



OPEN ACCESS

EDITED BY

Mingming Gong,
The University of Melbourne, Australia

REVIEWED BY

Yanwu Xu,
University of Pittsburgh, United States
Ziye Chen,
The University of Melbourne, Australia

*CORRESPONDENCE

Jinkun Yang
jinkunyang@hust.edu.cn
Chen Luo
luochen0902@163.com

SPECIALTY SECTION

This article was submitted to
Computer Vision,
a section of the journal
Frontiers in Computer Science

RECEIVED 09 May 2022

ACCEPTED 11 July 2022

PUBLISHED 25 August 2022

CITATION

Chen Z, Yang J, Luo C and Zhang C
(2022) A method for sperm activity
analysis based on feature point
detection network in deep learning.
Front. Comput. Sci. 4:861495.
doi: 10.3389/fcomp.2022.861495

COPYRIGHT

© 2022 Chen, Yang, Luo and Zhang.
This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

A method for sperm activity analysis based on feature point detection network in deep learning

Zhong Chen¹, Jinkun Yang^{1*}, Chen Luo^{2*} and Changheng Zhang¹

¹National Key Laboratory of Science and Technology on Multi-Spectral Information Processing, Key Laboratory for Image Information Processing and Intelligence Control of Education Ministry, Institute of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan, China, ²Center for Reproductive Medicine, Department of Obstetrics and Gynecology, Nanfang Hospital, Southern Medical University, Guangzhou, China

Sperm motility is an important index to evaluate semen quality. Computer-assisted sperm analysis (CASA) is based on the sperm image, through the image-processing algorithm to detect the position of the sperm target and track tracking, so as to judge the sperm activity. Because of the small and dense sperm targets in sperm images, traditional image-processing algorithms take a long time to detect sperm targets, while target-detection algorithms based on the deep learning have a lot of missed detection problems in the process of sperm target detection. In order to accurately and efficiently analyze sperm activity in the sperm image sequence, this article proposes a sperm activity analysis method based on the deep learning. First, the sperm position is detected through the deep learning feature point detection network based on the improved SuperPoint, then the multi-sperm target tracking is carried out through SORT and the sperm motion trajectory is drawn, and at last the sperm survival is judged through the sperm trajectory to realize the analysis of sperm activity. The experimental results show that this method can effectively analyze the sperm activity in the sperm image sequence. At the same time, the average detection speed of the sperm target detection method in the detection process is 65fps, and the detection accuracy is 92%.

KEYWORDS

sperm activity analysis, deep learning, convolutional neural network, feature point detection, tracking

Introduction

With the quickening pace of life and the increasing pressure of study and work, the quality of male sperm has become a topic of concern. Problems such as decreased sperm motility and sperm quality are occurring frequently. Semen quality analysis has become an important testing item in the medical field. Traditional semen detection methods require medical personnel to observe semen samples under a microscope and obtain information such as semen density and survival

rate. Because the analysis process depends on the experience level of medical personnel, this process is highly subjective. The emergence of Computer-Assisted Sperm Analysis (CASA; Wang et al., 2011) greatly reduced the difficulty and subjectivity of Sperm Analysis. However, CASA's calculation of sperm parameters requires accurate and reliable internal algorithms. For example, the detection of sperm activity can determine whether sperm survive by calculating the movement track of sperm in the field of vision. Therefore, CASA's internal algorithm is required to accurately locate the location of sperm and effectively track the movement track of sperm. At present, most of the sperm image processing in the CASA is adopted by the traditional image-processing algorithm. Traditional target detection algorithms are mostly based on the manual features to describe the target. Simple feature descriptions cannot find the target well, and complex design of feature description will lead to slow detection speed, so the traditional method cannot well meet the requirements of sperm detection accuracy and speed. With the development of computer vision technology, deep learning methods have made great breakthroughs in the visual tasks such as target detection, target tracking, semantic segmentation and instance segmentation. Compared with traditional methods, using the deep learning method to detect and track the target can obtain higher accuracy. But in the current mainstream target detection network for small target detection, there are still many missing cases. For example, the classical single-stage target detection network YOLOv3 (Redmon and Farhadi, 2018) has a good detection effect on the larger target in the image, while the smaller target is easy to be ignored. Two-stage target detection networks such as FasterRCNN (Ren et al., 2015) are better at detecting smaller targets, but not as fast as single-stage networks.

In order to accurately and efficiently detect sperm targets from images, this article proposes a method based on a deep learning feature point detection network to detect sperm targets. This method is based on the feature point detection network SuperPoint (DeTone et al., 2018). First, the branch network of the feature point descriptor is removed, and only the feature point detection branch network is retained, so as to detect sperm targets in the way of feature points, and improve the efficiency of network detection. Then, MagicPoint, the basic network of SuperPoint, is retrained to improve the ability of the basic network to detect sperm targets. Finally, fine-tune the SuperPoint network as a whole to achieve efficient and accurate detection of sperm targets in the image.

