Sperm Tracking Tool for Predicting sperm motion and Concentration using Deep Learning Models

Akshay Valsaraj
Bits Pilani Kk Birla Goa Campus
f20180608@goa.bits-pilani.ac.in

Ithihas Madala
Bits Pilani Kk Birla Goa Campus
f20180607@goa.bits-pilani.ac.in

Prof. Siddhartha Tripathi
Department of Mechanical Engineering
Bits Pilani Kk Birla Goa Campus
siddharthat@goa.bits-pilani.ac.in

Prof. Arnab Banerjee
Department of Biological Sciences
Bits Pilani Kk Birla Goa Campus
arnabb@goa.bits-pilani.ac.in

Abstract

There have been global concerns regarding falling rates of infertitlity among men and there has been an increase of couples seeking medical assistance for reproduction.The analysis of the semen sample is usually performed by clinician who calculates the sperm concentraion , motility, morphology , volumne , appearence ,pH, viscosoty, Percentage of motile and non motile sperms and other parameters.Manual evaluation of a sample using a microscope takes a lot of time and it requires extensive traning and practise for the clinican. Morever the results obtained from the clinician have limited reproducibilty and it can have lot of variations between different clinicians. To overcome these issues we have come up with a deep learning model that can predict the concentraion of the sperm sample along with the prediction of the path and other paramters of the most viable sperms using Deep learning algorithms.

1 Introduction

There has been increased attension over the decreased trend in male reproductivity over the years[2]. infertitlity analysis includes semen analysis but the process has been quite uncertain.According to the current standards, semen analysis is usually done according to the recommendations provided by the WHO which includes detailed analysis of assessing semen volume, sperm concentration, total sperm count, sperm motility, sperm morphology, and sperm vitality [1]. Sperm Concentraion and total number of spermatozoa are important factor for conception.Semen evalution is required to calculate the number of spermatozoa from the ejaculate.Sperm Concentration is not a measure of testicular function but it is determined by the seminal vesicles and prostrate secretions [9]. There has been a huge trend in developing computer aided softwares for sperm analysis as earlier it was very difficult to distingusih between spermatazoa from the debris. Hence these softwares could be used for routine diagnositic applications provided sufficent care has been taken in preparing the samples and using the instrument. Eventhough there have been algorithms to idenity the heads and the tail regions of the sperm [6] , there has been a shift of focus of using Deep learning methods for visual inspection of videos and images in the clincial institutes with a wide range of applications from celluar classification and tracking [12], microscopy image enhancement , cancer and disease diagnositics and prognosis [8] [4].
Recent Research have been used to predict the total number of sperms in a glass slide microscope and also predict the motion , motility and morphology using deep learning algorithms [3] [7] [11] [10]. While conventional CASA systems use digital microscopes with phase-contrast accessories,

Preprint. Under review.

producing higher contrast images, we have used raw semen samples (no staining materials) and a regular light microscope, with a digital camera directly attached to its eyepiece, to insure cost benefits and simple assembling of the system. However, since the accurate finding of sperms in the semen image is the first step in the examination and analysis of the semen, any error in this step can affect the outcome of the analysis. This article tends to show that traditional deep learning algorithms can work in low contrast images too. Acording to the WHO manual the parameters used for identifying the motion of the sperm are :

1. VCL, curvilinear velocity (nm/s). Time-averaged velocity of a sperm head along its actual curvilinear path, as perceived in two dimensions in the microscope. A measure of cell vigour.
2. VSL, straight-line (rectilinear) velocity (nm/s). Time-averaged velocity of a sperm head along the straight line between its first detected position and its last.

There are other parameters that exist such as VAP(average path velocity), ALH (amplitude of lateral head displacement (nm). Magnitude) and many more but we will focus on only these two parameters for now.

