Tracking and Prediction Of Human Spermatozoa Motility Using Yolov8n with Greedy Shape Geometry Technique

Muhammad Osaid[1,*,†], Abdul Samad[1,†], Omer Qureshi[1,†], Muhammad Atif Tahir[1,†] and Muhammad Nouman Durrani[1,†]

[1]*National University of Computer and Emerging Sciences, Karachi, Pakistan*

Abstract

In this paper, we present the two deep learning methods for efficient detection and tracking of spermatozoa. In this task, human recorded video of sperm was provided by the Mediaeval task organizers. Our goal is to detect the motility of spermatozoa, for which we use two deep learning approaches. The first approach is to detect and track human sperm using Yolov8n and Byte-Track Algorithm. Its tracking speed was 80.4ms and flops were 8.7B which is outstanding. Then, we predict the motility of sperm using the Greedy Shape Geometry Technique for detecting progressive, non-progressive, and immotile sperm. In the second approach, we predict sperm motility using the provided graph data structure. We train yolov8n algorithm from scratch for the detection of healthy and unhealthy sperm which shows outstanding Mean Average Precision (MAP50) of 0.965.

1. Introduction

Human sperm motility prediction is a complex and time-consuming task. Automation of this task can minimize the time for the patient to see their test results. In this paper, we automate this task by using computer vision techniques to get some accurate predictions of the human sperm motility rate.

Predicting sperm motility and morphology from video is a challenging task. The video dataset has been provided with ground truth values. There is a lot of work going around related to video classification [1], segmentation [2], and video generation [3]. The significance of computer-aided sperm analysis helps automate the sperm detection task[4].

Transparent tracking of spermatozoa involves the application of advanced technologies to precisely track human sperm. By using Computer vision techniques in predicting sperm motility rate, it will enhance the efficiency and accuracy. Deep Learning algorithms play crucial role in automating the detection process, it also allows real time analysis will speed up the process for pathologists. These types of automated AI based solutions will be available 24 hours for patients and Patients can get their instant report because of its efficiency and speed.

Analyzing sperm samples manually is a time-consuming process which requires skilled experts with substantial training and years of experience. Manual sperm analysis is not reliable due to limited reproducibility and susceptibility to high inter-personal variations. Tracking and identifying sperm count in fresh samples is a complex task. Current computer-aided systems are not reliable therefore more research is required in this area.

MediaEval'23: Multimedia Evaluation Workshop, February 1–2, 2024, Amsterdam, The Netherlands and Online

*Corresponding author.

†These authors contributed equally.

mosaid.vsl@nu.edu.pk (M. Osaid); k191396@nu.edu.pk (A. Samad); omer.qureshi@nu.edu.pk (O. Qureshi); atif.tahir@nu.edu.pk (M. A. Tahir); muhammad.nouman@nu.edu.pk (M. N. Durrani)

CEUR Workshop Proceedings (CEUR-WS.org)

The 2023 Medico task involves different challenges [5]. Detection and Tracking of sperm cells in videos. They provided a dataset containing videos from 20 participants. We perform efficient sperm detection and tracking, and prediction of motility on videos and graph data structures. For prediction of motility, we introduce our own algorithm which is discussed in 3.2 .

2. Related Work

This section provides a brief literature review of the previous work related to to Human Spermatozoa task. In the paper [6], the authors presented VISEM-Tracking, a multi-modal sperm dataset containing videos, biological analysis data, and participant data of 85 individuals. They conducted baseline analysis to predict the motility and morphology of sperm. The authors analyzed microscopic images of sperms to deduce indicators such as sperm count to better understand human fertility. The main problem with sperm-related data is that it is often restricted to share such information due to legal matters and researchers need to have sound subject knowledge about the matter to come up with reasonable conclusions.

There is an increasing use of machine learning to analyze the videos of spermatozoa [7] as it is difficult to study their motility due to the fast-moving view. In the paper [6], the authors provide a dataset called VISEM-Tracking which has 20 videos of 30 seconds each and comprised of 29,196 frames. In the videos, wet semen was observed with manual bounding boxes along with sperm characteristics. The VISEM-Tracking dataset is an extension of the VISEM dataset [7] and performs better for training supervised ML models due to the presence of annotated bounding boxes. In addition to the sperm tracking annotation, the sperms are categorized into three categories: "normal sperm", "pin-head" and "cluster". The pinhead category has small blackheads when studied under a microscope, whereas the cluster category consists of sperm that are grouped together.

In the paper [8], the authors used a CNN to analyze sequence of frames to predict sperm motility and categorize it into progressive, non-progressive, and immotile spermatozoa. Subsequently, the video recordings are integrated with the participant data to determine how it may improve performance while using different modalities.