On the basis of realizing sperm target detection, we track sperm through a multi-target tracking algorithm SORT (Bewley et al., 2016) to obtain the trajectory of each sperm. Then, the range of motion of the sperm trajectory is calculated and compared with the set threshold to determine whether each sperm is alive. Finally, the survival of all sperm in the process is counted to realize the analysis of sperm activity in the image sequence.

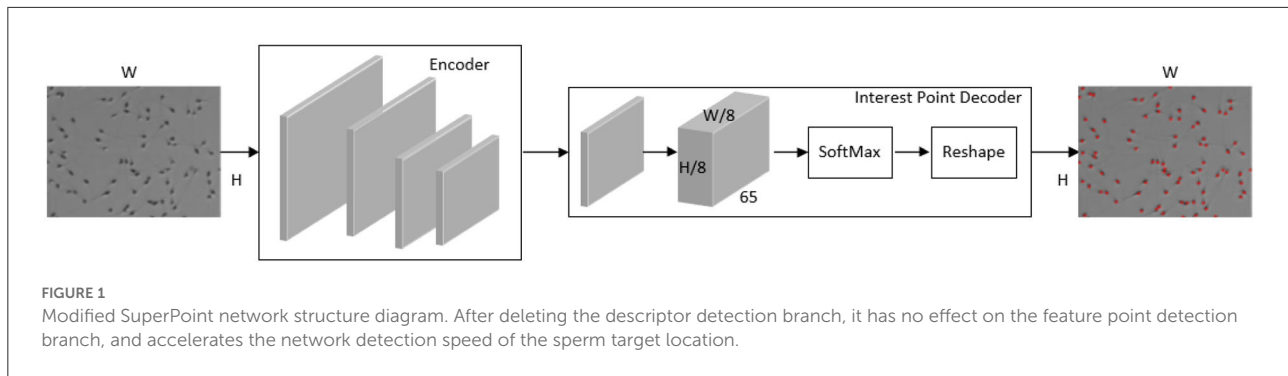
This article is organized as follows. The first section introduces the background of sperm activity analysis and the methods proposed in this article. The second section introduces the related work of sperm activity analysis. Section Methods describes our proposed approach, including how to improve SuperPoint for sperm target detection and multi-sperm target tracking using SORT. The fourth section introduces the experimental process, including data sets, evaluation index, and experimental results. The fifth section is the conclusion of this article.

Related work

Computer-assisted sperm analysis CASA mainly evaluates the density, concentration, activity, and morphology of sperm (Wang et al., 2011), among which evaluation and analysis of sperm activity is a very important link. CASA can determine the sperm activity by detecting the target location of sperm in the sperm image and tracking the trajectory to determine the survival of each sperm in the image and then analyze the sperm activity.

Initially, sperm targets were detected by threshold segmentation (Hidayatullah and Zuhdi, 2014) and other traditional image processing algorithms, and sperm were segmented from the background according to the gray level of the image. Obviously, this method is not accurate in the detection of sperm targets, nor can it distinguish between sperm targets and impurities in the background. Qi et al. (2019) proposed a sperm cell tracking recognition method based on the Gaussian mixture model in view of the problem that the threshold segmentation method is easy to lose sperm targets when sperm and impurities collide, leading to detection errors. Background modeling is used to detect and track moving sperm, so as to avoid the loss of tracking track when sperm and impurities collide. This method has a good detection and tracking effect on motile sperm, but when most sperm die and stop motile, it is easy to be confused with impurities in the detection. Hidayatullah et al. (2021) classified sperm movement through the machine learning method SVM and analyzed sperm activity by judging the curve speed, linear speed, and linear speed of sperm movement. This method is also needed to accurately detect the target location of sperm. In general, sperm target detection based on a traditional image-processing algorithm can achieve good results when the background and sperm are clearly distinguished and impurities are small, but it is difficult to distinguish impurities when there are more impurities. Compared with such traditional methods, the proposed method can better identify sperm targets and effectively distinguish sperm targets from impurities, so as to achieve a better tracking effect.