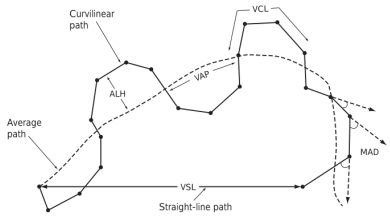

Figure 1: **Left** : Standard terminology for variables measured by CASA systems

Sperm Concentraion is measured with the help of a haemocytometer chamber and allowing it to settle. The sperm sample needs to be well mixed and diluted and the slide needs to be covered by a cover slip and the results needs to be taken within 10-15 minutes to minimize the effects of evaporation.Morever 200 spermatozoa needs to be counted per replicate. If the replicate counts are close they can be accepted otherwise we need to prepare new dilutions.After that the concentraions are calculated per ml. haemocytometer has two separate counting chambers, each of which has a microscopic 3 mm × 3 mm pattern of gridlines etched on the glass surface. It is used with a special thick coverslip (thickness number 4, 0.44 mm), which lies over the grids and is supported by glass pillars 0.1 mm above the chamber fl oor. Each counting area is divided into nine 1 mm × 1 mm grids.

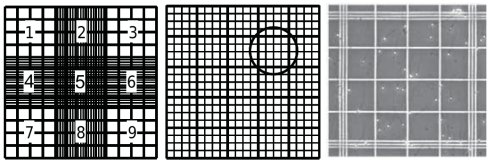

Figure 2: **Left** : Sketches of the inscribed area showing: all nine grids in one chamber of the haemocytometer (left panel); the central grid (number 5) of 25 large squares (middle panel); and a micrograph of part of a fi lled chamber (right panel), showing one of the 25 squares of the central grid (the circled square in the middle panel) bounded by triple lines and containing 16 smaller squares.

2 Methods

Datasets

In this project we used 4 videos for traning , 2 of the videos were taken directly from a sperm sample under a microscope covered under a cover slip. While the other 2 sperm videos are taken from a sperm sample under a microscope in a glassslide along with the hemocytometer.All the samples have been taken from one individual. Morever the validation data contains 1 data from the glass slide

sample and one more data from the glass slide along with the hemocytometer. The validation data is used to tune the deep learning model while its training.

Data type	Sperm sample on glass slid	Sperm sample on glass slide along with the hemocytometer grid
TRAIN	2	2
VALIDATION	1	1
TEST	1	6

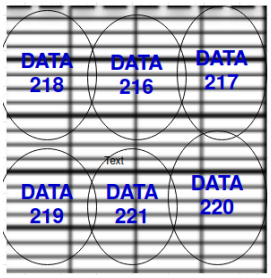

Figure 3: The videos taken from grid 4 to predict the sperm concentraion using deep learning methods

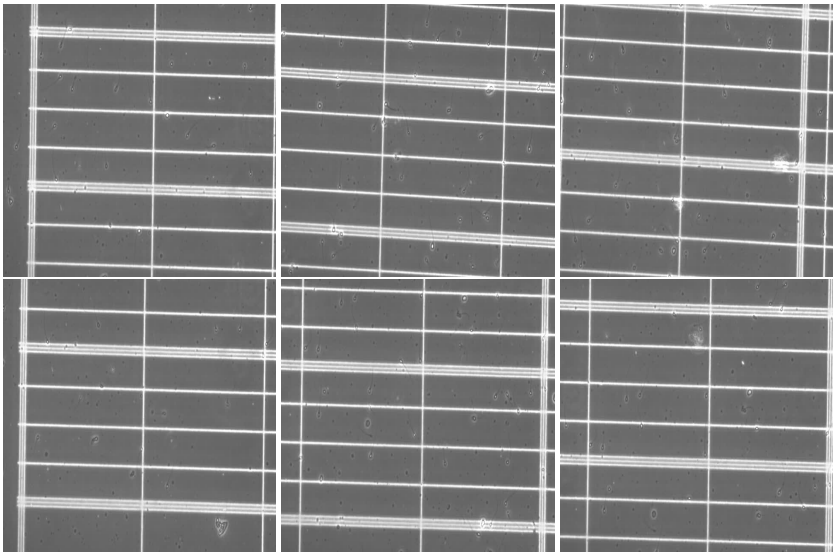

Figure 4: **Top**:First frame of the video of Data 218, 216 and 217. **Bottom** : First frame of the video of Data 219, 221 and 220.

