Machine Learning Based Tool for Automated Sperm Cell Tracking and Sperm Bundle Detection

Jakub Horenin[1], Veronika Magdanz[2], Islam S. M. Khalil[3], Anke Klingner[4], Alexander Kovalenko[1], and Miroslav Čepek[1(✉)]

[1] Faculty of Information Technology, Czech Technical University in Prague, Prague, Czech Republic
cepekmir@fit.cvut.cz

[2] Department of Systems Design Engineering, Waterloo Institute for Nanotechnology, University of Waterloo, Waterloo, Canada

[3] Department of Biomechanical Engineering, University of Twente, Twente, The Netherlands

[4] Department of Physics, German University in Cairo, Cairo, Egypt

Abstract. This study introduces a novel machine learning-based methodology for automated detection and tracking of sperm cells within microscopic video recordings, aiming to elucidate the dynamics and motion patterns of individual sperm cells as well as sperm cell bundles. At first, the method identifies sperm cells across successive frames within a video sequence, facilitating the reconstruction of each cell's trajectory over time. Subsequently, we introduce a classification algorithm that distinguishes between solitary sperm cells, clusters of adjacent cells, and cohesive sperm cell bundles, addressing a gap in existing methodologies. Finally, we employ three conventional metrics for velocity assessment: Straight Line Velocity (VSL) and Average Path Velocity (VAP) and Curvilinear velocity (VCL), to quantify the movement speed of both individual sperm cells and bundles. The approach represents a significant advancement in the automated analysis of sperm motility and aggregation phenomena, providing a robust tool for researchers to study sperm behavior with enhanced accuracy and efficiency. The integration of machine learning techniques in sperm cell detection and tracking offers promising insights into reproductive biology and fertility studies. https://gitlab.fit.cvut.cz/horenjak/sperm_cell_tracking_app https://apps.datalab.fit.cvut.cz/sperm_tracking/

Keywords: Sperm Cell Tracking · Motion Dynamics · Bundle Formation · Bundle Detection · Kalman Filter

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/978-3-031-70381-2_2.

A. Bifet et al. (Eds.): ECML PKDD 2024, LNAI 14950, pp. 19–32, 2024.
https://doi.org/10.1007/978-3-031-70381-2_2

1 Introduction

Advanced machine learning techniques in the era of data abundance offer numerous advantages across various fields. However, for almost any sort of efficient decision-making process or predictive analysis, the quality of data remains crucial. Due to its peculiar characteristics and complexity, biological data, including microscopic images and recorded videos, stand apart from data encountered in daily life like images of common objects, text, or music. Particularly, video recordings of motile sperm cells is a good example to expose biological data complexity. Due to the nature of video recording with a microscope, the data often appear to be noisy, containing artifacts and/or blurred objects. Certain attributes of the video, that are of interest to biologists are not labeled, therefore for many tasks, a straightforward approach such as supervised learning can't be applied. On the other hand, labeling of these data is extremely demanding, as it requires domain expert knowledge and some of the events that need to be labeled are extremely rare. Even though, numerous computer-assisted sperm analysis (CASA) systems have been developed [1,6,18], there are still under-explored fields where machine learning (ML) can bring unquestionable benefits.

One such field is an automated exploration and detection of sperm bundles [24,28,30]. Sperm bundle formation can be observed in various animals ranging from bovines [25] to ants [14], echidnas and others [5,24,25], and often associated with higher sperm motility [8,25,30] and higher reproduction rate [24]. Therefore, detecting and quantifying sperm bundles and their formation gives a direct insight into the quality of explored material, which is extremely important in various fields such as assisted reproductive technologies [24] including artificial insemination and *in-vitro* fertilization.