With the rapid development of deep learning, it has been widely used in medical image processing (Hidayatullah et al.,



2019; Tortumlu and Ilhan, 2020; Yüzkat et al., 2020), among which target detection networks and semantic segmentation networks have also been applied to detect sperm targets in images. Movahed and Orooji (2018) used a convolutional neural network to segment the sperm head, so as to obtain sperm targets from the background of the sperm image, and then used the SUPPORT vector machine SVM to classify the segmented head pixels. Similarly, Melendez et al. (2021) used a convolutional neural network based on U-NET structure to segment sperm targets. Such methods are based on the deep learning semantic segmentation network to find sperm targets. However, the semantic segmentation model always has a large model volume, including the encoding part and decoding part, which makes the network detection time longer than the target detection network. Rahimzadeh and Attar (2020) use deep fully convolutional neural network RetinaNet (Lin et al., 2017) as a target detector to detect sperm targets and improve detection accuracy by connecting some consecutive frames as network input. After the detection, the improved CSR-DCF algorithm is used to track the sperm trajectory and extract the sperm motion parameters. Similarly, the method is based on the universal target detection network to achieve sperm target detection, and this kind of universal target detection network is often inferior to the detection effect of small targets such as sperm. In general, target detection algorithms based on deep learning are still prone to miss detection when detecting small targets, and there are many small targets in sperm images, so these algorithms cannot give full play to their advantages. Compared with such deep learning methods based on target detection, the method presented in this article has fewer missed detection and is faster in the sperm target detection.

Considering that traditional image processing algorithms are difficult to distinguish sperm targets from impurities, and deep learning target detection network is also prone to miss detection for small targets like sperm, this article proposes a method to detect sperm targets based on deep learning feature point detection network. In this method, descriptor branches are deleted based on the SuperPoint network to achieve faster sperm target detection, and then the sperm image retraining network is used to make the network capable of sperm target detection.

TABLE 1 Comparison of sperm target detection results.

Methods	Precision	Recall	AP	FPS
FasterRCNN	0.796	0.937	0.909	13.256
YOLOv3	0.861	0.785	0.872	52.907
SuperPoint	0.736	0.848	0.834	42.337
Ours	0.815	0.949	0.920	65.688

Methods

Detection

Using a target detector to detect sperm targets is the first and most important step of the whole method. The sperm target in each frame image can be accurately and efficiently detected, and the subsequent correct tracking track and analysis of sperm activity can be achieved. On the basis of the feature point detection network SuperPoint, we improved it and retrained the two stages of SuperPoint to achieve accurate and efficient position detection of the sperm targets.

Modify the SuperPoint

SuperPoint is a self-supervising framework for training feature point detectors and descriptors. The full convolution model of SuperPoint runs on full-size images and computes pixel-level feature point positions and related descriptors in a forward transfer. The network also introduces homography adaptive, which is a multi-scale, multi-homography method to improve the repeatability of feature point detection and perform cross-domain adaptive.

SuperPoint network structure is mainly composed of three parts: MagicPoint, feature point extraction branch, and descriptor extraction branch. MagicPoint is a four-layer full convolutional network. In total, eight times down-sampling is performed on the whole network. The length and width of the final output feature graph are 1/8 of the original graph, respectively, and the number of channels is 65. The feature point

extraction branch conducts a softmax operation on the feature graph and upsamples to the size of the original graph, so as to output whether there are feature points on each pixel of the original graph. The descriptor extraction branch performs linear interpolation operation and L2 normalization operation on the feature graph, and also upsamples to the size of the original image, so as to output the corresponding descriptor for each feature point.

As shown in Figure 1, we delete the descriptor branch of SuperPoint and only retain the MagicPoint and feature point detection branch because the purpose of using the feature point extraction network SuperPoint is to obtain the position of sperm in the image without describing sperm. The advantage of deleting descriptor branches is that this operation removes unnecessary descriptor information, making sperm target extraction faster.

Training

The training of SuperPoint mainly consists of three parts: pre-training of feature point extraction, real image self-labeling, and joint training of feature point extraction and descriptor calculation.

Feature point extraction pre-training: For the feature point extraction task, it is difficult for us to judge which pixel point can be used as a feature point, so it is difficult to mark. Therefore, SuperPoint considers a rendered image with only line segments, triangles, rectangles, cubes, and other basic shape elements. This image has obvious feature points, such as endpoints and turning points, as feature points, so as to mark these feature points to train basic network MagicPoint.

Real image self-tagging: image self-tagging mainly uses the method of homography adaptation, which transfers the ability of MagicPoint above to the general image through homography transformation. Because the previous MagicPoint is only trained on images of basic shape elements, it may not be good for general image detection. Here, 100 homography transformations are performed on the general images to obtain 100 transformed images, and feature points are extracted on these images by the MagicPoint model. The feature points of these 100 images are superimposed as the standard feature points of this image, and then, the truth value of the feature points of this image is obtained, so as to realize the training of the whole SuperPoint network.

Joint training of feature point extraction and descriptor calculation: the aforementioned method has obtained the truth value of the corresponding feature point of the image, and the truth value of the descriptor is also needed below. Obtaining the truth value of the descriptor requires the perspective transformation of an image first. Since the transformation is known, the position correspondence of feature points before and after the transformation is also determined. On this basis,

we hope that the distance between the descriptors of the corresponding feature points should be as small as possible, and the distance between the descriptors of different feature points should be as large as possible. Thus, we can calculate the loss of any two pairs of feature points to optimize the matching distance between feature points, so as to find the appropriate descriptor.