we annotated 4 sperm videos each of 3-5 seconds and it was taken at 30frames per second and we took 100 frames from each and we annoted the sperm heads in all of them.There were about 20-30 sperm heads on an average in each video.So a total of 2000-3000 sperm were annotated and the sperm sample on glass slide along with the hemocytometer was used to predict the sperm concentraion while the sperm sample on the glass slides was used to predic the sperm motion. The test data for the sperm concentraion contains 6 videos and each form an approximate part of the hemocytometer while the test data for the sperm motion contains data from the sperm sample taken from a glass slide. The inital sperm concentration was very dense so we decided to go for a 1:5 dilution using PBS(phosphate-buffered saline) and according to the WHO manual , grids 4,5,6 was used for counting the sperms and a minimum of 200 was counted in it with 15 rows in total.While it

was difficult to take a video of grids 4,5 and 6 hence for the deep learning algorithm we decided to take only grid 4 as each videos were 1GB in size and data transfer was time consuming so we took 6 videos that provide an approximate overview of grid 4 which contains 5 rows.

The videos initally had the resoluton of $1920x1280$, they were resized using a python script and made into the resolution of $640x640$ by cubic interpolation to minimize the model computaion memory and to maintain consistency between all the data.

Models

For predicting the sperm concentration we decided to use the YOLOv5[5] algorithm , YOLO is an abbreviation for the term 'You Only Look Once'. This is an algorithm that detects and recognizes various objects in a picture (in real-time). Object detection in YOLO is done as a regression problem and provides the class probabilities of the detected images. YOLO algorithm employs convolutional neural networks (CNN) to detect objects in real-time. As the name suggests, the algorithm requires only a single forward propagation through a neural network to detect objects.This means that prediction in the entire image is done in a single algorithm run. The CNN is used to predict various class probabilities and bounding boxes simultaneously. The YOLO algorithm consists of various variants. Some of the common ones include tiny YOLO, YOLOv3, YOLOv4 and YOLOv5. For the purpose of this study, we have chosen to go with YOLOv5 for the object detection.

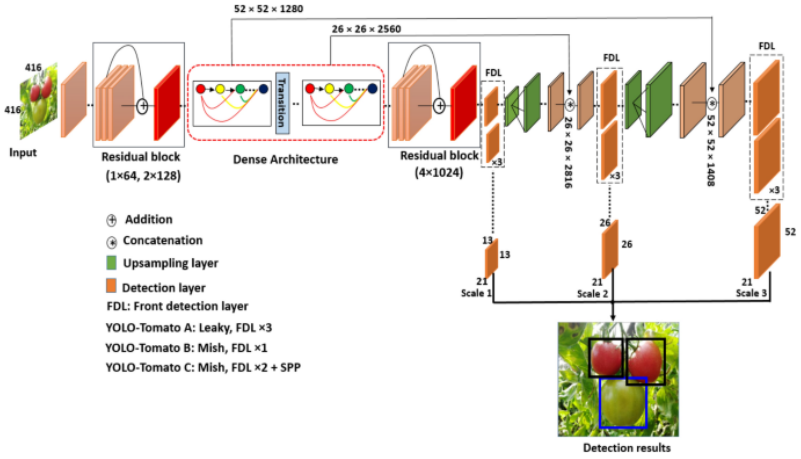

Figure 5: An example of the model used to create bounding boxes to detect fruits.In this case the bounding boxes would be created for the sperm heads.

3 Results

Sperm Concentraion

Using the yolo algorithm we predicted the sperm heads in each frame of the test set data containing the hemocytometer and took an average of the number of sperms in the video by counting the sperm heads in each frame and dividing the total number of frames which in this case is 100.

For the first case we manually calcuated the number of sperms in each of grids 4,5,6 using the method given below. For 1:5 diluton

Replicate 1 of total sperm count = 228
Replicate 2 of total sperm count = 200
Total number of rows$(n) = 30(15 + 15)$

Concentraion C is calcualted by the Formula :
$(N/(n) * (1/(v))) * (d)$
where
N = Total number of sperm including the 2 replicas
n = total rows used for measurement