To solve the problem of predicting morphology and motility from videos, [3] presents two methods: stacked pure video frames and dense optical flows of video frames. To address the regression task, stacked dense optical flows and extracted original frames from sperm videos were utilized in combination of modified CNNs. For modification, they included an additional MLP layer to address the overfitting problem. The authors conducted experiments using a pre-trained ResNet-34[9] for predicting sperm motility and morphology.

In another paper [4] the authors present two deep learning techniques for predicting sperm motility and morphology on a video dataset. First, they used autoencoder to extract temporal features from videos and then plot those images into image space. Secondly, they used these extracted features to perform transfer learning to predict the required morphology and motility of human sperm. Their two-step process is different from previous approaches[5].

3. Approach

The Medico 2023 involves different tasks, which are categorized below.

3.1. Detection and Tracking of spermatozoa in Videos

For the detection of sperm in videos, we apply the yolov8n algorithm [10]. For training we used High performance machine with Nvidia RTX 3080 GPU, 64GB RAM and Corei9-10900k CPU. We use Cuda Version 11.7 and used Windows OS. We trained our model for 30 epochs and 32 batch size. And it gives a detection accuracy of 96.5%. For tracking the motion of sperm we use the tracking algorithm "Byte Track". It gives motion tracking of each sperm, which will be helpful in predicting the motility rate.

3.2. Proposed Algorithm for Motility Prediction on Spermatozoa Videos

We proposed an algorithm for the prediction of motility rate on spermatozoa videos. First, we got the detections and tracking values of each sperm in videos using a detection and tracking algorithm as discussed in section 3.1. Then we applied greedy shape geometry technique for predicting motility rate using tracking values of each sperm. First, if the sperm moves in a circle, then algorithm will count it as non-progressive sperm. If the sperm is moving forward then it counts it as a progressive sperm and if the sperm is at rest and showing no movement then it will count it as immotile sperm. We track a sperm by its tracking values. The tracking algorithm Byte Track [11] gives positional values of each sperm in x and y points. By applying this simple logic we are able to get the motility predictions from the videos.

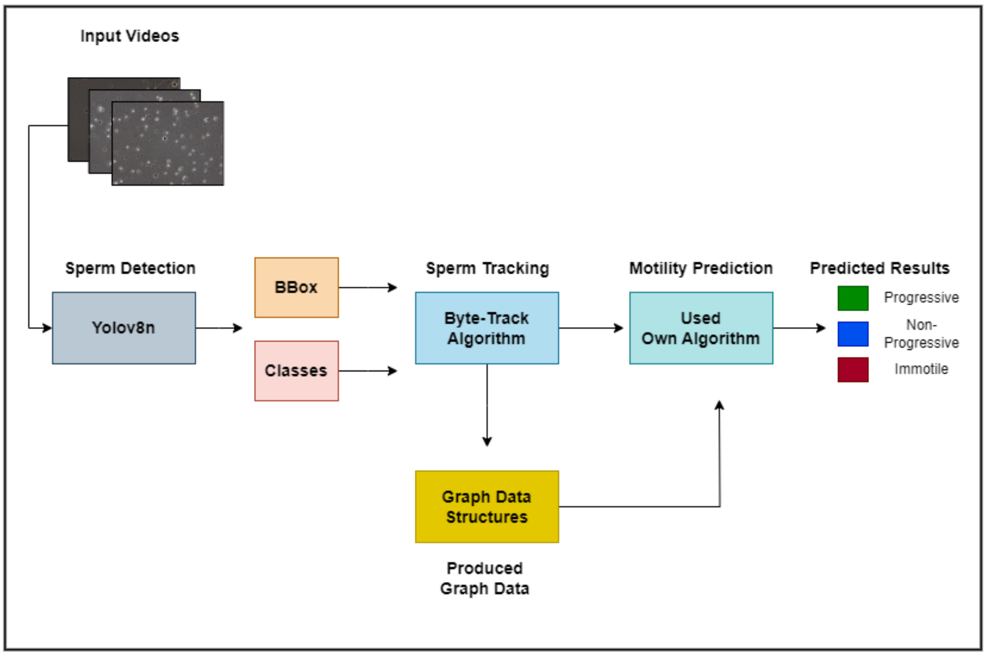

Figure 1: Proposed Model for Spermatozoa Tracking and Predicting Motility Rate.

3.2.1. Proposed Algorithm for Motility Prediction on Spermatozoa Graph Data Structures

For motility predictions on graph data structure, we used the same approach as before. We extract detections and tracking values of spermatozoa from graphs and then use the same approach for predicting motility rate.