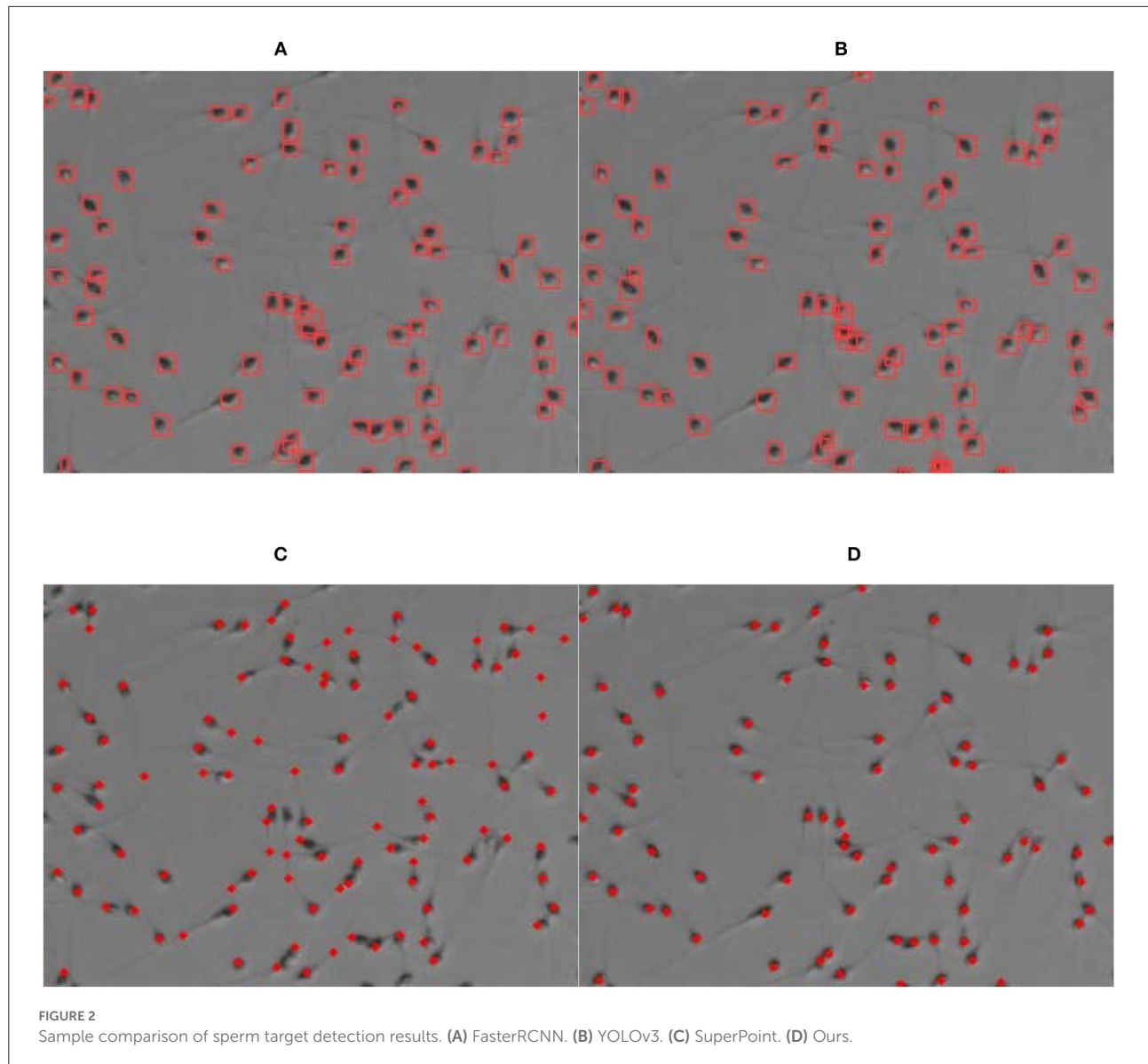
In order to enable the improved SuperPoint to better detect sperm targets, we retrained SuperPoint's basic network MagicPoint with labeled sperm images, and fine-tuned SuperPoint as a whole with unlabeled sperm images. The benefit of training for both phases is that SuperPoint can concentrate on detecting sperm targets and ignore surrounding impurities, improving sperm detection accuracy.

As for the annotation of the sperm target, we use labeling to mark the rectangular box where the target is located. What is written in the label is the position coordinate and size of the target in the upper left corner of the target. Then, we use the script to convert the coordinates of the upper left corner of the sperm and the size of the target into the coordinate position of the center of the sperm head. The reason for this is that MagicPoint, the basic network of SuperPoint, requires the coordinates of the center point of the sperm head, rather than using a rectangular box.