v = Volume of each row
d = dilution factor used

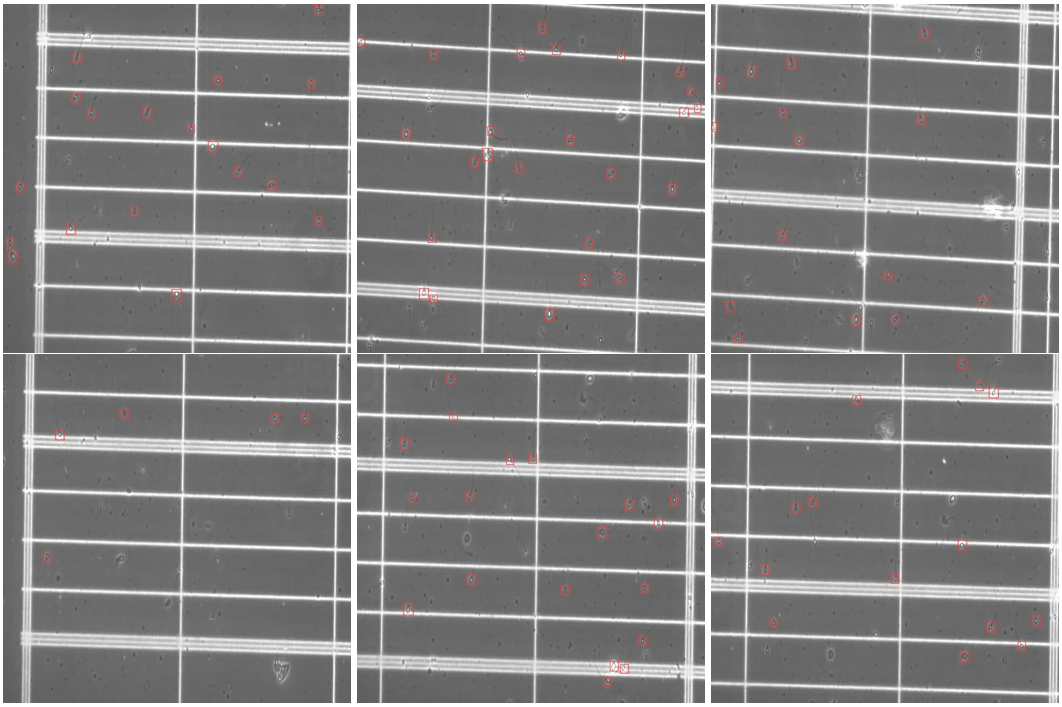

Figure 6: **Top**:First frame of the predicted sperm heads in the video of Data 218, 216 and 217. **Bottom** : First frame of the predicted sperm heads in the video of video of Data 219, 221 and 220.

Hence the calculation is :

$N = 228 + 200 = 428$
$d = 5$
$v = 20$

$C = (428/30) * (1/20) * (5)$
$C = 3.56/nl$
Hence the concentration by manual Calculation is 3.56x10^6 sperms per milli litre
To cross check if the values are correct or not , we take the difference between them
$228 - 220 = 28$

Sum	Acceptable Difference*
144–156	24
157–169	25
170–182	26
183–196	27
197–211	28
212–226	29
227–242	30
243–258	31
259–274	32
275–292	33
293–309	34
310–328	35

Sum	Acceptable Difference*
329–346	36
347–366	37
367–385	38
386–406	39
407–426	40
427–448	41
449–470	42
471–492	43
493–515	44
516–538	45
539–562	46
563–587	47

Figure 7: : Acceptable differences between two replicate counts for a given sum

and we look up at the table and since the calculated different is less that the acceptalble difference , Hence out values are correct.
We decided to use the same method to calculate the sperm concentration using the deep learning method .For this we decided to take a video of each of the

Hence from this we got the following values:
Sperms in Data218 : 18
Sperms in Data219 : 5
Sperms in Data216 : 25
Sperms in Data221: 19
Sperms in Data217: 15
Sperms in Data220: 15
Total Sperms = 97

For 1:5 diluton

Total number of rows(n) $= 5$

Concentraion C is calcualted by the Formula :
$(N/(n) * (1/(v))) * (d)$
where
N = Total number of sperm including the 2 replicas
n = total rows used for measurement
v = Volume of each row
d = dilution factor used

Hence the calculation is :

$N = 97$
$d = 5$
$v = 20$

$C = (97/5) * (1/20) * (5)$
$C = 4.85/nl$
Hence the concentration by the Deep Learning method is 4.85x10^6 sperms per milli litre , since the total number of sperms in close to 100 hence the results would have a sampling error of 10% according to the WHO manual.

Sperm Motion

For predicting the VCL and VSL for the top three most viable sperm , three videos of the sperm sample from the same individual were taken on a glass slide and two of the videos were annoted and trained and test on the third one.The current algorithm tracks the changes of the sperms present in the first frame to the last frame.

