Tail-Aware Sperm Analysis for Transparent Tracking of Spermatozoa

Tuan-Luc Huynh[1,2], Huu-Hung Nguyen[1,2], Xuan-Nhat Hoang[1,2], Thao Thi Phuong Dao[1,2,3], Tien-Phat Nguyen[1,2,3], Viet-Tham Huynh[1,2], Hai-Dang Nguyen[1,2], Trung-Nghia Le[1,2] and Minh-Triet Tran[1,2,3,*]

[1]University of Science, VNU-HCM, Vietnam

[2]Viet Nam National University Ho Chi Minh City, Vietnam

[3]John von Neumann Institute, VNU-HCM, Vietnam

Abstract

Semen analysis is crucial to determine men's fertility; however, microscope-based manual spermatozoa evaluation is time-consuming and costly. Therefore, it has become essential to develop computer-aided-semen-analysis systems. To facilitate automated spermatozoa analysis, we propose a simple yet efficient framework for tracking sperms and predicting their motility. Different from existing methods, our proposed framework centralizes a new paradigm, dubbed sperm having a tail. We develop a novel tail-aware sperm detection model to advance the detection ability of dense, tiny, and transparent sperm cells. Furthermore, to enhance sperm tracking, a scene change detection technique is utilized to suppress identity assignment errors of similar sperms, resulting in improved sperm motility measurement. Experimental results show that our framework works well with an insignificant trade-off in execution time, which is suitable for the real-time clinical setting requirement.

1. Introduction

Manual sperm quality assessment using microscopes is time-consuming and requires costly experts with extensive training. In addition, the validity of manual sperm analysis becomes unreliable due to limited reproducibility and high inter-personnel variations [1]. Therefore, Computer-Aided-Sperm-Analysis (CASA) systems have been introduced to reduce errors in identifying, tracking, and counting spermatozoa (i.e., living sperm) in fresh samples. As a result, CASA has been the most popular method for male fertility appraisal. However, CASA reportedly suffers from many drawbacks, such as inaccurately predicting the fertility of a given sample [2] and requiring expensive instruments [3].

To promote automated sperm analysis, artificial intelligence has been developed and substantially improved the assessment of the sperm [1, 4, 5, 6]. Particularly deep learning based methods [5, 6, 7, 8, 9] utilized various CNN architectures to improve the semen analysis.

The Medical Multimedia track of the MediaEval 2022 challenge, dubbed Medico 2022, tackles the challenge of automatically tracking sperm cells in video recordings of spermatozoa [10] using the new dataset VISEM-tracking [11]. We participate in the Medico 2022 challenge and solve issues of correctly detecting and tracking sperm cells, i.e., Subtask 1 and utilizing the tracking results to predict sperm motility of each video sample, i.e., Subtask 2.

✉ htluc@selab.hcmus.edu.vn (T. Huynh); nhhung19@apcs.fitus.edu.vn (H. Nguyen); thao.dao2020@ict.jvn.edu.vn (T. T. P. Dao); phat.nguyen2021@ict.jvn.edu.vn (T. Nguyen); hvtham@selab.hcmus.edu.vn (V. Huynh); nhdang@selab.hcmus.edu.vn (H. Nguyen); ltnghia@fit.hcmus.edu.vn (T. Le); tmtriet@fit.hcmus.edu.vn (M. Tran)

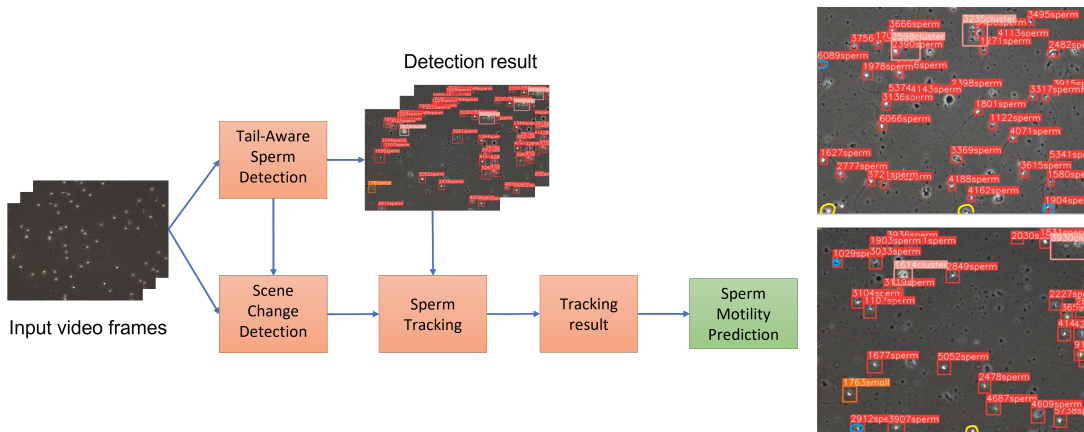CEUR Workshop Proceedings (CEUR-WS.org)

