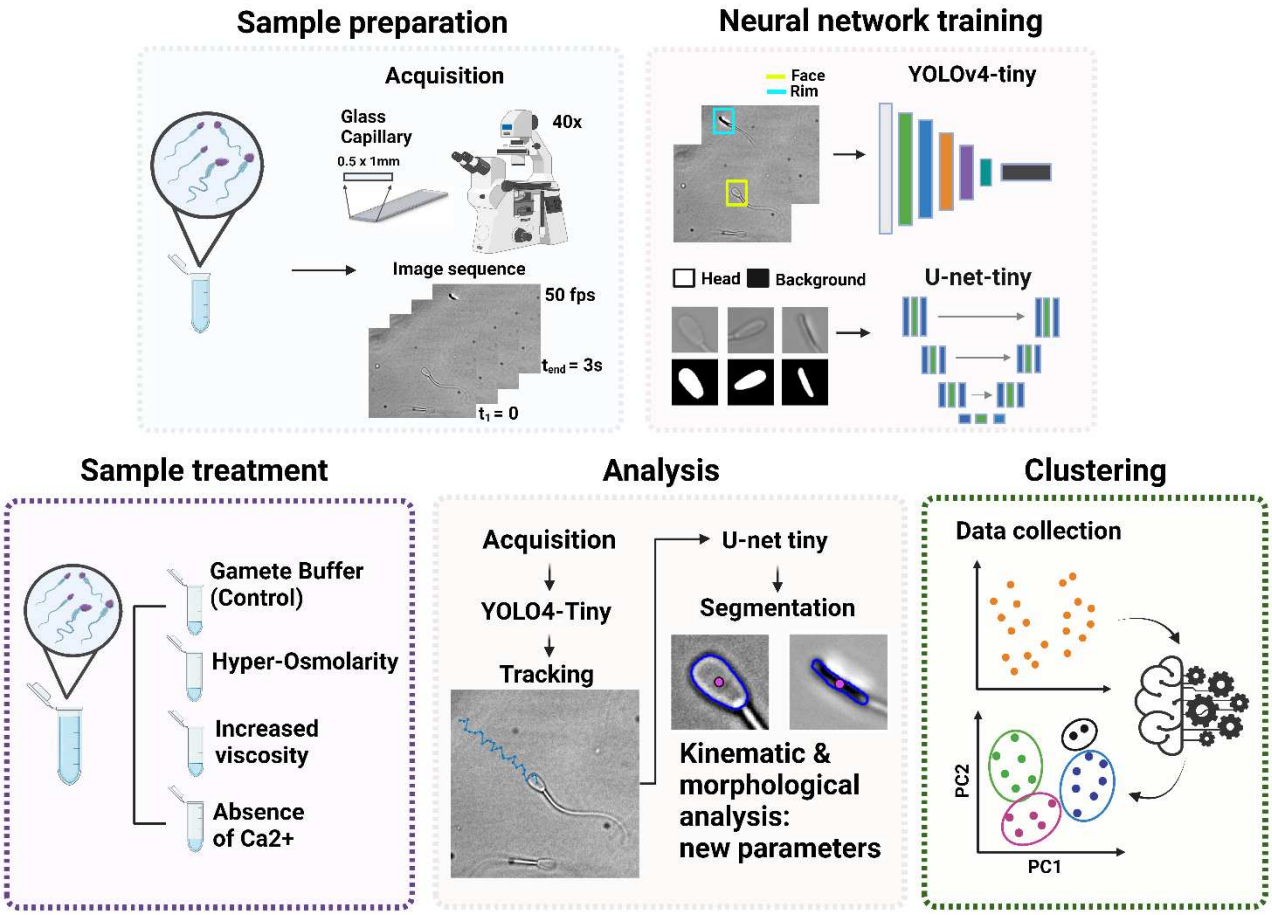

Figure 16: Scheme of the proposed approach, created with Biorender. Sperm were diluted in Gamete buffer and were acquired in brightfield with an inverted microscope, at 40x magnification. The images were manually labelled for a double task: sperm head detection and segmentation, performed with Yolov4-tiny and U-net-tiny respectively. Sperm were treated to induce alteration of motion. The proposed analysis comprises a cascade of neural networks to detect sperm heads, track them and segment them. The final goal is the retrieval of known CASA parameters and the introduction of new descriptors to characterize sperm motion. The obtained dataset was investigated with PCA and k-means algorithm to retrieve motility groups.

3.1 Tracking algorithm

A self-written MATLAB routine was optimised to investigate the behaviour of bovine sperm under unconfined conditions, analysing sperm motility and trajectories over extended periods. An example of output is reported in Figure 17. The routine was adapted to bridge the gap between missing detections, avoiding the interruption of the trajectory. An example is reported in Figure 17-4, where a missed detection on the Rim detection is bridged restoring the track. All parameters reported in Table 4 can be adapted to different acquisitions settings as frame rate and magnification.

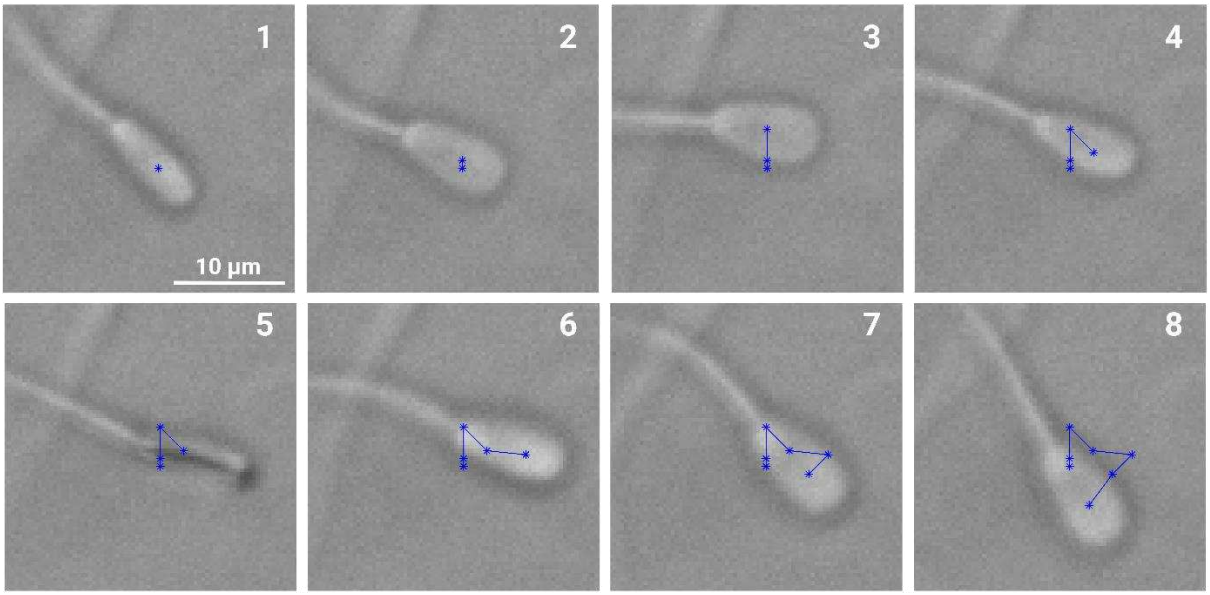

Figure 17: Algorithm optimization for sperm tracking at 20x magnification. The algorithm was optimised to deal with missing links (image 4), in order to restore the track.

Such self-written tracking routine was used to investigate the effect of a pesticide called cypermethrin (CYP) on sperm motion [150]. CYP is known to behave as an inhibitor of protein phosphatase type 2B (PP2B), which has a role in the sperm maturation process [159]. As a matter of fact, while sperm undergo physiological maturation process in the epididymis until becoming fully motile, PP2B activity is progressively being inhibited. The chosen concentration of CYP, close to the inhibitory concentration of PP2B ($0.1\,nM$), was investigated as a potentially inductor of increased sperm motility and progression. This was evaluated through the computation of VSL, VCL and LIN (Figure 18). Sperm were classified in motility classes according to CASA definitions reported in Table 9.

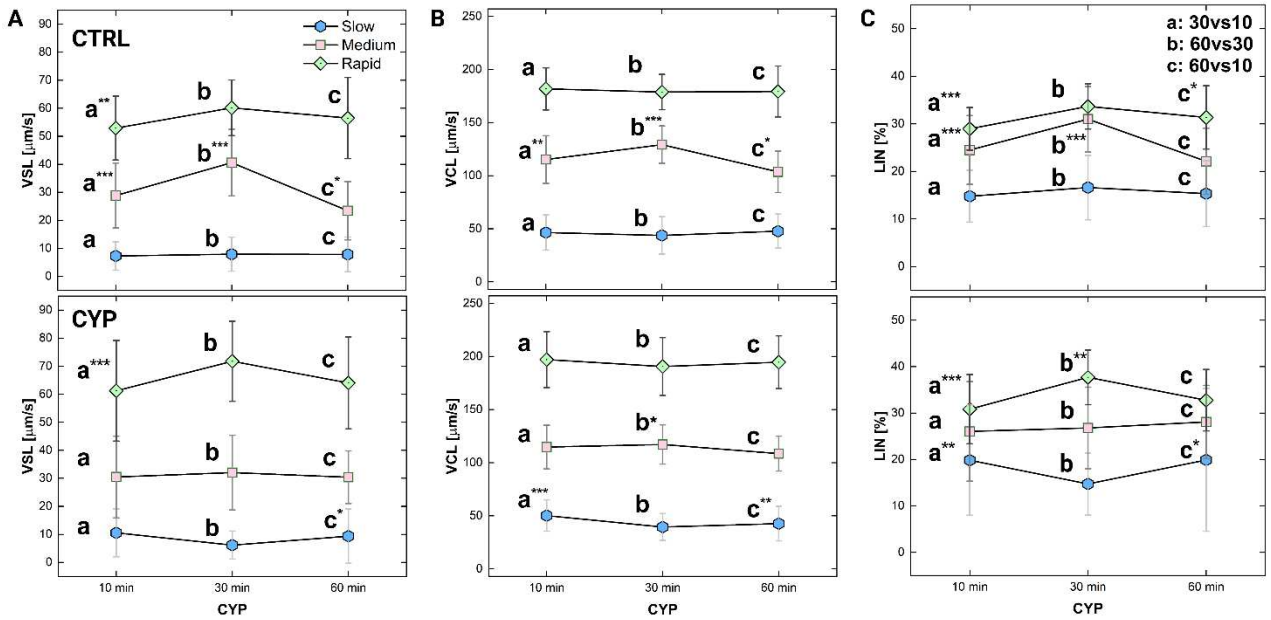

Figure 18: Motility parameter of sperm in both CTR and CYP for the three sub-populations and at the three time-step for both CTRL (on the top) and CYP (on the bottom) [150]. Each plot shows the motility parameter associated with each condition, for all the sperm classes encoded by different colors: blue for slow, pink for medium and green for rapid progressive. The statistical analysis has been performed for each parameter of a specific class between time-steps of the same experimental condition: 30vs10, 60vs30 and 60vs10. The sub-populations were not compared among each other. For each sub-population, a: 30vs10, b: 60vs30 and c: 60vs10. The degree of significance – evaluated with the Kruskal-Wallis criterion – is represented by the * associated to the letters: ***p-value<0.001, **p-value<0.01 and *p-value<0.05.

The motility analysis of sperm across the three time points is presented in terms of VSL, VCL, and LIN for both CTRL and CYP conditions (Figure 18). Rapid sperm in the CYP condition displays higher VSL and VCL compared to CTRL. Specifically, at 10 minutes, VSL and VCL for CYP were 16% and 8% greater than CTRL, respectively (Figures 18A, B). At 30 minutes, the increase reaches 19% and 7%, while at 60 minutes, it is 13% and 9%. When examining individual sperm classes, the VSL of rapid sperm increases between 10 and 30 minutes in both conditions (Figure 18A), with these variations being statistically significant ($p < 0.01$ for CTRL and $p < 0.001$ for CYP). This increase is more pronounced in the CYP condition. An overall raise in progressiveness over time is reflected in LIN ($p < 0.001$; Figure 18C), as VCL remains relatively stable between time points. For the medium sperm class, progressiveness (LIN) increases significantly in the CTRL condition ($p < 0.001$), showing a stronger dependence on time (Figure 18C–top). Conversely, the slow sperm class is minimally affected by both incubation time and treatment, with VSL, VCL, and LIN values remaining consistent throughout the experiment, except for a slight reduction in LIN at 30 minutes ($p < 0.01$; Figure 18C–bottom). At 60 minutes, an opposite trend emerges for the rapid sperm class, where both VSL and VCL decreased in both CTRL and CYP conditions. The decline is more pronounced for VSL, resulting in a reduction in LIN (Figure 18C). The medium sperm class, however, maintains constant VSL and LIN at 60 minutes under the CYP condition (Figure 18C–bottom). The observed changes in VSL at CYP, which are more substantial than those in VCL (Figure 18B), directly influences LIN (Figure 18C). LIN increases at 30 minutes for the rapid sperm class ($p < 0.001$) but decreases for the slow class ($p < 0.01$), while it remains unchanged for the medium class (Figure 18C–bottom). Thanks to the optimised tracking, it was possible to assess CYP effect on sperm cells acquired at 20x magnification in brightfield, in unconfined conditions [150].

3.2 Performance of sperm head detection and segmentation

3.2.1 Sperm detection

YOLOv4-tiny was trained with four different combinations of batch size and input size, namely 32 and 64 and 512x512 and 640x640 pixels (Figure 19). The second input size was chosen as seen in literature, although this cropping size has been applied to sperm images acquired at 100x magnification [114]. The training time for 1 epoch incremented from 15 mins to 17-18 mins with the increasing input size. The fixed settings for the training are reported in Table 5. The network performances were evaluated in terms of precision, recall, AP^{50} and F1score (Table 6). Precision is the ratio of true positive prediction with respect to all positive predictions, whether true of false. Consequently, high precision indicates that the network false detections are low. Recall on the other hand evaluates the true positive predictions on all true positive predictions, both true positive and false negatives. High recall means that the model is correctly predicting all the true positive in the test set. Precision-Recall curves for both Face and Rim detection are reported in Figure 19 A-D. As can be appreciated from Figure 19A-B, the networks perform well for Face detection when the batch size is equal to 32, without a relevant influence of the input size. When the batch size increases, Face detection decreases at high recalls, meaning that although most of the objects are correctively detected (high recall), the number of false positive increases (Figure 19C-D). With regards to Rim detection, the influence of the input size is more relevant, as it is to be expected due to the smaller dimension of the sperm head in Rim configuration. The highest precision-recall values for this class are in fact obtained for input size equal to 640x640 pixels. Increasing input size performs well on both classes in terms of precision and recall.

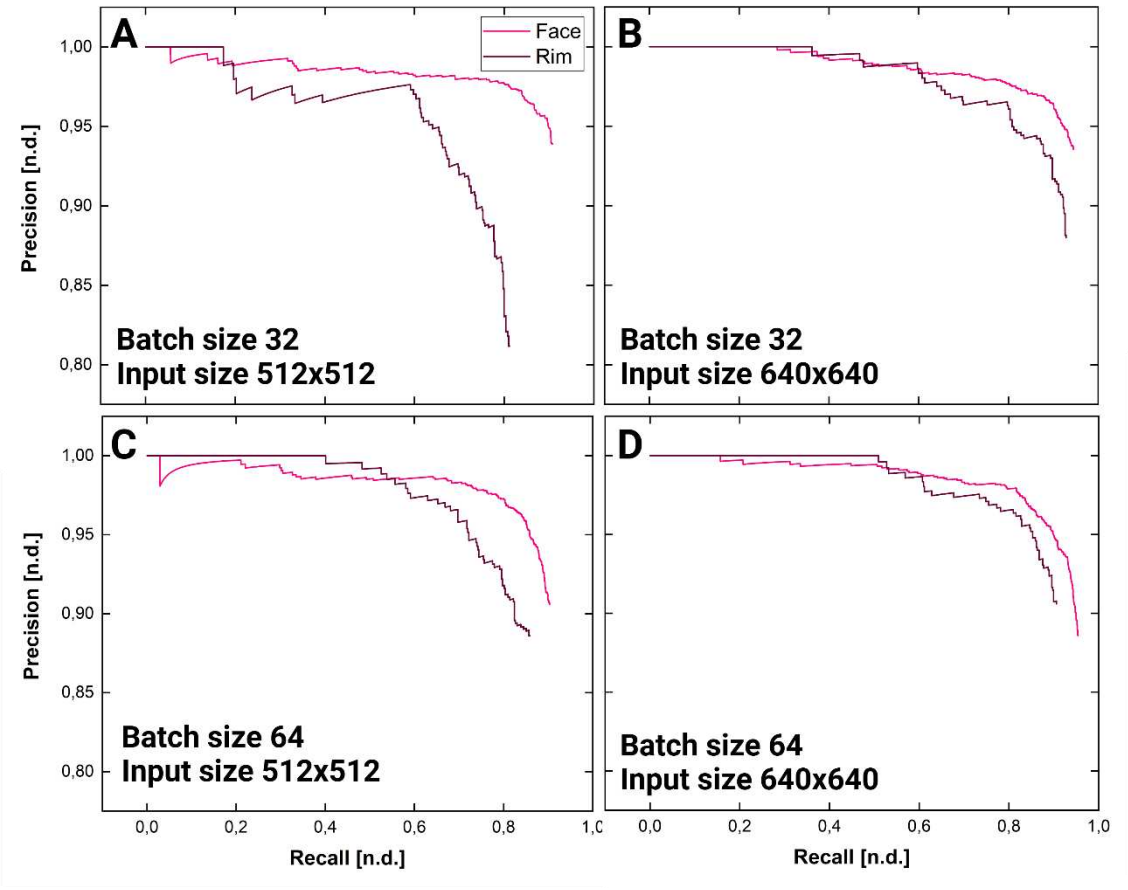

Figure 19: Yolov4-tiny performances in terms of Precision-Recall curves for an IoU=0.5. For each configuration both Face and Rim detection performances have been reported.

The F1score can be defined as a measure of classification accuracy, represented as a harmonic mean of the previous metrics, which, when combined in this way, provide a better summary of the model's predictive performance. The F1 value ranges from 0 to 1, with the upper limit indicating maximum model performance. The AP^{50} is the precision for an IoU of 0.5%. The network having batch size equal to 32 and input size 640x640 pixels was chosen as the best performing one, given that both AP^{50} and F1 score are the highest for both classes, as can be seen in Table 11.

Table 11: Performance metrics for sperm detection. The highlighted box represents the best performing network.

	CLASS	AP^{50}	F1SCORE
Batch size 32, Input size = 512x512 pixels	Face Rim	0.89 0.78	0.58 0.57
Batch size 32, Input size = 640x640 pixels	Face Rim	0.93 0.91	0.61 0.62
Batch size 64, Input size = 512x512 pixels	Face Rim	0.88 0.84	0.6 0.58
Batch size 64, Input size = 640x640 pixels	Face Rim	0.94 0.89	0.62 0.59

These results are further confirmed by the validation loss curve acquired during the training (Figure 20). As previously mentioned, the learning rate was fixed for the first 20 epochs, then scaled twice for 10 additional epochs. For each training session, the best validation loss configuration was selected

for further training with different learning rate. This is reflected by Figure 20A, where the validation curves do not have the same length for the tested networks. As a matter of fact, the network configuration having batch size equal to 32 and input size 640x640 pixels (purple line) reached the lowest validation loss among the tested configurations. Figure 20B reports an example of sperm head detection with the best performing network.

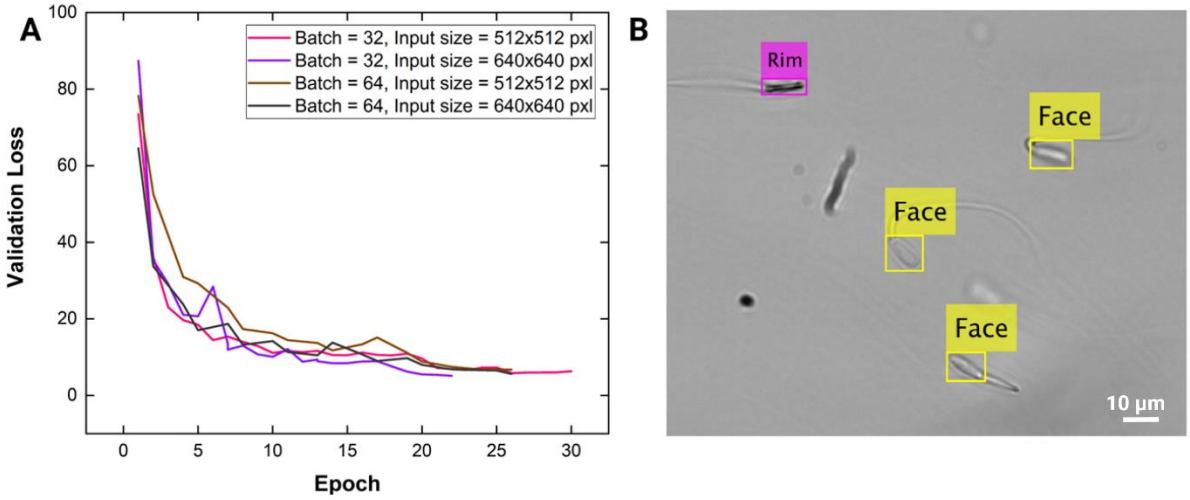

Figure 20: Validation loss and inference of the best performing network configuration for sperm head detection. A) Validation loss curves for the tested configurations. B) Zoomed-in field of view showing the output of the detection network. Yellow bounding boxes identify Face class and purple ones Rim class.

3.2.2 Sperm segmentation

For live sperm head segmentation, a lightweight version of the U-net was chosen [128]. Three different dataset combinations were analysed and compared. Given that the sperm head exhibits varying intensity and contrast when transitioning from Rim to Face, the dataset was split in three portions: one dataset (Net FR) consisted of all labelled images, including Face, Rim, and Intermediate configurations (Figure 21E). Another dataset (Net F) contained only Face and Intermediate configurations (Figure 21A), totalling 1482 images. The third dataset (Net R) consisted of 1058 images, containing only Rim configurations (Figure 21C). The dataset split had the aim to assess the network performances on swimming sperm, in order to guarantee reliable results. The three networks were trained using a fixed learning rate of 0.0001 for the U-net while varying the batch size between 16, 32, and 64 (Table 7).

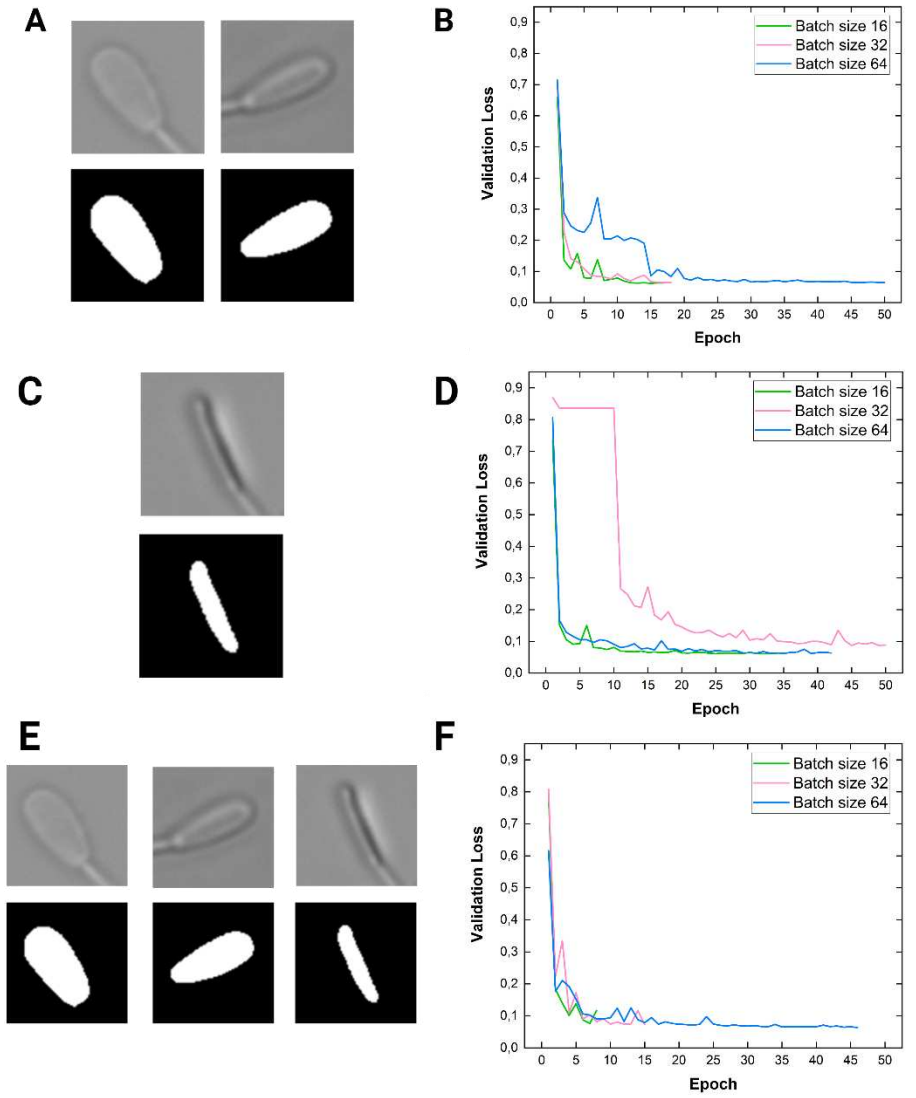

Figure 21: Combinations of datasets used for U-net-tiny training and validation loss. A) Dataset comprising Face and Intermediate configurations with the corresponding binary masks. B) Validation losses for the different configurations of networks trained with different batch sizes. The network with batch size 16 reached the minimum validation loss during the training (Green curve). C) Dataset comprising only Rim configurations. D) Validation losses for the different configurations of networks trained with different batch sizes. In this case as well batch size 16 gave the best results in terms of validation loss. E) Dataset comprising Face, Intermediate and Rim configurations. F) In terms of validation loss, batch size 64 gave the best results.

Looking at the validation loss, the best configuration for NET F is the network trained with a batch size equal to 16 (Green curve in Figure 21B). In a similar manner, batch size equal to 16 allows to reach the minimum validation loss for NET R. The network trained with the whole dataset has the best validation loss when the batch size was set to 64. To assess the performances on the sperm head in all the possible configurations acquired during the motion, the networks were tested on a dataset divided in three parts, each of them containing only Face, intermediate or Rim heads. The performances of the networks were verified with the Dice score, which compares the overlap between the predicted segmentation and the ground truth (Table 8) [160]. Results are reported in the table below.

Table 12: Segmentation performances in terms of Dice score. Net F was trained and tested only on sperm heads in Face and Intermediate configurations. Net R was trained and tested only on Rim. NET FR was trained and tested on the whole dataset. Highlighted boxes represent the best performing models for each training set.

BATCH SIZE	16	32	64
NET F			
Face	94.9%	94.4%	94.1%
Intermediate	93.9%	94.3%	94.1%
NET R			
Rim	89.2%	86.7%	88.8%
NET FR			
Face	92.5%	91.9%	93%
Intermediate	88.0%	93.4%	93.8%
Rim	93.0%	88.7%	88.7%

Although the validation loss of NET F reports lower values for the network having batch size equal to 16, the split of the test set reveals lower performances on the Intermediate configuration with respect to the one having a batch size of 32. Hence this configuration is chosen as the best NET F.
The best performing networks for each training dataset combination are reported in Table 12. Comparing the Dice scores, separating the Face and Intermediate from Rim show higher scores for all classes, although slightly for Intermediate and Rim (Figure 22A). For this reason, the segmentation is operated by NET F (BS=32) and NET R (BS=16) based on the labels predicted by the previous detection step done with YOLOv4. Once the sperm head has been segmented, it is possible to retrieve the sperm head contour and centroid (Figure 22B). Knowing this information for every point of the trajectory, it is possible to compute kinematic and morphometric parameters.

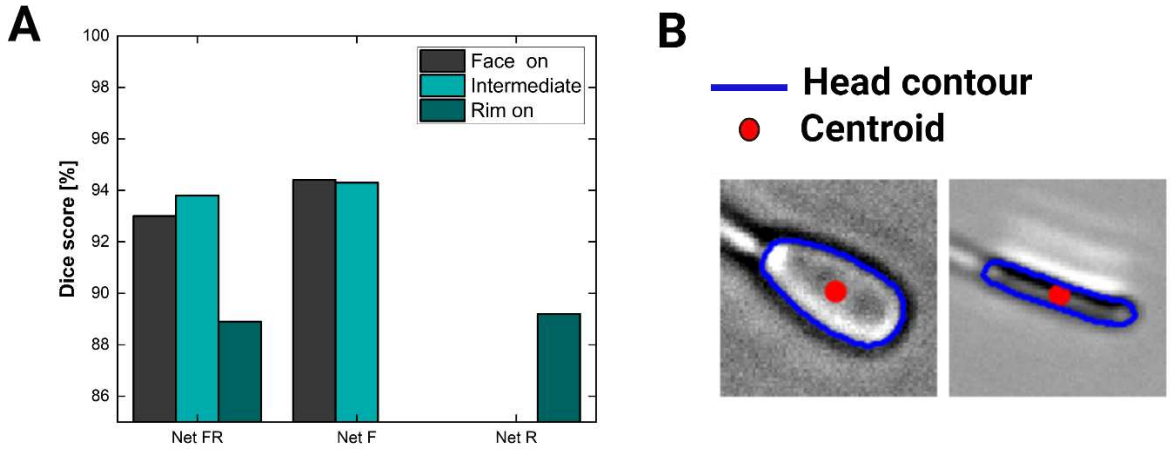

Figure 22: Comparison of the networks' performances in terms of Dice score and example of segmentation. A) Dice score of the best performing networks for sperm head segmentation. Net FR was trained on the whole dataset; Net F only on heads in Face and Intermediate configuration; Net R only on Rim configurations. The networks were tested on a dataset split in accordance with the training set. Separating the Face and Intermediate from Rim gave slightly better performances than the network trained on the whole dataset. B) Head contour (blue) retrieved from the output of the segmentation network. Through the contour the centroid is detected (red dot), allowing the precise computation of the single-sperm trajectory.

3.3 Tracking and Validation

The algorithm parameters were adjusted to track sperm head bounding boxes generated by YOLOv4 as reported in Table 4. Once the region containing the head has been identified (Figure 23A) and segmented along the whole trajectory, both kinematic and morphological parameters are computed. Figure 23B reports an example of final output of the approach, where cells are color-coded according to the WHO progressive definitions shown in Table 7. To validate the presented sperm analysis approach, Area, VSL and VCL were manually validated by tracking 24 cells from different video acquisitions for at least 1 sec and computing the mean error with respect to the automated analysis. The evaluation comprised both normally and abnormally moving sperm cells. The mean value for each parameter, measured both manually and with the NNs approach, is reported in Figure 23C. The results were in good agreement, as can be also appreciated from the mean percentage errors (Figure 23C). Morphology evaluation of sperm cells with normal motility are in good agreement with literature, although the provided values were computed on fixed semen samples (Table 13).

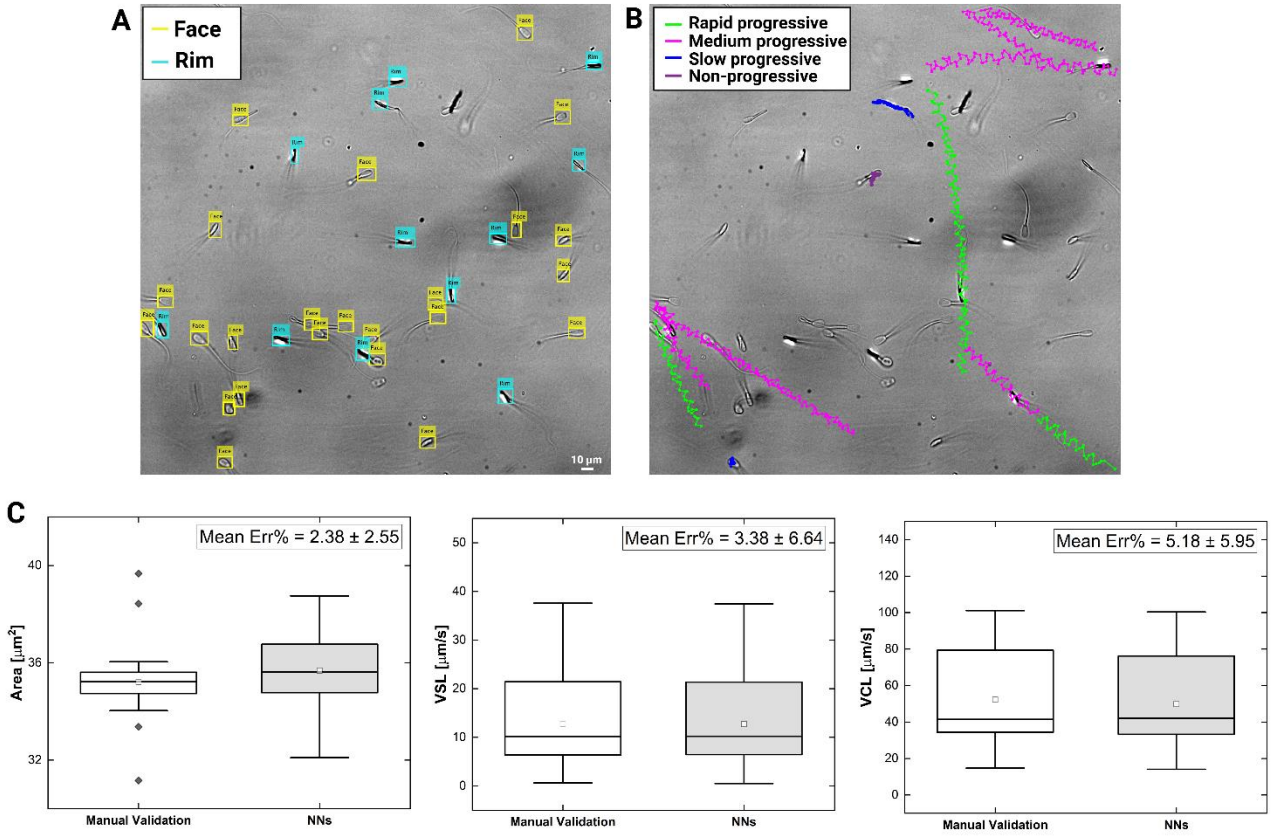

Figure 23: Output of sperm head detection and tracking. A) Example of output of YOLOv4-tiny. Sperm head in the field of view are detected and classified as Face (Yellow) or Rim (cyan). B) Kinematic analysis of single-sperm cells. Once bounding boxes are tracked, each sperm is cropped and fed to the U-net for segmentation. Computing the head centroid for all the frames, it is possible to perform the kinematic and morphometric analysis. Different colours encode for motility classes defined by CASA and reported in Table 7. C) Validation of the Area, VSL and VCL, computed on 24 cells which were manually tracked for at least 1sec and compared with the network results. Each plot reports the mean percentage error.