Track

When the sperm target in each frame of the image can be well-detected, tracking the sperm movement trajectory is relatively simple. Sort is a classical real-time multi-target tracking algorithm, which first uses IoU as the indicator of the measurement of the target relation between the front and back frames, then predicts the current position by using the Kalman filter, and then correlativity the matching between the detection frame and the target through The Hungarian algorithm, and finally removes the matching whose overlap with the target is less than the IoU threshold.

Unlike the multi-target tracking scenario applied by SORT, the detected target describes sperm location in the form of points, so some modifications to SORT are needed here to accommodate the detection results in this article. In Sort, IoU is mainly used as the measurement index to evaluate whether two targets match, and there is no IoU between two points. Therefore, we change the measurement index here to the distance between two points, and achieve the same effect as IoU by adjusting the threshold of the distance. In addition, another way is to transform the position point into a rectangular target box centered on the point, and the size of the target box is the size of the sperm head in the image. However, under pictures of different resolutions, the size of sperm will be different, so it is difficult to set the size of the target box, so we adopted the first method here.



Experiment

Our experiment was conducted under Ubuntu18.04.5 with NVIDIA RTX3090 GPU model. The deep learning framework we use is Pytorch1.7.0, the programming language is Python3, and the CUDA version is 11.2.152.

Dataset

We took semen samples under the microscope with Leica camera to obtain the original semen video. The resolution of the video is 1920×976 pixels and the frame rate is 30fps. Next, we use semen video to make multiple image sequences. First,

TABLE 2 Comparison of sperm target tracking results.

Methods	MOTA (%)	MOTP (%)
FasterRCNN	92.5	23.0
YOLOv3	94.0	26.2
SuperPoint	89.0	26.9
Ours	94.8	23.7

we select a window with a size of 320×240 , then truncate the video every 10 s, and save the images within 10 s into an image sequence. At last, we obtain 20 image sequences, each of which has 300 images with a size of 320×240 . The reason for choosing the 320×240 size window here is that the original resolution size