The algorithm goes through each frame and records the cordinates of each sperm in the frame as it traverses through the video If the distance between a sperm in frame i to the frame $i + 1$ is more than 0.05 then its a new sperm or the sperm went out of frame or it wasnt detected by the algorithm.

Algorithm 1 Sperm motin capture

initialize $SpermDictonary \leftarrow$ Contains the Cordinates of all the sperms of frame0 along with their distances initialised to 0 in $totaldistance$.
for i = 1 to $frameN$ **do**
  for All $sperm_cordinates$ in $frame[i]$ **do**
    initialize $max \leftarrow 100$
    for key in $SpermDictonary$ **do**
      $distance$ = function to calculate distance between last cordinates of sperm in the key and $sperm_cordinates$
      if $distance \leq max$ **then**
        $max \leftarrow distance$
        $index \leftarrow key$
        $predictedx \leftarrow x$
        $predictedy \leftarrow y$
      end if
    end for
    if $max \leq 0.05$ **then**
      Append the sperm in the $SpermDictonary$ with the new cordinate and add max to the $totaldistance$.
    end if
  end for
end for

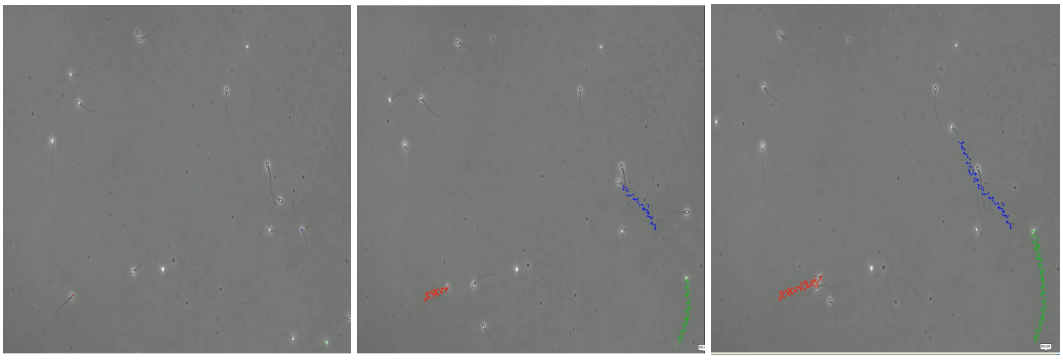

Figure 8: **Left**:Tracked motion of the top 3 sperms in frame 0, **Center** : Tracked motion of the top 3 sperms in frame 50, **Right** : Tracked motion of the top 3 sperms in frame 100

Each pixel in the frame was considered to be 0.5 microns and let the Top 3 viable sperms be named as sperm red, blue and green.

Sperm type	VCL(in micrometer per second)	VSL(in micrometer per second)
Red	94.63	9.06
Green	68.29	33.3
Blue	81.94	36.14

4 Conclusion

Through this study we tend to show that deep learning tool can be used to predict parameters of sperm viablity. These results should be still considered as preliminary until wide scale generalizability of these models are established. This study provides a proof of concept that needs careful consideration before clinical implementation. It should be noted that application of these models should be limited to helping the clinical team in decision making, and not to replace their role entirely.

5 Future Work

For our current study we have worked with only one subject hence multi subject studies can improve the generalizability of the models , morever the sperm motion algorithm may fail to detect sperms that enter after the first frame and this has to be improved , morever we cannot always relay on the model to predict the sperm heads at all times hence a fail safe needs to be implemented such that the algorithm can track the sperm head even if the model didnt detect it based on the previous frames and successive frames.
Morever our model currently uses only the sperm head , this could be expanded to include the dynamics of the tail which could help in improving the predictability of sperm motion and dynamics. We also provide the data for the annotated frames so that future researchers can continue on the project.

References

[1] *WHO laboratory manual for the examination and processing of human semen*.

[2] Elisabeth Carlsen, Aleksander Giwercman, Niels Keiding, and Niels E. Skakkebæk. Evidence for decreasing quality of semen during past 50 years. *BMJ: British Medical Journal*, 305 (6854):609–613, 1992. ISSN 09598138, 17561833. URL `http://www.jstor.org/stable/29716905`.