Table 13: Bull sperm head dimension reported in literature according to different evaluation methods.

AREA [μm^2]	MAJOR AXIS [μm]	MINOR AXIS [μm]	REF.
37.7 ± 1.76	9.4 ± 0.27	5.2 ± 0.31	Our results
32.01± 0.438	8.99 ± 0.1	4.52 ± 0.038	[161]
38.08 ± 0.055	9.49 ± 0.017	4.68 ± 0.007	[161]
39.97±0.17	9.43 ± 0.02	5.13 ± 0.01	[162]
36.20±0.09	9.00 ± 0.01	4.82 ± 0.009	[162]
43.19 ± 4.47	10.11 ± 0.57	5.21 ± 0.43	[163]
41.97 ± 2.90	10.27 ± 0.57	5.78 ± 0.31	[164]

3.4 Influence of hyperosmolarity on sperm motion

The osmolarity of the environment has significant effects on sperm motility and function [28]. Osmotic stress can alter the integrity and motility of sperm by affecting the movement of water across the sperm membrane, causing either swelling or shrinking, which subsequently impacts cellular activity. Exposure to hypoosmotic conditions (low osmolarity, below $150\ mOsm/kg$) causes sperm swelling, leading to impaired membrane integrity, compromised mitochondrial function, and reduced motility in different mammals. The swelling effect is however visible mainly in the tail, due to the reduced tail membrane tightness with respect to the head [165]. The same is reported for sperm in hyperosmolar conditions, although the resistance to permanent damages is higher with respect to hypoosmotic conditions [166]. In hyperosmotic environments sperm cells lose water, leading to cell shrinkage and subsequent activation of regulatory volume mechanisms. The plasma membrane undergoes hyperpolarization as a response, and ion channels such as sodium-hydrogen exchangers (NHE) and aquaporins (AQPs) are activated to counteract the osmotic imbalance [27]. As previously mentioned, the ability of the sperm to adapt to surrounding fluid osmolarity is fundamental for the sperm journey through the FRT [27]. This is particularly relevant for bovine sperm, which experience an increasing gradient of osmolarity going from the seminal plasma to the oviduct ($280-300\ mOsm/Kg$ to $355\ mOsm/Kg$). While the opposite occurs for humans, it has been seen that human seminal plasma osmolarity is altered in pathological conditions such as asthenozoospermia, where the higher osmolarity negatively impacts sperm motility [167].
Bovine sperm cells were exposed to a medium osmolarity comparable to the oviduct (HYP1) and a hyperosmotic condition at $420\ mOsm/Kg$. Further osmolarities were not explored to preserve vitality [166, 167]. The latter was created in two different ways, that is enriching the sperm medium with NaCl in one case (HYP2) and Fructose in another (FR). Three timesteps were analysed, namely 0, 30 and 60 min after contact with the hyperosmolar media.

3.4.1 NaCl

To verify sperm head volume alterations due to hyperosmolarity an investigation of the area, major and minor axes, and height was performed though confocal analysis. By using z-stacks, it was possible to estimate the head third dimension analysing the head fluorescence intensity among the slices (Figure 14). Globally, Area, Height and Major axis are lower than ISO for both HYP1 and HYP2 (Figure 24), although not statistically significant. Further experiments would be needed to improve sample numerosity. In literature, 60 min is the time required for volume regulation after

exposure to hyperosmotic conditions [168]. Further timepoints were not explored due to the mortality increase occurring at longer times, which would have been detrimental to motility investigation.

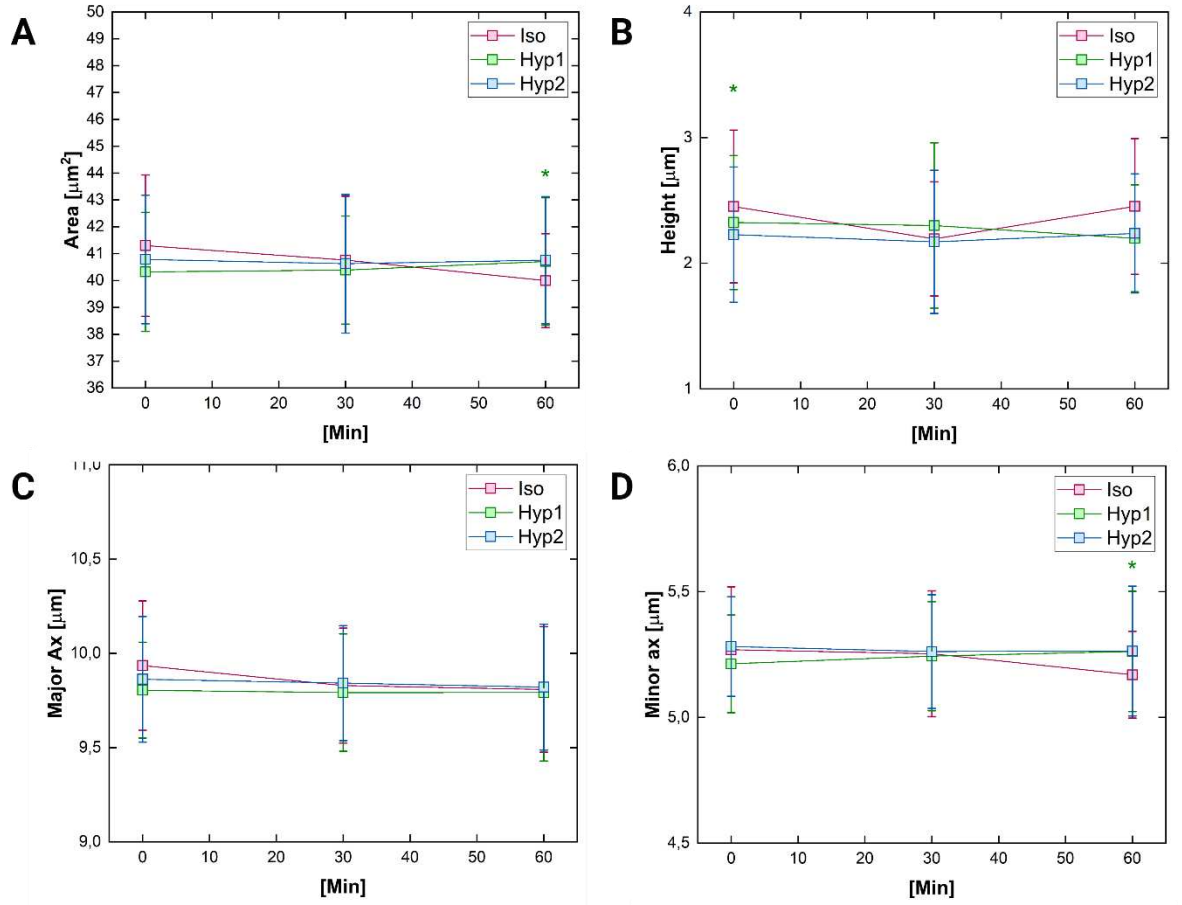

Figure 24: Confocal analysis of bovine sperm in hyperosmolar conditions. A) Head area, evaluated at 0, 30 and 60 min. Different colours encode for the studied conditions. B) Height C) Major axis and D) Minor axis. ISO numerosity: $N_{0min} = 39, N_{30min} = 30, N_{60min} = 27$. HYP1: $N_{0min} = 32, N_{30min} = 34, N_{60min} = 32$. HYP2: $N_{0min} = 32, N_{30min} = 31, N_{60min} = 35$. Statistical analysis was performed between hyperosmolar conditions and ISO at fixed timesteps. (*: p<0.05; **:p<0.005; ***:p<0.001*).*

The resulting kinematic parameters for HYP1 and HYP2 conditions are reported below. Figure 25A reports the summary of the CASA parameters definitions. Both hyperosmolar treatments had a detrimental effect on sperm motility with increasing time as can be seen from VSL, VCL and LIN decrease (Figure 25B-D). HYP2, in particular, shows higher detrimental effect at 30 min already with respect to HYP1, which is in accordance with the inverse relationship of motility and hyperosmolarity [167]. No significant differences are found in MAD between ISO, HYP1 and HYP2. It was seen that in hyperosmolar conditions many sperm cells showed abnormal midpiece curvature, resulting in abnormal motion. This is reflected by the altered VCL and fractal dimension. FD in particular shows significant differences at 30 and 60 min among all conditions. Higher FD are indicative of irregular, hyperactivated patterns [55], indicating the prevalence of altered motion patterns. These alterations are due to NaCl as ionic compounds influence the membrane potential and intracellular pH, both of which are essential for processes like capacitation and hyperactivation [169]. The presence of NaCl in the environment is thought to change the ion channels' activity, including sodium-hydrogen exchangers (NHEs), which are critical for maintaining the internal pH during the journey through the

female reproductive tract. This ion unbalance ultimately impacts sperm motility and fertilization capability [169].

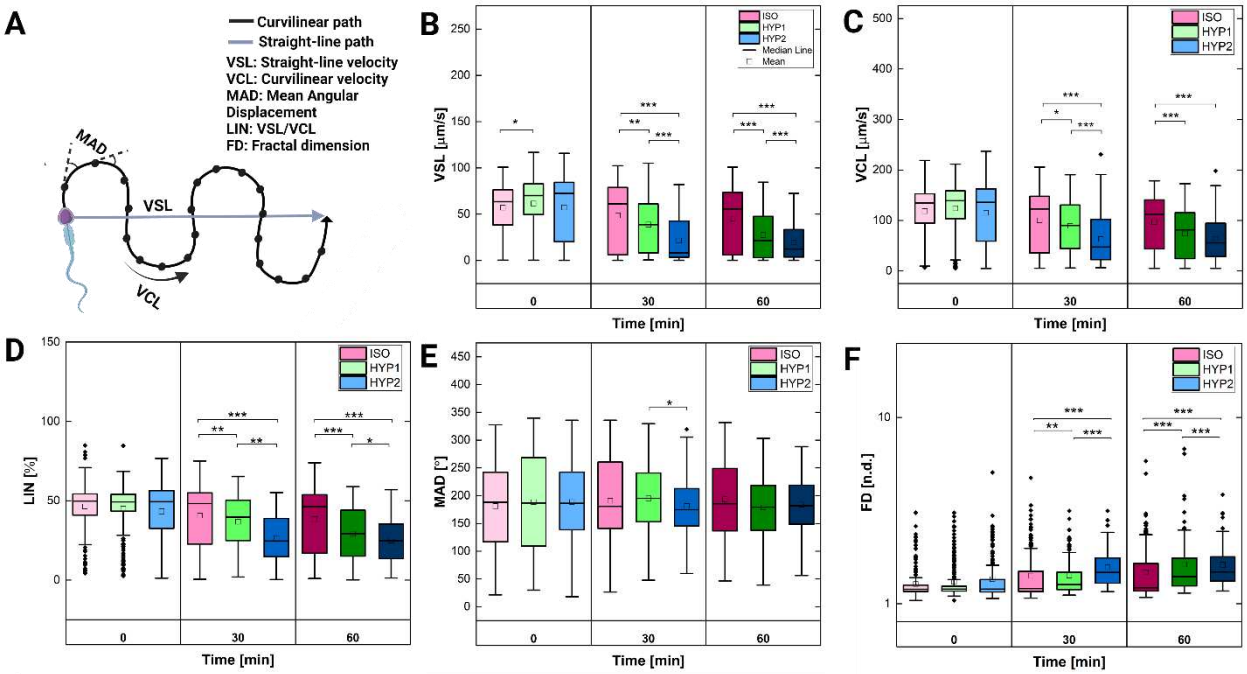

Figure 25: Kinematic analysis of bovine sperm in hyperosmolar conditions (HYP1 and HYP2). A) Sketch of the CASA parameters definition. B-F) CASA parameters comparison between control condition (ISO), HYP1 and HYP2 at the three timesteps, 0, 30 and 60 min. FD is reported in the logarithmic scale to better appreciate variations among the timesteps. Numerosity for ISO:$N_{0min} = 226, N_{30min} = 241, N_{60min} = 230$. Numerosity for HYP1: $N_{0min} = 212$, $N_{30min} = 141$, $N_{60min} = 157$. Numerosity of HYP2: $N_{0min} = 222, N_{30min} = 144, N_{60min} = 157$. Statistical dependence among conditions at fixed timesteps was assessed with Kruskal-Wallis analysis. (*: p<0.05; **: p<0.005; ***: p<0.001).

3.4.2 Fructose

The confocal analysis was repeated for FR as well. A statistically significant decrease of the Area in FR with respect to ISO was found at 30 and 60 min. The same holds true for the major axis at 60 min. The trend was more significant than that induced by NaCl as osmolyte. As discussed in the previous section, longer times were not investigated to avoid alterations in the motility evaluation due to increased mortality.

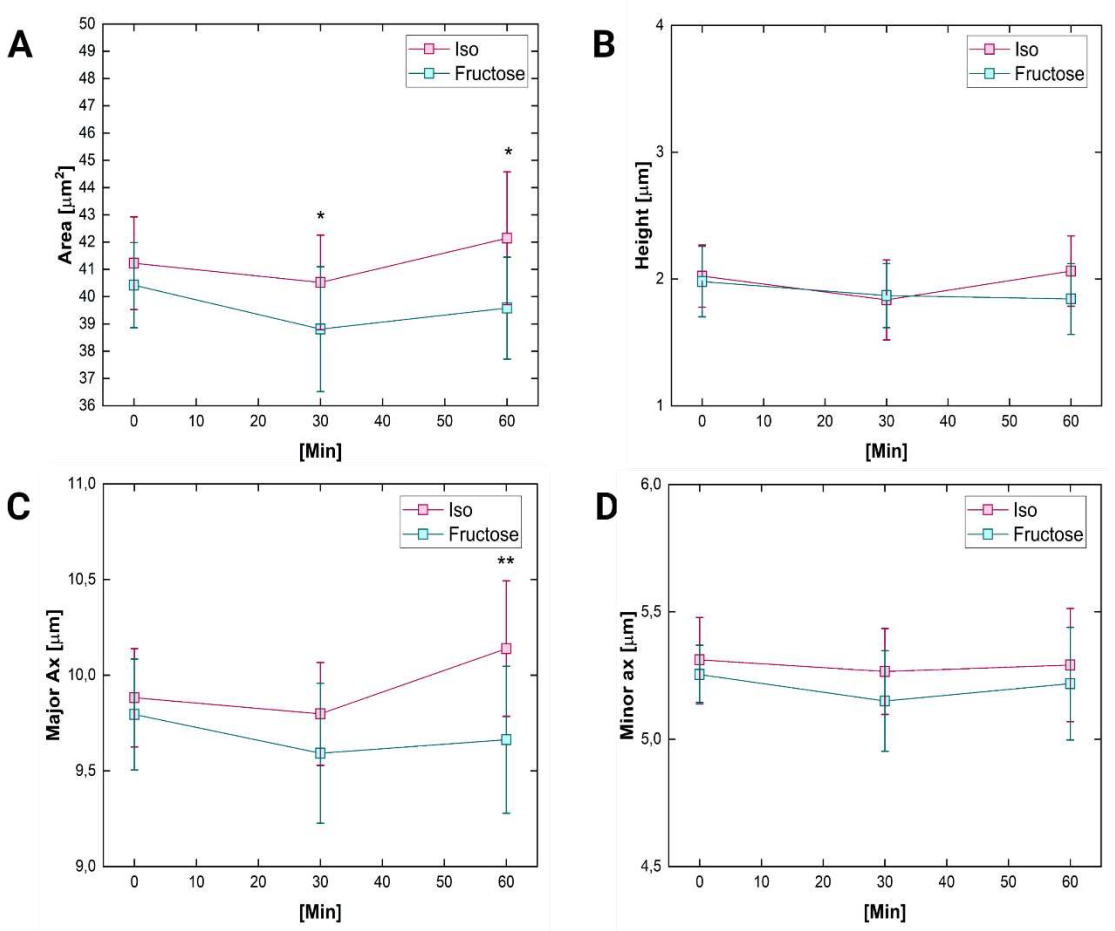

Figure 26: Confocal analysis of bovine sperm in hyperosmolar conditions. A) Head area, evaluated at 0, 30 and 60 min. Different colours encode for the studied conditions. B) Height C) Major axis and D) Minor axis. ISO numerosity: $N_{0min} = 13, N_{30min} = 14, N_{60min} = 8$. FR: $N_{0min} = 15, N_{30min} = 13, N_{60min} = 14$. Statistical analysis was performed between hyperosmolar conditions and ISO at fixed timesteps. (*: p<0.05; **:p<0.005; ***:p<0.001)

With regards to the effect of hyperosmolarity induced by increased fructose on sperm kinematics (FR), the results are reported in Figure 27. The influence of fructose-induced hyperosmolarity had less impact on sperm kinematic parameters with respect to HYP2 (Figure 25). The only significant differences were found for VCL and FD at 0mins, but it could be due to the scarce numerosity of ISO. However, from the statistical point of view, the result is consistent since the minimum sample size recommended for the Kruskal-Wallis test is respected [170]. Being an organic, non-penetrating osmolyte, fructose affects sperm cells by creating a hypertonic environment that induces osmotic stress primarily through water efflux, leading to cell shrinkage. Since it does not penetrate the cell membrane, its osmotic action tends to be less aggressive on the intracellular environment and ionic balance compared to inorganic salts.

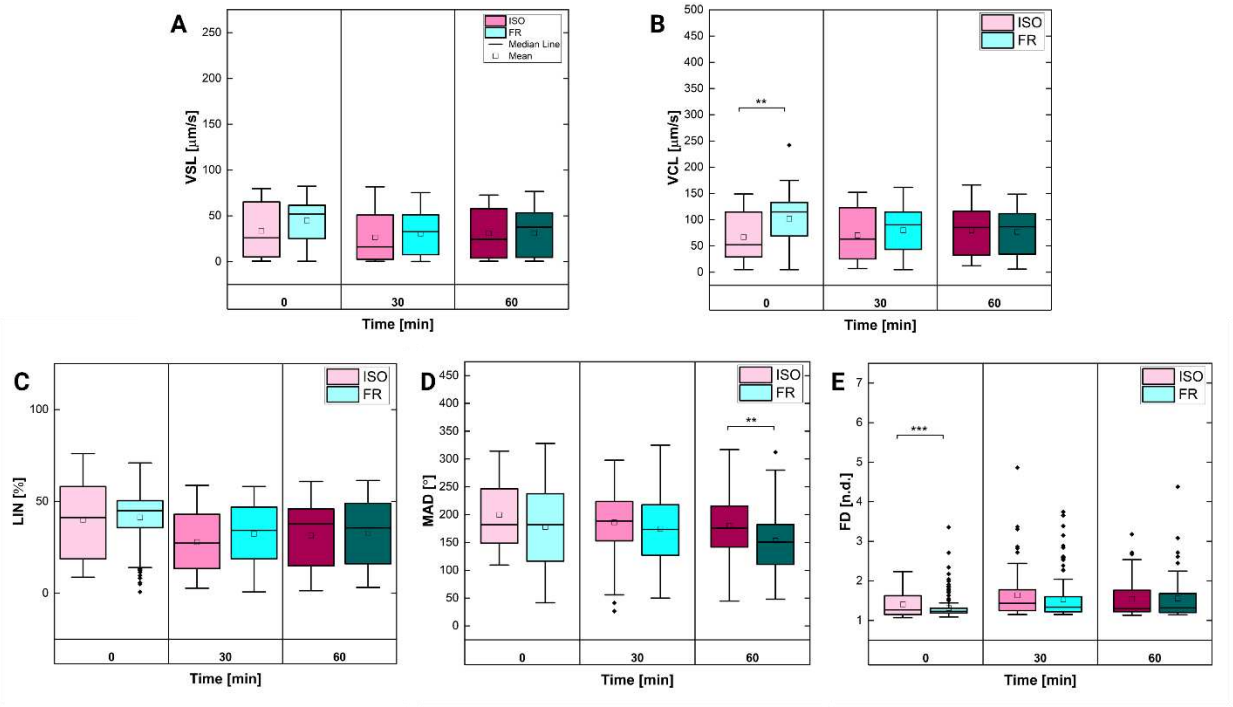

Figure 27: Kinematic analysis of bovine sperm in hyperosmolarity obtained adding fructose to sperm buffer (FR). ISO numerosity: $N_{0min} = 20, N_{30min} = 81, N_{60min} = 73$. FR: $N_{0min} = 153, N_{30min} = 136, N_{60min} = 66$. (*: p<0.05; **:p<0.005; ***:p<0.001).

3.4 Caffeine effect on sperm kinematic parameters

Caffeine is known in literature for having an enhancing effect on sperm motility in different species [157], [171], [172], [173]. This is thought to be attributed to the phosphodiesterase enzyme inhibition, which lead to increased intracellular cyclic adenosine monophosphate (cAMP) levels. This rise in cAMP activates protein kinase A (PKA), which then stimulates the processes necessary for sperm flagellar beating, thus enhancing motility [172]. Human sperm exposed to $10mM$ of caffeine exhibited a positive increase of motility percentages [174]. For this reason, it was chosen to expose bovine sperm to an equal caffeine concentration. Caffeine appears to cause a reduction in linear progressivity immediately after contact, as can be seen from both VSL and LIN (Figure 28 A and C). Furthermore, FD also shows significant differences at all timepoints (Figure 28D). High FD is associated with highly irregular, hyperactivated motion [55]. This is in accordance with a literature finding where ram sperm, exposed to a medium containing $10mM$ of caffeine, showed significantly increased hyperactivation [175]. The study found that caffeine induces hyperactivation independently of PKA pathway, suggesting the involvement of other mechanisms such as calcium influx.

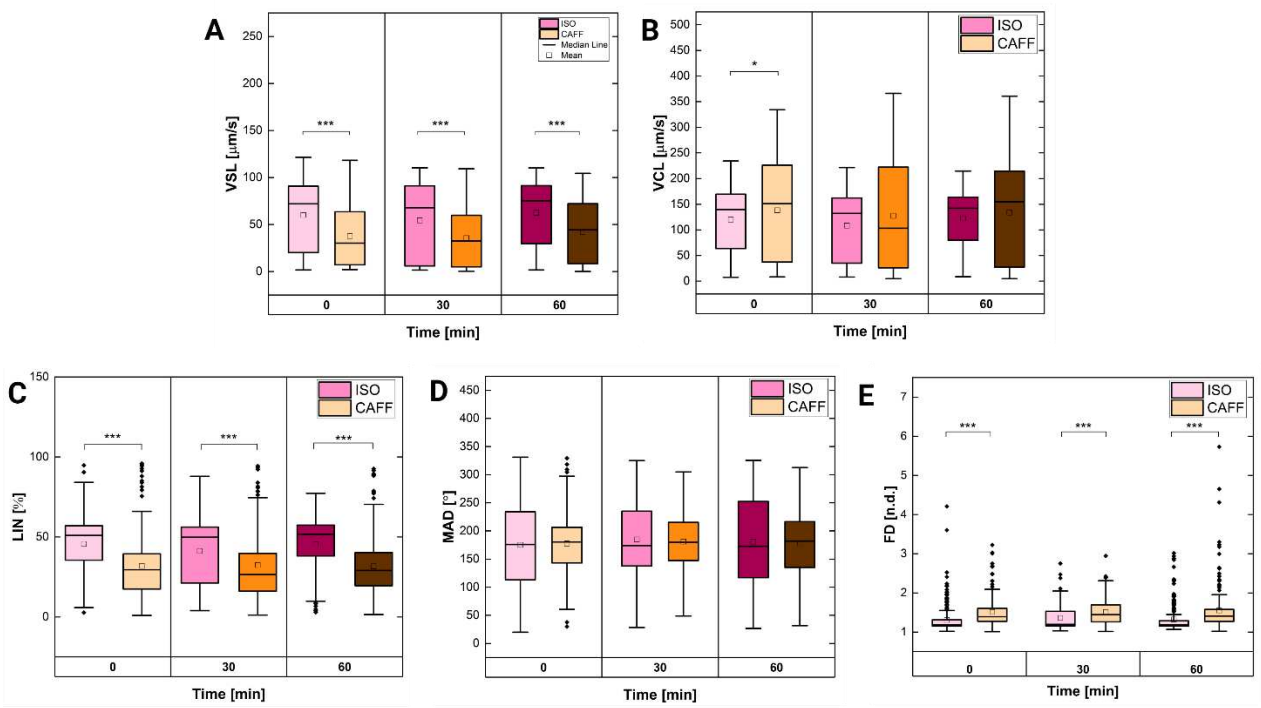

Figure 28: Effect of caffeine on sperm kinematic parameters. A-E) CASA parameters comparison between control condition (ISO) at the three timesteps, 0, 30 and 60 min. Numerosity of ISO: $N_{0\,min} = 271$, $N_{30\,min} = 201, N_{60\,min} = 215$. Numerosity of CAFF: $N_{0\,min} = 149$, $N_{30\,min} = 109, N_{60\,min} = 180$. (*: p<0.05; **:p<0.005; ***:p<0.001).

3.5 Absence of calcium on sperm motility

Calcium plays a crucial role in sperm motility by acting as a major secondary messenger that regulates processes such as capacitation**,** hyperactivation**,** and the acrosome reaction [176]. The impact of the extracellular concentration of Ca^{2+} on motility has been well-debated. Several studies have reported that extracellular Ca^{2+} enhances sperm motility, whereas others have also reported that a high concentration of Ca^{2+} inhibits sperm motility [15]. When Ca^{2+} was removed from the external medium, sperm motility of bovine sperm was significantly reduced [177]. Another study found similar results with human sperm, and hypothesized that calcium depletion in the external medium impairs the voltage-gated channels inducing membrane depolarization [158]. This depolarization impaired sperm motility by stimulating Na+/K+-ATPase activity, which in turn depleted intracellular ATP reserves, crucial for powering the flagellar movement necessary for motility. To assess sperm motility in absence of extracellular calcium, sperm were diluted in PBS. The kinematic parameters resulting from the analysis are reported below in Figure 29. A decrease in kinematic parameters is in accordance with the motility reduction seen in literature. Both VSL, VCL and FD are significantly lower at all timesteps with respect to ISO. The higher FD suggest the prevalence of in-situ/irregularly swimming sperm. No significant differences were seen for MAD.

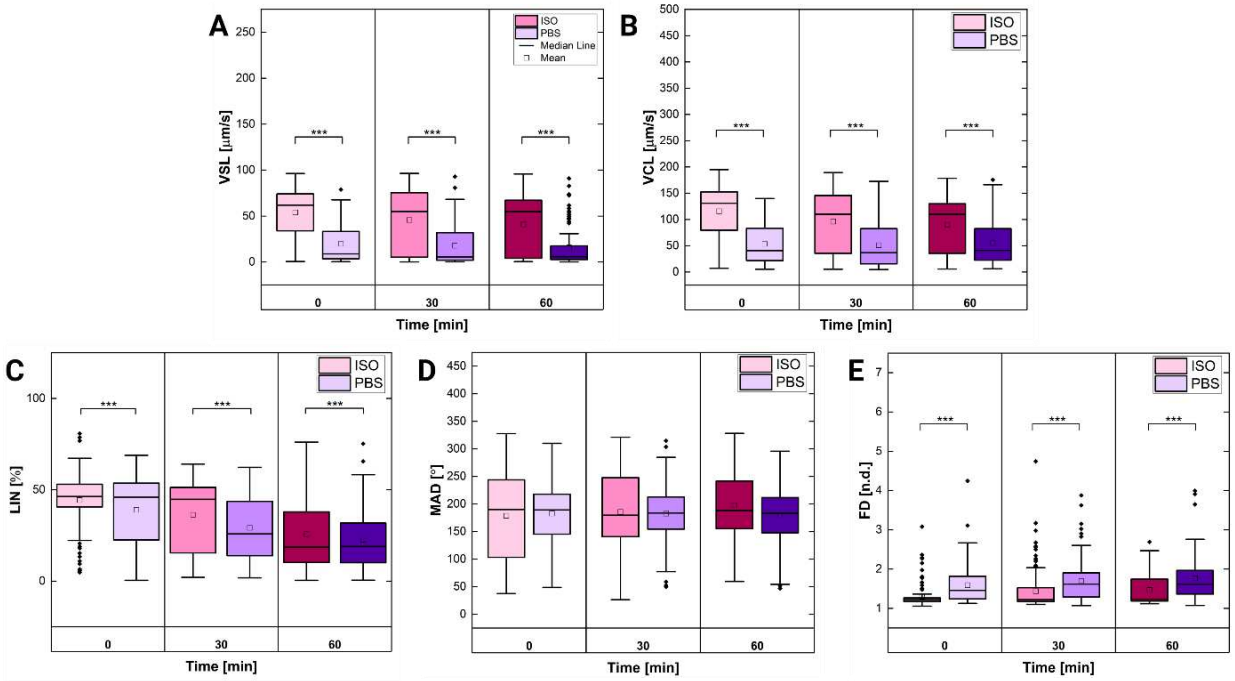

Figure 29: Effect of a medium without Ca^{2+} (PBS) on sperm kinematic parameters. A-E) CASA parameters comparison between control condition (ISO) at the three timesteps, 0, 30 and 60 min. Numerosity of ISO: $N_{0\,min} = 118$, $N_{30\,min} = 151, N_{60\,min} = 86$. Numerosity of CAFF: $N_{0\,min} = 108$, $N_{30\,min} = 128, N_{60\,min} = 84$. (*: p<0.05; **:p<0.005; ***:p<0.001)

3.6 Sperm kinematics in altered viscosity

Viscosity is one of the physical barriers of the female reproductive tract [24]. It has been estimated that the cervical mucus viscosity in humans varies from 0.1 $Pa \cdot s$ to 1 $Pa \cdot s$, changing along the menstrual cycle [24], [178]. The main role of cervical mucus is to block all the sperm having reduced motility and abnormal morphology. Furthermore, altered viscosity in human seminal plasma is indicative of pathological conditions as sperm motility is highly impaired [179]. It has been seen that high viscosity induces a transition from a 3D helical swimming pattern to a 2D slithering mode in both human and bovine sperm [180]. In high viscosity, the flagellum and the sperm head align parallel to the surface, reducing yaw movement and resulting in a straighter swimming trajectory**.** To investigate motility alterations in different viscosities, sperm were diluted in PEO at two concentrations, 0.5 and 0.9%wt. As can be seen from Figure 30, viscosity lowers kinematic parameters. In accordance with literature findings on sperm motion planarization, VCL results significantly lower in both PEO conditions with respect to control. The motion planarization induced a loss of progressiveness manifesting in circular trajectories, which were strongly present in both conditions. As a matter of fact, this is suggested by the increased FD in PEO0.5. This does not happen in PEO0.9 since higher viscosity translates in shorter tracks, which appear more linear. Alike PEO0.9, PEO0.5 still presented both rolling and no-rolling sperm cells while all cells tracked and analysed in PEO0.9 showed only planar swimming patterns. Kinematic parameters alone are not able to account for this phenomenon.

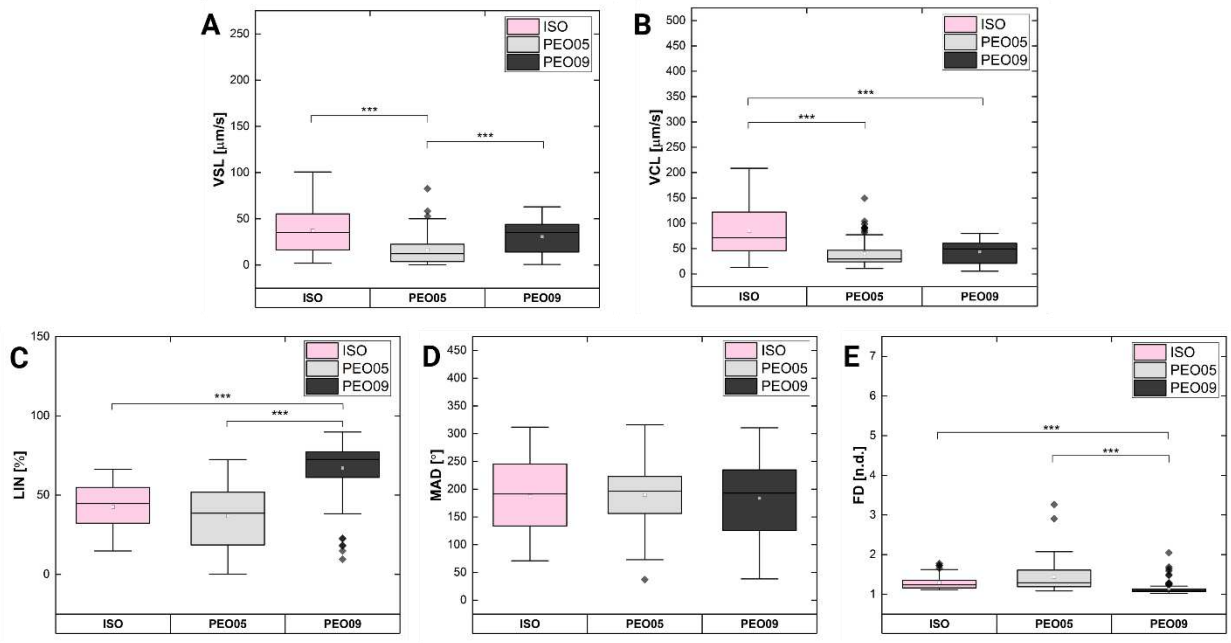

Figure 30: Kinematic parameters of sperm cells exposed to media with different viscosities. A-E) CASA parameters for control (ISO) and increased viscosity conditions (PEO0.5 and PEO0.9). Numerosity: $N_{ISO} = 47, N_{PEO0.5} = 60, N_{PEO0.9} = 92$. (*: p<0.05; **:p<0.005; ***:p<0.001)

3.7 New parameters

As mentioned before, mean CASA parameters do not account for sperm motion variability. Moreover, there is an insufficiency in the standardization of parameters such as VAP, ALH and BCF due to the variability of the average path computation [43], [44], [46]. The remaining CASA parameters do not account for the complexity of sperm trajectories patterns, since they are related only to displacement in time. For this reasons, three novel parameters were proposed, namely the prominence, rolling frequency and radius of circular fit.