Figure 1: (Left) Overview of our proposed tail-aware sperm analysis framework. (Right) Our observation that sperms near video frame borders might be ignored (Yellow circles), while annotated sperms often have tails (Blue circles).

Unlike previous works, which only focus on using different feature engineering techniques and CNN architectures to improve the efficiency of CASA, we aim to tackle the challenges of the small number of data and significant class imbalance. We propose a novel framework for tracking sperms and predicting their motility. Our framework is concentrated on the biology-inspired paradigm "sperm having a tail." Our contributions can be summarized as follows: (1) We develop a simple yet efficient framework for tracking sperm cells and predicting sperm motility. Our framework can meet the real-time clinical requirements. (2) We propose a novel tail-aware sperm detection model. Our detection model can suppress uncertain detections near the borders of video frames. We also empirically evaluate different implementation settings of our detection model using YoloV7 [12]. (3) We introduce a detect change detection method based on Peak Signal-to-Noise Ratio (PSNR) to prevent trailing of undesired tracking from the previous unrelated video frame.

2. Methodology

Figure 1 (Left) illustrates our proposed tail-aware sperm analysis framework for the Medico 2022 challenge, consisting of four main components: tail-aware sperm detection, scene change detection, sperm tracking, and sperm motility prediction.

2.1. Subtask 1: Sperm Cell Tracking

Our preliminary detection results using YoloV5[13] indicate that the data of the Medico 2022 challenge is exceptionally challenging to train detection models due to the small number of samples and the heavy class imbalance where the major class sperm accounts for over 93% sperm cells. Thus, we solved these issues using several simple strategies, including weighted sampling and background oversampling. We then doubled the video resolution to tackle the issues of relatively small and dense ground truth bounding boxes.

By exploiting the underlying properties of the data, we found that sperms on the border of the videos tended to be ignored since annotators were uncertain (see Fig. 1(Right). Hence, we propose suppressing the bounding box near the video frame borders. To achieve this goal, we devise a biology-based solution that leverages the hypothesis that spermatozoa have tails (see Fig. 1(Right). We propose a novel **tail-aware sperm detection**, which focuses on detecting only sperm cells whose head is surrounded by a black line - *the sperm tail*. Particularly, we

combine various edge detection algorithms [14, 15, 16] to extract sperm tails. Detected sperms without tails are then eliminated to avoid false alarm detection. We empirically conducted experiments to compare the performance of edge-cutting object detection models in terms of the trade-off between accuracy and efficiency (i.e., running time) and finally employed and trained YoloV7 [12] as the core of our tail-aware sperm detection model.

As in the dataset description [10], there are view changes in videos because the sample is moved below the microscope to observe the complete sample area. Therefore, incorrect sperm tracking usually happens due to the similar appearance of sperms. To thoroughly mitigate the propagation of movement trails to new scenes, we force the tracking model to assign new identities to sperms whenever the view changes. We introduce a **scene change detection method based on Peak Signal-to-Noise Ratio (PSNR)** to identify abrupt changes in the video. Candidate video frames, which have abnormal PSNR values, are further filtered out if there is an abrupt change in the number of detected sperms. In our implementation of **sperm cell tracking**, we employed the SORT method [17] to fulfill the real-time requirement of the Medico 2022 challenge. We followed an $80\%/20\%$ stratified train-test split strategy.