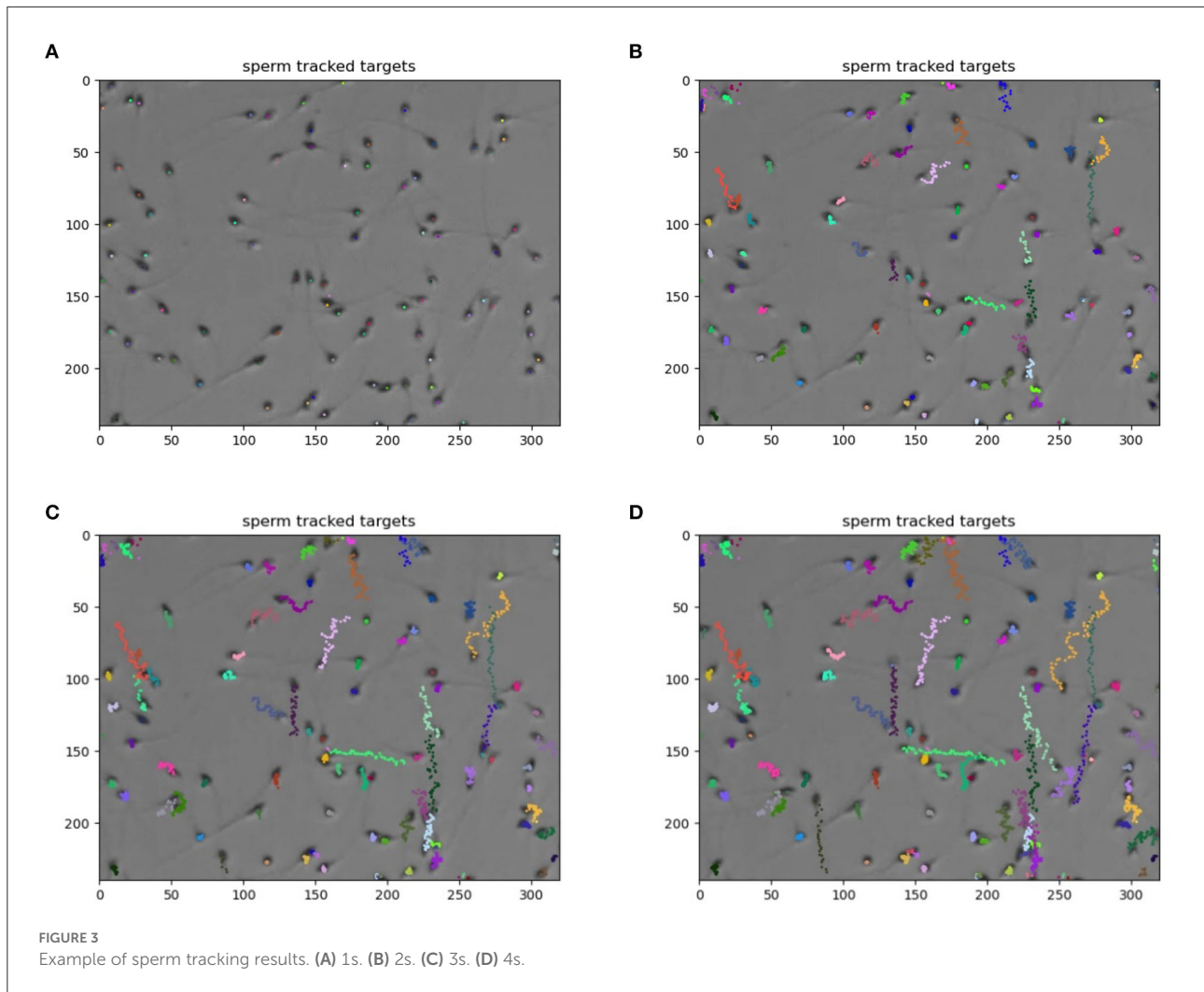


TABLE 3 Sperm activity test results.

Number	All sperm	Live sperm	Dead sperm	Survival rate (%)
1	116	57	59	49.138
2	121	59	62	48.760
3	116	59	57	50.862
4	130	71	59	54.615
5	123	68	55	55.285

image is too large and contains too many sperm, which makes the labeling work too difficult, and it will be difficult to track all the sperms.

We use half of the image sequence as the training set and the other half as the verification set. For the annotation of the training set, we use LabelImg to label the position of the sperm target, obtain the coordinate position and sperm size of the upper left corner of the sperm target, and then convert it into

the position of the center point of the sperm head through the script as the annotation of the sperm target.

Evaluation

Detection evaluation

In the process of evaluating the multi-target detection algorithm, the IoU between the detected target and the real target is usually calculated and compared with the threshold value to judge whether the target is detected correctly. Since our algorithm outputs the center point position of the sperm target, it can be simplified to judge the distance between the detected target and the real target to determine whether the detection is correct. Since there is only one category of sperm to be tested, mAP is simplified to AP here. AP is obtained by calculating the integral under the PR curve:

$$AP = \int_0^1 p(r)dr \tag{1}$$

Where $p(r)$ is pr curve, in the actual process, we should smooth pr curve, that is, take the maximum Precision value on the right side. The formula is used to describe:

$$P_{smooth}(r) = \max_{r' \geq r} P(r') \quad (2)$$

Calculation of PR curve requires calculation of Precision and Recall, and the formula is as follows:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (4)$$

Where TP is True Positives, FP is False Positives, TN is True Negatives, and FN is False Negatives.

Tracking evaluation

Among the evaluation indexes of the multi-target tracking algorithm, multiple object tracking precision (MOTP) and multiple object tracking accuracy (MOTA) are two commonly used indexes for price comparison (Bernardin and Stiefelhagen, 2008). MOTP is defined as accuracy in determining the location of a target, while MOTA is defined as accuracy in determining the number of targets, as well as related attributes of the target. Both of them jointly measure the ability of the algorithm to continuously track the target, that is, it can accurately judge the number of the target in consecutive frames and accurately delimit its position, so as to achieve uninterrupted continuous tracking.

Assuming that the total number of frames of the test video is N , the tracking matching sequence is constructed from the first frame, and the initial matching sequence is M_0 . We calculate three kinds of errors for each frame to calculate MOTP, in which the unmatched detection position is recorded as the number of misjudgments, the unmatched real position is recorded as the number of missing, and the detection position matches the real position, but the matching error is recorded as the number of mismatches. For each frame T , sperm target $\{h_1, h_2 \dots h_n\}$ and real sperm target $\{o_1, o_2 \dots o_n\}$ were detected in frame T . M_{t-1} assumes that there is a matching (o_i, h_j) , and we judge whether the position of the two exceeds the threshold. If not, the matching is maintained in M_t . If o_i is not in the current matching sequence, a new match is looked for o_i , and if a false match occurs in the new match, the match is logged and updated. In this way, we constructed the matching sequence at t time, using c_t to represent the number of matches at T time, and d_t^i to represent the matching error of o_i under t frame. The calculation of d_t^i can be measured by IoU or the Euclidean distance. If IoU is used, the larger the MOTP is, the better. If the Euclidean distance is used, the smaller the MOTP is, the better. Finally, we use fp_t and m_t to represent misjudged numbers and missing numbers.

g_t indicates the number of targets tracked by t frames. The calculation of MOTP can be calculated by the following formula:

$$\text{MOTP} = \frac{\sum_{i,t} d_t^i}{\sum_t c_t} \quad (5)$$

The calculation of MOTA can be calculated by the following formula:

$$\text{MOTA} = 1 - \frac{\sum_t (m_t + fp_t + mme_t)}{\sum_t g_t} \quad (6)$$

Where m_t , fp_t and mme_t represent missing number, misjudged number and mismatched number, respectively.