3.7.1 Prominence

Prominence aims at quantifying the spatial excursion the sperm computes going from a Face to a Rim configuration. The definition is represented in Figure 31. The computation relies on the sperm area quantification of the sperm head, identifying the centroids of the maximum Face and of the minimum Rim respectively along the curvilinear path (magenta and orange points in Figure 31). The prominence is computed as the distance between the Rim on centroid and the intersection of the perpendicular line passing for that point and the segment joining two consecutive Face on centroids. In this way the head amplitude oscillation is computed bypassing the limitation of the average path computation [44]. The computation is repeated along the track and all the prominences are mediated to output the mean prominence.

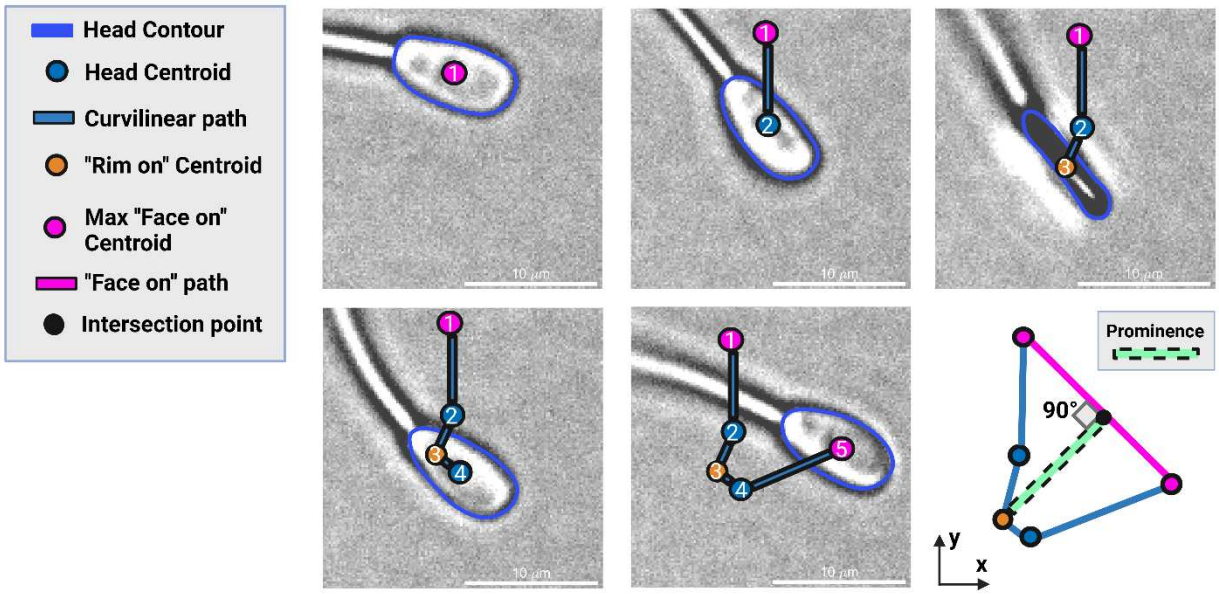

Figure 31: Graphical representation of the prominence computation. The centroids corresponding to a maximum Face on configuration (represented as magenta circles) and Rim on configuration (orange points) are identified along the curvilinear path of a single cell. The prominence is defined as the distance between the Rim on point and the intersection point of the perpendicular line passing by the Rim on point and the segment joining two consecutive Face on points (magenta line) along the trajectory.

To directly link this parameter to the rolling motion, Face on and Rim on points were estimated using the MATLAB 'findpeaks' function, set with a 'MinPeakProminence' of 20. The detailed code is reported in Appendix A. This parameter is set to 0 for no rolling and irregularly rolling sperm, where their alternation between Face and Rim is not detectable.

The implementation of this parameter in the explored experimental conditions were useful to gain more insight into altered sperm patterns. With regards to osmolarity, HYP1 and HYP2 have a very strong negative impact on kinematic parameters as shown in Figure 25. Looking at the prominence (Figure 32A), it can be seen that HYP2 is again the one having the most relevant altering effect in terms of motion especially at 30 min, indicating the presence of oscillating sperm patterns. The lack of statistical differences at 60 min can be explained by the fact that no rolling sperm have null prominence, lowering both mean and median values. Hyperosmolarity induced by fructose show on the other hand low impact on kinematic parameters with the exception of VCL and FD at 0min (Figure 32B and E). The prominence indicates the presence of oscillating patterns at 0 and 30 min (Figure 32B) similar to HYP2, suggesting the presence of highly oscillating motion patterns on earlier timesteps.

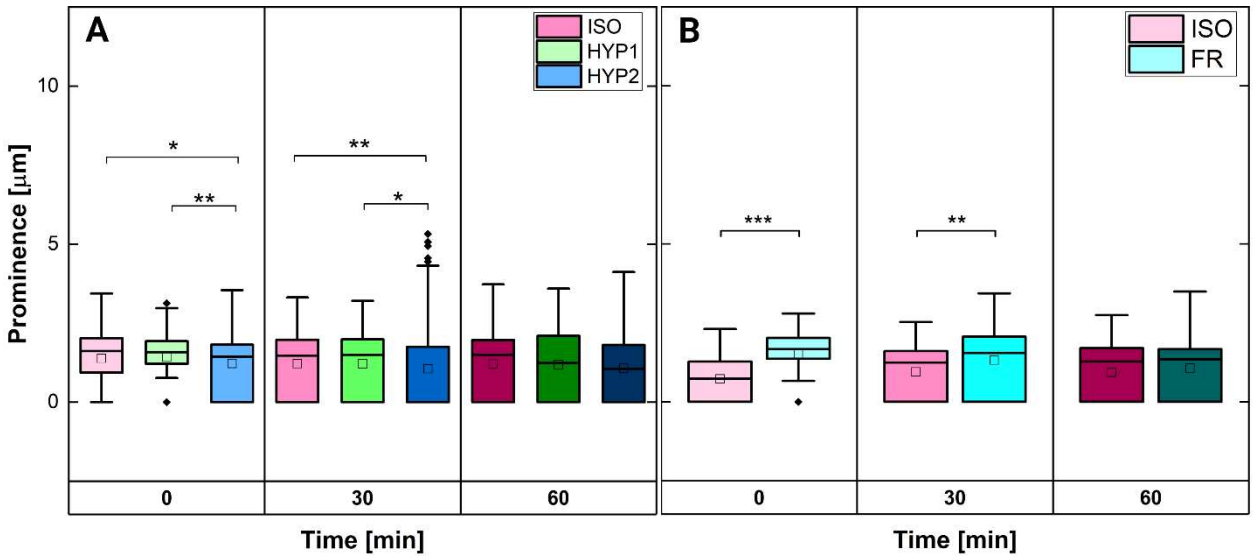

Figure 32: Prominence evaluation in experiments with altered osmolarity. A) Prominence of HYP1 and HYP2 and B) Prominence in FR. (*: p<0.05; **: p<0.005; ***: p<0.001)

Prominence was plotted as a function of VCL as the latter is the parameter that defines motile classes according to CASA systems (Table 9). Linear fitting was performed on Origin and slope and coefficient of determination (R^2) were reported (Figure 33). By looking at ISO (Figure 33A), the changing on slope is in accordance with a positive effect on motility induced by the prolonged stay at $37°C$. The opposite occurs for HYP1, HYP2 and FR (Figure 33B-D), where the slope of the linear fit decreases for longer times. While normal prominence values in ISO reside from 1 and 3 μm, the prominence distribution moves towards higher values in HYP2 at 30 and 60min and FR at 60min. The R^2 value indicates a moderate linear dependency among the observed variables.

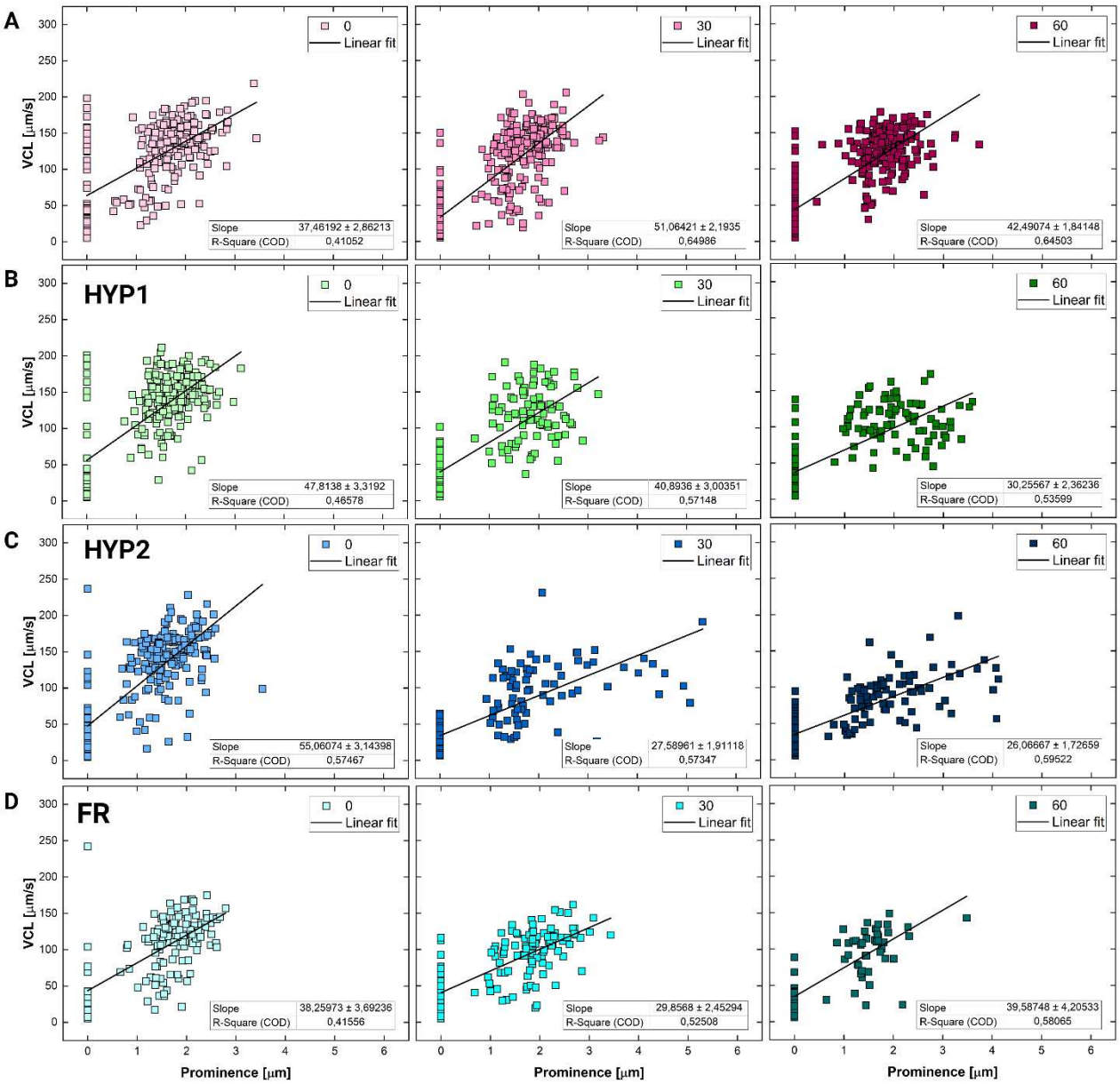

Figure 33: Scatter plots presenting the prominence for (A) ISO, (B) HYP1, (C) HYP2 and (D) FR. Liner fitting was performed with OriginLab and the resulting slope and R-squared are reported in the plots. Each condition is colour coded, and gradients of the same colour encode for different timesteps.

Prominence of bovine sperm diluted in medium with caffeine is higher than that of control (Figure 34A). Although no significant statistical differences were retrieved, the distributions were wider with respect to control at every timestep. It has been found in literature that hyperactivation may have a double pattern, one progressive and one star-like [33]. The progressive hyperactivation is characterized by high lateral head displacements and high VCL. This effect is captured by the prominence, which is higher for more oscillating trajectories. The coefficients of correlation are around 0.5 for all conditions and timesteps, indicating a moderate linear dependence between the parameters. As a matter of fact, there are cells with high VCL and prominence, indicating high velocity along the undulatory path. However, there are cells having high VCL but low prominence, indicating more linear trajectories. Null prominences, as mentioned before, indicate no rolling or irregularly rolling sperm, where the oscillation between Face and Rim is not detected. Further

investigation of these fractions in each treated condition will be of interest for future work. With regards to experiments in PBS, the decrease of prominence at all timesteps is in accordance with that of kinematic parameters (Figure 34B). The absence of calcium in the extracellular medium induces a motility reduction in terms as head oscillation as well as progressivity.

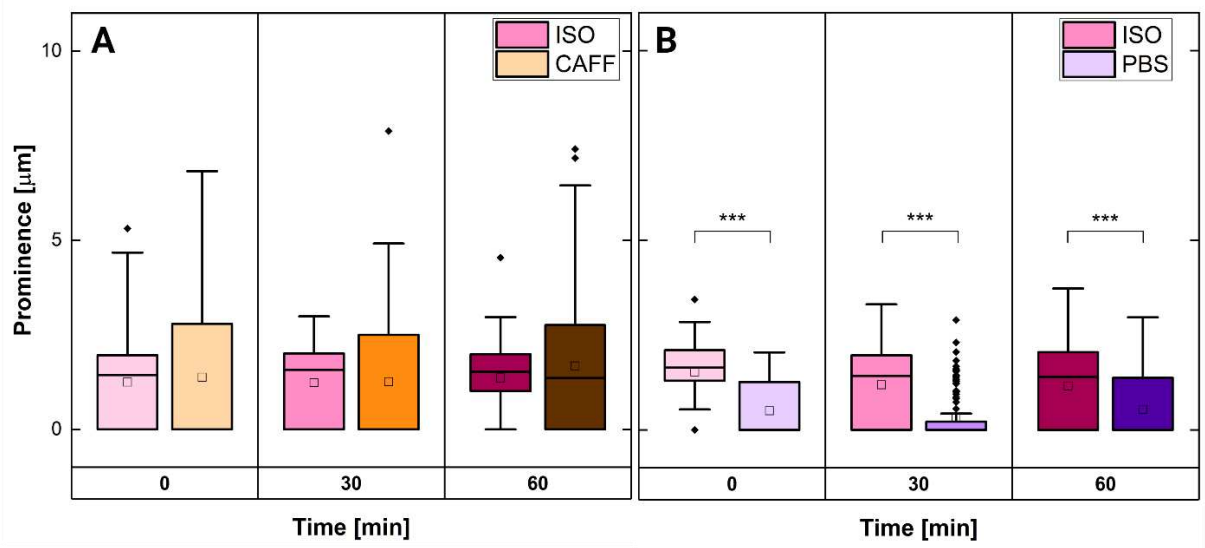

Figure 34: Prominence evaluation in experiments with caffeine and absence of Ca^{2+}. A) Prominence in CAFF and B) Prominence in PBS. (*: p<0.05; **: p<0.005; ***: p<0.001)

These findings are confirmed by looking at Figure 35, where prominence is presented as a function of VCL. The highest values for prominence are in fact reached in CAFF, especially at 60min (Figure 35B). Further alterations are suggested by the number of cells having null prominence and high VCL in CAFF, higher with respect to ISO. This indicates the presence of irregular behaviours similar to hyperactivation, that is characterized by very oscillating trajectories that not always result in progressiveness [30]. The opposite occurs in PBS, where low prominence corresponds with low VCL, highlighting the motility loss discussed before (Figure 35C). With respect to hyperosmolar conditions discussed before, much higher prominences are reached in sperm treated with caffeine, underlying a strong motility alteration. This is reflected by the low values of R^2 in the treatments with respect to ISO.

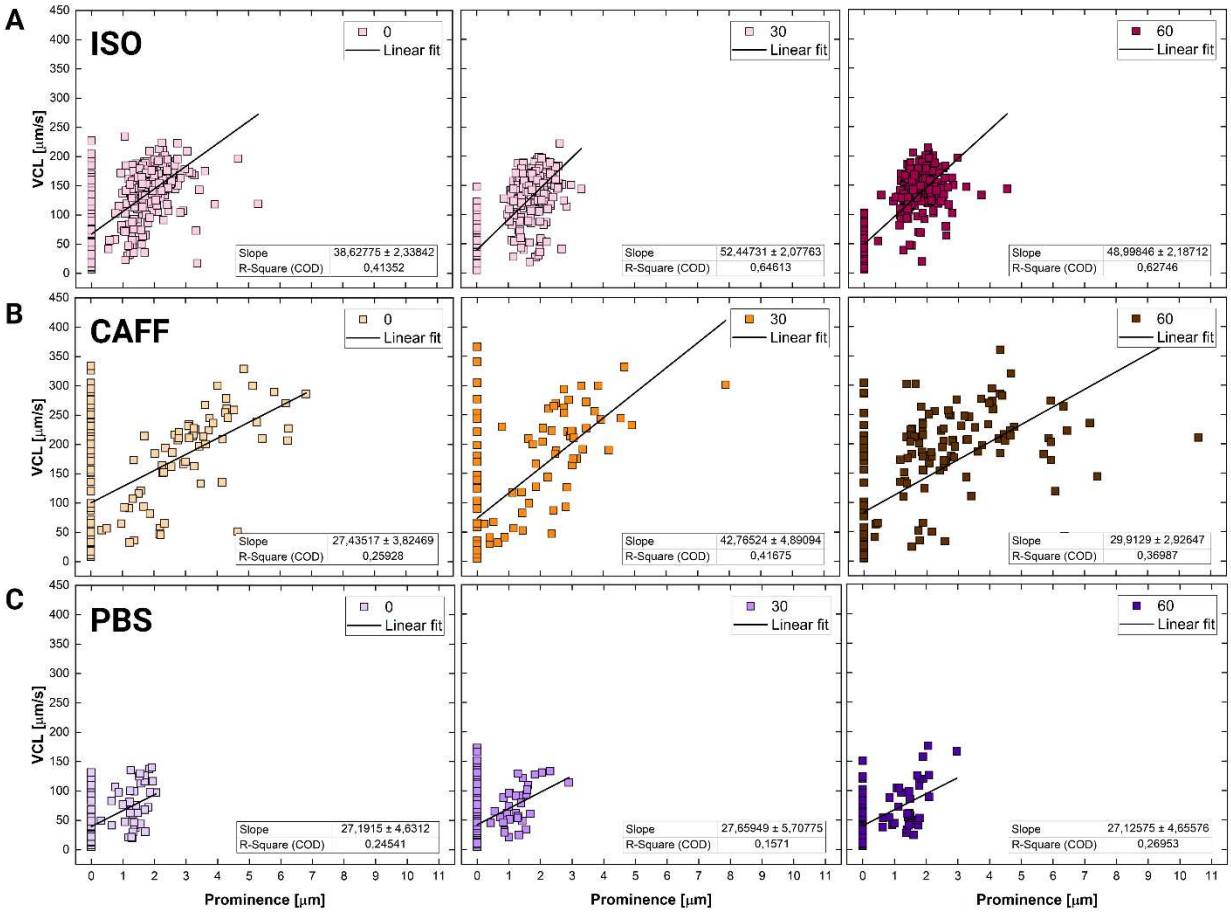

Figure 35: Scatter plots presenting VCL as a function of prominence for (A) ISO, (B) CAFF, (C) PBS. Liner fitting was performed with OriginLab and the resulting slope and R-squared are reported in the plots. Each condition is colour coded, and gradients of the same colour encode for different timesteps.

The prominence of sperm in altered viscosities confirmed the planarization of motion occurring at high viscosities (Figure 36). As a matter of fact, such parameter decreased with increasing viscosity. PEO0.5 show some non-zero prominence values due to the fact that viscosity in this condition has a modest planarization effect with respect to PEO0.9, where all cells show no-rolling behaviour.

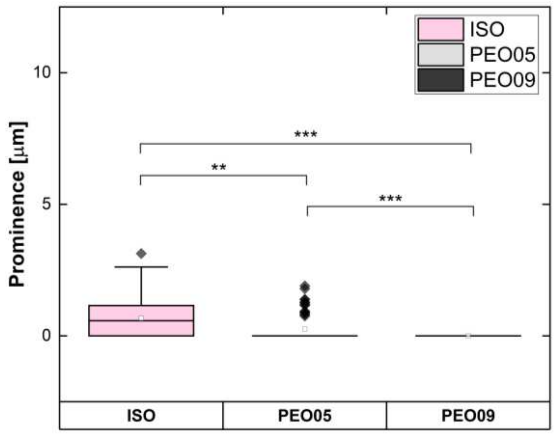

Figure 36: Mean prominence of bovine sperm in three different viscosities, ISO, PEO0.5 and PEO0.9.

Same results can be appreciated from Figure 37, where both VCL and prominence decrease with increasing viscosity. The linear fit for PEO0.9 did not gave results due to the data distribution along 0.

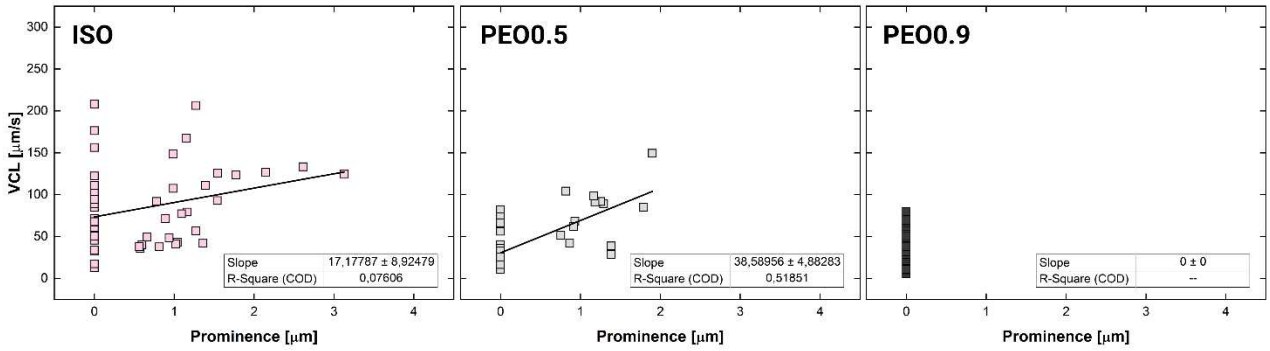

Figure 37: Scatter plots presenting VCL as a function of prominence for experiments with altered viscosity. Liner fitting was performed with Origin and the resulting slope and R-squared are reported in the plots.

3.7.2 Rolling frequency

Sperm swim through a complex 3D rolling motion, and thus alternate different head configurations. Bovine sperm, due to the shape resembling a paddle, oscillate between Face on and Rim on configurations (Figure 38A). The evaluation of sperm head area on swimming sperm allowed the computation of the rolling frequency, which was computed applying the Fourier Fast Transform (FFT) on the area sequence. Since missing points can happen along the trajectory, the null points in the area sequence are interpolated with the spline algorithm in MATLAB. An example of area sequence is reported in Figure 38B. Then, the continuous component is removed from the signal with the "detrend" function. The signal is transformed with the "fft" function, using Hanning as the windowing method, and the peak in the amplitude of the transformed signal is detected (Figure 38C). Detailed code is reported in Appendix A.

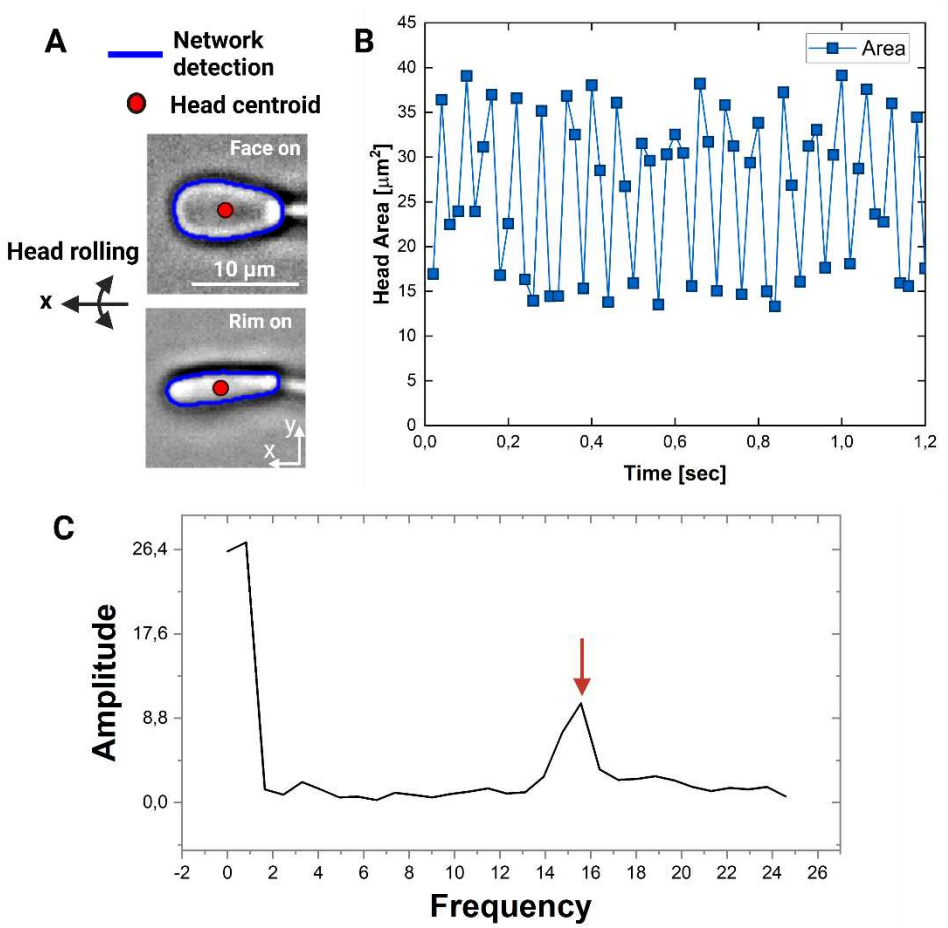

Figure 38: Sperm rolling motion and acquired area sequence. A) The rolling motion sees the alternation of Face on and Rim on configuration. U-net-tiny acquires sperm head area along the whole trajectory, allowing the identification of the head contour (blue) and centroid (red dot). B) Example of sperm head area acquisition in time. C) Amplitude of the FFT transformation on the head area signal. The frequency corresponding to the peak (red arrow) is the rolling frequency.

In hyperosmolar conditions induced with NaCl (HYP1 and HYP2), the rolling frequency mirrors the motility loss seen at 30 and 60 min (Figure 39A). The significant increase seen at 0 min between HYP1 and ISO may be due to an initial rolling alteration that quickly decreases over time. Such decrease over time indicates a loss of progressivity, in accordance with the prominence increase seen in Figure 32. The effect of fructose-induced osmolarity appear less detrimental with respect that induced by NaCl. As a matter of fact, the rolling frequency results higher than ISO, although not significantly except for $T = 0\ min$.

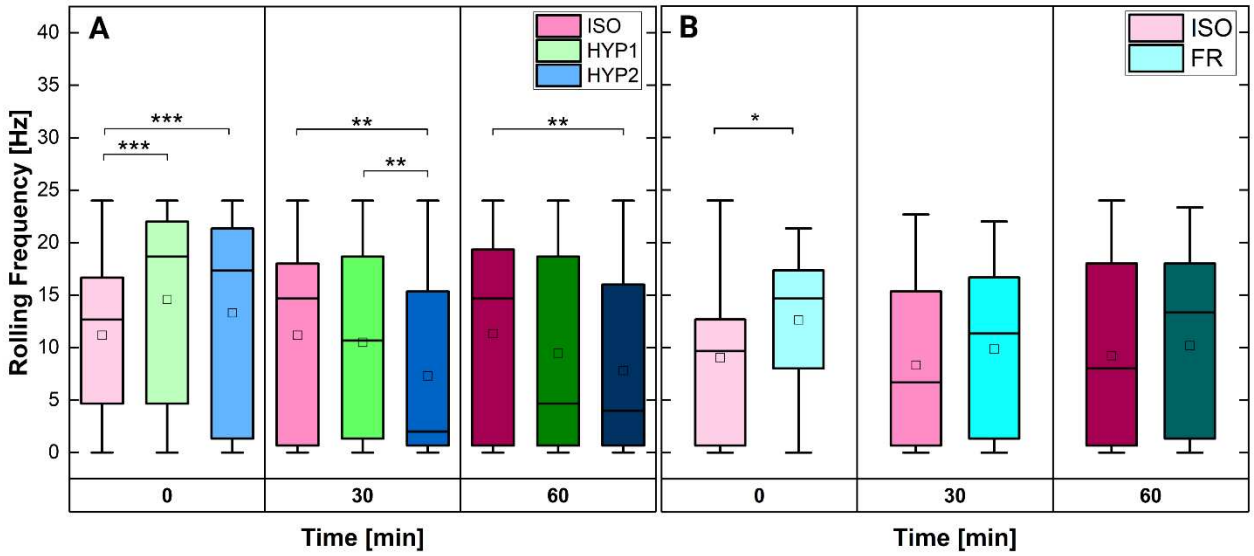

Figure 39: Rolling frequency in hyperosmolar conditions where A) NaCl and B) Fructose were used as osmolytes. (*: p<0.05; **:p<0.005; ***:p<0.001)

Plotting the rolling frequency as a function of VCL, a stronger correlation can be seen between the parameters with respect to prominence. As a matter of fact, R^2 of the linear fit indicates a stronger linear correlation between the parameters (Figure 40A-C), indicating that cells rolling at high frequency will most likely have high VCL and vice versa. The decrease in slope with increase time in HYP2 underlines the decrease of sperm motility already seen from the kinematic parameters. It is interesting to point out that for longer times the scattered points divide in two groups, one with low and the other with high rolling frequency. This suggests the onset of two subpopulations with different grades of motility. This parameter does not correlate with VCL thresholds defined by CASA.

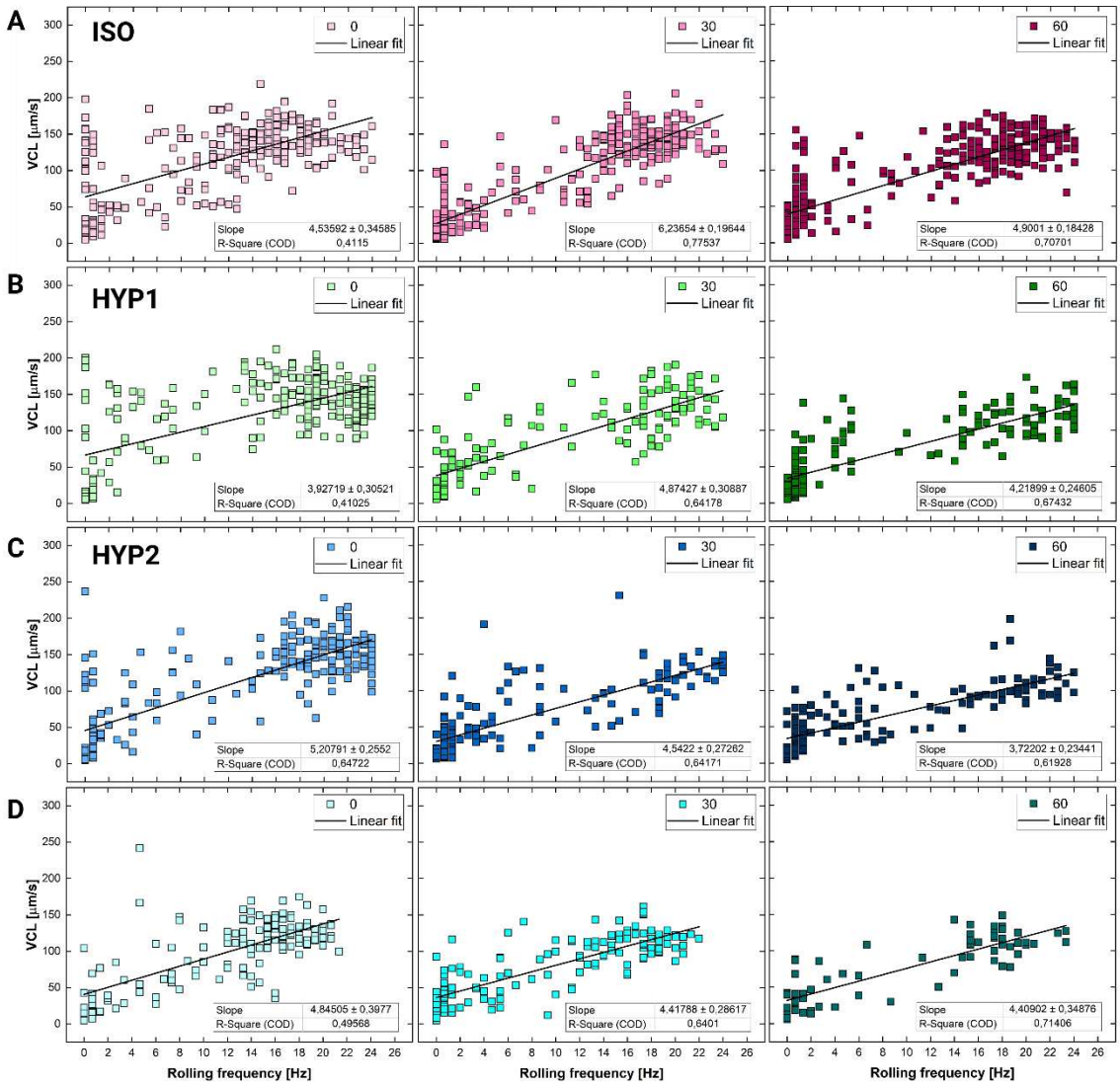

Figure 40: Scatter plots presenting the Rolling frequency for (A) ISO, (B) HYP1, (C) HYP2 and (D) FR. Liner fitting was performed with Origin and the resulting slope and R-squared are reported in the plots. Each condition is colour coded, and gradients of the same colour encode for different timesteps.

The rolling frequency for experiments in caffeine and PBS show a significant decrease with respect to ISO condition (Figure 41A-B). As for CAFF, the decrease in rolling frequency goes along with the increase in the prominence, indicating a change of motion pattern that is less rolling and more oscillating. The lack of calcium in the sperm medium on the other hand induced a decrease in all kinematic parameters including the rolling frequency (Figure 41B).