4. Results and Analysis

The sperm motility prediction results is categorized into three classes: Progressive, Non Progressive and Immotile sperms. Table 1 shows the prediction results of sperm motility on videos using the proposed Greedy Shape Geometry Algorithm. As we can see progressive sperm on video ID 66 is 3.24%. Similarly in Video ID 80 the The ration of Immotile sperm is too high therefore Progressive sperm Count is almost zero. We also calculated the MAE, to check the our model predicted values accuracy as shown in Table 1 and Table 2. As we can see Progressive and Non-Progressive Motility predicted values MAE (Mean Absolute Error) is less than Immotile Sperm Means our model is good at predicting Progressive and Non-Progressive Sperm Motility.

ID	**Progressive motility** (%)	**Non-progressive sperm motility** (%)	**Immotile sperm** (%)
66	3.247	37.094	59.658
68	3.631	39.709	56.658
73	0.888	34.444	64.666
76	1.754	33.333	64.912
80	0	26.956	73.043
MAE	3.096	12.669	59.987

Table 1
Motility Predictions Rate on Videos.

Table 2 shows the prediction results of sperm motility on graph data structures using the proposed algorithm. We can see that MAE of Progressive and Non-Progressive Sperm Motility are less than the Immotile Sperm, Means Model is predicting is Nearly Accurate.

ID	**Progressive motility** (%)	**Non-progressive sperm motility** (%)	**Immotile sperm** (%)
66	2.564	7.521	89.914
68	1.694	3.631	94.673
73	0.222	17.555	82.222
76	10.638	11.347	78.014
80	0.865	56.277	42.857
MAE	5.083	24.044	73.736

Table 2
Motility Predictions Rate on Graph Data Structures.

5. Discussion and Outlook

Our proposed algorithm shows quite impressive results on Videos. As there is a lot of research scope in this area, we can make a lot of improvements using different machine learning based approaches like Regression, K-Nearest Neighbors (KNN) or Support Vector Machines (SVM) based Techniques for motility prediction part. To do more accurate predictions we can use some hybrid approaches and other SOTA Deep Learning models.

In Future we will advance our algorithm by using hybrid approach in which we will concatenate greedy shape geometry with convex hull in combination to Regression based approaches to get better prediction results. Furthermore, tracking algorithms can be optimized using Gaussian Mixture and Kalman Filter.

References

[1] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, L. Fei-Fei, Large-scale video classification with convolutional neural networks, in: Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, 2014, pp. 1725–1732.

[2] I. Koprinska, S. Carrato, Temporal video segmentation: A survey, Signal processing: Image communication 16 (2001) 477–500.

[3] V. Thambawita, P. Halvorsen, H. Hammer, M. Riegler, T. B. Haugen, Stacked dense optical flows and dropout layers to predict sperm motility and morphology, arXiv preprint arXiv:1911.03086 (2019).

[4] V. Thambawita, P. Halvorsen, H. Hammer, M. Riegler, T. B. Haugen, Extracting temporal features into a spatial domain using autoencoders for sperm video analysis, arXiv preprint arXiv:1911.03100 (2019).

[5] V. Thambawita, A. M. Storås, T.-L. Huynh, H.-D. Nguyen, M.-T. Tran, T.-N. Le, P. Halvorsen, M. A. Riegler, S. Hicks, Medico Multimedia Task at MediaEval 2023: Transparent Tracking of Spermatozoa, in: Proceedings of MediaEval 2023 CEUR Workshop, 2023.

[6] V. Thambawita, S. A. Hicks, A. M. Storås, T. Nguyen, J. M. Andersen, O. Witczak, T. B. Haugen, H. L. Hammer, P. Halvorsen, M. A. Riegler, Visem-tracking, a human spermatozoa tracking dataset, Scientific Data 10 (2023) 1–8.

[7] T. B. Haugen, S. A. Hicks, J. M. Andersen, O. Witczak, H. L. Hammer, R. Borgli, P. Halvorsen, M. Riegler, Visem: A multimodal video dataset of human spermatozoa, in: Proceedings of the 10th ACM Multimedia Systems Conference, 2019, pp. 261–266.

[8] S. A. Hicks, J. M. Andersen, O. Witczak, V. Thambawita, P. Halvorsen, H. L. Hammer, T. B. Haugen, M. A. Riegler, Machine learning-based analysis of sperm videos and participant data for male fertility prediction, Scientific reports 9 (2019) 16770.

[9] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.

[10] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, You only look once: Unified, real-time object detection, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 779–788.

[11] Y. Zhang, P. Sun, Y. Jiang, D. Yu, F. Weng, Z. Yuan, P. Luo, W. Liu, X. Wang, Bytetrack: Multi-object tracking by associating every detection box, in: European Conference on Computer Vision, Springer, 2022, pp. 1–21.