2.2. Subtask 2: Sperm Motility Prediction

We propose a simple yet efficient feature extraction followed by a shallow network to **predict sperm motility** (i.e., progressive and non-progressive). We first aggregate movement statistics of every tracked sperm resulting in features of each patient. For each sperm, we exploit three movement characteristics: (1) Total movement distance between neighboring frames; (2) Average velocity between two neighboring frames; (3) Distance between the first and the last location. These measurements are then averaged over the number of sperm, and the number of sperm cells is considered the fourth feature.

To avoid overfitting when training on limited data, we implemented a shallow network instead of using popular deep neural networks. The network contains one hidden layer with four hidden units, four input units, and three output units, followed by an extra soft-max layer. The network architecture corresponds to four feature inputs and three category outputs (e.g., motility rate, progressive rate, and non-progressive rate). We used L1 distance to calculate loss.

3. Experimental Results

3.1. Medico 2022 Challenge Results

Table 3.1 shows our promising results on the Medico 2022 challenge. Our best submission, denoted by YoloV7_1280+Tail, achieved the highest scores in various metrics. The high resolution of input video frames improves sperm detection and sperm motility prediction. Furthermore, the proposed tail-aware sperm detection boosts the Det_Pr metric with a large margin of 3% compared to not having it. Finally, we argue that the drop in HOTA might be related to Det_Re.

3.2. Ablation Study

We tested our tail-aware sperm detection on the validation set. Table 2 demonstrates only results of the "sperm" category since the organizers concentrate on evaluating this weighty class. In our initial attempts, we naively suppressed all detections in the region near the borders, which led to a degradation in performance. Tail-aware sperm detection circumvents the mentioned problem by detecting sperms with tails, but it only improves precision. We argue that the model was not robust enough; thus, we added a confidence threshold to prevent it from discarding

Table 1

Our results on the test set of the Medico 2022 challenge. Yolov7c denotes Yolov7 using a custom configuration; 1280 stands for using high resolution of input video frames.

Submission method	Det_AP	Det_Re	Det_Pr	Det_F1	HOTA[18]	MAE	RMSE
YoloV7	0.738	**1.000**	0.738	0.849	**0.364**	36.473	38.602
YoloV7c	0.743	**1.000**	0.743	0.853	0.362	30.356	36.925
YoloV7c_1280	0.794	0.991	0.801	0.886	0.342	23.300	**26.203**
YoloV7c+Tail	0.796	0.977	0.814	0.888	0.352	30.912	37.961
YoloV7c_1280+Tail	**0.808**	0.959	**0.843**	**0.897**	0.334	**23.218**	26.885

Table 2

Sperm detection results on val set

Method	Pr	Re	mAP@.5	FPS
No Tail	0.499	**0.666**	**0.515**	**200**
Tail	0.501	0.626	0.496	192
Tail+Thresh@.7	0.502	0.642	0.507	192
Tail+Thresh@.7+Sperm	**0.524**	0.639	0.507	192

Table 3

Sperm tracking results on val set

Method	HOTA[18]	IDF1[19]
No Tail	**42.263**	44.432
No Tail+SCD	42.206	**44.584**
Tail	**40.379**	42.885
Tail+SCD	40.263	**43.003**

highly confident bounding boxes. Due to the class imbalance, we did not discard any cluster or small-pin-head sperm to prevent overuse of the tail-aware mechanism, which may lead to false positive results. This gives the class sperm a substantial 2.2% boost in terms of precision. Finally, we added the confidence threshold and the "only sperm filter" to our tail-aware sperm detection. The extra cost of tail-aware sperm detection is negligible since it only lowers the FPS by a margin of less than ten.

We verified our scene change detection (SCD) using the latest HOTA [18] and IDF1 [19] metrics. We chose IDF1 since we aim to eliminate the potential trailing ID problem from previous irrelevant frames, which might degrade the tracking performance. Further investigation shows that the tracking without SCD has two issues: (1) associating new sperms with previous unrelated sperms causing unwanted trailings; (2) failure to associate trailings. We confirm that the suppression does mitigate the above issues. Indeed, as seen in Table 3, increases in IDF1 prove the effectiveness of our proposed method. We also conclude that the proposed tail-aware sperm detection does not favor HOTA, likely due to the decreases in the recall.