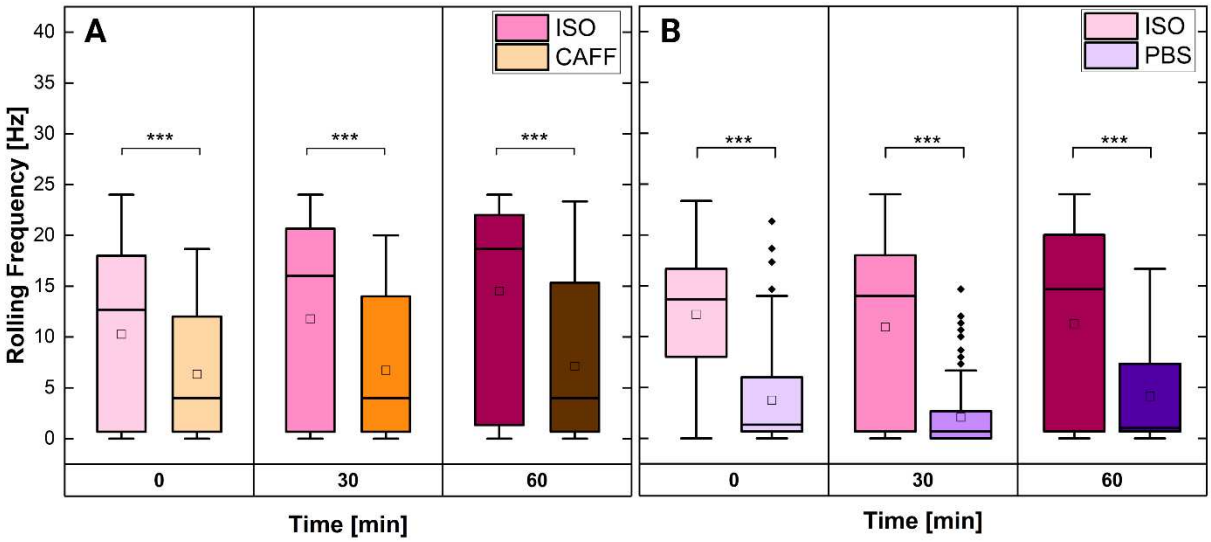

Figure 41: Rolling frequency in conditions that unbalance sperm motility. A) CAFF and B) PBS. (*: p<0.05; **:p<0.005; ***:p<0.001)

Alike the experiments in osmolarity, the linear fit of rolling frequency and VCL gave low values of R^2, indicating a scarce linear correlation among the parameters (Figure 42A-C). As time progresses, similarly as what seen for hyperosmolar conditions, the rolling frequency suggest the presence of subpopulations in the sample. As a matter of fact, especially in ISO and CAFF two groups can be identified at 30 and 60 min, on the lower and higher part of the rolling frequency axis (Figure 42A-B). This is less pronounced in PBS, where the loss of motility seems more homogeneous. The presence of points with lower rolling frequency and very high VCL in CAFF (Figure 42B) at all times is indicative of altered motion patterns, in accordance with other parameters.

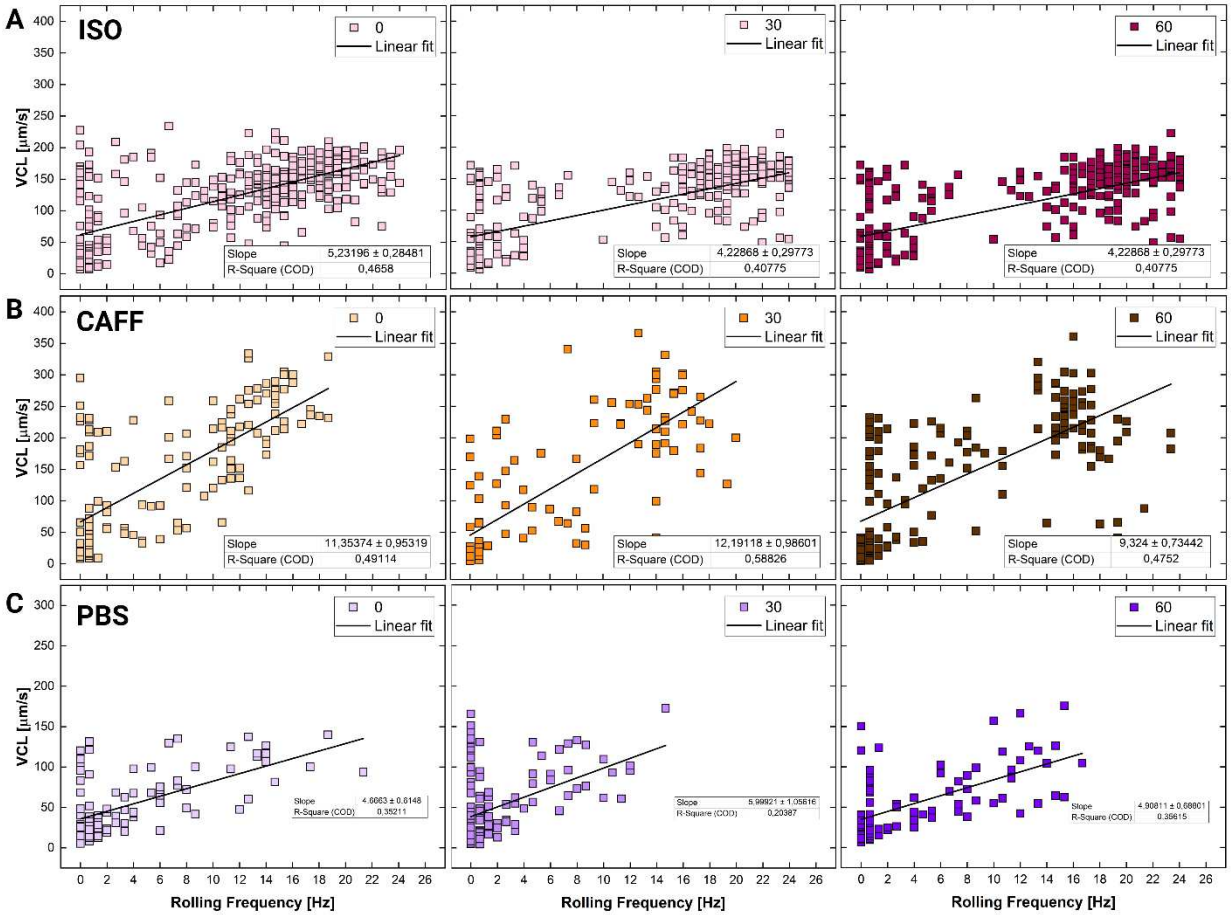

Figure 42: Scatter plots presenting the Rolling frequency for (A) ISO, (B) CAFF, (C) PBS. Liner fitting was performed with Origin and the resulting slope and R-squared are reported in the plots. Each condition is colour coded, and gradients of the same colour encode for different timesteps.

Rolling frequency in PEO conditions is very useful to confirm the planarization phenomenon induced by the increase of viscosity (Figure 43A-B). As a matter of fact, this parameter significantly decreased with respect to ISO and between PEO0.5, while in PEO0.9 rolling frequencies are close or equal to 0 (Figure 43A). A very weak if no linear correlation results from linearly fitting rolling frequency and VCL (Figure 43B).

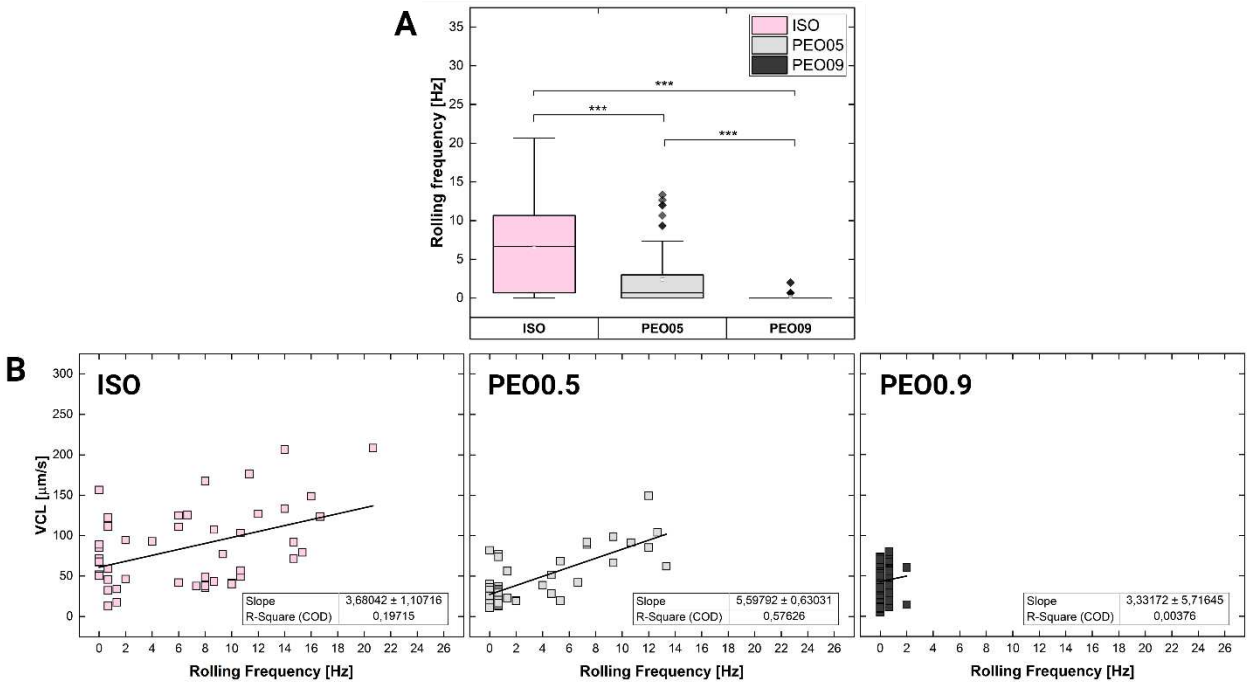

Figure 43: Rolling frequency in conditions of altered viscosity. A) Mean rolling frequency in ISO, PEO0.5 and PEO0.9; B) VCL as a function of the rolling frequency. Liner fitting was performed with Origin and the resulting slope and R-squared are reported in the plots. (*: $p<0.05$; **:$p<0.005$; ***:$p<0.001$)

3.7.3 Radius of circular fit

Circular motion has been found to be correlated to fertilization. As a matter of fact, sperm switch from a straight, symmetrical motion to an asymmetrical one, promoting circular motion in the egg's proximity. This change in behaviour helps them thoroughly search the egg's surroundings, increasing their chances of finding it [181]. The early onset of circular motion is an indicator of poor progressivity. To characterize this altered motion pattern the radius of circular fit (R) was proposed (Figure 44A-B). It is computed starting from the single sperm trajectory, captured for a time long enough to guarantee a good reconstruction (Figure 44A). Then, the MATLAB function "CircleFitByPratt" is applied (Figure 44B) [182]. Such parameter was first introduced in a study conducted by our research group on bovine sperm treated with Cypermethrin (CYP), a pesticide known to induce motility alterations (Figure 44C-E). Circular trajectories were identified only in treated sperm with respect to control (CTRL, Figure 44D-E), and they were classified according to CASA motility classes definition [150]. The computation of R was then implemented to sperm acquired at 40x magnification to correlate this parameter with the morphological analysis as well as with kinematic parameters. An upper limit of 1000 was set for R to characterize very straight tracks.

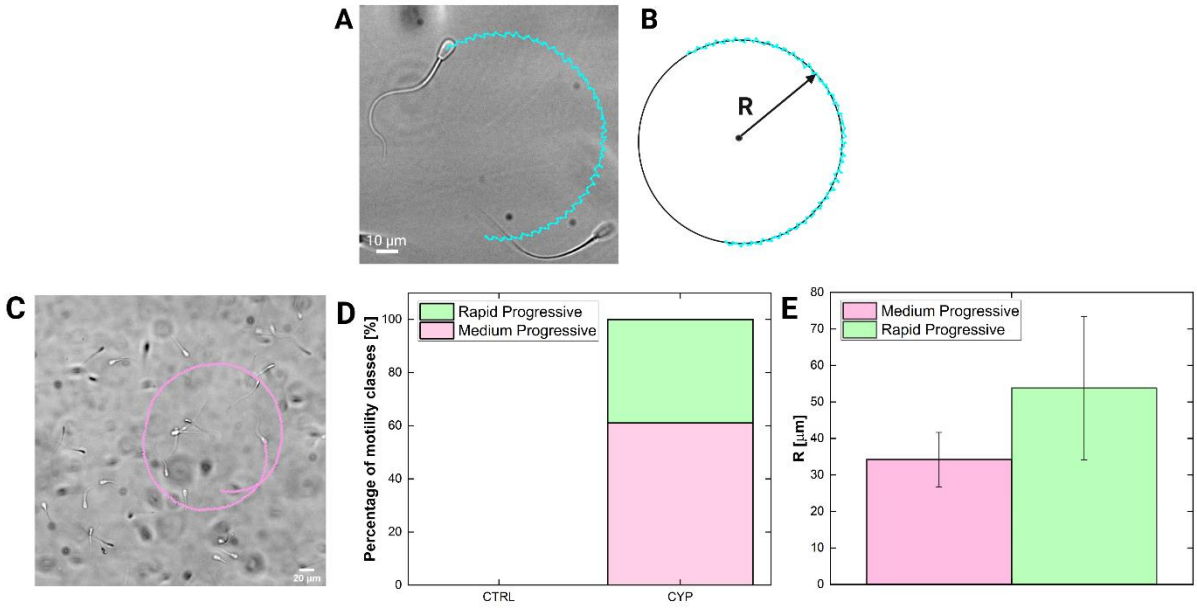

Figure 44: Radius of circular fit. A) Tracking circular sperm at 40x magnification and B) computation of R through a circular fit on the track. C) Circular fit radius computation on 20x-magnified bovine sperm, treated with CYP. The treatment induced circular swimming sperm (D), although the radius of such trajectories was not significantly different among different motility classes (E).

The radius of circular fit reflects the alteration of motility induced by hyperosmolar conditions in the parameters presented before (Figure 45). As a matter of fact, HYP1 confirms its milder effect with respect to HYP2. In fact it is significantly decreases with respect to ISO at 60 min while HYP2 has a stronger effect already at 30min. The milder effect of fructose with respect to HYP2 seems confirmed by the radius as well. The higher R, the straighter the trajectories, thus lower values of R indicate in-situ or circular swimming sperm.

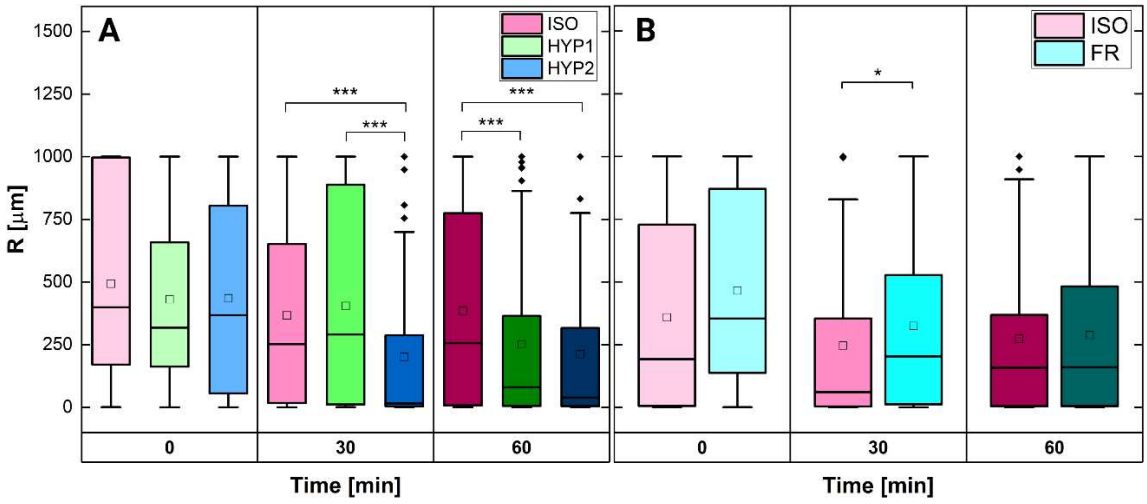

Figure 45: Radius of circular fit (R) in hyperosmolar conditions where A) NaCl and B) Fructose were used as osmolytes. (*: p<0.05; **:p<0.005; ***:p<0.001)

Applying a linear fit on VCL as a function of R, it can be seen that there is very low linear relationship between the two parameters (Figure 46A-D). With respect to ISO, R acquires lower values for all

hyperosmolar treatments at 30 and 60min, highlighting the detrimental effect of osmolarity on sperm motion.

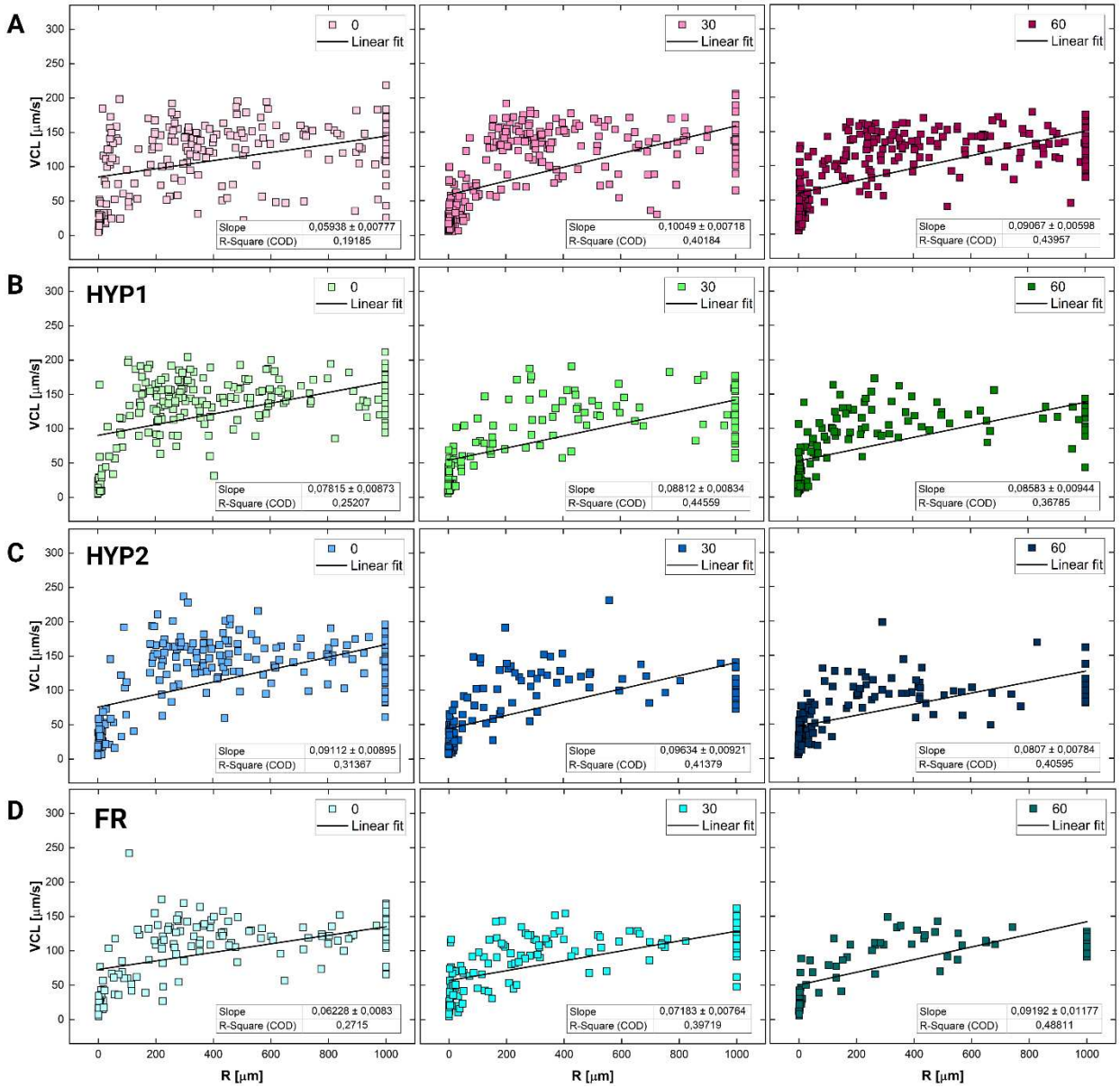

Figure 46: Scatter plots presenting the radius of circular fit R for (A) ISO, (B) HYP1, (C) HYP2 and (D) FR. Liner fitting was performed with Origin and the resulting slope and R-squared are reported in the plots. Each condition is colour coded, and gradients of the same colour encode for different timesteps.

The mean values of R for CAFF and PBS are reported in Figure 47. In both cases, the parameters assume much lower values with respect to ISO at the same timestep indicating that most sperm cells are moving in small circles or in-situ.

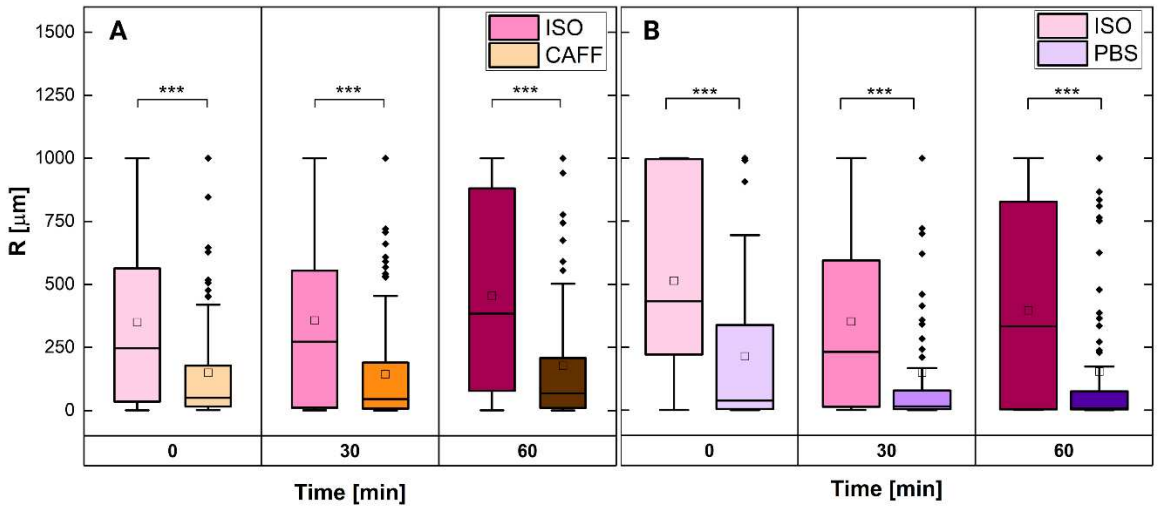

Figure 47: Circular fit radius R in conditions that unbalance sperm motility. A) CAFF and B) PBS. (*: p<0.05; **:p<0.005; ***:p<0.001)

Observing the VCL as a function of R, a non-linear relationship can be seen between the two variables (Figure 48). In CAFF experiments R did not significantly change among timesteps as can be seen from the slope and R^2 of the linear fit. This result, combined with the information regarding the increase of VCL and Prominence, underlines the onset of a hyperactivated non-progressive motion [33]. The low R^2 in these conditions highlights the altered motion of sperm cells, since there a relevant distribution of cells that move in place with high VCL already at 0min with respect to ISO. The distribution of points at very low values of R confirms the presence of non-progressive sperm in both CAFF and PBS (Figure 48B ad C).

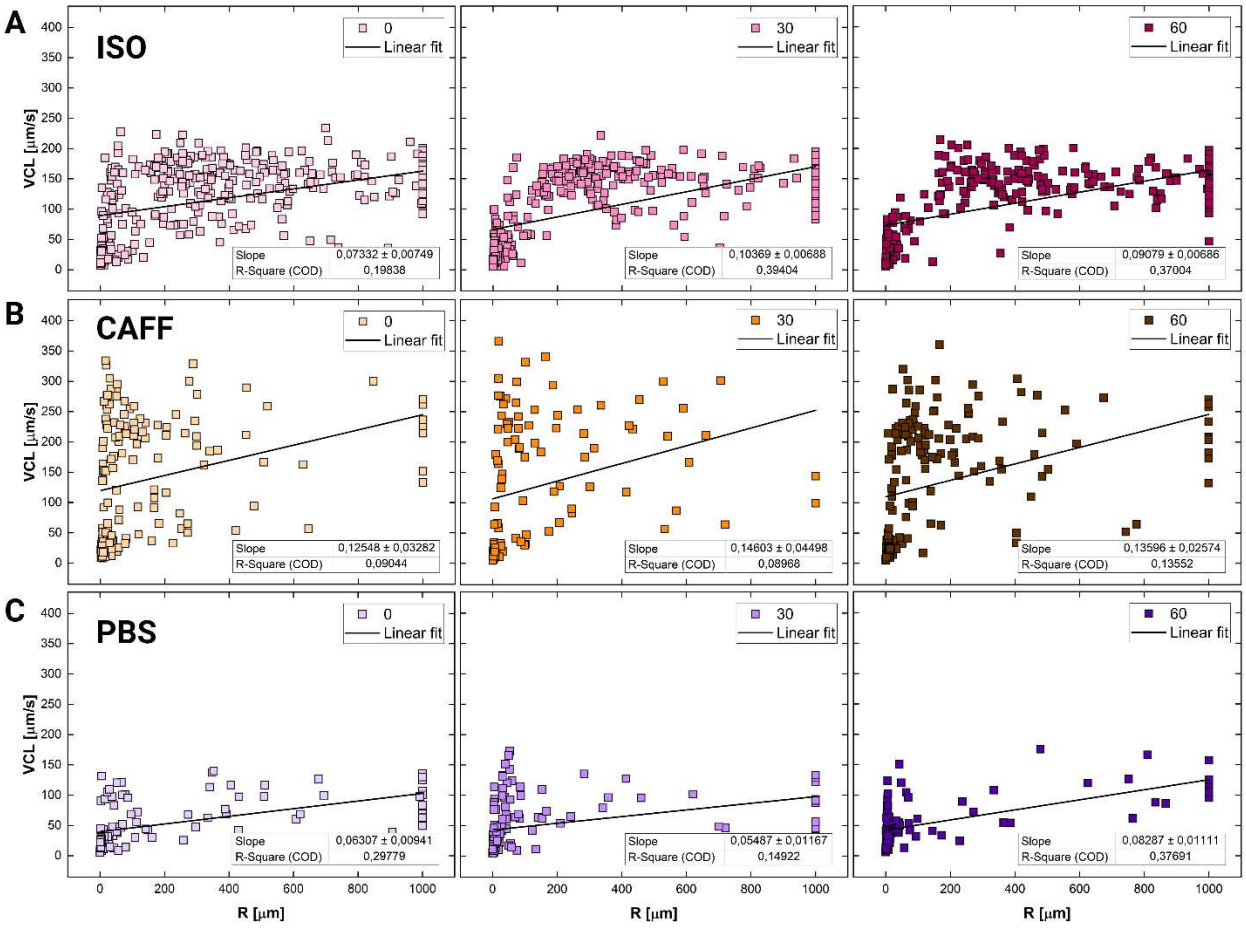

Figure 48: Scatter plots presenting the radius of circular fit R for (A) ISO, (B) CAFF, (C) PBS. Liner fitting was performed with Origin and the resulting slope and R-squared are reported in the plots. Each condition is colour coded, and gradients of the same colour encode for different timesteps.

The radius R in PEO conditions gives an insight into the loss of progressiveness occurring at increased viscosities (Figure 49A-B). As a matter of fact, this parameter significantly decreased with respect to ISO and between PEO0.5 and PEO0.9 as well (Figure 49A). The different effects of viscosity are reflected by the difference in slope and R^2 among the viscous conditions. The cells having high R but low VCL in PEO0.9 are a result of the increased drag force exerted by the viscous fluid, which limits sperm motion. A very weak if no linear correlation results from linearly fitting rolling frequency and VCL (Figure 49B).

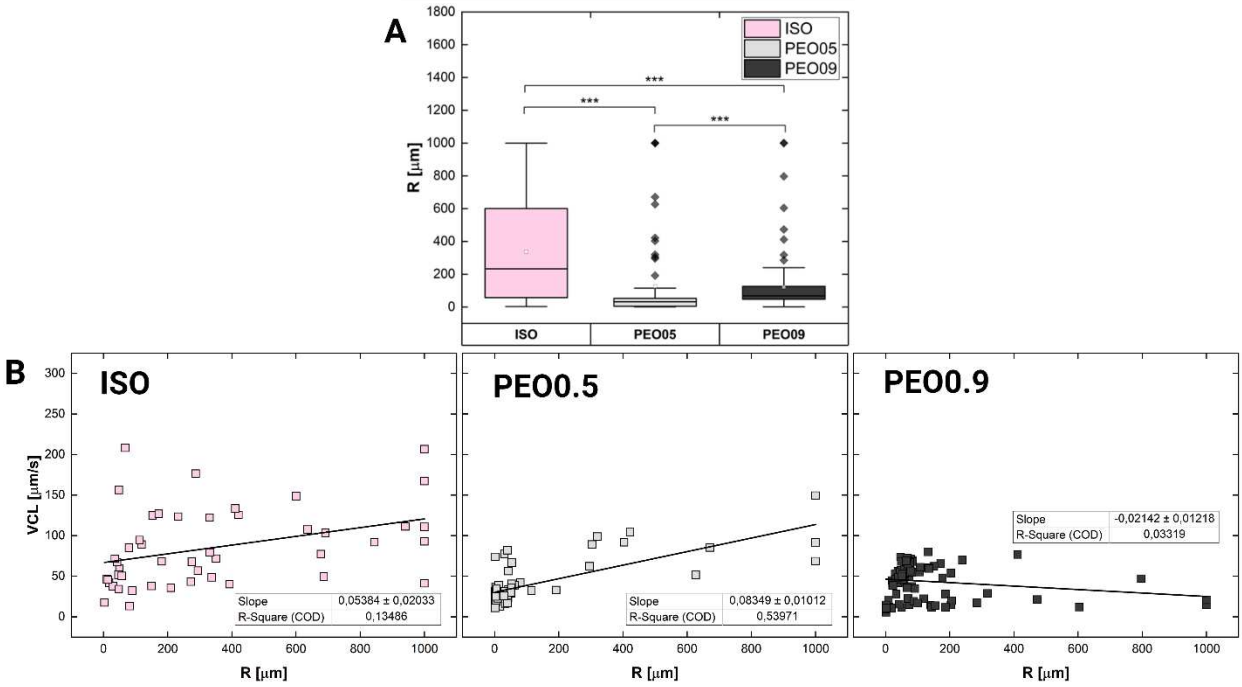

Figure 49: Radius of circular fit R in conditions of altered viscosity. A) Mean rolling frequency in ISO, PEO0.5 and PEO0.9; B) VCL as a function of the rolling frequency. Liner fitting was performed with Origin and the resulting slope and R-squared are reported in the plots. (*: p<0.05; **:p<0.005; ***:p<0.001)

The parameters of the linear fitting applied on the other kinematic parameters, VSL, MAD and FD, are reported in Appendix B.

The proposed parameters aim at enriching the pool of sperm descriptive parameters, gaining more insight into sperm motility patterns. Prominence helps in gaining information on the sperm spatial excursion from Face to Rim, identifying a parameter specifically dependent on the single-cell motion. It reaches high values in sperm treated with caffeine, confirming the onset of hyperactivated motion. Rolling frequency, on the other hand, allows the characterization of planar motion seen in high viscosity conditions, namely PEO0.5 and PEO0.9. Low values of rolling frequency further hint at the loss of motility seen in PBS and CAFF, confirming the expected alteration of motion due to unbalance of the membrane potential [158], [175]. Finally, the radius R gives a description of sperm trajectories straightness, assuming low values in case of in-situ, non-progressive sperm. These were induced particularly by HYP2, CAFF, PBS and PEO. This parameter is particularly useful to characterize sperm swimming in circles, whose presence may be induced by increased viscosity or motility impairment due to treatments such as caffeine. The proposed parameters describe motility patterns that would be otherwise uncaptured by standard CASA parameters. Moreover, none of them showed a strong linear correlation with VCL, underlying the need to go beyond the static thresholds defined for sperm motility.

3.8 Clustering analysis

The data acquired from the conditions at all timepoints discussed above were joined to build a dataset that included the highest possible variability, induced by both treatment and time. To investigate the presence of motility groups in the sample a clustering procedure was performed based on PCA and k-means clustering. Before applying the PCA, the Spearman correlation coefficient was evaluated on

the whole resulting dataset to remove uncorrelated variables that could hinder the effectiveness of the analysis. LIN was excluded a priori since it is a linear combination of VSL and VCL. Based on the correlation coefficient, MAD was removed from the dataset as it was not correlated to any variable as can be seen from Table 14.

Table 14: Spearman correlation coefficient to measure correlation between sperm kinematic and morphometric parameters.

	VSL $[\mu m/s]$	**VCL** $[\mu m/s]$	**ROLLING FREQUENCY** $[Hz]$	**PR** $[\mu m]$	**R** $[\mu m]$	**FD** $[n.d.]$	**MAD** [°]
VSL	1	0,89826	0,73536	0,69835	0,78912	-0,78748	0,05505
VCL	0,89826	1	0,70857	0,71022	0,68566	-0,50409	0,0376
Rolling f	0,73536	0,70857	1	0,71631	0,69768	-0,5011	0,03703
Pr	0,69835	0,71022	0,71631	1	0,66254	-0,37661	0,05464
R	0,78912	0,68566	0,69768	0,66254	1	-0,70041	0,04263
FD	-0,78748	-0,50409	-0,5011	-0,37661	-0,70041	1	-0,07136
MAD	0,05505	0,0376	0,03703	0,05464	0,04263	-0,07136	1

The moderate linear dependency of prominence, rolling frequency and R from VCL were already investigated. Figure 50 reports these variables as a function of VCL, color-coded according to the motility classes definitions reported in Table 10. None of the parameters define distinctive groups based on these definitions. The fractal dimension, however, tends to scale differently with increasing VCL, and the spread of data points shows a sharp rise initially that then starts to taper off at higher values of VCL. FD captures the geometric complexity of sperm movement and can differentiate between sperm that follow irregular, looping paths and those that move more linearly independently of velocity. Moreover, it is known from literature that a $FD < 1.3$ in humans defines the onset of hyperactivated motion [55]. Looking at Figure 50D, FD equal to 1.5 marks a division in the datasets (black dashed line). Presumably sperm with an FD above 1.5 are likely to exhibit non-progressive or slow motility, characterized by very highly irregular movement. Sperm whose FD are below this threshold reflects a simpler, more linear swimming pattern, as can be seen from the presence of medium and rapid sperm in this group. For these reasons, this threshold of FD was chosen as a pre-screening parameter to separate the dataset and perform the clustering separately.