[3] Priyanto Hidayatullah, Xueting Wang, Toshihiko Yamasaki, Tati L.E.R. Mengko, Rinaldi Munir, Anggraini Barlian, Eros Sukmawati, and Supraptono Supraptono. Deepsperm: A robust and real-time bull sperm-cell detection in densely populated semen videos. *Computer Methods and Programs in Biomedicine*, 209:106302, 2021. ISSN 0169-2607. doi: https://doi.org/10.1016/j.cmpb.2021.106302. URL `https://www.sciencedirect.com/science/article/pii/S016926072100376X`.

[4] Hyungsoon Im, Divya Pathania, Philip J McFarland, Aliyah R Sohani, Ismail Degani, Matthew Allen, Benjamin Coble, Aoife Kilcoyne, Seonki Hong, Lucas Rohrer, Jeremy S Abramson, Scott Dryden-Peterson, Lioubov Fexon, Misha Pivovarov, Bruce Chabner, Hakho Lee, Cesar M Castro, and Ralph Weissleder. Design and clinical validation of a point-of-care device for the diagnosis of lymphoma via contrast-enhanced microholography and machine learning. *Nat. Biomed. Eng.*, 2(9):666–674, September 2018.

[5] Glenn Jocher, Alex Stoken, Ayush Chaurasia, Jirka Borovec, NanoCode012, TaoXie, Yonghye Kwon, Kalen Michael, Liu Changyu, Jiacong Fang, Abhiram V, Laughing, tkianai, yxNONG, Piotr Skalski, Adam Hogan, Jebastin Nadar, imyhxy, Lorenzo Mammana, AlexWang1900, Cristi Fati, Diego Montes, Jan Hajek, Laurentiu Diaconu, Mai Thanh Minh, Marc, albinxavi, fatih, oleg, and wanghaoyang0106. ultralytics/yolov5: v6.0 - YOLOv5n 'Nano' models, Roboflow integration, TensorFlow export, OpenCV DNN support, October 2021. URL `https://doi.org/10.5281/zenodo.5563715`.

[6] Clement Leung, Zhe Lu, Navid Esfandiari, Robert F. Casper, and Yu Sun. Detection and tracking of low contrast human sperm tail. In *2010 IEEE International Conference on Automation Science and Engineering*, pages 263–268, 2010. doi: 10.1109/COASE.2010.5584613.

[7] Christopher McCallum, Jason Riordon, Yihe Wang, Tian Kong, Jae Bem You, Scott Sanner, Alexander Lagunov, Thomas G Hannam, Keith Jarvi, and David Sinton. Deep learning-based selection of human sperm with high DNA integrity. *Commun. Biol.*, 2(1):250, July 2019.

[8] Nicolas Pavillon, Alison J. Hobro, Shizuo Akira, and Nicholas I. Smith. Noninvasive detection of macrophage activation with single-cell resolution through machine learning. *Proceedings of the National Academy of Sciences*, 115(12):E2676–E2685, 2018. ISSN 0027-8424. doi: 10.1073/pnas.1711872115. URL `https://www.pnas.org/content/115/12/E2676`.

[9] JOHN P. PRYOR. 3 - seminal analysis. *Clinics in Obstetrics and Gynaecology*, 8(3):571–585, 1981. ISSN 0306-3356. doi: https://doi.org/10.1016/S0306-3356(21)00169-2. URL `https://www.sciencedirect.com/science/article/pii/S0306335621001692`.

[10] Mohammad reza Mohammadi, Mohammad Rahimzadeh, and Abolfazl Attar. Sperm detection and tracking in phase-contrast microscopy image sequences using deep learning and modified csr-dcf, 2020.

[11] Jason Riordon, Christopher McCallum, and David Sinton. Deep learning for the classification of human sperm. *Computers in Biology and Medicine*, 111:103342, 2019. ISSN 0010-4825. doi: https://doi.org/10.1016/j.compbiomed.2019.103342. URL `https://www.sciencedirect.com/science/article/pii/S0010482519302112`.

[12] D P Sullivan, C F Winsnes, and L Åkesson. Deep learning is combined with massive-scale citizen science to improve large-scale image classification. *Nat Biotechnol*, 2018;36(9):820-828: 4225, 1038/nbt.