4. Discussion and Outlook

In this work, we proposed a novel tail-aware sperm analysis framework for transparent tracking of spermatozoa in the Medico 2022 challenge. Leveraging the biology-inspired paradigm "sperm having a tail," we proposed a tail-aware sperm detection to reduce false alarm detection of non-tail spermatozoa and a PSNR-based scene change detection to suppress incorrectly tracked trails. Experimental results on the Medico 2022 challenge demonstrated the potential of our proposed framework. We firmly believe that "sperm having a tail" is a strong cue for developing promising techniques for semen analysis.

In the future, we aim to investigate robust metrics for detecting scene changes to remove the performance dependent on Yolo. We also will study approaches that can overcome the limitation of currently implemented tail-aware sperm detection, such as hard-coded confidence thresholds and limited to regions near video frame borders.

References

[1] R. Finelli, K. Leisegang, S. Tumallapalli, R. Henkel, A. Agarwal, The validity and reliability of computer-aided semen analyzers in performing semen analysis: a systematic review, Translational Andrology and Urology (2021).

[2] R. P. Amann, D. Waberski, Computer-assisted sperm analysis (casa): Capabilities and potential developments, Theriogenology (2014).

[3] S.-Y. Wei, H.-H. Chao, H.-P. Huang, C. F. Hsu, S.-H. Li, L. Hsu, A collective tracking method for preliminary sperm analysis, Biomedical engineering online (2019).

[4] K. Y. Chu, D. E. Nassau, H. Arora, S. D. Lokeshwar, V. Madhusoodanan, R. Ramasamy, Artificial intelligence in reproductive urology, Current urology reports (2019).

[5] S. Javadi, S. A. Mirroshandel, A novel deep learning method for automatic assessment of human sperm images, Computers in biology and medicine (2019).

[6] V. Valiuškaitė, V. Raudonis, R. Maskeliūnas, R. Damaševičius, T. Krilavičius, Deep learning based evaluation of spermatozoid motility for artificial insemination, Sensors (2020).

[7] 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).

[8] 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).

[9] S. Hicks, T. B. Haugen, P. Halvorsen, M. Riegler, Using deep learning to predict motility and morphology of human sperm, in: MediaEval, 2019.

[10] V. Thambawita, S. Hicks, A. M. Storås, J. M. Andersen, O. Witczak, T. B. Haugen, H. Hammer, T. Nguyen, P. Halvorsen, M. A. Riegler, Medico Multimedia Task at MediaEval 2022: Transparent Tracking of Spermatozoa, in: Proceedings of MediaEval 2022 CEUR Workshop, 2022.

[11] 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: Human spermatozoa tracking dataset, arXiv preprint arXiv:2212.02842 (2022).

[12] C.-Y. Wang, A. Bochkovskiy, H.-Y. M. Liao, Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors, ArXiv (2022).

[13] G. Jocher, A. Stoken, J. Borovec, NanoCode012, ChristopherSTAN, L. Changyu, Laughing, tkianai, A. Hogan, lorenzomammana, yxNONG, AlexWang1900, L. Diaconu, Marc, wanghaoyang0106, ml5ah, Doug, F. Ingham, Frederik, Guilhen, Hatovix, J. Poznanski, J. Fang, L. Yu, changyu98, M. Wang, N. Gupta, O. Akhtar, PetrDvoracek, P. Rai, ultralytics/yolov5: v3.1 - Bug Fixes and Performance Improvements, 2020.

[14] N. Dalal, B. Triggs, Histograms of oriented gradients for human detection, 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) (2005).

[15] J. F. Canny, A computational approach to edge detection, IEEE Transactions on Pattern Analysis and Machine Intelligence (1986).

[16] R. O. Duda, P. E. Hart, Use of the hough transformation to detect lines and curves in pictures, Commun. ACM (1972).

[17] A. Bewley, Z. Ge, L. Ott, F. T. Ramos, B. Upcroft, Simple online and realtime tracking, 2016 IEEE International Conference on Image Processing (ICIP) (2016).

[18] J. Luiten, A. Osep, P. Dendorfer, P. H. S. Torr, A. Geiger, L. Leal-Taixé, B. Leibe, Hota: A higher order metric for evaluating multi-object tracking, International Journal of Computer Vision (2021).

[19] E. Ristani, F. Solera, R. S. Zou, R. Cucchiara, C. Tomasi, Performance measures and a data set for multi-target, multi-camera tracking, in: ECCV Workshops, Springer, 2016.