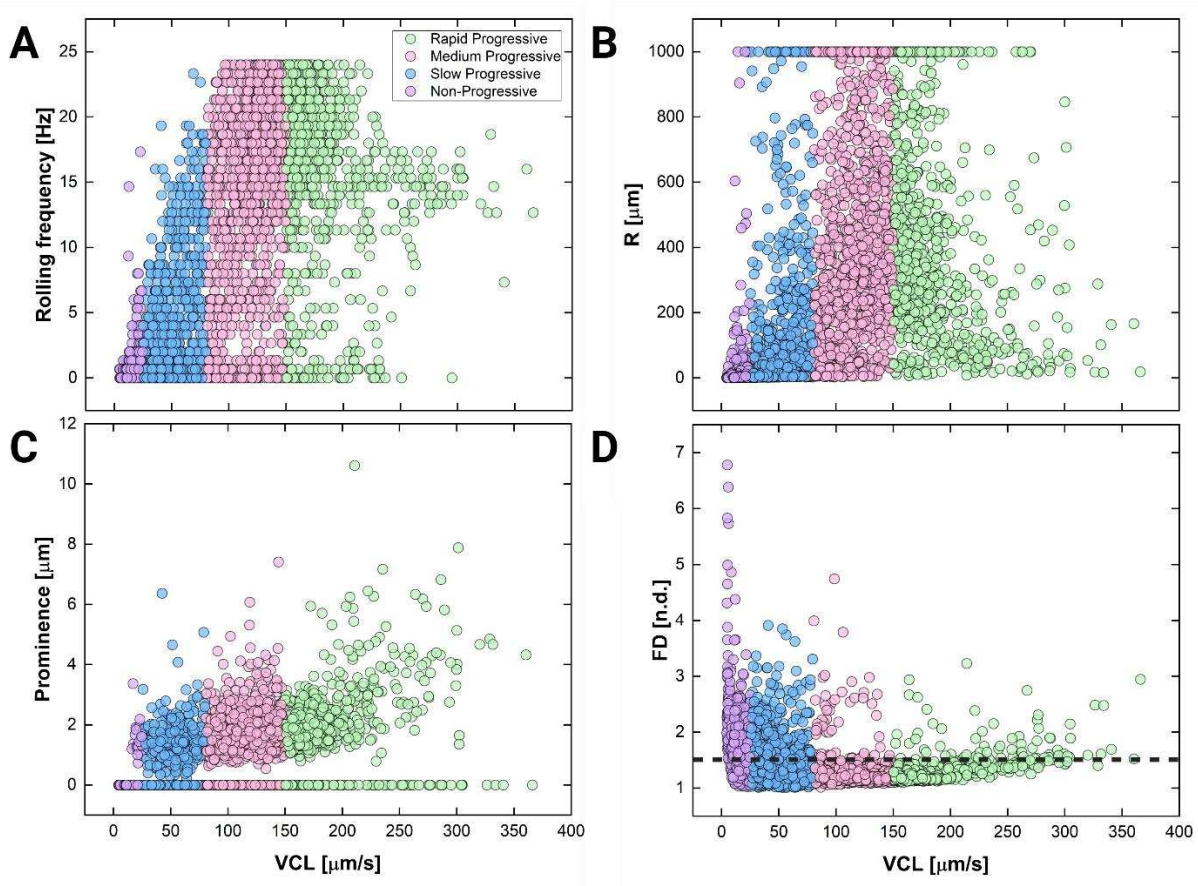

Figure 50: Motility characterization according to CASA motility classes definitions [42]. Non-progressive class: $VCL < 25\ \mu m/s$; Slow progressive:25 $25\ \mu m/s < VCL < 80\ \mu m/s$; Medium progressive: $80\ \mu m/s < VCL < 150\ \mu m/s$; Rapid progressive: $VCL > 150\ \mu m/s$. A) Rolling frequency; B) Radius of circular fit; C) Prominence and D) Fractal dimension as a function of VCL. The dashed line represents the chosen threshold for the dataset split.

The PCA was chosen to reduce the dataset dimensionality, identifying the parameters that capture most variance. The number of principal components was selected in order to guarantee a variance of at least 80%. With regards to sperm having $FD < 1.5$, three PCs were selected. The coefficients for each sperm parameter are reported in Table 15. The first component contributes to 62.7% of the variance, and captures velocity-based parameters such as VSL, VCL and Rolling frequency, even to a lesser degree. PC2, having a contribution of 15.1% of variance, is strongly positively related to the radius R, suggesting that this component captures most of the variance associated with the sperm trajectory curvature. Lastly PC3 explains11.6% of dataset variation and is strongly related with the prominence Pr, indicating that sperm with high PC3 scores have very oscillatory trajectories and low VSL.

Table 15: Coefficients of the kinematic and morphometric parameters in the space of the principal components (FD<1.5).

	COEFFICIENTS OF PC1	**COEFFICIENTS OF PC2**	**COEFFICIENTS OF PC3**
VSL	0,502	-0,136	-0,490
VCL	0,501	-0,325	-0,289
Rolling f	0,461	0,0826	0,0772
Pr	0,411	-0,269	0,816
R	0,340	0,892	0,0588

The loading plot representing the contribution of each parameter in the PC space is reported in Figure 51A. Four subpopulations were defined from these PCs (Figure 51B-D). Rolling frequency and VSL separates well the cluster 1 from the rest, indicating that sperm isolated in this cluster supposedly are very progressive (high VSL and Rolling frequency). Cluster 2 (green) is well described by prominence and VCL, indicating that cells in this cluster have a strongly undulatory motion. Cluster 3 and 4 are characterized by slow and in situ sperm.

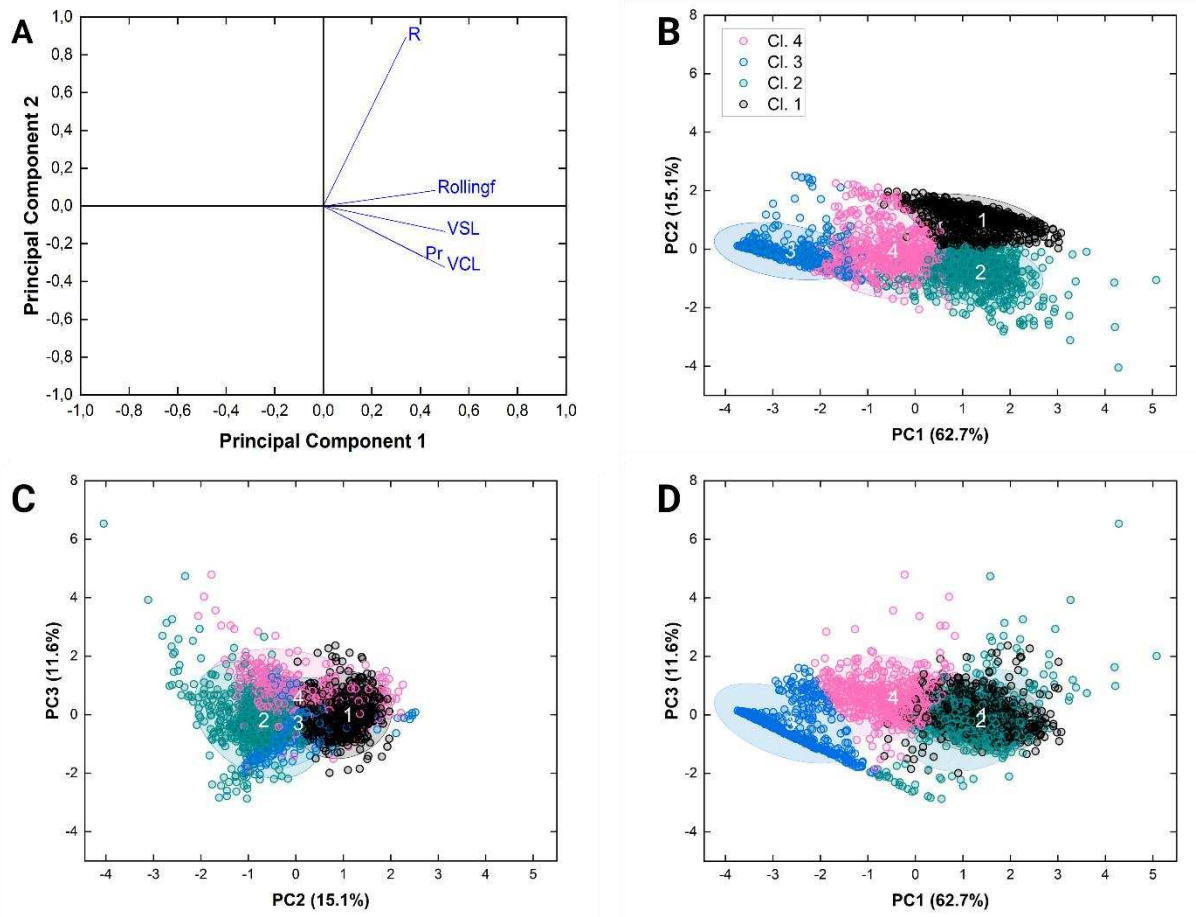

Figure 51: PCA and K-means results for sperm cells having a FD < 1.5. A) Loading plot representing the contribute of each variable to the principal components. B) k-means clustering results in the space of PC1 and PC2. C) k-means clustering results in the space of PC2 and PC3. C) k-means clustering results in the space of PC1 and PC3. The numbers in brackets represent the total variance accounted by the PC.

This is quantitively confirmed by the parameters descripting each cluster, whose mean values and standard deviations are reported in Table 16.

Table 16: Descriptive parameters of the 4 clusters resulting from the PCA for sperm having FD<1.5. Groups sharing the same letter are not significantly different (p>0.05).

CLUSTER	VSL [$\mu m/s$]	VCL [$\mu m/s$]	ROLLING FREQUENCY [Hz]	PR [μm]	R [μm]	MAD [°]
1	70.9 ± 17.3	138.0 ± 27.0	17.0 ± 4.8	1.8 ± 0.6	919.9 ± 136.8	187.2 ± 85.2[c]
2	79.1 ± 14.2	165.6 ± 33.4	17.5 ± 5.7	2.0 ± 0.9	317.3 ± 151.6[b]	185.9 ± 86.5[c]
3	22.3 ± 17.7	49.4 ± 36.0	1.4 ± 2.6	0.2 ± 0.4	81.7 ± 147.9	182.4 ± 62.4[c]
4	39.7 ± 13.7	92.0 ± 25.7	11.4 ± 6.1	1,7 ± 0.7	348.2 ± 253.4[b]	182.7 ± 67.5[c]

MAD did not show significant differences among the clusters. Cluster 1 is the one characterized by very progressive, straight trajectories. This fast and linear class is in accordance with literature findings [144]. Cluster 2 as well is characterized by high VCL, but lower R with respect to cluster 1. The high mean prominence indicates that what would have been incorrectly classified as rapidly progressive is more likely to be a spermatozoon with a less linear movement pattern. Categorizing the cells in each cluster according to CASA motility definitions, it was seen that clusters 1 and 2 are composed of both medium and rapid progressive sperm (Table 17). Non-progressive and slow sperm are found mainly in cluster 3, while in cluster 4 mainly slow and medium sperm are found. The lack of a distinctive distribution of motility classes in the resulting clusters underscores the insufficiency of these parameters to describe sperm motility patterns.

Table 17: Distribution of motility classes according to CASA categorization in the clusters obtained with k-means.

CLUSTERS	NON-PROGRESSIVE	SLOW	MEDIUM	RAPID
1	0	0	68	31
2	0	0	36	64
3	28	57	12	3
4	0	32	67	1

The three proposed parameters are represented in Figure 52 as a function of VCL, and the points are color-coded according to the clustering results. Based on the results, an alternative parametrization of motility classes may be proposed:

- Non progressive/Slow: $0\,Hz < Rolling\ frequency < 5\,Hz$, $0\,\mu m < R < 500\,\mu m$, $0\,\mu m < Pr < 1\,\mu m$
- Medium progressive: $5\,Hz < Rolling\ frequency < 15\,Hz$, $R > 500\,\mu m$, $1\,\mu m < Pr < 2\,\mu m$
- Rapid progressive: $Rolling\ frequency > 15\,Hz$, $R > 500\,\mu m$, $2\,\mu m < Pr < 3\,\mu m$
- Rapid Abnormal: $15\,Hz < Rolling\ frequency < 25\,Hz$, $100\,\mu m < R < 600\,\mu m$, $2\,\mu m < Pr < 5\,\mu m$

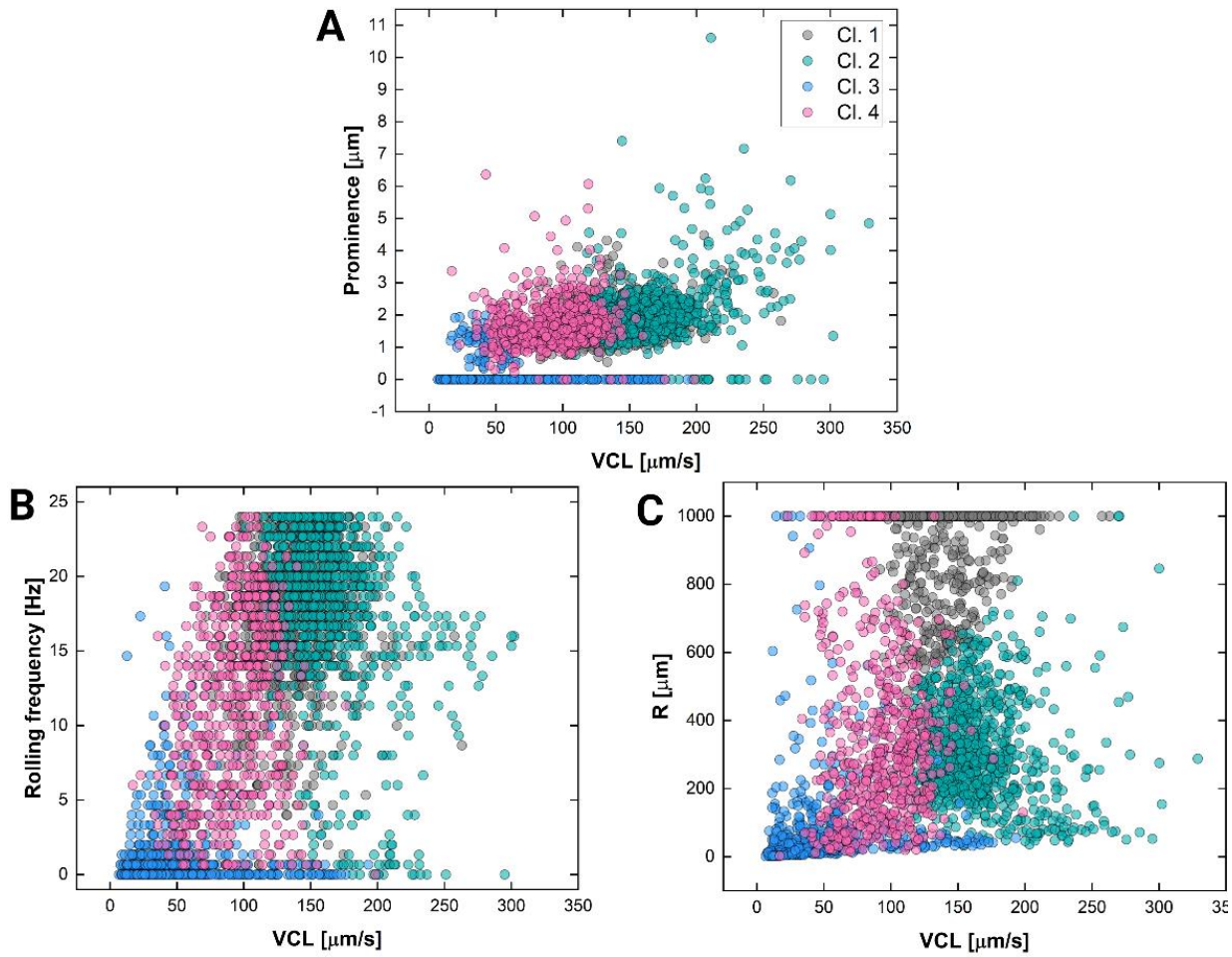

Figure 52: Prominence, radius of circular fit and rolling frequency as a function of VCL, color-coded according to the k-means clustering output.

As for sperm having $FD > 1.5$, the same procedure was followed. This pre-screening however did a first separation between sperm having more regular patterns with ones moving more erratically. Two PCs were enough to capture at least 80% of the dataset variation. PC1 has high scores of VSL, VCL and Rolling frequency, suggesting that PC1 capture the overall speed and oscillatory movement of the sperm combining velocity and rolling behaviour into one primary component (Table 18). As for PC2, the high contribution of R suggests that this component primarily represents curvature or path tightness. Furthermore, the second highest correlation to PC2 is given by the rolling frequency, indicating that sperm with higher curvature tend to have lower rolling frequencies, or slower head rotations.

Table 18: Coefficients of the kinematic and morphometric parameters in the space of the principal components (FD>1.5).

	COEFFICIENTS OF PC1	**COEFFICIENTS OF PC2**
VSL	0.511	-0.043
VCL	0.489	-0.225
Rolling f	0.456	-0.406
Pr	0.431	-0.157
R	0.324	0.934

The loading plot representing the parameters distribution in the PC space is reported in Figure 53A. K-means algorithm was applied on the reduced dataset to investigate the presence of 4 clusters,

reported in Figure 53B. Cluster 1 (blue) is identified by high R values, indicating potentially straighter paths but with low progressivity. Cluster 2 (green) has negative components on PC2, thus are characterize by high curvature radii but higher VSL, VCL and Rolling frequency. Cluster 3 (purple) and 4 (cyan) are composed of highly non-progressive sperm.

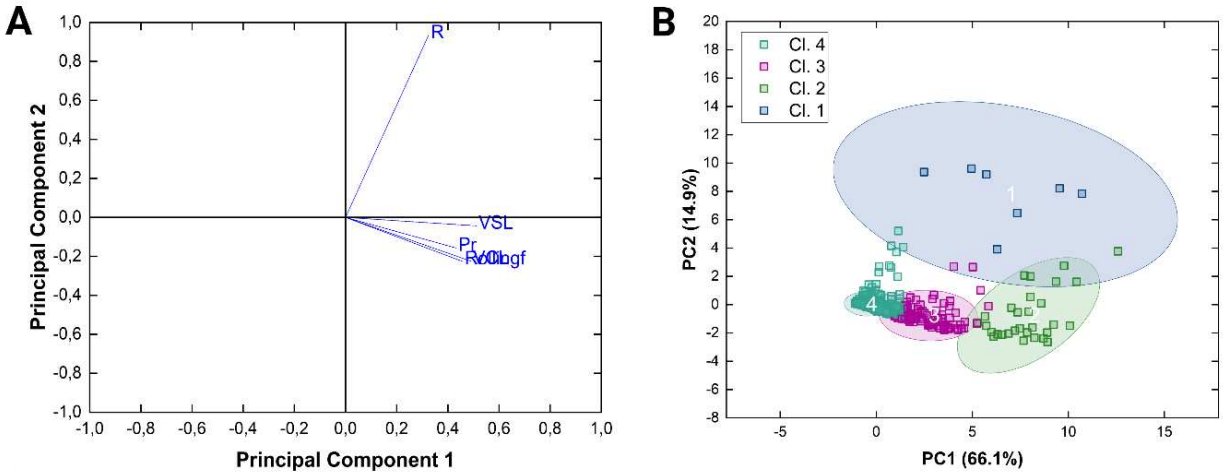

Figure 53: PCA and K-means results for sperm cells having a FD > 1.5. A) Loading plot representing the contribute of each variable to the principal components. B) k-means clustering results in the space of PC1 and PC2.

The cluster characterization in terms of descriptive parameters is reported in Table 19.

Table 19: Descriptive parameters of the 4 clusters resulting from the PCA for sperm having FD<1.5. Statistical analysis was performed between clusters for each parameter. Groups sharing the same letter are not significantly different (p>0.05).

CLUSTER	VSL $[\mu m/s]$	VCL $[\mu m/s]$	ROLLING FREQUENCY $[Hz]$	PR $[\mu m]$	R $[\mu m]$	MAD [°]
1	32.0 ± 17.3 [a]	156.0 ± 27.0 [b]	11.0 ± 4.8 [c]	1.2 ± 0.6 [e]	892.3 ± 136.8	177.8 ± 85.2 [f]
2	54.0 ± 14.2	268.9 ± 33.4	14.0 ± 5.7 [c,d]	4.2 ± 0.9	182.0 ± 151.6	177.6 ± 86.5 [f]
3	18.9 ± 17.7 [a]	156.7 ± 36.0 [b]	9.4 ± 2.6 [d]	0.9 ± 0.4 [e]	45.6 ± 147.9	182.6 ± 62.4 [f]
4	3.5 ± 13.7	29.0 ± 25.7	1.0 ± 6.1	0.1 ± 07	11.2 ± 253.4	178.0 ± 67.5 [f]

The low values of all parameters characterizing cluster 3 and 4 identify these clusters as composed of non-progressive sperm. Cluster 1 is characterized by very straight trajectories with respect to the other clusters, but low VSL and rolling frequency suggest that these cells could be outliers. Cluster 2 is characterized by very high VCL and ALH and low R, indicating highly oscillatory motion patterns. Moreover, this cluster is composed of 31 cells, all of them coming from the conditions treated with caffeine (Table 20). This supports the literature findings addressing caffeine as an inductor of hyperactivated motility, and highlights the ability of our approach to identify SPs in the sample. Cluster 3 is composed by in situ sperm, potentially as can be seen by the high VCL but very low PR and VSL. Cluster 4 is characterized by non-progressive sperm.

Table 20: Characterization of each cluster according to the different treatments.

K-MEANS	CAFF	ISO	HYP1	HYP2	PEO05	PEO09	FR	PBS
Cl. 1	43	29	0	14	0	0	14	0
Cl. 2	100	0	0	0	0	0	0	0
Cl. 3	52	16	9	9	0	0	9	6
Cl. 4	8	33	12	18	2	0	8	18

4. Human sperm analysis: preliminary results

Human sperm was donated by the University of Naples Federico II, and was prepared as reported in 2.1.2 of Materials and methods.
The human sperm head is more rounded with respect to bovine sperm cells, thus Face and Rim configuration are more resembling one another. Nonetheless, the cascade of networks trained on bovine cells was able to track and segment human sperm cells, allowing to retrieve preliminary results on the tracking (Figure 54B).

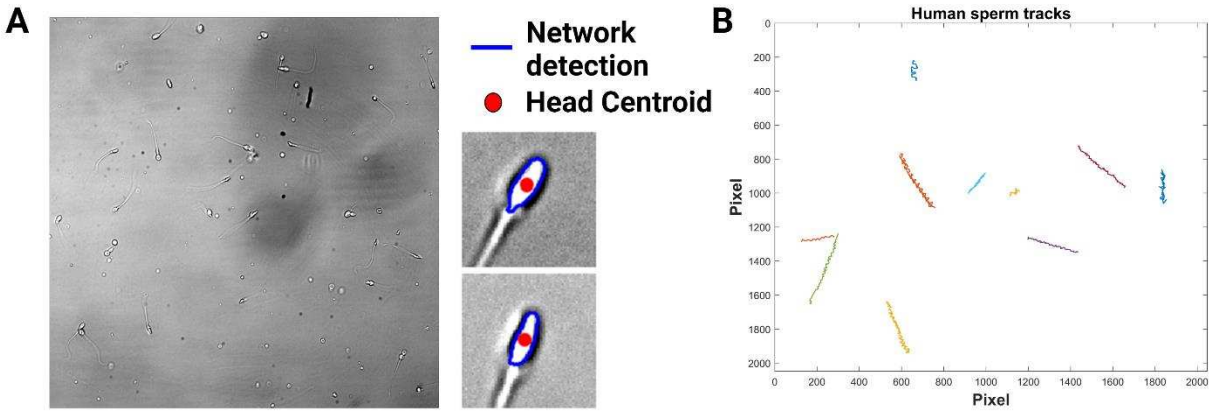

Figure 54: Human sperm analysis with neural networks. A) Acquisition and example of segmentation. The blue line represents the head contour retrieved through the U-net-tiny segmentation, and the red point is the resulting centroid. B) Preliminary tracking outcome on human sperm.

Due to the different morphology, the proposed approach requires fine tuning of the neural networks to achieve precise tracking, ensuring a good sperm characterization. Furthermore, one of the main challenges in human semen analysis is the high intrinsic variability of the sample. For this reason, the dataset of both detection and segmentation networks should account for the maximum possible variability of human sperm configurations. However, these preliminary results prove that the approach can be optimised for human semen analysis. The investigated label-free single-cell kinematic and morphological characterization could help gaining more insight into sperm analysis, enriching current diagnostic approaches.

Conclusion

Unravelling cell complexity is still a major challenge in biomedical research. Single-cell analysis provides detailed insights that bulk analysis cannot, revealing unique properties within individual cells that contribute significantly to diagnostic precision. Single-cell tracking in particular holds great diagnostic potential as it allows the precise monitoring of cell movement and morphology, generating detailed data on kinematics and structural features. This process is however computationally challenging due to sample numerosity and dataset dimensions. In this context the use of staining techniques, despite enhancing cell visualization, could potentially alter results and negatively impact sample viability. For this reasons, computer-assisted tracking techniques are widely used to implement label-free approaches for high throughput analysis.

Single-cell tracking is particularly relevant in infertility diagnostics, where motility is a key indicator of sperm quality. Despite male infertility being a significant clinical issue widely, current strategies lack a comprehensive approach to sperm characterization. Although sperm morphology and kinematics are interconnected phenomena, standard semen analysis evaluates them separately, preventing correlation between the two. Commercially available alternatives called CASA systems characterize sperm motility computing a set of kinematic parameters, although some of them lack standardization. This limits comparability of results across different instruments. Furthermore, these methods are not accessible and do not allow semen analysis in non-standard settings. For these reasons, various proposals, both open access and proprietary, have emerged in literature. Algorithms for sperm tracking across different experimental contexts have been developed, allowing the characterization of sperm motility in different conditions such as chemotaxis or present of a fluid flow. The complexity of sperm motion poses a great challenge, and despite the effort made in the development of complex tracking algorithms, image processing techniques alone do not provide complete semen characterization at the single-cell level.

With the increasing use of Artificial Intelligence in the medical field, several applications arise that incorporate machine learning and deep learning in semen analysis. These systems can efficiently handle large datasets, allowing the identification of trends that go beyond the capabilities of traditional analyses. Computer vision tasks in particular are taking a hold in this field. As a matter of fact, convolutional neural networks can transform semen analysis by delivering consistent, efficient, and highly accurate results. Tasks like detection and segmentation are fundamental to face the challenges related to sperm motility and morphology computation.

This thesis proposes a single-cell tracking approach for semen analysis based on neural networks, using bovine sperm as a model for human cells. It utilizes two cascaded neural networks: the first detects sperm heads within the observation field, and the second performs segmentation. The method is based on label-free acquisition in brightfield using an inverted microscope equipped with a 40x objective lens. The simplicity of this setup ensures reproducibility. The first neural network implemented is a lightweight version of the YOLO4 network. Thanks to the one-stage approach, the network efficiently outputs object positions and classes. The achieved results in terms of AP^{50} for Face and Rim detection are 93% and 91% respectively, comparable with other literature findings despite the lower magnification (Table 1). Once detected, the bounding boxes are then tracked using a self-written MATLAB algorithm, optimized first for a 20x objective and later applied to 40x. Each sperm is segmented throughout its trajectory using a lightweight version of the U-net neural network. The Dice scores for the best performing networks were 94.3% for Face and 89.2% for Rim segmentation. Despite the challenges posed by the segmentation of moving cells, our approach has good performances with respect to other literature findings (Table 1 and 2). Both kinematic and morphological parameters are computed from swimming cells, without the use of staining that can alter the analysis outcome.

Moreover, three novel parameters were proposed, namely the Rolling frequency, Radius of curvature fit (R) and Prominence (Pr). These parameters are calculated individually for each sperm cell, without making assumptions about the type of movement. In this way issues related to CASA parameters, particularly those tied to the computation of the average path, are resolved. To demonstrate the strength of the approach and the efficacy of the enriched dataset of descriptive parameters, bovine sperm were exposed to different conditions known in literature to induce motility alterations. The exposure of sperm to three conditions of hyperosmolarity had a strong impact on sperm motility, leading to a decrease of all descriptive parameters. Low values of rolling frequency and radius suggest the onset of circular-in situ sperm at longer times. The addition of caffeine to the buffer medium instead induce irregular motility pattern characterized by high VCL and prominence, which are typical of hyperactivated motion. The absence of calcium, reproduced by diluting sperm in PBS, shows a decrease in motility that underlines the importance of calcium in the process of flagellar beating and motility regulation. Finally, sperm was analysed in a buffer enriched with PEO, a highly biocompatible polymer, determining a viscosity in the range of cervical mucus. Viscosity is in fact one of the biophysical barriers of the FRT, devoted to favouring the passage of only motile cells. It is known from literature that higher viscosity induces planarization of sperm 3D motion. The low rolling frequency and radius of circular fit R captured this behaviour, since both parameters acquired very low values in these conditions.

The resulting data was put together to build a wide, heterogeneous dataset of sperm descriptive parameters in different conditions, totalling 3900 cells. CASA analysis usually derives information on motility classes in sperm samples based on fixed thresholds of VSL and VCL, according to the specie. This approach is limited since it does not account for the variability of sperm motility patterns, too complex to be accounted for using one parameter. The fractal dimension was used to perform a first distinction between sperm based on trajectory regularity. A threshold of 1.5 was set to divide the dataset. To gain more insight into sperm dynamics, an unsupervised machine learning approach was proposed based on principal component analysis (PCA) and k-means algorithm. Both datasets were clustered in four groups. Sperm having a FD>1.5 were mainly outliers and non-progressive sperm. Cluster 2 in particular emerged as composed of very distinctive motion patterns, characterized by very high VCL and low PR, and low VSL and R. This indicates in-situ cells characterized by a strongly oscillatory motion, typical of non-progressive hyperactivation. The resulting clusters for sperm having FD<1.5 were populated for the major part by slow, medium and rapid sperm according to the VCL classification. Cluster 1 particular was composed on 68% of medium and 31% of rapid sperm cells, while Cluster 2 of 36% of medium and 64% of rapid progressive sperm. However, high prominence and rolling frequency of the latter suggest an altered, highly oscillating motion with respect to the progressivity classification according to VCL, providing more information on the sample composition. These results are in accordance with literature findings, which identify the presence of fast and linear and non-linear clusters in bull sperm , , [139], [146], [148] (Table 3).

On the basis of these results, a new parametrization for slow, medium, rapid and abnormal sperm was proposed, correlating both motility and head morphology. The provided motility characterization is therefore more effective than standard procedures in capturing sperm motion alterations.

The codes provided in Appendix A allow the adaptation of this approach to other cells as well, although training parameters such as image input size, batch size and number of epochs should be tuned according to cells dimensions and dataset features. The tracking algorithm can be adapted as well by optimizing the settings parameters according to the acquisition modalities. As future perspective, the dataset will be shared as public in a FAIR mode.

The proposed tracking workflow with neural networks was tested on human sperm cells. The preliminary results confirm the feasibility of the approach on human semen, highlighting its potential as enrichment to current diagnostic approaches.