Results

Detection results

We evaluated our sperm target detector through the aforementioned evaluation indexes, namely, Precision, Recall, AP, and FPS. At the same time, we set up a comparison experiment with the commonly used target detection algorithm and feature point detection algorithm for comparison. Target detection algorithm uses FasterRCNN and YOLOv3, and feature point detection algorithm uses SuperPoint. We conducted fine-tuning training of 800 EPOCH using sperm images with the rectangular target frame labels for the two target detection algorithms. Since the training method of SuperPoint is self-supervised, we directly used unlabeled sperm images to fine tune training of 800 EPOCH.

When training our sperm target detector, we first trained 500 EPOCH with sperm images tagged with position points for the basic network MagicPoint, and then trained 300 EPOCH with sperm images untagged for the whole network. After the training of the method in this paper and the comparison method, the validation set samples are used to test and generate evaluation results, as shown in Table 1.

It can be seen from the results that compared with the target detection network based on deep learning, our method has better detection accuracy and faster detection speed in the sperm target detection. Since SuperPoint is not a special target detection network, it achieves relatively low-detection accuracy when it is directly used for the sperm target detection after training. Our method eliminates unnecessary descriptor detection branches based on SuperPoint, thus, achieving faster detection speed. At the same time, the basic network MagicPoint is trained with labeled sperm pictures, which enabled the network to have the ability to detect sperm targets, thus, achieving high-detection accuracy.

As shown in Figure 2, the following images show the detection effects of different networks. After training, the target detection network can better detect sperm targets, but because sperm targets are small and dense, there will still be a small

number of false detection. SuperPoint, a feature point detection network, is not targeted at sperm targets, so there are many errors in detection. In contrast, our method effectively reduces error detection after retraining the basic network MagicPoint.

Tracking results

After using the aforementioned sperm detection network to detect the sperm samples of the verification set, the multi-target tracking algorithm SORT is used for tracking. The results of tracking calculate MOTA and MOTP through the evaluation code of mot Benchmark (Leal-Taixé et al., 2015). The results are shown in Table 2.

Because of the higher accuracy of our method in detection, higher MOTA scores are obtained in tracking. A higher MOTA score indicates higher tracking accuracy, and fewer false detection, missed detection, and trajectory tracking errors. At the same time, our method also has a good MOTP score, which indicates that the predicted position in the tracking process is well-matched with the labeled position, and also reflects that our algorithm has a good detection accuracy.

As shown in Figure 3, the following pictures show the tracking effect of sperm image sequence. The first picture shows the tracking result of the first frame, and the following three pictures are the results after tracking 1s, 2s, and 3s, respectively. It can be seen from the tracking results that the moving sperm keeps moving in the image, and its track is constantly bending line, while the track of the dead sperm is basically in a small circle in the image. Therefore, we can judge whether each sperm in the image is alive according to the sperm movement track in the image sequence, so as to obtain the sperm activity statistically.

Sperm activity test results

Sperm image sequences from five validation sets were used to analyze sperm activity in the dataset. We detect and track the trajectories using the proposed method, and then determine whether sperm survive by calculating the maximum distance in each trajectory and comparing it with the threshold. Here, we set a threshold of 5 pixels. When the maximum distance of the trajectory is greater than or equal to the threshold, the sperm is considered as living sperm, while when the maximum distance of the trajectory is <5 pixels, the sperm is then considered as dead sperm. The experimental results are shown in Table 3.

We obtained the sperm survival rate of each group of sperm image sequences through experiments. The error can be further reduced by detecting multiple groups of sperm image sequences and averaging the results. At the same time, the threshold can be

modified according to the jitter situation in the video shooting process, so as to obtain the result more consistent with the actual situation. At last, we can calculate that the survival rate of sperm in this data set is 51.732%.

Conclusion

In this paper, we propose a method to analyze sperm activity based on the sperm image. We detect sperm targets by network detection based on feature points, track the detection results to obtain sperm trajectory, judge sperm survival through trajectory, and then analyze sperm activity. Compared with the detection of sperm targets by target detection network, the improved feature point detection network can detect sperm targets with higher detection accuracy and speed, and make the follow-up tracking achieve better results. Finally, under the condition of accurate detection of sperm targets and effective tracking, we can get the survival of sperm by judging the trajectory of each sperm.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