Appendix A

MATLAB code for YOLOv4 training and testing

Definition of training and validation datasets

```
% gTruth2 is a table containing the ground truth training data in a table form, where
% the first column contains the image paths (imageFilename), the second contains the
% bounding boxes coordinates of Face detections (Face) and the third the Rim bounding
boxes coordinates (Rim).
% tempdir should be the folder path where it is desired to save the checkpoints along
the training process.

%Split in training and validation
indices = crossvalind('Kfold',size(gTruth2.imageFilename,1),5); %Generate validation
indices as 20% of the training set
validation_i = (indices == 1);
t4 = table([gTruth2.imageFilename(validation_i)], [gTruth2.Face(validation_i)],
[gTruth2.Rim(validation_i)], 'VariableNames', {'imageFilename'; 'Face'; 'Rim'}); %the
validation data are organized
%in a table with the same structure as gTruth2
t5 = table(t4.Face, t4.Rim, 'VariableNames', {'Face'; 'Rim'}) %Validation Labels
blds1 = boxLabelDatastore(t5); %Create a datastore containing the bounding box labels
imds1 = imageDatastore(t4.imageFilename); %images
validation = combine(imds1, blds1); %Validation dataset, defined as a combination of
image datastore and label datastore

gTruth2(validation_i,:) = []; %delete validation images from the training dataset
t3 = table([gTruth2.Face], [gTruth2.Rim], 'VariableNames', {'Face'; 'Rim'}); %Labels
blds = boxLabelDatastore(t3);
imds = imageDatastore(gTruth2.imageFilename);
start_dataset = combine(imds, blds); %Training dataset
```

Data Augmentation

```
%Agumentation techniques to improve the training set dimensions%
reflected_data = transform(start_dataset, @reflection); %Flips the input image with 50%
probability in each dimension
translated_data = transform(start_dataset, @translation); %Random translation of 1024 -
1024 pixel
rot1_basedata = transform(start_dataset,@rotation2); %Random 25 325 gradi
rot1_horflip = transform(reflected_data,@rotation2);
rot1_verflip = transform(translated_data,@rotation2);

dataset_base_1 = combine(start_dataset, reflected_data, translated_data, rot1_basedata,
rot1_horflip, rot1_verflip, 'ReadOrder','sequential'); %combination the the augmented
datasets
Scaled = transform(dataset_base_1, @Scaling_Unet2); %Random scaling
final_trainingdata_1 = combine(dataset_base_1, Scaled, 'ReadOrder','sequential');
Intensity_transf = transform(final_trainingdata_1, @Intensity_transformation); %Random
shift and scale of intensity
final_trainingdata_2 = combine(start_dataset, reflected_data, translated_data,
rot1_basedata, rot1_horflip, rot1_verflip, Scaled, Intensity_transf,
'ReadOrder','sequential'); %combination
%of the augmented datasets to create the final training dataset
```

Preprocessing the training set

```
%Settings of the network
```

```
inputSize = [640 640 3]; %Image input size
numClasses = 2; %the classes to detect are Face and Rim
preprocessedTrainingData = transform(final_trainingdata_2,
@(data)preprocessData(data,inputSize)); %Preprocessing the images to the final format
for YOLO4 training
preprocessedTrainingData.shuffle() %Shuffling the training dataset
```

Anchors definition

```
%Yolo4 requires the definition of anchor boxes for object detection. They
%must closely represent the sizes and aspect ratios of the objects in the
%training data. The number for the training is established evaluating the
%IoU between the anchor boxes and the trainining dataset bounding boxes
%%Estimation of anchors for Yolo heads
allBoxes = vertcat(preprocessedTrainingData.Face{:}, preprocessedTrainingData.Rim{:});
labels = cell2mat(allBoxes); %Labels of the preprocessed training dataset
blds_1 = boxLabelDatastore(table(mat2cell(labels, size(labels,1))));
maxNumAnchors = 50; %Max number of anchor boxes to investigate.
meanIoU = zeros([maxNumAnchors,1]);
anchorBoxes = cell(maxNumAnchors, 1);
for k = 1:maxNumAnchors
    % Estimate anchors and mean IoU.
    [anchorBoxes{k},meanIoU(k)] = estimateAnchorBoxes(blds_1,k); %Estimates the
Intersection over Union of the training data bounding boxes and the specified number of
anchor boxes
end
numAnchors = 18; %this number is chosen based on the best value of IoU estimated by the
previous function.
anchorBoxes_ = anchorBoxes{numAnchors, 1}  %Set of anchor boxes for the training

%%Assingment of the anchor boxes to the YOLO4 detection heads. In this
%%application, 2 detection heads are chosen. They should be organized in a
%%way that the first head works with bigger bounding boxes, which outputs
%%the lower resolution feature maps. The smaller ones instead are assigned
%%to the second one, which focuses on higher resolution feature maps.
area = anchorBoxes_(:,1).*anchorBoxes_(:,2); %Area
[~,idx] = sort(area,"descend"); %sorting from bigger to smaller
sortedAnchors = anchorBoxes_(idx,:);
anchorBoxes = {sortedAnchors(1:9,:)
    sortedAnchors(10:18,:)}; %Anchores for the two detection heads: bigger to the first
head, smaller to the second
```

Building the network and training options

```
%%YOLOv4 Tiny definition%%
className =  {'Face'; 'Rim'};
detector = yolov4ObjectDetector("tiny-yolov4-
coco",className,anchorBoxes,InputSize=inputSize); %This function builds the YOLO
network according to the number of classes, input size
%and anchore boxes.

%division of the dataset in training and validation
TRAIN = preprocessedTrainingData;
VALIDATION = transform(validation, @(data)preprocessData(data,inputSize));

options = trainingOptions("adam",... %Settings for the training process
    SquaredGradientDecayFactor=0.999,... %Default
    InitialLearnRate=0.001,... %Learning rate
    MiniBatchSize=16,... %Batch size
    Shuffle='every-epoch',...
    MaxEpochs=20,...
```

```
    BatchNormalizationStatistics="moving",...
    DispatchInBackground=true,...
    ResetInputNormalization=false,...
    VerboseFrequency=300,...
    ValidationFrequency=1200, ...
    CheckpointPath=tempdir, ...
    CheckpointFrequency=1,...
    ValidationData=VALIDATION,...
    OutputNetwork="best-validation-loss",...
    L2Regularization=0.0005);

[detector,info] = trainYOLOv4ObjectDetector(TRAIN,detector,options); %Training
```

Testing

```
%%Testing%%
bb_test = boxLabelDatastore(table(test.Face, test.Rim, 'VariableNames', {'Face',
'Rim'})); %The test dataset has the same structure as gTruth2
imds_test = imageDatastore(test.imageFilename);
com_test = combine(imds_test, bb_test);
preprocessedTestData = transform(com_test,@(data)preprocessData(data, inputSize));
[detectionResults] = detect(detector, preprocessedTestData);
[ap,recall,precision] = evaluateDetectionPrecision(detectionResults,
preprocessedTestData);
X = evaluateObjectDetection(detectionResults, preprocessedTestData, .50) %Test metrics
for IoU = 0.5
```

Auxiliary functions for the data augmentation process

```
function B = reflection(A)
% Apply random horizontal and/or vertical flipping %
%A is a cell array containing the image, the corresponding bounding boxes
%and the labels
I = A{1};

tform = randomAffine2d(XReflection=true,YReflection=true); %Definition of the affine
transformation
outputView = affineOutputView(size(I),tform);
B{1} = imwarp(I,tform,OutputView=outputView);

% Apply same transform to the bounding boxes.
[B{2}, indices] = bboxwarp(A{2},tform,outputView,'OverlapThreshold',0.25);
B{3} = A{3}(indices); %Store the labels of the transformed bounding boxes

% Return original data only when all boxes are removed by warping.
if isempty(indices)
    B = A;
end
end

function B = translation(A)
% Apply random translation along the x and Y axis%
I = A{1};
tform = randomAffine2d(XTranslation=[-1024 1024],YTranslation=[-1024 1024]); %distance
selected randomly from the range.
outputView = affineOutputView(size(I),tform);
B{1} = imwarp(I,tform,OutputView=outputView, FillValues=mean(I(:)));

% % Apply same transform to boxes.
[B{2}, indices] = bboxwarp(A{2},tform,outputView,'OverlapThreshold',0.25);
```

```
B{3} = A{3}(indices); %Store the labels of the transformed bounding boxes

% Return original data only when all boxes are removed by warping.
if isempty(indices)
    B = A;
end
end

function B = rotation2(A)
% Apply random rotation to both images ad bounding boxes %
B = cell(size(A));
I = A{1};

%Random rotation
tform = randomAffine2d(Rotation=[25 335]);
rout = affineOutputView(size(I),tform);
B{1}  = imwarp(I,tform,OutputView=rout, FillValues=mean(I(:)));

% Apply same transform to boxes.
[B{2},indices] = bboxwarp(A{2},tform,rout,'OverlapThreshold',0.25);
B{3} = A{3}(indices); %Store the labels of the transformed bounding boxes

% Return original data only when all boxes are removed by warping.
if isempty(indices)
    B = A;
end
end

function [B] = Scaling_Unet2(data)
%Random scaling%
B = cell(size(data));
I = data{1};

tform = randomAffine2d(Scale=[0.8,1.8]); %Define transformation
outputView = affineOutputView(size(I),tform);
B{1} = imwarp(I,tform,OutputView=outputView, FillValues=mean(I(:))); %Apply to image
[B{2},indices] = bboxwarp(A{2},tform,outputView,'OverlapThreshold',0.25); %Apply to
bounding boxes
B{3} = A{3}(indices);

if isempty(indices)
    B = A;
end
end

function [B] = Intensity_transformation(data)
%Random intensity transformation%
B = cell(size(data));

I = jitterIntensity(im2single(data{1}),Brightness=[-0.1 0.2], Contrast = [1 2]);
%Define brightness and contrast interval for random transformation

B{1} = I;
B{2} = data{2};
B{3} = data{3};

end

function data = preprocessData(data, targetSize)
% Resize the images and scale the pixels to the desired input size. Also scale the
```

```
% corresponding bounding boxes.
    I = data{1};
    imgSize = size(I);

    % Convert an input image with single channel to 3 channels sigce YOLOv4
    % wants rgb images as input%
    if size(I,3)<3
       I = repmat(I, [1 1 3]);
    end

    bboxes = data{2};
    scale = targetSize(1:2)./imgSize(1:2); %Define the scaling factor

    I = imresize(I,targetSize(1:2)); %Image resize
    bboxes = bboxresize(bboxes,scale); %Bounding boxes resize
    data{1} = I;
    data{2} = bboxes;

end
```

MATLAB code for U-net training and testing

Definition of training and validation datasets

```
%% Dataset split %%
%path_brightfield is the path where brightfield images are saved.
% path_mask is the path for the corresponding masks
inputSize = [80 80 1]; %Image input size
numClasses = 2; %number of classes for the training
classNames = ["Head","background"]; %Pixel Labels
labelIDs   = [1 0]; %Label values
BS = 16; %Selection of the desired batch size for the training
%CheckpointPath is the folder path where it is desired to save the trained
%network

imds = imageDatastore(path_brightfield); %Datastore for the brightfield images
pxds = pixelLabelDatastore(path_mask,classNames,labelIDs); %Label datastore for the
binarized images
cds_tot = combine(imds, pxds); %Combined datastore
N = size(imds.Files, 1); %Dimension of the training set
indices = crossvalind('Kfold', N, 4); %Generation of random indices for dataset split.
The method uses K-fold cross-validation to generate indices.
%To separate the dataset in 75% training - 25% validation, the selection parameter is
set to 4, so that 3 folds are used for training and the last fold for evaluation.
validation_i = (indices == 1); %Validation indices
train_i = ~validation_i; %Training indices
cds_train = subset(cds_tot, train_i); %The dataset it split among train and validation
cds_validation = subset(cds_tot, validation_i); %Validation dataset
```

Data augmentation

```
%% Data augmentation %%
Scaling = transform(cds_train, @Scaling_Unet); %Scaling
Translation = transform(cds_train, @Translation_Unet);  %Random translation
Int = transform(cds_train, @Intensity_transformation_Unet); %Intensity transformation
flip = transform(cds_train, @reflection_Unet); %Random horizontal and vertical flip
blur = transform(cds_train, @blurring); %Blurring

C5 = combine(cds_train, flip, Translation, Int, Scaling, 'ReadOrder','sequential');
```

```
Rotation = transform(C5, @rotation_Unet); %Random rotation
dataset_fin = combine(C5,Rotation, 'ReadOrder','sequential');
```

Dataset preprocessing and training

```
preprocessedTrainData = transform(dataset_fin, @(data)preprocessData_Unet(data,
inputSize)); %Preprocessing to scale the training dataset
preprocessedValidationData = transform(cds_validation, @(data)preprocessData_Unet(data,
inputSize)); %Preprocessing to scale the validation dataset

%% U net training %%
lgraph = unetLayers(inputSize, numClasses, 'EncoderDepth', 3, 'NumFirstEncoderFilters',
16); %U-net-tiny architecture. EncorderDepth is the order of down-sampling, that is the
dept of the encoder part of the U-net. NumFirstEncoderFilters is the number of kernels
of the first layer of the decoder part. The kernels double at each layer of the encoder
part.
lgraph = removeLayers(lgraph,'Segmentation-Layer'); %Removal of the last layer
layerlast =
dicePixelClassificationLayer('Classes',classNames,'Name','New_segmentation_Layer');
%Generalized Dice Classification layer
layer_to_add = layerlast;
lgraph = addLayers(lgraph,layer_to_add); %Addition of the classification layer to the
network
lgraph = connectLayers(lgraph, 'Softmax-Layer' , 'New_segmentation_Layer'); %Connect
the network to the final classification layer

%Training options
options = trainingOptions("rmsprop", ...
    'InitialLearnRate',0.0001, ...
    'MaxEpochs', 50, ...
    'MiniBatchSize', BS, ...
    'VerboseFrequency', 100, ...
    SquaredGradientDecayFactor=0.99, ...
    ValidationData=preprocessedValidationData, ...
    ValidationPatience=40, ...
    ValidationFrequency=50,...
    Shuffle="every-epoch",...
    ExecutionEnvironment="gpu", ...
    OutputNetwork="best-validation-loss", ...
    CheckpointFrequency=10,  Plots="training-progress");

save(fullfile(CheckpointPath, '\Complete training workspace.mat')); %Saving the
workspace
```

Test

```
%% Testing %%
%% Face %%
test = transform(imageDatastore(test_Face_images),
@(x)preprocessTestData_Unet(x,inputSize)); %Preprocessing input data
pxds = pixelLabelDatastore(test_Face_mask,classNames,labelIDs); %Label datastore
YPred = semanticseg(test,net); %Inference on test set
ssm = evaluateSemanticSegmentation(YPred,pxds); %Evaluation

for ii = 1:100
    prova = read(YPred);
    prova3 = read(pxds);
    Dicescore_f(ii,:) = dice(prova{1}=='Head',(prova3{1}=='Head')); %Dice score to
confront the predicted with the true value for each pixel
end
mean(Dicescore_f)
```

```
%% Intermediate %%
testroll = transform(imageDatastore(test_Intermediate_images),
@(x)preprocessTestData_Unet(x,inputSize));
pxds_ = pixelLabelDatastore(test_Intermediate_masks,classNames,labelIDs);

YPredroll = semanticseg(testroll,net);
ssm_faceroll = evaluateSemanticSegmentation(YPredroll,pxds_);

close all
for ii = 1:100
    prova = read(YPredroll);
    prova3 = read(pxds_);
    Dicescore_froll(ii,:) = dice(prova{1}=='Head',(prova3{1}=='Head'));
end
mean(Dicescore_froll)

%% Rim %%
test_rim = transform(imageDatastore(test_Rim_images),
@(x)preprocessTestData_Unet(x,inputSize));
pxds_rim = pixelLabelDatastore(test_Rim_masks,classNames,labelIDs);

YPred_rim = semanticseg(test_rim,net);
ssm_rim = evaluateSemanticSegmentation(YPred_rim,pxds_rim);
close all
figure
for ii = 1:100
    prova = read(YPred_rim);
    prova3 = read(pxds_rim);
    Dicescore_r(ii,:) = dice(prova{1}=='Head',prova3{1}=='Head');
end
mean(Dicescore_r)
```

Auxiliary functions for the data augmentation process

```
function [B] = Intensity_transformation_Unet(data)
%Apply random contrast and brightness transformation to the training set%
B = cell(size(data));
I = data{1};

I = jitterIntensity(im2single(I),Brightness=[-0.1 0.2], Contrast = [1 2]); %Brightness
era -0.1 0.2 prima del 13 3 2024 3 contrast era [1 2]

B{1} = I;
B{2} = data{2};
end

function [B] = Contrast_transformation_Unet(data)
%Apply contrast transformation through adaptive histogram equalization%
B = cell(size(data));
I = data{1};
I = adapthisteq(I);

B{1} = I;
B{2} = data{2};
end

function B = Translation_Unet(A)
% Apply random translation along x and y%
I = A{1};
```

```
tform = randomAffine2d(XTranslation=[-20 20],YTranslation=[-20 20]); %distance selected
randomly from the range.
outputView = affineOutputView(size(I),tform);
B{1} = imwarp(I,tform,OutputView=outputView, FillValues=mean(I(:))); %Application on
the brightfield image

B{2} = imwarp(A{2},tform,OutputView=outputView,  FillValues='background'); %Application
on the binary masks
end

function [B] = Scaling_Unet2(data)
%Random scaling
B = cell(size(data));
I = data{1};

tform = randomAffine2d(Scale=[0.8,1.8]); %Transformation definition
outputView = affineOutputView(size(I),tform);

imAugmented = imwarp(I,tform,OutputView=outputView, FillValues=mean(I(:)));
maskAugmented = imwarp(data{2},tform,OutputView=outputView);

B{1} = imAugmented;
B{2} = maskAugmented;
end

function data = preprocessData_Unet(data, targetSize)
% Resize the images as desired. Also scale the
% corresponding masks.
I = data{1};
mask = data{2};

I = imresize(I,targetSize(1:2)); %Resize brightfield
mask =  imresize(mask,targetSize(1:2)); %Resize binary mask

data{1} = I;
data{2} = mask;
end
```

Routine for sperm tracking at 20x magnification

```
Pixlsize = 6.5/20; %Scale factor for pixel dimensions
L = 15; %Linking distance
L_gap_closing = 30; %Gap closing linking distance
S = 0.6; %Sensitivity of the imfindcircles function
FR = 100; %Frame rate
C = 0; %Control for the existence of Valid tracks
se = strel('disk',5,4); %Disk for morphological filtering
G = 8; %Size of the gaussian blur kernel
min_track_l = 100; %Threshold for track length
min_vsl = 2; %Threshold for VSL
Gap_closing = 6; %Number of frames for the gap closing step

%Tracking algorithm%%
imgPATH = uigetdir('Select the folder','MATLAB Root Directory'); %Select the folder
where the images are stored
if imgPATH ~= 0 %Check if empty

    cc = dir(fullfile(imgPATH, '*.tif')); %list TIF files in the selected folder.
```

```
    if isempty(cc) == 0
        imds = imageDatastore(imgPATH, "FileExtensions", ".tif"); %create datastore
        imds.Files = natsortfiles(imds.Files); %Sort the filenames by ascending order
        Posslash = find(imgPATH=='\'); %Locate the slash
        Filename = imgPATH(Posslash(end)+1:end); %Name of the selected folder
        ImageNUM = size(imds.Files,1); %Number of images
        Im_1 = readimage(imds,1); %Load the first image
        dimx = size(Im_1,1); dimy = size(Im_1,2); %Retrieve image dimensions and store
them
        M_img = Mean_image(ImageNUM, imds, dimx, dimy); %Calculate mean image
    end

else
    msgbox('User selected Cancel'); %No folder was selected
end

[particle_track] = particle_find(imds, parf_ind, M_img, G, se, S); %This function
applies pre-processing filters and locates the head centroids
%in the frames

if isempty(particle_track) == 0

    [track, adjacency_tracks] = simpletracker(particle_track, L_gap_closing,
'MaxLinkingDistance', L, 'MaxGapClosing', Gap_closing); %%Frame-to-frame linking of
points in consecutive frames
    % adjacency_tracks contains one track in each cell, but the indices in each track
are the global indices of the concatenated points array
    isEm = cellfun( @isempty, particle_track ); %Detect empty cells
    x=~isEm==1;
    points_for_track = particle_track(x); %Voids are excluded
    iii=1;  %Index for loop
    concatenated_points = [];
    for i=1:length(points_for_track)
        x1 = size(points_for_track{i},1);
        x2 = size(points_for_track{i},2);
        for ii=1:x1
            for j=1:x2
                concatenated_points(iii,j)=points_for_track{i}(ii,j); %to concatenate
in an array the cell array points_for_track
            end
            iii=iii+1;
        end
    end

    if isempty(track) == 0
        i = 1;
        for i_track = 1 : numel(track)
            track_points = concatenated_points(adjacency_tracks{i_track}, :);
            xx = find(~isnan(track{i_track}));
            if size(track_points,1) > min_track_l %to eliminate short tracks
                travelled_distance = pdist([track_points(1,1:2);
track_points(end,1:2)], 'euclidean');
                vsl(i) = travelled_distance*Pixlsize*FR/(parf_ind(xx(end)) -
parf_ind(xx(1))) ; %Straight-line velocity
                if vsl(i) > min_vsl %Eliminate objects with VSL bigger than the
threshold
                    Valid_tracks{i} = [track_points, parf_ind(xx)']; %Contains the
tracks which satisfy the thresholds for the length and the VSL
                    C = 1; %control variable
                end
```

```
                i = i+1;
            end
        end

        if C == 1 %Valid tracks exist
            %Global parameters for CASA parameters evaluation
            Valid_tracks_fin = Valid_tracks(~cellfun('isempty', Valid_tracks)); %Remove
empty cells and saves the final tracks for the velocities computation
        else
            msgbox('No valid tracks were found', 'Error');
        end
    else
        msgbox('No centroids were found', 'Error');
    end
else
    msgbox('No centroids were found', 'Error');
end

%%CASA Parameters%%
factor_for_cleaning = 5; %multiplicative factor for the niose removal

parfor i_track = 1 : size(Valid_tracks_fin,2)
    track_points = (Valid_tracks_fin{i_track}(1:end, 1:2)); %Track
    frames = Valid_tracks_fin{i_track}(1:end, 3); %Frame sequence
    N = size(track_points, 1); %Track length

    d = pdist([track_points(1,:); track_points(N,:)], 'euclidean'); %initial to final
point distance
    VSL__ = d*(FR/((frames(end)-frames(1))-1))*Pixlsize;

    if VSL__< 5

        out = size(track_points,1); %we don't check on tracks with VSL < 5um/s

    else

        %%Cleaning to remove noise at the end
        %%of a trajectory
        d5 = zeros(1,N-1);
        for i=1:N-1
            d5(i) = pdist([track_points(i,:); track_points(i+1,:)], 'euclidean');
        end
        out = find(d5 > mean(d5)*factor_for_cleaning & d5 ~= 0); %the points which are
distant more than (mean distance*multiplicative factor) are discarded
        %to separate two trajectories that might have been wrongly linked together

    end

    if isempty(out)==1
        out = size(track_points,1);
    end

    if out(1) > min_track_l %Check if the first portion of the trajectory cut from the
previous step satisfies the length requirement
        %if it does, it is stored in a new variable for velocity computation
        frames_= frames(1:out(1));
        track_points_ = track_points(1:out(1),:);
        N = size(track_points_,1);
        cut_tracks{i_track} = [track_points(out(1)+1:end,:), frames(out(1)+1:end)];
%the cut part of the track is stored and analysed separately
```

```
        %Straight-line velocity  VSL - Straight line distance between the
        %first and last point of the trajectory corrected for time.
        d = pdist([track_points_(1,:); track_points_(N,:)], 'euclidean'); %Distance
between first and last track points
        VSL_(i_track) = d*(FR/((frames_(N)-frames_(1))-1))*Pixlsize;

        if VSL_(i_track) > min_vsl %Threshold on VSL

            %Curvilinear velocity  VCL - Distance travelled by the sperm cell
            %along the curvilinear path. Calculated by finding the sum of the
            %distances along the trajectory, corrected for time.
            p=0;
            for i=1:size(track_points_,1)-1
                d = pdist([track_points_(i,:); track_points_(i+1,:)], 'euclidean');
%Distance between consecutive points
                p = p+d;
            end
            VCL_(i_track) = p*(FR/(N-1))*Pixlsize;

            %Linearity LIN
            LIN_(i_track) = round((VSL_(i_track)/VCL_(i_track)) * 100,2);

            if track_points_(1,1) ~= track_points_(N,1) || track_points_(1,2) ~=
track_points_(N,2)  %Check if the first and last position of the sperm cell are the
same

                VCL(i_track) = round(VCL_(i_track),2);
                VSL(i_track) = round(VSL_(i_track),2);
                LIN(i_track) = LIN_(i_track);

                %Radius of circular fit
                Par = CircleFitByPratt(track_points_); %Fit the trajectory to a circle
                Circle_radius(i_track) = round(Par(1,3)*Pixlsize,3); %Radius in microns

            end
        end
    end
end

i_track = size(Valid_tracks_fin,2)+1;

%CASA parameters are evaluated in the same way for the tracks that have
%been cut in the previous computation
for i = 1:size(cut_tracks,2)
    if isempty(cut_tracks{i}) == 0
        track_points = cut_tracks{i}(:,:);
        N = size(track_points,1);
        frames_ = track_points(1:end, 3);

        if N > min_track_l %Threshold on the track length
            %Straight-line velocity  VSL - Straight line distance between the
            %first and last point of the trajectory corrected for time.
            d = pdist([track_points(1,:); track_points(N,:)], 'euclidean'); %Distance
between first and last track points
            VSL_(i_track) = d*(FR/((frames_(end)-frames_(1))-1))*Pixlsize*beta;

            if VSL_(i_track) > min_vsl
                %Curvilinear velocity  VCL - Distance travelled by the sperm cell
                %along the curvilinear path. Calculated by finding the sum of the
```

```
                %distances along the trajectory, corrected for time.
                p=0;
                for i=1:size(track_points,1)-1
                    d = pdist([track_points(i,:); track_points(i+1,:)], 'euclidean');
%Distance between consecutive points
                    p = p+d;
                end
                VCL_(i_track) = p*(FR/(N-1))*Pixlsize;

                %Linearity LIN
                LIN_(i_track) = round((VSL_(i_track)/VCL_(i_track)) * 100,2);

                if track_points(1,1) ~= track_points(N,1) || track_points(1,2) ~=
track_points(N,2)  %Check if the first and last position of the sperm cell are the same
                        VCL(i_track) = round(VCL_(i_track),2);
                        VSL(i_track) = round(VSL_(i_track),2);
                        LIN(i_track) = LIN_(i_track);

                        %Radius of circular fit
                        Par = CircleFitByPratt(track_points(:,1:2)); %Fits the
trajectory to a circle
                        Circle_radius(i_track) = round(Par(1,3)*Pixlsize,3); %Radius in
microns
                end
            end
        end
    end
end

x_tot = find(VSL~=0 & VSL~=Inf);
T_total = table(VSL(x_tot).',VCL(x_tot).', LIN(x_tot).', Circle_radius(x_tot).');

function [outputArg1] = particle_find(imds, parf_ind, M_img, G, se, S)
%particle_find works with images aquired with 20x magnification.
%This function applies pre-processing filters and locates the head centroids
%in the frames.

%%INPUT:
% imds: image datastore
% parf_ind: indexes for the loading of the images inside the parfor loops
% M_img: Mean image
% G: Dimension of the kernel for the Gaussian filter
% se: disk object for morphological filtering
% S: sensitivity of 'imfindcircles'

parfor ii = 1:length(parf_ind)
    Im = (wiener2(readimage(imds,parf_ind(ii)) - M_img)); %M
    ImageSeq_filtered = imgaussfilt(Im,G); %Gaussian Filter
    ImageSeq_filtered = imerode(ImageSeq_filtered,strel('disk',5,4)); %Erosion
    ImageSeq_filtered = imdilate(ImageSeq_filtered,se); %Dilation

    [centers] = imfindcircles((ImageSeq_filtered),[6 50],'Sensitivity', S, 'Method',
'TwoStage'); %Location of head apparent centroids

    if isempty(centers) == 0
        particle_track{ii} = round(centers(:,1:2));
    else
        particle_track{ii} = [];
    end
end
```

```
outputArg1 = particle_track;
end


function [outputArg1] = Mean_image(ImageNUM, imds, dimx, dimy)
%%Mean_image calculates the mean image of the image sequence for a further
%%step of noise removal
%%INPUT:
% imds: image datastore
% ImageNUM: number of images
% dimx: image x dimension
% dimy: image y dimension
Image = uint64(zeros(dimx,dimy,1)); %Create empty image
parf_ind = 1:ImageNUM; %indexes for parfor loop

parfor j = 1:length(parf_ind)
    Im = uint64(readimage(imds,parf_ind(j))); %loading image from datastore
    Image = imadd(Image,Im); %Create sum of subsequent images
end

M_img = Image./ceil(length(parf_ind));  %Mean calculation

outputArg1 = uint8(M_img);
end
```

Sperm analysis code

This code performs sperm head detection, tracking, segmentation and motility and morphology analysis. To use the following code, the external "simpletracker" module is required (available at the following link: https://it.mathworks.com/matlabcentral/fileexchange/34040-simpletracker). The code was modified to include the maximum gap closing distance (GC_distance in the code below), which regulates the gap closing operations. It is given as input to the "nearestneighbour" supplementary function in line 242 of the "simpletracker" code. The detection network should be stored in a variable called 'net' in the workspace; the same holds true for the Face segmentation and Rim segmentation networks, which should be saved as ReteFace and Reterim respectively in the workspace.

```
%The code is thought to work on multiple folders stored in an
%external folder. After detection and segmentation, the resulting tracks
%are saved in a Folder called Tracks, and the motility and morphology
%analysis module works on the tracks stored in such folder.

%% External Folder Loading %%
[imgPath] = uigetdir('Select the folder','MATLAB Root Directory'); % load folder
FolderTYPE = '/';     imgTYPE = '*.tif';
Folder  = dir([imgPath FolderTYPE]); %Load multiple folders
NumbFOLDER = length(Folder);
inputSize = [640 640 3]; %Detection network input size
scale = 2048/640; %to size up the bounding boxes
CS = 0.5; %Threshold for confidence Score
Pixlsize = 6.5/40; %conversion factor from pixel to microns. It depends on the camera
and magnification used
FR = 50; %Frame rate
CASA_folder = [];
Motility_tracks = {};
tracks_label = [];
MinTrackL = 75; %Desired minimum track length
```

```
factor_for_cleaning = 5; %Multiplicative factor for detecting tracks erroneously linked
together.
minvsl = 0; %Minimum Threshold for VSL; setting this parameter, cells having a vsl
lower than the threshold are discarded.
T = 1/FR; % Sampling period
GC_distance = 35; %Maximum distance for gap closing in pixels. If a gap is detected,
the research for a successive linking point is done within a max distance indicated by
this parameter.
Linking_distance = 70; %Maximum linking distance for two consecutive points. This
parameter should be carefully set according to the frame rate, as lower frame rate will
determine higher distances in sperm movement across frames.
GC = 3; %Maximum Gap closing frame interval. If a gap is detected, the possible linking
point is searched within a number of successive frames determined by this parameter.
Higher GC may cause incorrect trajectories reconstructions.

newSubFolder_tracks = sprintf('%s/Tracks%s', imgPath); %Create folder to save all
tracks
mkdir(newSubFolder_tracks);
pathname = [newSubFolder_tracks '/'];
```

Detection and tracking

Images are scaled down to 640x640 pixels and given to YOLOv4, which outputs bounding boxes around the sperm heads and labels them as Face or Rim. The coordinates of the bounding boxes are then tracked with the “simpletracker” function.

```
figure
for j=3:NumbFOLDER %the script is made to work on a folder contanining multiple video
acquisitions, each stored in a sub-folder

    File = char(Folder(j).name);
    path= [imgPath '\' Folder(j).name];
    video = imageDatastore(path); %The image sequence is stored in a datastore

    final_video = transform(video, @(data)preprocessData2(data,inputSize)); %input pre-
processing for detection with YOLO4

    newSubFolder = sprintf('%s/YOLO4%s', path); %create folder for saving detection
output
    if ~exist(newSubFolder, 'dir')
        mkdir(newSubFolder);
    end

       for i = 1:size(final_video.Files,1) %Each image is given to YOLO4 for sperm head
detection

        [detectedBoxes, detectedScores, detectedLabels] = detect(net,
read(final_video), "Threshold", 0.3); %Object detection with YOLO4
        [selectedBoxes, selectedScores, index] =
selectStrongestBbox(detectedBoxes,detectedScores, "OverlapThreshold",0.5); %Delete
boxes overlapping more than 50%

        bboxes = bboxresize(selectedBoxes, scale); %Scale up bounding boxes
        scores = selectedScores; %Storing scores
        labels = string(detectedLabels(index)); %Storing labels

        %%Plot and save output%%
        for ii =1:size(labels,1) %Colormap for plot labels, whether they are Face or
Rim
            if labels(ii) == 'Face'
```

```
                colormap_{ii} = ['yellow']; %Colors for plot
            else
                colormap_{ii} = ['magenta'];
            end
        end

        if isempty(bboxes) == 0

            I = imread(video.Files{i}); %Plot and Save output
            I = insertObjectAnnotation(I,'rectangle',bboxes,  string(scores),
'LineWidth',5,'FontSize',30, AnnotationColor=colormap_); %Insert bounding boxes and
label annotations
            %on each image
            imshow(I, 'Parent', gca);
            exportgraphics(gca, fullfile(newSubFolder, [num2str(i) '.jpg'])); %saving
            clf
            x = scores >= CS;
            if isempty(x) == 0 %Storing in cell array for further elaboration
                centroids{i,:} = [bboxes(x,1)+bboxes(x,3)/2 bboxes(x,2)+bboxes(x,4)/2];
%Storing centerpoints of bounding boxes sequence
                frame{i,:} =  repmat(i, [size(bboxes,1), 1]); %Frames
                Scores{i,:} = scores(x); %Confidence scores
                Label_Class{i,:} = labels(x); %Labels
            end

        else %if no bounding boxes are detected
            centroids{i,:} = [];
            frame{i,:} = 0;
            Scores{i,:} = [];
            Label_Class{i,:} = [];
        end
        clear x colormap_
    end

    %% Tracking %%
    [track adjacency_tracks] = simpletracker(centroids, GC_distance,
'MaxLinkingDistance', Linking_distance, 'MaxGapClosing', GC); %Tracks points closer
than Linking_distance pixels. Closes gaps spanning max GC frames and
    %distant less than GC_distance pixels
    isEm = cellfun( @isempty, centroids ); %detect empty cells
    x=find(~isEm==1);
    points_for_track = centroids(x); %voids are excluded
    labels_for_track = Label_Class(x);
    iii=1; concatenated_points = []; %index for loop
    for i=1:length(points_for_track)
        x1 = size(points_for_track{i},1);
        x2 = size(points_for_track{i},2);
        for ii=1:x1
            for j=1:x2
                concatenated_points(iii,j)=points_for_track{i}(ii,j); %to concatenate
in an array the cell array points_for_track
            end
            concatenated_labels(iii,1)=labels_for_track{i}(ii);
            iii=iii+1;
        end
    end

    figure %Plotting the trajectories
    p = 1;
    for i_track = 1 : numel(track)
```

```
        track_points = concatenated_points(adjacency_tracks{i_track}, :);
        labels_final = concatenated_labels(adjacency_tracks{i_track}, :);
        xx = find(~isnan(track{i_track}));
        d = pdist([track_points(1,:); track_points(end,:)]);

        if size(track_points,1) > MinTrackL %Tracks are cut at 1.5 secs
            vsl(p) = d*Pixlsize*FR/((xx(end) -xx(1)) -1) %cells with VSL < 1 are
discarded

            if vsl(p) > 1 %changing this threshold slow cells can be discarded if they
are not of interest
                plot(track_points(:,1), track_points(:,2), '*-')
                hold on
                text(gca,track_points(ceil(end/2),1), track_points(ceil(end/2),2),
num2str(p), 'Clipping','on');
                final_tracks{p} = [track_points(1: MinTrackL,1), track_points(1:
MinTrackL,2) xx string(labels_final)]; %contains the tracks that are above the lenght
and VSL thresholds
                p = p+1;
            end
        end

        clear xx track_points
    end
    xlim([0 2048]);
    ylim([0 2048]);
    set(gca, 'Ydir', 'reverse')
    xlabel(gca,'x');
    ylabel(gca, 'y');
    title('Tracking bounding boxes')
    saveas(gca, [newSubFolder '\Traiettorie_.fig']);

    newSubFolder2 = sprintf('%s/Unet%s', newSubFolder); %Create subfolder to save the
segmentation results
    if ~exist(newSubFolder2, 'dir')
        mkdir(newSubFolder2);
    end
```

Segmentation

This part of the code crops the image following the coordinates of the bounding boxes detected in the previous section so that the sperm head is centred in the cropped image. Then, based on the label associated to the detected head, the image is segmented by a U-net trained on only Face or Rim configurations respectively. An elliptical fit on the head contour allows to compute the x and y coordinates of the head centre, allowing a precise motility computation. Morphological parameters such as Major and Minor axes are computed with “regionprops”.

```
    %% Segmentation %%
    p =1;
    figure
    if exist('final_tracks', 'var')
        for i_track = 1:size(final_tracks,2)

            centroids = str2double(final_tracks{i_track}(:,1:2)); %Retrieve the track
points, frames, labels
            frames = str2double(final_tracks{i_track}(:,3));
            labels = (final_tracks{i_track}(:,4));
```

```
            frames_nogapclosing = frames(1):1:frames(end); %If the gap closing has
occurred, the frames sequence contained in frames has gaps. This variable fill these
gaps

            for i = 1:size(frames_nogapclosing,2)
                ind_real = find(frames_nogapclosing(1,i) == frames); %To put zero when
there is a gap closing
                if isempty(ind_real) == 0 %Cells are segmented only in frames where the
bounding box has been tracked
                    I1 = (readimage(video, frames(ind_real,1)));
                    I_cropped = imcrop(I1, [centroids(ind_real,:)-40 79 79]); %Crop
80x80

                    if size(I_cropped) > [69,69] %tolerance on the dimension of the
input images, that can change for objects detected near the image borders

                        xscale = 80/size(I_cropped,1);
                        yscale = 80/size(I_cropped,2);

                        I_cropped = imresize(I_cropped, [80 80]); %Sets the dimension
to 80x80 pixels

                        switch labels(ind_real) %According to the  label detected by
YOLO, the image is  segmented with one U-net-tiny trained entirely on Face
configurations or Rim respectively
                            case 'Face'
                                YPred = semanticseg(I_cropped, ReteFace.net);
                                I = YPred == 'Head';
                                I_back = YPred == 'background';

                            case 'Rim'
                                YPred = semanticseg(I_cropped, Reterim.net);
                                I = YPred == 'Head';
                                I_back = YPred == 'background';
                        end

                        I1 = bwareaopen (I, 10); %Opening filter to remove small
connected noise
                        I3 = imclearborder(bwareaopen(imfill(I1, 'holes'), 100));
%Close holes and further remove small connected components

                        s = regionprops(logical(I3),
'Area','Orientation','MajorAxisLength', 'MinorAxisLength', 'Centroid'); %Evaluation of
morphological features with regionprops
                        aree_crop = zeros(1,size(s,1));
                        for kk = 1:size(s,1)
                            aree_crop(kk) = s(kk).Area;
                        end
                        x2 = find(aree_crop == max(aree_crop)); %Stores the bigger
detected value of Area in case there are multiple heads in the same crop or there is
noise
                        %which has not been removed from previous filtering
                        %steps

                        if isempty(s) == 0

                            B = cell2mat(bwboundaries(bwlabel(I3)==x2(1))); %Detection
of Head Contour

                            %%Parametric fit of the ellipse
```

```
                            phi = linspace(0,2*pi,50);
                            cosphi = cos(phi) ;
                            sinphi = sin(phi);
                            xbar = s(x2(1),1).Centroid(1); %Coordinate x
                            ybar = s(x2(1),1).Centroid(2); %Coordinate y
                            a = s(x2(1)).MajorAxisLength/2; %Major ax
                            b = s(x2(1)).MinorAxisLength/2; %Minor ax
                            theta = pi*s(x2(1)).Orientation/180;
                            R = [ cos(theta) sin(theta)
                                -sin(theta) cos(theta)];
                            xy = [a*cosphi; b*sinphi];
                            xy = R*xy;
                            x = xy(1,:) + xbar;
                            y =xy(2,:)+ ybar;

                            xbar = xbar*xscale; %Center x coordinate
                            ybar = ybar*yscale; %Center y coordinate

                            Area(i,1) = s(x2(1)).Area;
                            Orientation(i,1) = s(x2(1)).Orientation;
                            Perimeter(i,1) = s(x2(1)).Perimeter.*Pixlsize; %Converted
to  microns and stored
                            MajAx(i,1) = s(x2(1)).MajorAxisLength.*Pixlsize;
                            MinAx(i,1) = s(x2(1)).MinorAxisLength.*Pixlsize;
                            Centroide_xy(i,1:2) = [centroids(ind_real,:)-40]+[xbar,
ybar]; %Centroid coordinates retrieved from the elliptical fit on the head contour,
transformed
                            %from the crop reference frame to the global reference
frame
                            Final_label(i,1) = labels(ind_real);

                        else

                            Area(i,1) = 0; %0 in case nothing is segmented
                            Orientation(i,1) = 0;
                            Perimeter(i,1) = 0;
                            MajAx(i,1) = 0;
                            MinAx(i,1) = 0;
                            Centroide_xy(i,1:2) = [0,0];
                            Final_label(i,1) = convertCharsToStrings('null');

                        end

                    else %0 in case the bounding boxes dimensions are below the chosen
threshold
                        Area(i,1) = 0;
                        Orientation(i,1) = 0;
                        Perimeter(i,1) = 0;
                        MajAx(i,1) = 0;
                        MinAx(i,1) = 0;
                        Centroide_xy(i,1:2) = [0,0];
                        Final_label(i,1) = convertCharsToStrings('null');

                    end

                else %0 if gap closing has occurred and there is no detection for that
frame
                    Area(i,1) = 0;
                    Orientation(i,1) = 0;
                    Perimeter(i,1) = 0;
```

```
                    MajAx(i,1) = 0;
                    MinAx(i,1) = 0;
                    Centroide_xy(i,1:2) = [0,0];
                    Final_label(i,1) = convertCharsToStrings('null');
                    Circularity(i,1) = 0;

                end
                clear aree_crop x2
            end

            TF = find(isoutlier(Area, "quartiles")==1); %Remove outlilers in the Area
sequence, where
            % an oulier is a point more than 1.5 interquartile ranges above the upper
quartile (75 percent) or below the lower quartile (25 percent)

            p = p+1;
            Area =  Area.*Pixlsize^2; %Scaling to micron
            AR = MajAx./MinAx;
            symbols_face = find(Final_label=='Face');
            symbols_rim = find(Final_label=='Rim');

            Area(TF) = 0; %0 is put where Area outliers have been detected
            MajAx(TF)=0;
            MinAx(TF)=0;
            Centroide_xy(TF,:) = 0;
            Orientation(TF) = 0;

            %% Saving results %%
            varnames = {'X';'Y';'Frames';'Labels'; 'Area'; 'MajorAx'; 'MinorAx';
'Perimeter'; 'Orientation'; 'Circularity'};
            T = table(Centroide_xy(:,1), Centroide_xy(:,2),frames_nogapclosing',
Final_label, Area,MajAx, MinAx,Perimeter ,Orientation,  'VariableNames',varnames);
            writetable(T, fullfile(newSubFolder_tracks, ['\Results_Folder_'
num2str(File) '_Track_' num2str(i_track) '.xlsx'])) %Each track is saved in a xlsx file

            close all
            clear Area MajAx MinAx Perimeter Orientation Centroide_xy
frames_nogapclosing Final_label symbols_face symbols_rim centroids frames labels
frames_nogapclosing
        end
    end

    save(fullfile(path, '\Complete tracking and morphometrical analysis.mat'));
    clear final_tracks
end
```

Motility and morphometry analisis

This part of the code computes the motility values presented in Table 6 together with the morphometrical parameters defined in paragraph 3.8.

```
%% CASA and MORPHOMETRIC EVALUATION %%
dinfo = dir([newSubFolder_tracks '/*.xlsx']); %Load track files
filenames = {dinfo.name};
pathname = [newSubFolder_tracks '/'];
filenames = cellstr(filenames);
files_to_delete = zeros(1,size(filenames,2));

newSubFolder = sprintf('%s/CASA%s', pathname); %Create folder for CASA files
% Finally, create the folder if it doesn't exist already.
if ~exist(newSubFolder, 'dir')
```

```
    mkdir(newSubFolder);
end
CASA_folder = newSubFolder;

close all
for k=1:numel(filenames)
    opts = detectImportOptions(fullfile(pathname, filenames{k}), 'FileType',
'spreadsheet'); %Import options for track files
    opts.SelectedVariableNames = {'Labels', 'Area','MajorAx', 'MinorAx', 'Orientation',
'Frames', 'X', 'Y'}; %Columns to import

    Morphology = readtable(fullfile(pathname, filenames{k}), opts); %Morphological
parameters
    [Area_original, TF] = rmmissing(Morphology.Area); %Removal of NaN values
    Morphology(TF==1, :) = [];

    index = find(Area_original ~= 0,1, 'first'); %Check for zeros at the beginning of
the data column
    index2 = find(Area_original ~= 0, 1, 'last'); %Check for zeros at the end of the
data column
    Morphology2 = Morphology(index:index2, :); %It may happen that null detections
cause 0 to appear at the beginning and at the end of the column vector.

    Area_sequence = Area_original(index : index2); %Storing variables
    Minor_ax = rmmissing(Morphology2.MinorAx);
    labels_yolo{k} = rmmissing(Morphology2.Labels);
    frame = rmmissing(Morphology2.Frames);
    Tracks{k} = [rmmissing(Morphology2.X), rmmissing(Morphology2.Y)];
    Frames{:,k} = rmmissing(Morphology2.Frames);
    Area1{:,k} = rmmissing(Morphology2.Area);
    labels_yolo{:,k} = rmmissing(Morphology2.Labels);

    %% Morphometrical analysis %%
    %Rolling Frequency %
    %Computed as the FFT of the Area sequence%
    vo = interp1(frame(Area_sequence>0), Area_sequence(Area_sequence>0),
frame(Area_sequence==0), 'spline'); %Spline interpolation to fill the null
    % points in correspondence of gaps in the trajectory
    Area_sequence(Area_sequence==0) = vo;
    Minor_ax(find(Minor_ax==0)) = interp1(frame(Minor_ax>0), Minor_ax(Minor_ax>0),
frame(Minor_ax==0), 'spline'); %Same is done for the minor ax

    figure
    plot(Area_sequence); %plot the area sequence
    title('Area sequence')
    plot(detrend(Area_sequence,0)); %Removes the continuos component from the Area
signal
    legend('Original', 'spline interpolation', 'detrended');

    L = size(Area_sequence,1); % Length of signal
    t = (0:L-1)*T;        % Time vector
    Y = fft((detrend(Area_sequence(:,:), 0).*hann(L)).*1/mean(hann(L))); %FFT of the
Area signal, detrended with respect to the continuous component, adopting the hanning
window
    f = Fs*(0:(L/2))/L; %frequency ax
    P2 = abs(Y/L); %Amplitude spectrum
    P1 = P2(1:L/2+1); %Positive amplitude spectrum
    P1(2:end-1) = 2*P1(2:end-1);

    subplot(1,2,2)
```

```
    % P1(1) = 0;
    plot(f,P1)
    title("Single-Sided Amplitude Spectrum of Area sequence")
    xlabel("f (Hz)")
    ylabel("|P1(f)|")
    [m,pos] = findpeaks(P1, 'MinPeakProminence',1); %To identify the peak in the
amplitude spectrum
    hold on

    if isempty(m) == 1
        Rolling_f(k) = 0;
    else
        Rolling_f(k) = f(find(P1 == max(m))); %the frequency corresponding to the peak
is the rolling frequency
    end

    x1 = find(Area_sequence == 0);
    Morphology2(x1,:) = [];
    lab=find(cellfun(@(v)isequal(v, 'Face'), Morphology2.Labels)==0); %find Rim labels

    if size(lab,1) == size(Morphology2,1) %IF there are no Face configurations we don't
evaluate Area, Major and Minor ax

        Area_fin(k) = 0;
        Major_Ax(k) = 0;
        Minor_Ax(k) = 0;

    else

        Morphology2(lab,:) = []; %Rim removal for area computation

        [~, n2] = maxk(Morphology2.Area, 15); %Take the first 15 biggest face
configurations
        Biggest_Areas_Face = Morphology2(n2,:);

        L = mean(maxk(Biggest_Areas_Face.MajorAx, 5)); %Search among this the biggest
five major axes
        W = mean(maxk(Biggest_Areas_Face.MinorAx, 5));%Search among this the biggest
five minor axes

        for kk = 1:size( Biggest_Areas_Face, 1)
            d_i(kk) = abs(Biggest_Areas_Face.MajorAx(kk) - L) +
abs(Biggest_Areas_Face.MinorAx(kk) - W);
        end

        [~, n_fin] = mink(d_i, 10); %Select the 10 Area measurements that are the
closest to L and W defined above. In correspondence of these points
        %mean area, major and minor ax are computed.
        Area_fin(k) = mean(Biggest_Areas_Face.Area(n_fin));
        Major_Ax(k) = mean(Biggest_Areas_Face.MajorAx(n_fin));
        Minor_Ax(k) = mean(Biggest_Areas_Face.MinorAx(n_fin));

        clear x1 d_i lab L Q
    end

    %% Motility analysis %%
    track_points = tracks{k};
    frames = Frames{k};

    x = find(track_points(:,1) == 0 & track_points(:,2) == 0);%Gap closing points
```

```
    track_points(x,:) = []; %remove points in correspondence of frames where there are
no information on the sperm head
    frames(x,:) = [];

    %Straight-line velocity VSL%
    %Distance between the first and last track point, divided by time
    N = size(track_points, 1);
    d = pdist([track_points(1,:); track_points(end,:)], 'euclidean'); %Distance between
first and last point
    VSL(k) = d*(FR/((frames(end)-frames(1))))*Pixlsize; %Straight-line velocity

    %Curvilinear velocity VCL%
    %sum of the curvilinear path with respect to time
    p=0;
    for i=1:size(track_points,1)-1
        d = pdist([track_points(i,:); track_points(i+1,:)], 'euclidean'); %Sum of
curvilinear path
        p = p+d;
    end
    VCL(k) = p*(FR/((frames(end)-frames(1))))*Pixlsize;

    %%Fractal dimension FD%%
    for i=1:size(track_points,1)-1
        d(i) = pdist([track_points(1,:); track_points(i+1,:)]);
    end
    L = max(d); %Distance from the track origin and the most distant point from the
origin
    FD(k) = log(N-1)/(log(N-1)+log(L/p));

    %Mean angular displacement - MAD - Measure of the trajecotry
    %curvature, mediated in an interval of
    %Delta points
    Delta = 5; z_x = zeros(1,N); z_y = zeros(1,N); theta = zeros(1,N);
    for i=1:N-Delta
        z_x(i) = (track_points(i+Delta,1) - track_points(i,1));
        z_y(i) = -(track_points(i+Delta,2) - track_points(i,2));
        theta(i) = mod(360 + (atan2d(z_y(i), z_x(i))),360);
    end
    c = find(isnan(theta)); z = find(~isnan(theta));
    N1 = (N-length(c));
    MAD(k) = round(sum(theta(z))/(N1-Delta),2);

    if isnan(MAD(k))
        MAD(k) = 0;
    end

    %Radius of circular fit%
    Par = CircleFitByPratt(track_points(:,1:2)); %fits the trajectory to a circle
    Circle_radius(k) = round(Par(1,3)*Pixlsize,3);
    if Circle_radius(k) > 1000
        Circle_radius(k) = 1000;
    end

    %%Prominence%%
    [pks, locs, width_max, prominence_max] = findpeaks(Area_sequence,
'MinPeakProminence', 20,'Annotate','extents',
'Annotate','extents','WidthReference','halfheight'); %Find Max values
    %of the Area Sequence
    [Minima,MinIdx, width_min, prominence_min] = findpeaks(-Area_sequence,
'MinPeakProminence', 20,'Annotate','extents'); %find minima in the area sequence
```

```
    peak = sort([locs; MinIdx]);
    for ii=1:size(peak,1)-1
        d = pdist([track_points(peak(ii),:);  track_points(peak(ii+1),:)],
'euclidean'); %Distance between minima peaks
        dist(ii) = d*Pixlsize;
    end

    avg_path = [track_points(locs,1), track_points(locs,2)]; %Trajectory of only peaks
of max area build an average path
    frame_avgpath = frames(locs); %Corresponding frames

    if size(avg_path,1) > 2
        for ii = 1:size(avg_path,1)-1
            dd(ii) = pdist([avg_path(ii,:); avg_path(ii+1,:)]); %Distance among points
building the average path
        end
        APV(i) = sum(dd)*(Pixlsize*FR)/(frame_avgpath(end)-frame_avgpath(1)); %Velocity
along the average path defined by the coordinate points where the area sequence has a
maximum peak
    else
        APV(i) = 0;
    end

    peak_min = track_points(MinIdx,1:2); %x-y coordinates corresponding to min peaks

    if APV(i) > 0 %To avoid in-situ sperm cells, we check if the average path velocity
is bigger than 0
        for kk = 1:size(peak_min,1)

            for  k = 1:length(avg_path)-1 %The evaluation is done on all average points

                mAB = (avg_path(k+1,2)-avg_path(k,2))/(avg_path(k+1,1)-avg_path(k,1));
%SLope of the segment formed by consecutive average path points
                mABPerp = -1/mAB; %Angular coefficient of the perpendicular line
                A = (avg_path(k+1,1)-avg_path(k,1)); %X-axis projection of the segment
joininng two consecutive points of the average path
                B = (avg_path(k+1,2)-avg_path(k,2)); %Y-axis projection

                t1 = double((mABPerp*((avg_path(k,1) - peak_min(kk,1))) +
peak_min(kk,2) - avg_path(k,2))/(B-mABPerp*A)); %Parameter obtained substituting the
parametric representation
                %of the segment in the equation of the line
                %perpendicular to AB passing for the kk min peak

                if t1>0 & t1<1
                    xSol(k) = avg_path(k,1) + t1*A; %X coordinate of the intersection
point
                    ySol(k) = avg_path(k,2) + t1*B; %y coordinate
                end
            end


            if exist('xSol', 'var')
                xSol = xSol';
                ySol = ySol';

                Soluzions = [xSol(find(xSol > 0 & ySol > 0)), ySol(find(xSol > 0 & ySol
> 0))]; %non zero solutions to the equations
```

```
                for ii = 1:size(Soluzions,1) %Distance from intersection points and the
min peak
                    dist_peaks(ii) = pdist([peak_min(kk,:); Soluzions(ii,:)]);
                end
                c = find(dist_peaks==min(dist_peaks)); %Selection of the smallest
distance since the evaluation is done on the entire average path

                d=0; x=[]; y=[]; A=[]; f=1;
                xx = []; yy = []; r=1;
                %since trajectories are not always straight, we
                %evaluate the intersection of the prominence line
                %with the actual track to avoid errors
                for j=1:size(track_points,1)-1 %Evaluation of the number of
intersections between the segment and the trajectory
                    [x y] = polyxpoly([track_points(j,1) track_points(j+1,1)],
[track_points(j,2) track_points(j+1,2)], [peak_min(kk,1); Soluzions(c,1)],
[peak_min(kk,2); Soluzions(c,2)],'unique');
                    %get the intersection between the segment connecting two
consecutive track points and the segment
                    %connecting the corresponding point
                    %of the average path
                    if isempty(x) == 0
                        d=d+1;
                        A(f:f+length(x)-1,1) = x;
                        A(f:f+length(x)-1,2) = y;
                        f = f+1;
                    end
                end

                PR(kk,1) = d_picchi(c)*0.162;
                intersections(kk,1) = size(unique(A, 'rows', 'stable'),1);  %Save the
number of intersections

                clear c dist_peaks
            else
                PR(kk,1) = 0;
                intersections(kk,1) = 0;
            end

            clear xSol ySol t1

        end

        PR_within_intersection_limit = PR(find(intersections<5)); %Prominences with
more than 5 intersections with the track are discarded
        PROMINENCE(i) =
mean(PR_within_intersection_limit(PR_within_intersection_limit>0));

    else
        PROMINENCE(i) = 0;
    end

    close all
    Motility_tracks{k} = [track_points , frames];
    selected_tracks{k} = k;
    clear x p d smoothed_track_points theta M
end

    %% Saving variables %%
```

```
    varNames = {'Track','VSL', 'VCL', 'Area', 'MajAx', 'MinAx', 'Fractal dimension',
'Mad', 'Rolling Freq'};
    T = table(sheet_name', VSL.',VCL.',  Area_fin', Major_Ax', Minor_Ax',FD', MAD',
Rolling_f', 'VariableNames', varNames);
    writetable(T,[CASA_folder '\Casa_Parameters_.xlsx']);

    %Mat FILE
    save([pathname '\Backup workspace.mat'],'tracks','frame', 'FR',
'MinTrackL','Pixlsize', 'MinTrackL', 'factor_for_cleaning'); %Backup file for workspace
```

Appendix B

The slope and coefficient of correlation for the linear fitting on VSL, MAD and FD with VCL are reported below.

Table 19: Linear fitting parameters for VSL, MAD and FD in relationship with VCL.

VSL-VCL			
ISO	**0 min**	**30 min**	**60 min**
Slope	1.64 ± 0.056	1.6 ± 0.035	a. 0.038
R^2	0.79	0.91	0.88
HYP1			
Slope	1.66 ± 0.041	1.65 ± 0.048	1.79 ± 0.063
R^2	0.87	0.89	0.85
HYP2			
Slope	1.64 ± 0.038	1.99 ± 0.057	1.98 ± 0.067
R^2	0.89	0.89	0.85
FR			
Slope	1.74 ± 0.074	1.61 ± 0.067	1.6 ± 0.079
R^2	0.78	0.81	0.86
CAFF			
Slope	2.04 ± 0.21	2.35 ± 0.26	2.59 ± 0.13
R^2	0.39	0.44	0.68
PEO 0.5			
Slope	1.52 ± 0.1		
R^2	0.79		
PEO 0.9			
Slope	1.18 ± 0.037		
R^2	0.92		
MAD-VCL			
ISO	**0 min**	**30 min**	**60 min**
Slope	-0.0032 ± 0.043	0.13 ± 0.017	0.091 ± 0.047
R^2	0.00025	0.029	0.017
HYP1			
Slope	0.0253 ± 0.041	0.088 ± 0.068	0.032 ± 0.7
R^2	0.0016	0.012	0.0015
HYP2			
Slope	0.076 ± 0.055	0.13 ± 0.075	-0.0092 ± 0.069
R^2	0.0083	0.02	0.00011
FR			

Slope	-0.044 ± 0.05	0.028 ± 0.055	-0.047 ± 0.093
R^2	0.0051	0.002	0.0041
CAFF			
Slope	-0.052 ± 0.14	0.15 ± 0.18	-0.14 ± 0.12
R^2	0.00095	0.00659	0.0067
PEO 0.5			
Slope	-0.043 ± 0.063		
R^2	0.0081		
PEO 0.9			
	0.013 ± 0.029		
	0.0024		
FD-VCL			
ISO	**0 min**	**30 min**	**60 min**
Slope	-85.1 ± 11.5	-69.0 ± 7.09	-53.6 ± 4.71
R^2	0.20	0.31	0.36
HYP1			
Slope	-108.8 ± 7.93	-87.2 ± 9.9	-30.8 ± 4.7
R^2	0.44	0.36	0.23
HYP2			
	-88.9 ± 7.5	-79.2 ± 8.7	-55.4 ± 6.09
	0.38	0.34	0.34
FR			
	-75.6 ± 10.6	-44.0 ± 5.7	-50.4 ± 7.2
	0.25	0.31	0.43
CAFF			
	-64.0 ± 19.3	55.4 ± 27.0	-40.7 ± 12.0
	0.17	0.53	0.38
PEO 0.5			
	-12.7 ± 9.05		
	0.033		
PEO 0.9			
	-55.8 ± 12.9		
	0.17		

References

[1] M. N. Mascarenhas, S. R. Flaxman, T. Boerma, S. Vanderpoel, and G. A. Stevens, "National, Regional, and Global Trends in Infertility Prevalence Since 1990: A Systematic Analysis of 277 Health Surveys," *PLoS Med*, vol. 9, no. 12, Dec. 2012, doi: 10.1371/journal.pmed.1001356.

[2] E. Babakhanzadeh, M. Nazari, S. Ghasemifar, and A. Khodadadian, "Some of the factors involved in male infertility: A prospective review," 2020, *Dove Medical Press Ltd.* doi: 10.2147/IJGM.S241099.

[3] P. R. Brezina, N. Ning, E. Mitchell, H. A. Zacur, T. A. Baramki, and Y. Zhao, "Recent Advances in Assisted Reproductive Technology," *Curr Obstet Gynecol Rep*, vol. 1, no. 4, pp. 166–173, Dec. 2012, doi: 10.1007/s13669-012-0019-2.

[4] D. Zheng *et al.*, "Intracytoplasmic sperm injection (ICSI) versus conventional in vitro fertilisation (IVF) in couples with non-severe male infertility (NSMI-ICSI): Protocol for a multicentre randomised controlled trial," *BMJ Open*, vol. 9, no. 9, Sep. 2019, doi: 10.1136/bmjopen-2019-030366.

[5] H. Levine *et al.*, "Temporal trends in sperm count: A systematic review and meta-regression analysis," *Hum Reprod Update*, vol. 23, no. 6, pp. 646–659, Nov. 2017, doi: 10.1093/humupd/dmx022.

[6] WHO, "WHO laboratory manual for the examination and processing of human semen Sixth Edition," 2021.

[7] R. A. Hess and L. Renato De Franca, "CHAPTER 1 Spermatogenesis and Cycle of the Seminiferous Epithelium."

[8] C. Staub and L. Johnson, "Review: Spermatogenesis in the bull," Jun. 01, 2018, *Cambridge University Press*. doi: 10.1017/S1751731118000435.

[9] D. M. De Kretser, K. L. Loveland, A. Meinhardt, D. Simorangkir, and N. Wreford, "Spermatogenesis." [Online]. Available: https://academic.oup.com/humrep/article/13/suppl_1/1/788755

[10] D. Mortimer, "The functional anatomy of the human spermatozoon: relating ultrastructure and function," *Mol Hum Reprod*, vol. 24, no. 12, pp. 567–592, Dec. 2018, doi: 10.1093/molehr/gay040.

[11] S. Pesch and M. Bergmann, "Structure of mammalian spermatozoa in respect to viability, fertility and cryopreservation," Oct. 2006. doi: 10.1016/j.micron.2006.02.006.

[12] A. Touré, "Importance of slc26 transmembrane anion exchangers in sperm post-testicular maturation and fertilization potential," 2019, *Frontiers Media S.A.* doi: 10.3389/fcell.2019.00230.

[13] H. M. Florman and R. A. Fissore, "Fertilization in Mammals," in *Knobil and Neill's Physiology of Reproduction: Two-Volume Set*, vol. 1, Elsevier Inc., 2015, pp. 149–196. doi: 10.1016/B978-0-12-397175-3.00004-1.

[14] N. H. Gu, W. L. Zhao, G. S. Wang, and F. Sun, “Comparative analysis of mammalian sperm ultrastructure reveals relationships between sperm morphology, mitochondrial functions and motility,” *Reproductive Biology and Endocrinology*, vol. 17, no. 1, Aug. 2019, doi: 10.1186/s12958-019-0510-y.

[15] S. Chakraborty and S. Saha, “Understanding sperm motility mechanisms and the implication of sperm surface molecules in promoting motility,” Dec. 01, 2022, *Springer Science and Business Media Deutschland GmbH*. doi: 10.1186/s43043-022-00094-7.

[16] M. B. R. Alves, E. C. C. Celeghini, and C. Belleannée, “From Sperm Motility to Sperm-Borne microRNA Signatures: New Approaches to Predict Male Fertility Potential,” Aug. 21, 2020, *Frontiers Media S.A.* doi: 10.3389/fcell.2020.00791.

[17] P. E. Visconti, D. Krapf, J. L. De La Vega-Beltrán, J. J. Acevedo, and A. Darszon, “Ion channels, phosphorylation and mammalian sperm capacitation,” May 2011. doi: 10.1038/aja.2010.69.

[18] A. P. Singh and S. Rajender, “CatSper channel, sperm function and male fertility,” Jan. 01, 2015, *Elsevier Ltd*. doi: 10.1016/j.rbmo.2014.09.014.

[19] S. Young *et al.*, “Human fertilization in vivo and in vitro requires the CatSper channel to initiate sperm hyperactivation,” *Journal of Clinical Investigation*, vol. 134, no. 1, Jan. 2024, doi: 10.1172/JCI173564.

[20] D. T. Carrell and B. R. Emery, “Fertilization,” 2004. doi: Carrell, D. T., & Emery, B. R. (2004). Fertilization. Encyclopedia of Endocrine Diseases, 58–62. doi:10.1016/b0-12-475570-4/00477-7.

[21] S.-K. Jin and W.-X. Yang, “Factors and pathways involved in capacitation: how are they regulated?,” 2017. [Online]. Available: www.impactjournals.com/oncotarget/

[22] F. Y. Nakano, R. de B. F. Leão, and S. C. Esteves, “Insights into the role of cervical mucus and vaginal pH in unexplained infertility,” *Medical Express*, vol. 2, no. 2, 2015, doi: 10.5935/medicalexpress.2015.02.07.

[23] S. Hugentobler, D. G. Morris, M. T. Kane, and J. M. Sreenan, “In situ oviduct and uterine pH in cattle,” *Theriogenology*, vol. 61, no. 7–8, pp. 1419–1427, May 2004, doi: 10.1016/j.theriogenology.2003.08.008.

[24] F. Y. Nakano, R. de B. F. Leão, and S. C. Esteves, “Insights into the role of cervical mucus and vaginal pH in unexplained infertility,” *Medical Express*, vol. 2, no. 2, 2015, doi: 10.5935/medicalexpress.2015.02.07.

[25] L. J. Fauci and R. Dillon, “Biofluidmechanics of Reproduction,” *Annu Rev Fluid Mech*, 2005, doi: 10.1146/annurev.fluid.

[26] B. T. Hinton, J. P. Pryor, A. V. Hirsh, and B. P. Setchell, “The concentration of some inorganic ions and organic compounds in the luminal fluid of the human ductus deferens,” *Int J Androl*, vol. 4, no. 1–6, pp. 457–461, 1981, doi: 10.1111/j.1365-2605.1981.tb00730.x.

[27] M. Lavanya, S. Selvaraju, B. Krishnappa, N. Krishnaswamy, G. Nagarajan, and H. Kumar, “Microenvironment of the male and female reproductive tracts regulate sperm fertility: Impact

of viscosity, pH, and osmolality," Jan. 01, 2022, *John Wiley and Sons Inc*. doi: 10.1111/andr.13102.

[28] Q. Chen and E. K. Duan, "Aquaporins in sperm osmoadaptation: An emerging role for volume regulation," in *Acta Pharmacologica Sinica*, Jun. 2011, pp. 721–724. doi: 10.1038/aps.2011.35.

[29] M. Eisenbach and L. C. Giojalas, "Sperm guidance in mammals - An unpaved road to the egg," Apr. 2006. doi: 10.1038/nrm1893.

[30] D. Mortimer *et al.*, *A Practical Guide to Basic Laboratory Andrology*. Cambridge University Press, 2022. doi: 10.1017/9781009181648.

[31] S. T. Mortimer, "A critical review of the physiological importance and analysis of sperm movement in mammals*," 1997.

[32] T.-W. Su, L. Xue, and A. Ozcan, "High-throughput lensfree 3D tracking of human sperms reveals rare statistics of helical trajectories," *PNAS*, 2012, doi: 10.1073/pnas.1212506109/-/DCSupplemental.

[33] T.-W. Su, L. Xue, and A. Ozcan, "High-throughput lensfree 3D tracking of human sperms reveals rare statistics of helical trajectories", doi: 10.1073/pnas.1212506109/-/DCSupplemental.

[34] E. T. Donnelly, S. E. M. Lewis, J. A. Mcnally, and W. Thompson, "In vitro fertilization and pregnancy rates: the influence of sperm motility and morphology on IVF outcome," *Fertil Steril*, 1998.

[35] N. Gatimel, J. Moreau, J. Parinaud, and R. D. Léandri, "Sperm morphology: assessment, pathophysiology, clinical relevance, and state of the art in 2017," Sep. 01, 2017, *Blackwell Publishing Ltd*. doi: 10.1111/andr.12389.

[36] A. Agarwal *et al.*, "Sperm morphology assessment in the era of intracytoplasmic sperm injection: Reliable results require focus on standardization, quality control, and training," Jun. 17, 2021, *Korean Society for Sexual Medicine and Andrology*. doi: 10.5534/WJMH.210054.

[37] A. García-Molina *et al.*, "Effect of incubation and analysis temperatures on sperm kinematics and morphometrics during human semen analysis," *Rev Int Androl*, vol. 21, no. 2, Apr. 2023, doi: 10.1016/j.androl.2023.100350.

[38] G. & D. P. S. Van Der Horst, "Not just the Marriage of Figaro: But the Marriage of WHO/ESHRE Semen Analysis Criteria with Sperm Functionality," *Advances in Andrology*, 2017, [Online]. Available: http://www.postepyandrologii.pl

[39] R. Finelli, K. Leisegang, S. Tumallapalli, R. Henkel, and A. Agarwal, "The validity and reliability of computer-aided semen analyzers in performing semen analysis: a systematic review," Jul. 01, 2021, *AME Publishing Company*. doi: 10.21037/tau-21-276.

[40] "SCA® Morphology sperm analysis with a CASA System - HT CASA".

[41] S. T. MORTIMER, "CASA—Practical Aspects," *J Androl*, vol. 21, no. 4, pp. 515–524, Jul. 2000, doi: 10.1002/j.1939-4640.2000.tb02116.x.

[42] G. Van Der Horst, L. Maree, and S. S. Du Plessis, “Current perspectives of CASA applications in diverse mammalian spermatozoa,” *Reprod Fertil Dev*, vol. 30, no. 6, pp. 875–888, 2018, doi: 10.1071/RD17468.

[43] L. Larsen, T. Scheike, and T. K. Jensen, “Computer-assisted semen analysis parameters as predictors for fertility of men from the general population,” 2000.

[44] R. O. DAVIS, P. W. NISWANDER, and D. F. KATZ, “New Measures of Sperm Motion I. Adaptive Smoothing and Harmonic Analysis,” *J Androl*, vol. 13, no. 2, pp. 139–152, 1992, doi: 10.1002/j.1939-4640.1992.tb01645.x.

[45] J. C. Lu, Y. F. Huang, and N. Q. Lü, “Computer-aided sperm analysis: Past, present and future,” *Andrologia*, vol. 46, no. 4, pp. 329–338, 2014, doi: 10.1111/and.12093.

[46] A. Valverde, V. Barquero, and C. Soler, “The application of computer-assisted semen analysis (CASA) technology to optimise semen evaluation. A review,” *J Anim Feed Sci*, vol. 29, no. 3, pp. 189–198, Sep. 2020, doi: 10.22358/JAFS/127691/2020.