ZC provided important method guidance. JY was the main person in charge. CL provided data support. CZ provided experimental help. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Bernardin, K., and Stiefelwagen, R. (2008). Evaluating multiple object tracking performance: the clear mot metrics. *EURASIP J. Image Video Proces.* 2008, 1–10. doi: 10.1155/2008/246309
- Bewley, A., Ge, Z., Ott, L., Ramos, F., and Upcroft, B. (2016). “Simple online and realtime tracking,” in *2016 IEEE International Conference on Image Processing (ICIP)*. (Piscataway, NJ: IEEE), 3464–3468. doi: 10.1109/ICIP.2016.7533003
- DeTone, D., Malisiewicz, T., and Rabinovich, A. (2018). “Superpoint: self-supervised interest point detection and description,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (Piscataway, NJ: IEEE), 224–236. doi: 10.1109/CVPRW.2018.00060
- Hidayatullah, P., Mengko, T. E. R., Munir, R., and Barlian, A. (2019). “A semiautomatic sperm cell data annotator for convolutional neural network,” in *2019 5th International Conference on Science in Information Technology (ICSITech)* (Piscataway, NJ: IEEE), 211–216. doi: 10.1109/ICSITech46713.2019.8987471
- Hidayatullah, P., Mengko, T. L., Munir, R., and Barlian, A. (2021). Bull sperm tracking and machine learning-based motility classification. *IEEE Access* 9, 61159–61170. doi: 10.1109/ACCESS.2021.3074127
- Hidayatullah, P., and Zuhdi, M. (2014). “Automatic sperms counting using adaptive local threshold and ellipse detection,” in *2014 International Conference on Information Technology Systems and Innovation (ICITSI)* (Piscataway, NJ: IEEE), 56–61. doi: 10.1109/ICITSI.2014.7048238
- Leal-Taixé, L., Milan, A., Reid, I., Roth, S., and Schindler, K. (2015). Motchallenge 2015: towards a benchmark for multi-target tracking. *arXiv preprint arXiv:1504.01942*. doi: 10.48550/arXiv.2002.04034
- Lin, T. Y., Goyal, P., Girshick, R., He, K., and Dollár, P. (2017). “Focal loss for dense object detection,” in *Proceedings of the IEEE International Conference on Computer Vision* (Piscataway, NJ: IEEE), 2980–2988. doi: 10.1109/ICCV.2017.324
- Melendez, R., Castañón, C. B., and Medina-Rodríguez, R. (2021). “Sperm cell segmentation in digital micrographs based on convolutional neural networks using U-Net architecture,” in *2021 IEEE 34th International Symposium on Computer-Based Medical Systems (CBMS)* (Piscataway, NJ: IEEE), 91–96. doi: 10.1109/CBMS52027.2021.00084
- Movahed, R. A., and Orooji, M. (2018). “A learning-based framework for the automatic segmentation of human sperm head, acrosome and nucleus,” in *2018 25th National and 3rd International Iranian Conference on Biomedical Engineering (ICBME)* (Piscataway, NJ: IEEE), 1–6. doi: 10.1109/ICBME.2018.8703544
- Qi, S., Nie, T., Li, Q., He, Z., Xu, D., and Chen, Q. (2019). “A sperm cell tracking recognition and classification method,” in *2019 International Conference on Systems, Signals and Image Processing (IWSSIP)* (Piscataway, NJ: IEEE), 163–167. doi: 10.1109/IWSSIP.2019.8787312
- Rahimzadeh, M., and Attar, A. (2020). Sperm detection and tracking in phase-contrast microscopy image sequences using deep learning and modified csr-dcf. *arXiv preprint arXiv:2002.04034*. doi: 10.48550/arXiv.1504.01942
- Redmon, J., and Farhadi, A. (2018). Yolov3: an incremental improvement. *arXiv preprint arXiv:1804.02767*. doi: 10.48550/arXiv.1804.02767
- Ren, S., He, K., Girshick, R., and Sun, J. (2015). Faster r-cnn: towards real-time object detection with region proposal networks. *Adv. Neural Informat. Proces. Syst.* 28, 91–99. doi: 10.1109/TPAMI.2016.2577031
- Tortumlu, O. L., and İlhan, H. O. (2020). “The analysis of mobile platform based CNN networks in the classification of sperm morphology,” in *2020 Medical Technologies Congress (TIPTEKNO)* (Piscataway, NJ: IEEE), 1–4. doi: 10.1109/TIPTEKNO50054.2020.9299281
- Wang, Y., Jia, Y., Yuchi, M., and Ding, M. (2011). “The computer-assisted sperm analysis (CASA) technique for sperm morphology evaluation,” in *2011 International Conference on Intelligent Computation and Bio-Medical Instrumentation* (Piscataway, NJ: IEEE), 279–282. doi: 10.1109/ICBMI.2011.21
- Yüzkat, M., İlhan, H. O., and Aydın, N. (2020). “Morphological classification of low quality sperm images using deep learning networks,” in *2020 Medical Technologies Congress (TIPTEKNO)* (Piscataway, NJ: IEEE), 1–4. doi: 10.1109/TIPTEKNO50054.2020.9299318