[47] M. K. Hoogewijs, S. P. De Vliegher, J. L. Govaere, C. De Schauwer, A. De Kruif, and A. Van Soom, “Influence of counting chamber type on CASA outcomes of equine semen analysis,” *Equine Vet J*, vol. 44, no. 5, pp. 542–549, Sep. 2012, doi: 10.1111/j.2042-3306.2011.00523.x.

[48] 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?,” *Fertil Steril*, vol. 96, no. 1, pp. 24–27, Jul. 2011, doi: 10.1016/j.fertnstert.2011.04.096.

[49] M. Elsayed, T. M. El-Sherry, and M. Abdelgawad, “Development of computer-assisted sperm analysis plugin for analyzing sperm motion in microfluidic environments using Image-J,” *Theriogenology*, vol. 84, no. 8, pp. 1367–1377, Nov. 2015, doi: 10.1016/j.theriogenology.2015.07.021.

[50] C. Alquézar-Baeta *et al.*, “OpenCASA: A new open-source and scalable tool for sperm quality analysis,” *PLoS Comput Biol*, vol. 15, no. 1, 2019, doi: 10.1371/journal.pcbi.1006691.

[51] M. T. Gallagher, G. Cupples, E. H. Ooi, J. C. Kirkman-Brown, and D. J. Smith, “Rapid sperm capture: High-throughput flagellar waveform analysis,” *Human Reproduction*, vol. 34, no. 7, pp. 1173–1185, 2019, doi: 10.1093/humrep/dez056.

[52] J. N. Hansen, S. Rassmann, J. F. Jikeli, and D. Wachten, “SpermQ-A simple analysis software to comprehensively study flagellar beating and sperm steering,” *Cells*, vol. 8, no. 1, Jan. 2019, doi: 10.3390/cells8010010.

[53] J. G. Wilson-Leedy and R. L. Ingermann, “Development of a novel CASA system based on open source software for characterization of zebrafish sperm motility parameters,” *Theriogenology*, vol. 67, no. 3, pp. 661–672, Feb. 2007, doi: 10.1016/j.theriogenology.2006.10.003.

[54] E. Giaretta *et al.*, “Implementing an open-access CASA software for the assessment of stallion sperm motility: Relationship with other sperm quality parameters,” *Anim Reprod Sci*, vol. 176, pp. 11–19, Jan. 2017, doi: 10.1016/j.anireprosci.2016.11.003.

[55] S. T. Mortimer, M.-A. Swan, and D-Mortimer, “Fractal analysis of capacitating human spermatozoa*,” 1996. [Online]. Available: https://academic.oup.com/humrep/article/11/5/1049/667467

[56] L. F. Urbano, P. Masson, M. Vermilyea, and M. Kam, “Automatic Tracking and Motility Analysis of Human Sperm in Time-Lapse Images,” *IEEE Trans Med Imaging*, vol. 36, no. 3, pp. 792–801, Mar. 2017, doi: 10.1109/TMI.2016.2630720.

[57] G. Tumuklu Ozyer, B. Ozyer, F. Negin, İ. Alarabi, and S. Agahian, “A hybrid IMM-JPDAF algorithm for tracking multiple sperm targets and motility analysis,” *Neural Comput Appl*, vol. 34, no. 20, pp. 17407–17421, Oct. 2022, doi: 10.1007/s00521-022-07390-3.

[58] Imani Y, Teyfouri N, Ahmadzadeh MR, and Golabbakhsh M., “A New Method for Multiple Sperm Cells Tracking,” *J Med Signals Sens*, Jan. 2015.

[59] H. O. Ilhan, M. Yuzkat, and N. Aydin, “Sperm Motility Analysis by using Recursive Kalman Filters with the smartphone based data acquisition and reporting approach,” *Expert Syst Appl*, vol. 186, Dec. 2021, doi: 10.1016/j.eswa.2021.115774.

[60] J. Yániz *et al.*, “Expanding the limits of computer-assisted sperm analysis through the development of open software,” *Biology (Basel)*, vol. 9, no. 8, pp. 1–16, Aug. 2020, doi: 10.3390/biology9080207.

[61] T. S. Santra and F. G. Tseng, “Single-Cell Analysis,” Aug. 29, 2020, *NLM (Medline)*. doi: 10.3390/cells9091993.

[62] P. Shinde *et al.*, “Current trends of microfluidic single-cell technologies,” Oct. 12, 2018, *MDPI AG*. doi: 10.3390/ijms19103143.

[63] M. Suárez Korsnes and R. Korsnes, “Initial refinement of data from video-based single-cell tracking,” Apr. 26, 2022. doi: 10.1101/2022.04.26.489486.

[64] A. Nousi, M. T. Søgaard, M. Audoin, and L. Jauffred, “Single-cell tracking reveals super-spreading brain cancer cells with high persistence,” *Biochem Biophys Rep*, vol. 28, Dec. 2021, doi: 10.1016/j.bbrep.2021.101120.

[65] D. Dannhauser, M. I. Maremonti, V. Panzetta, D. Rossi, P. A. Netti, and F. Causa, “Mechanical phenotyping of breast cell lines by in-flow deformation-dependent dynamics under tuneable compressive forces,” *Lab Chip*, vol. 20, no. 24, pp. 4611–4622, Dec. 2020, doi: 10.1039/d0lc00911c.

[66] Y. Ai, M. Liang, D. Yang, Y. Zhou, P. Li, and J. Zhong, “Single-cell stretching in viscoelastic fluids with electronically triggered imaging for cellular mechanical phenotyping,” *Anal Chem*, vol. 93, no. 10, pp. 4567–4575, Mar. 2021, doi: 10.1021/acs.analchem.0c05009.

[67] R. Gómez-Villafuertes *et al.*, “Live imaging followed by single cell tracking to monitor cell biology and the lineage progression of multiple neural populations,” *Journal of Visualized Experiments*, vol. 2017, no. 130, Dec. 2017, doi: 10.3791/56291.

[68] J. A. Cornwell *et al.*, “TrackPad: Software for semi-automated single-cell tracking and lineage annotation,” *SoftwareX*, vol. 11, Jan. 2020, doi: 10.1016/j.softx.2020.100440.

[69] S. Chen, S. Zhang, and R. Zhu, “Computer-Vision-Based Dielectrophoresis Mobility Tracking for Characterization of Single-Cell Biophysical Properties,” *Anal Chem*, vol. 94, no. 41, pp. 14331–14339, Oct. 2022, doi: 10.1021/acs.analchem.2c02935.

[70] N. Apichitsopa, A. Jaffe, and J. Voldman, “Multiparameter cell-tracking intrinsic cytometry for single-cell characterization,” *Lab Chip*, vol. 18, no. 10, pp. 1430–1439, May 2018, doi: 10.1039/c8lc00240a.

[71] R. A. Ghayda *et al.*, “Artificial Intelligence in Andrology: From Semen Analysis to Image Diagnostics,” *World Journal of Men’s Health* , vol. 42, no. 1, pp. 39–61, Jan. 2024, doi: 10.5534/wjmh.230050.

[72] S. Sharma, S. Sharma, and A. Athaiya, “ ACTIVATION FUNCTIONS IN NEURAL NETWORKS,” 2020. [Online]. Available: http://www.ijeast.com

[73] R. Pascanu, T. Mikolov, and Y. Bengio, “On the difficulty of training Recurrent Neural Networks.” Accessed: Oct. 01, 2024. [Online]. Available: arXiv:1211.5063

[74] K. Janocha and W. M. Czarnecki, “On Loss Functions for Deep Neural Networks in Classification,” *arXiv:1211.5063*, Feb. 2017, [Online]. Available: http://arxiv.org/abs/1702.05659

[75] M. Andrychowicz *et al.*, “Learning to learn by gradient descent by gradient descent.”

[76] S. Ruder, “An overview of gradient descent optimization algorithms,” Sep. 2016, [Online]. Available: http://arxiv.org/abs/1609.04747

[77] N. Qian, “On the momentum term in gradient descent learning algorithms,” *Neural Networks*, no. 12, 1999.

[78] F. Zou *et al.*, “A Sufficient Condition for Convergences of Adam and RMSProp,” 2019.

[79] “Gradient Descent for Machine Learning (ML) 101 with Python Tutorial _ Towards AI”.

[80] L. Wang and D. Sng, “Deep Learning Algorithms with Applications to Video Analytics for A Smart City: A Survey,” Dec. 2015, [Online]. Available: http://arxiv.org/abs/1512.03131

[81] K. Maharana, S. Mondal, and B. Nemade, “A review: Data pre-processing and data augmentation techniques,” *Global Transitions Proceedings*, vol. 3, no. 1, pp. 91–99, Jun. 2022, doi: 10.1016/j.gltp.2022.04.020.

[82] C. Shorten and T. M. Khoshgoftaar, “A survey on Image Data Augmentation for Deep Learning,” *J Big Data*, vol. 6, no. 1, Dec. 2019, doi: 10.1186/s40537-019-0197-0.

[83] E. Goceri, “Medical image data augmentation: techniques, comparisons and interpretations,” *Artif Intell Rev*, vol. 56, no. 11, pp. 12561–12605, Nov. 2023, doi: 10.1007/s10462-023-10453-z.

[84] O. Rainio, J. Teuho, and R. Klén, “Evaluation metrics and statistical tests for machine learning,” *Sci Rep*, vol. 14, no. 1, Dec. 2024, doi: 10.1038/s41598-024-56706-x.

[85] J. R. Terven and D. M. Cordova-Esparza, “A COMPREHENSIVE REVIEW OF YOLO: FROM YOLOV1 AND BEYOND A PREPRINT,” 2023.

[86] R. G. L. da Silva, “The advancement of artificial intelligence in biomedical research and health innovation: challenges and opportunities in emerging economies,” *Globalization and Health* , vol. 20, no. 1, Dec. 2024, doi: 10.1186/s12992-024-01049-5.

[87] E. al. Syed Qamrul Kazmi, “Image Retrieval Using Auto Encoding Features In Deep Learning,” *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 11, no. 10, pp. 155–171, Nov. 2023, doi: 10.17762/ijritcc.v11i10.8478.

[88] W. B. Verschoof-Van der Vaart and K. Lambers, “Learning to look at LiDAR: The use of R-CNN in the automated detection of archaeological objects in lidar data from the Netherlands,” *Journal of Computer Applications in Archaeology*, vol. 2, no. 1, pp. 31–40, 2019, doi: 10.5334/jcaa.32.

[89] M. Yani, B. Irawan, and C. Setiningsih, “Application of Transfer Learning Using Convolutional Neural Network Method for Early Detection of Terry’s Nail,” in *Journal of Physics: Conference Series*, Institute of Physics Publishing, May 2019. doi: 10.1088/1742-6596/1201/1/012052.

[90] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection.” [Online]. Available: http://pjreddie.com/yolo/

[91] N. Ibtehaz and M. S. Rahman, “MultiResUNet : Rethinking the U-Net Architecture for Multimodal Biomedical Image Segmentation,” 2019.

[92] X. Kang, B. Song, and F. Sun, “A deep similarity metric method based on incomplete data for traffic anomaly detection in IoT,” *Applied Sciences (Switzerland)*, vol. 9, no. 1, Jan. 2019, doi: 10.3390/app9010135.

[93] P. Chhabra and S. Goyal, “A Thorough Review on Deep Learning Neural Network,” in *2023 International Conference on Artificial Intelligence and Smart Communication, AISC 2023*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 220–226. doi: 10.1109/AISC56616.2023.10085166.

[94] H. Gholamalinezhad and H. Khosravi, “Pooling Methods in Deep Neural Networks, a Review.”

[95] J. Ren and Y. Wang, “Overview of Object Detection Algorithms Using Convolutional Neural Networks,” *Journal of Computer and Communications*, vol. 10, pp. 115–132, 2022, doi: 10.4236/jcc.2022.101006.

[96] A. Neubeck ETH Zurich and L. Van Gool ETH Zurich, “Efficient Non-Maximum Suppression,” 2006.

[97] U. Sirisha, S. P. Praveen, P. N. Srinivasu, P. Barsocchi, and A. K. Bhoi, “Statistical Analysis of Design Aspects of Various YOLO-Based Deep Learning Models for Object Detection,” Dec. 01, 2023, *Springer Science and Business Media B.V.* doi: 10.1007/s44196-023-00302-w.

[98] F. Sultana, A. Sufian, and P. Dutta, “Evolution of Image Segmentation using Deep Convolutional Neural Network: A Survey.”

[99] N. N. Ali, N. A. Kako, and A. S. Abdi, “Review on Image Segmentation Methods Using Deep Learning,” in *ICOASE 2022 - 4th International Conference on Advanced Science and*

Engineering, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 7–12. doi: 10.1109/ICOASE56293.2022.10075607.

[100] Y. Guo, Y. Liu, T. Georgiou, and M. S. Lew, "A review of semantic segmentation using deep neural networks," *Int J Multimed Inf Retr*, vol. 7, no. 2, pp. 87–93, Jun. 2018, doi: 10.1007/s13735-017-0141-z.

[101] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," May 2015, [Online]. Available: http://arxiv.org/abs/1505.04597

[102] H. H. Lee *et al.*, "Foundation Models for Biomedical Image Segmentation: A Survey," Jan. 2024, [Online]. Available: http://arxiv.org/abs/2401.07654

[103] M. K. Panner Selvam, A. K. Moharana, S. Baskaran, R. Finelli, M. C. Hudnall, and S. C. Sikka, "Current Updates on Involvement of Artificial Intelligence and Machine Learning in Semen Analysis," Feb. 01, 2024, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/medicina60020279.

[104] P. Cherouveim, C. Velmahos, and C. L. Bormann, "Artificial intelligence for sperm selection—a systematic review," Jul. 01, 2023, *Elsevier Inc.* doi: 10.1016/j.fertnstert.2023.05.157.

[105] S. Shahzad, M. Ilyas, M. I. U. Lali, H. T. Rauf, S. Kadry, and E. A. Nasr, "Sperm Abnormality Detection Using Sequential Deep Neural Network," *Mathematics*, vol. 11, no. 3, Feb. 2023, doi: 10.3390/math11030515.

[106] J. Riordon, C. McCallum, and D. Sinton, "Deep learning for the classification of human sperm," *Comput Biol Med*, vol. 111, Aug. 2019, doi: 10.1016/j.compbiomed.2019.103342.

[107] A. Butola *et al.*, "High spatially sensitive quantitative phase imaging assisted with deep neural network for classification of human spermatozoa under stressed condition," *Sci Rep*, vol. 10, no. 1, Dec. 2020, doi: 10.1038/s41598-020-69857-4.

[108] T. B. Haugen, O. Witczak, S. A. Hicks, L. Björndahl, J. M. Andersen, and M. A. Riegler, "Sperm motility assessed by deep convolutional neural networks into WHO categories," *Sci Rep*, vol. 13, no. 1, Dec. 2023, doi: 10.1038/s41598-023-41871-2.

[109] Priyansi, B. Bhattacharjee, and J. Rahim, "Predicting Semen Motility using three-dimensional Convolutional Neural Networks," Jan. 2021, [Online]. Available: http://arxiv.org/abs/2101.02888

[110] L. Noy, I. Barnea, S. K. Mirsky, D. Kamber, M. Levi, and N. T. Shaked, "Sperm-cell DNA fragmentation prediction using label-free quantitative phase imaging and deep learning," *Cytometry Part A*, vol. 103, no. 6, pp. 470–478, Jun. 2023, doi: 10.1002/cyto.a.24703.

[111] C. McCallum *et al.*, "Deep learning-based selection of human sperm with high DNA integrity," *Commun Biol*, vol. 2, no. 1, Dec. 2019, doi: 10.1038/s42003-019-0491-6.

[112] R. S. Kumar, S. Sharma, A. Halder, and V. Gupta, "Deep Learning-Based Robust Automated System for Predicting Human Sperm DNA Fragmentation Index," *J Hum Reprod Sci*, vol. 16, no. 1, pp. 16–21, Jan. 2023, doi: 10.4103/jhrs.jhrs_4_23.

[113] R. Kumar, S. Sharma, A. Halder, and V. Gupta, “Deep learning-based robust automated system for predicting human sperm DNA fragmentation index,” *J Hum Reprod Sci*, vol. 16, no. 1, p. 16, 2023, doi: 10.4103/jhrs.jhrs_4_23.

[114] P. Hidayatullah *et al.*, “DeepSperm: A robust and real-time bull sperm-cell detection in densely populated semen videos,” *Comput Methods Programs Biomed*, vol. 209, Sep. 2021, doi: 10.1016/j.cmpb.2021.106302.

[115] J. Redmon and A. Farhadi, “YOLOv3: An Incremental Improvement,” Apr. 2018, [Online]. Available: http://arxiv.org/abs/1804.02767

[116] P. Hidayatullah, T. L. E. R. Mengko, R. Munir, and A. Barlian, “Bull Sperm Tracking and Machine Learning-Based Motility Classification,” *IEEE Access*, vol. 9, pp. 61159–61170, 2021, doi: 10.1109/ACCESS.2021.3074127.

[117] S. Zou *et al.*, “TOD-CNN: An effective convolutional neural network for tiny object detection in sperm videos,” *Comput Biol Med*, vol. 146, Jul. 2022, doi: 10.1016/j.compbiomed.2022.105543.

[118] T. Sato *et al.*, “A new deep-learning model using YOLOv3 to support sperm selection during intracytoplasmic sperm injection procedure,” *Reprod Med Biol*, vol. 21, no. 1, Jan. 2022, doi: 10.1002/rmb2.12454.

[119] V. Valiuškaitė, V. Raudonis, R. Maskeliūnas, R. Damaševičius, and T. Krilavičius, “Deep learning based evaluation of spermatozoid motility for artificial insemination,” *Sensors (Switzerland)*, vol. 21, no. 1, pp. 1–14, Jan. 2021, doi: 10.3390/s21010072.

[120] M. Yuzkat, H. O. Ilhan, and N. Aydin, “Detection of sperm cells by single-stage and two-stage deep object detectors,” *Biomed Signal Process Control*, vol. 83, May 2023, doi: 10.1016/j.bspc.2023.104630.

[121] R. Zhu *et al.*, “YOLOv5s-SA: Light-Weighted and Improved YOLOv5s for Sperm Detection,” *Diagnostics*, vol. 13, no. 6, Mar. 2023, doi: 10.3390/diagnostics13061100.

[122] X. Pan *et al.*, “A Sperm Quality Detection System Based on Microfluidic Chip and Micro-Imaging System,” *Front Vet Sci*, vol. 9, Jun. 2022, doi: 10.3389/fvets.2022.916861.

[123] R. A. Movahed, E. Mohammadi, and M. Orooji, “Automatic segmentation of Sperm’s parts in microscopic images of human semen smears using concatenated learning approaches,” *Comput Biol Med*, vol. 109, pp. 242–253, Jun. 2019, doi: 10.1016/j.compbiomed.2019.04.032.

[124] Q. Lv, X. Yuan, J. Qian, X. Li, H. Zhang, and S. Zhan, “An Improved U-Net for Human Sperm Head Segmentation,” *Neural Process Lett*, vol. 54, no. 1, pp. 537–557, Feb. 2022, doi: 10.1007/s11063-021-10643-2.

[125] R. Melendez, C. B. Castanon, and R. Medina-Rodriguez, “Sperm cell segmentation in digital micrographs based on convolutional neural networks using U-net architecture,” in *Proceedings - IEEE Symposium on Computer-Based Medical Systems*, Institute of Electrical and Electronics Engineers Inc., Jun. 2021, pp. 91–96. doi: 10.1109/CBMS52027.2021.00084.

[126] P. Hernández-Herrera, V. Abonza, J. Sánchez-Contreras, A. Darszon, and A. Guerrero, “Deep Learning-Based Classification and Segmentation of Sperm Head and Flagellum for Image-

Based Flow Cytometry," *Computacion y Sistemas*, vol. 27, no. 4, pp. 1133–1145, 2023, doi: 10.13053/CyS-27-4-4772.

[127] Kandel E. M., "Reproductive outcomes predicted by phase imaging with computational specificity of spermatozoon ultrastructure," *PNAS*, 2020, doi: 10.1073/pnas.2001754117/-/DCSupplemental.

[128] G. Liu *et al.*, "Fast Noninvasive Morphometric Characterization of Free Human Sperms Using Deep Learning," *Microscopy and Microanalysis*, vol. 28, no. 5, pp. 1767–1779, Oct. 2022, doi: 10.1017/S1431927622012132.

[129] H. Yang *et al.*, "Multidimensional morphological analysis of live sperm based on multiple-target tracking," *Comput Struct Biotechnol J*, vol. 24, pp. 176–184, Dec. 2024, doi: 10.1016/j.csbj.2024.02.025.

[130] K. Ben-Yehuda *et al.*, "Simultaneous Morphology, Motility, and Fragmentation Analysis of Live Individual Sperm Cells for Male Fertility Evaluation," *Advanced Intelligent Systems*, vol. 4, no. 4, Apr. 2022, doi: 10.1002/aisy.202100200.

[131] W. Dai *et al.*, "Automated Non-Invasive Analysis of Motile Sperms Using Sperm Feature-Correlated Network," *IEEE Transactions on Automation Science and Engineering*, 2024, doi: 10.1109/TASE.2024.3404488.

[132] A. Valverde, O. Castro-Morales, M. Madrigal-Valverde, and C. Soler, "Sperm kinematics and morphometric subpopulations analysis with CASA systems: a review," 2019.

[133] A. Quintero-Moreno, T. Rigau, and J. E. Rodríguez-Gil, "Multivariate cluster analysis regression procedures as tools to identify motile sperm subpopulations in rabbit semen and to predict semen fertility and litter size," *Reproduction in Domestic Animals*, vol. 42, no. 3, pp. 312–319, Jun. 2007, doi: 10.1111/j.1439-0531.2006.00785.x.

[134] J. L. Yániz, I. Palacín, K. S. Caycho, C. Soler, M. A. Silvestre, and P. Santolaria, "Determining the relationship between bull sperm kinematic subpopulations and fluorescence groups using an integrated sperm quality analysis technique," 2018, *CSIRO*. doi: 10.1071/RD17441.

[135] F. Martinez-Pastor, A. R. Diaz-Corujo, E. Anel, P. Herraez, L. Anel, and P. De Paz, "Post mortem time and season alter subpopulation characteristics of Iberian red deer epididymal sperm," *Theriogenology*, vol. 64, no. 4, pp. 958–974, Sep. 2005, doi: 10.1016/j.theriogenology.2005.01.003.

[136] I. Núñez-Martínez, J. M. Moran, and F. J. Peña, "A three-step statistical procedure to identify sperm kinematic subpopulations in canine ejaculates: Changes after cryopreservation," *Reproduction in Domestic Animals*, vol. 41, no. 5, pp. 408–415, Oct. 2006, doi: 10.1111/j.1439-0531.2006.00685.x.

[137] I. Yánez-Ortiz, W. García, A. Tabarez, and M. J. Palomo, "The contribution of different sperm parameters to better explain ram semen cryoresistance," *Veterinary Journal*, vol. 306, Aug. 2024, doi: 10.1016/j.tvjl.2024.106194.

[138] F. J. Peña *et al.*, "Identification of sperm morphometric subpopulations in two different portions of the boar ejaculate and its relation to postthaw quality," *J Androl*, vol. 26, no. 6, pp. 716–723, Nov. 2005, doi: 10.2164/jandrol.05030.

[139] L. Víquez, V. Barquero, C. Soler, E. R. S. Roldan, and A. Valverde, “Kinematic subpopulations in bull spermatozoa: A comparison of classical and bayesian approaches,” *Biology (Basel)*, vol. 9, no. 6, pp. 1–16, Jun. 2020, doi: 10.3390/biology9060138.

[140] A. Valverde, O. Castro-Morales, M. Madrigal-Valverde, and C. Soler, “Sperm kinematics and morphometric subpopulations analysis with CASA systems: a review,” 2019.

[141] C. Soler *et al.*, “Sperm kinematic, head morphometric and kinetic-morphometric subpopulations in the blue fox (Alopex lagopus),” *Asian J Androl*, vol. 19, no. 2, pp. 154–159, 2017, doi: 10.4103/1008-682X.188445.

[142] F. Murtagh and P. Contreras, “Algorithms for hierarchical clustering: An overview,” *Wiley Interdiscip Rev Data Min Knowl Discov*, vol. 2, no. 1, pp. 86–97, Jan. 2012, doi: 10.1002/widm.53.

[143] A. M. Ikotun, A. E. Ezugwu, L. Abualigah, B. Abuhaija, and J. Heming, “K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data,” *Inf Sci (N Y)*, vol. 622, pp. 178–210, Apr. 2023, doi: 10.1016/j.ins.2022.11.139.

[144] F. Martínez-Pastor, E. J. Tizado, J. J. Garde, L. Anel, and P. de Paz, “Statistical Series: Opportunities and challenges of sperm motility subpopulation analysis,” Mar. 15, 2011. doi: 10.1016/j.theriogenology.2010.11.034.

[145] F. Martínez-Pastor, E. J. Tizado, J. J. Garde, L. Anel, and P. de Paz, “Statistical Series: Opportunities and challenges of sperm motility subpopulation analysis,” Mar. 15, 2011. doi: 10.1016/j.theriogenology.2010.11.034.

[146] I. Ibanescu, M. Siuda, and H. Bollwein, “Motile sperm subpopulations in bull semen using different clustering approaches – Associations with flow cytometric sperm characteristics and fertility,” *Anim Reprod Sci*, vol. 215, Apr. 2020, doi: 10.1016/j.anireprosci.2020.106329.

[147] T. Peng *et al.*, “Machine learning-based clustering to identify the combined effect of the DNA fragmentation index and conventional semen parameters on in vitro fertilization outcomes,” *Reproductive Biology and Endocrinology*, vol. 21, no. 1, Dec. 2023, doi: 10.1186/s12958-023-01080-y.

[148] S. Vicente-Fiel, I. Palacín, P. Santolaria, and J. L. Yániz, “A comparative study of sperm morphometric subpopulations in cattle, goat, sheep and pigs using a computer-assisted fluorescence method (CASMA-F),” *Anim Reprod Sci*, vol. 139, no. 1–4, pp. 182–189, Jun. 2013, doi: 10.1016/j.anireprosci.2013.04.002.

[149] A. García-Molina *et al.*, “Human kinematic and morphometric sperm subpopulation analysis using CASA technology: A new approach to spermatozoa classification,” *Rev Int Androl*, vol. 20, no. 4, pp. 257–265, Oct. 2022, doi: 10.1016/j.androl.2021.05.003.

[150] L. F. Canonico *et al.*, “Exploring altered bovine sperm trajectories by sperm tracking in unconfined conditions,” *Front Vet Sci*, vol. 11, 2024, doi: 10.3389/fvets.2024.1358440.

[151] C. Yuan *et al.*, “Effects of permethrin, cypermethrin and 3-phenoxybenzoic acid on rat sperm motility in vitro evaluated with computer-assisted sperm analysis,” *Toxicology in Vitro*, vol. 24, no. 2, pp. 382–386, Mar. 2010, doi: 10.1016/j.tiv.2009.11.001.

[152] Jean-Yves Tinevez, “Simpletracker,” 2023.

[153] Z. Jiang, L. Zhao, S. Li, Y. Jia, and Z. Liquan, “Real-time object detection method for embedded devices.”

[154] S. L. N. and E. A. B. da S. R. Padilla, “A Survey on Performance Metrics for Object-Detection Algorithms,” in *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*, IEEE, 2020.

[155] “COCO - Common Objects in Context,” https://cocodataset.org/#detection-eval.

[156] C. H. Sudre, W. Li, T. Vercauteren, S. Ourselin, and M. J. Cardoso, “Generalised Dice overlap as a deep learning loss function for highly unbalanced segmentations,” Jul. 2017, doi: 10.1007/978-3-319-67558-9_28.

[157] J. Barkay, H. Zuckerman, M. D. D. Sklan, S. Gordon, and J. Barkay, “EFFECT OF CAFFEINE ON INCREASING THE MOTILITY OF FROZEN HUMAN SPERM,” 1977.

[158] V. Torres-Flores, G. Picazo-Juárez, Y. Hernández-Rueda, A. Darszon, and M. T. Gonzlez-Martínez, “Sodium influx induced by external calcium chelation decreases human sperm motility,” *Human Reproduction*, vol. 26, no. 10, pp. 2626–2635, 2011, doi: 10.1093/humrep/der237.

[159] A. F. Ferreira, J. Santiago, J. V. Silva, P. F. Oliveira, and M. Fardilha, “PP1, PP2A and PP2B Interplay in the Regulation of Sperm Motility: Lessons from Protein Phosphatase Inhibitors,” Dec. 01, 2022, *MDPI*. doi: 10.3390/ijms232315235.

[160] K. H. Zou *et al.*, “Statistical Validation of Image Segmentation Quality Based on a Spatial Overlap Index 1 : Scientific Reports,” 2004. [Online]. Available: http://www.slicer.org

[161] S. Vicente-Fiel *et al.*, “A comparative study of the sperm nuclear morphometry in cattle, goat, sheep, and pigs using a new computer-assisted method (CASMA-F),” *Theriogenology*, vol. 79, no. 3, pp. 436–442, Feb. 2013, doi: 10.1016/j.theriogenology.2012.10.015.

[162] J. L. Rubio-Guillén *et al.*, “HEAD DIMENSIONS OF BRAHMAN AND THEIR CROSSBRED BULL SPERMATOZOA ARE AFFECTED BY CRYOPRESERVATION. Las Dimensiones de la Cabeza del Espermatozoide de Toros Brahman y sus Mestizos son Afectadas por la Criopreservación,” 2007.

[163] S. Kondracki, D. Banaszewska, A. Wysokińska, and J. Chomicz, “Sperm morphology of cattle and domestic pigs,” 2006.

[164] J. Yániz, S. Capistrós, S. Vicente-Fiel, C. Hidalgo, and P. Santolaria, “A comparative study of the morphometry of sperm head components in cattle, sheep, and pigs with a computer-assisted fluorescence method,” *Asian J Androl*, vol. 18, no. 6, pp. 840–843, Nov. 2016, doi: 10.4103/1008-682X.186877.

[165] M. Zubair, M. Ahmad, and H. Jamil, “Review on the screening of semen by hypo-osmotic swelling test,” Sep. 01, 2015, *Blackwell Publishing Ltd*. doi: 10.1111/and.12335.

[166] Z. Liu and R. H. Foote, “Bull Sperm Motility and Membrane Integrity in Media Varying in Osmolality,” *J Dairy Sci*, vol. 81, no. 7, pp. 1868–1873, 1998, doi: 10.3168/jds.S0022-0302(98)75757-1.

[167] M. Rossato, G. Balercia, G. Lucarelli, C. Foresta, and F. Mantero, “Role of seminal osmolarity in the regulation of human sperm motility,” *Int J Androl*, vol. 25, no. 4, pp. 230–235, 2002, doi: 10.1046/j.1365-2605.2002.00353.x.

[168] M. Lavanya *et al.*, “Sperm preparedness and adaptation to osmotic and pH stressors relate to functional competence of sperm in Bos taurus,” *Sci Rep*, vol. 11, no. 1, Dec. 2021, doi: 10.1038/s41598-021-01928-6.

[169] E. Pinart, “Ion Channels of Spermatozoa: Structure, Function, and Regulation Mechanisms,” Jun. 01, 2022, *MDPI*. doi: 10.3390/ijms23115880.

[170] R. H. RIFFENBURGH, “Statistics in Medicine (Second edition),” 2006. doi: 10.1016/B978-0-12-088770-5.50067-8.

[171] J. R. Pariz and J. Hallak, “Effects of caffeine supplementation in post-thaw human semen over different incubation periods,” *Andrologia*, vol. 48, no. 9, pp. 961–966, Nov. 2016, doi: 10.1111/and.12538.

[172] N. de C. Alves *et al.*, “Addition of caffeine to equine thawed sperm increases motility and decreases nitrite concentration,” *Andrologia*, vol. 53, no. 2, Mar. 2021, doi: 10.1111/and.13918.

[173] N. Nabavi, F. Todehdehghan, A. Shiravi, and C. Author, “Effect of caffeine on motility and vitality of sperm and in vitro fertilization of outbreed mouse in T6 and M16 media,” 2013.

[174] S. A. Banihani and H. J. Khaled, “Caffeine increased progressive motility of human spermatozoa in normozoospermic and asthenozoospermic semen samples and enhanced activity of seminal creatine kinase,” *Andrologia*, vol. 53, no. 6, Jul. 2021, doi: 10.1111/and.14052.

[175] C. Colás, J. A. Cebrián-Pérez, and T. Muiño-Blanco, “Caffeine induces ram sperm hyperactivation independent of cAMP-dependent protein kinase,” *Int J Androl*, vol. 33, no. 1, Feb. 2010, doi: 10.1111/j.1365-2605.2009.00991.x.

[176] J. Correia, F. Michelangeli, and S. Publicover, “Regulation and roles of Ca2+ stores in human sperm,” Aug. 01, 2015, *BioScientifica Ltd.* doi: 10.1530/REP-15-0102.

[177] C. Riquelme, P. Olivares, A. Ramirez, and J. Parodi, “Regulation of Cellular Calcium, for Control of Motility and Reduction of Mortality of Bovine Sperm,” *J Vet Sci Technol*, vol. 09, no. 02, 2018, doi: 10.4172/2157-7579.1000521.

[178] D. P. Wolf, Phd. Luis Blasco, Md. A. Mohammad Khan, and M. Liti, “HUMAN CERVICAL MUCUS. II. CHANGES IN VISCOELASTICITY DURING THE OVULATORY MENSTRUAL CYCLE*,” 1977.

[179] J. Elia *et al.*, “Human semen hyperviscosity: Prevalence, pathogenesis and therapeutic aspects,” *Asian J Androl*, vol. 11, no. 5, pp. 609–615, 2009, doi: 10.1038/aja.2009.46.

[180] F. Yazdan Parast *et al.*, “Viscous Loading Regulates the Flagellar Energetics of Human and Bull Sperm,” *Small Methods*, vol. 8, no. 7, Jul. 2024, doi: 10.1002/smtd.202300928.

[181] M. Zaferani, S. S. Suarez, and A. Abbaspourrad, “Mammalian sperm hyperactivation regulates navigation via physical boundaries and promotes pseudo-chemotaxis,” *Biology, biophysics*

and computational biology, vol. 118, Oct. 2021, doi: 10.1073/pnas.2107500118/-/DCSupplemental.

[182] N. Chernov, "CircleFitbyPratt," MATLAB Central. Accessed: Oct. 21, 2024. [Online]. Available: https://it.mathworks.com/matlabcentral/fileexchange/22643-circle-fit-pratt-method