Additionally, it was reported that sperm cells were used as blueprints [15] or templates [16] for microrobots that can be used for *in-vivo* targeted therapy. In this case, understanding the motion of living cells can serve as an inspiration for the microrobot design [17]. Moreover, understanding sperm cell motion and bundle formation refers to the collective behavior of these microorganisms. By observing and analyzing the collective behaviors of cells, it is possible to glean significant statistical data that can inform and refine our understanding of the biosphere at the microscopic level, as from the collective behavior emerges the concept of swarm intelligence. Unraveling these intricacies could shed light on the decentralized, self-organized systems that can lead to the emergence of complex behavior from simple individual entities.

Therefore, in the present article, we focused on detecting, classifying, and evaluating motile sperm cells and sperm bundles. As a result, we present a novel CASA tool that is designed to detect, track, and classify parameters from the above-mentioned microorganisms. The system is currently being developed and tested using bovine sperm samples but is designed in a way that allows for simple expenditure for other types of sperm samples.

2 Background

2.1 Computer-Assisted Sperm Analysis Systems

The analysis of sperm motility is pivotal in assessing male fertility and understanding reproductive behaviors [4,9]. It aids in diagnosing male infertility, guiding clinical decisions in assisted reproductive technologies [21], and evaluating the effects of environmental and lifestyle factors on sperm health. The motility of sperm cells is indicative of their viability and fertilizing potential, making it a fundamental parameter in both clinical and research settings.

CASA systems have revolutionized the study and clinical assessment of sperm motility. These systems utilize advanced imaging and computational algorithms to automate the process of tracking and analyzing sperm movement. CASA technologies provide objective, repeatable, and high-throughput analysis, overcoming the limitations of manual microscopy methods. Prominent systems include the Hamilton-Thorne analyzer [22], Sperm Class Analyzer (SCA) [7], and Microptic's SCA system [26], among others. These systems vary in their imaging techniques, analysis algorithms, and the range of motility parameters they measure.

CASA systems employ a combination of video microscopy and digital image processing to capture and analyze sperm movement. The core steps involve:

- Capturing capture sequences of video frames by a high-resolution camera, documenting sperm movement over time.
- Identifying sperm cells within the video frames, often using thresholding and morphology-based techniques.
- Tracking sperm cells across frames, employing methods such as particle tracking or optical flow to calculate their trajectories.
- Analyzing trajectories to compute various motility parameters, including but not limited to average path velocity (VAP), straight-line velocity (VSL), and curvilinear velocity (VCL), as shown in Fig. 1.

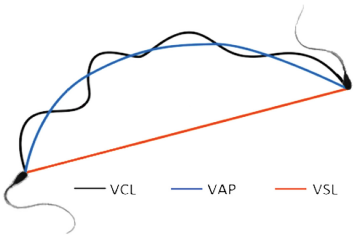


Fig. 1. Sperm Velocity Parameters (VCL, VAP, VSL)

CASA systems have provided invaluable tools for the detailed analysis of sperm motility, contributing to advances in reproductive medicine and biological research. Despite their benefits, ongoing development is necessary to address

existing limitations, enhance detection algorithms, and ensure consistency across different platforms. Therefore machine learning has a vast potential to bring an additional value of intelligence into the CASA systems by improving analysis accuracy and developing models that more closely mimic the natural environment of sperm migration.

3 Methodology

The methodology employed in this study is structured into several sequential steps, each designed to address distinct aspects of sperm cell analysis through video recordings. This systematic approach enables analysis from detection to classification and quantitative assessment. Below, we outline the key components of the methodology:

- Sperm Cell Detection: Initial analysis focuses on isolating live sperm cells from the background. This process involves employing contour detection techniques to accurately identify and delineate individual sperm cells, subsequently surrounding each identified cell with a bounding box for further analysis.
- Path Reconstruction: The trajectory of each sperm cell throughout the video sequence is reconstructed by examining the overlap of bounding boxes across consecutive frames. In instances where direct measurements are unavailable or measurements are noisy, the Kalman filter is applied to estimate the sperm cell's position, ensuring continuous tracking.
- Bounding Box Classification: Upon successful detection and tracking, each bounding box is subjected to classification using a residual neural network. This classification categorizes the contents of each box into one of four distinct groups: a single sperm cell, a bundle of cells, a group of nearby cells, or other entities.
- Velocity Calculation: With the paths of individual sperm cells established, we proceed to calculate key motility parameters such as the Straight-Line Velocity (VSL), the Average Path Velocity (VAP) and Curvilinear velocity (VCL). These calculations are performed using the reconstructed paths and are essential for assessing the straightness and linearity of the cells' path.
- Final Analysis: The culmination of this research involves an evaluation of the collected data, applying a set of predefined rules to assess and interpret the behavior and characteristics of the sperm cells.

3.1 Sperm Cell Detection

The initial step of data preprocessing involves the elimination of all static background components through the application of background subtraction techniques, with a focus on mean filtering that entails averaging all frames within the video and subsequently subtracting this average value from each frame. Given the brief duration of the videos and the stability of the environment, this method proved to be effective. This technique is instrumental in removing static entities,

predominantly non-viable sperm cells, and noise, thereby enhancing the visibility and distinction of dynamic objects such as viable sperm cells and other particulates in the fluid. For foundational insights into the approach, references include [12] and [23]. An illustration of a frame prior to background removal is presented in Fig. 2a.

In Fig. 2b, the subtracted background is illustrated. Following background removal, the resulting image foreground is depicted in Fig. 2c. After the background subtraction, the image is converted to a binary format using a dynamically determined threshold selected via Otsu's method [20]. To further refine the image and reduce noise, a series of erosion and dilation operations are applied, as detailed in [10], with the results shown in Fig. 2d. Using the binary image, bounding boxes are delineated around each distinct cluster of white pixels, ensuring each box meets predefined minimum size criteria to exclude background interference. Concurrently, the "center of mass" for each detected object is calculated based on the binary image (Fig. 2d), and this data is used to refine the Kalman filter's movement predictions as discussed in Sect. 3.2. The bounding boxes are then mapped back onto the original frame for visual evaluation, with the results presented in Fig. 2e.

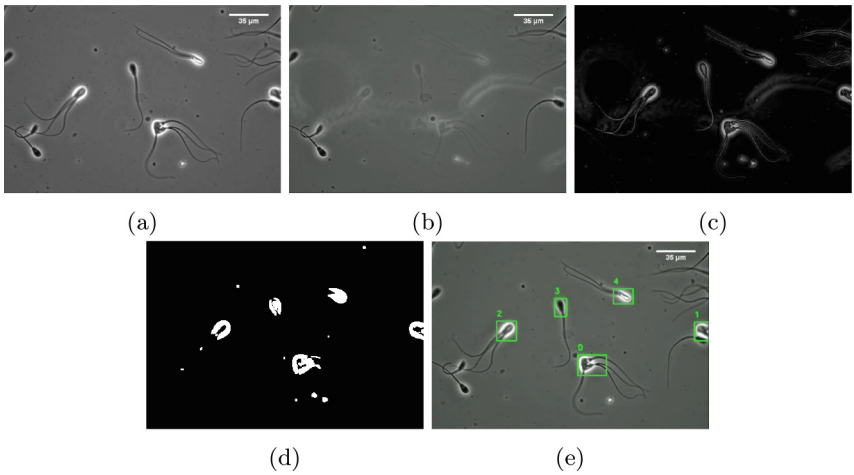


Fig. 2. Demonstration of a step-by-step process of sperm cell detection. (a) – A single input frame in a recording and the starting point of detection. (b) – An image of calculated background from the whole video that will be used to isolate the foreground. (c) -Static component removed by mean filtering. (d) – Thresholded, image derived from the frame after mean filtering. (e) – Final bounding boxes projected back to the original image.

3.2 Path Reconstruction

In the process of analyzing the trajectories of detected sperm cells within a video sequence, an important step involves the temporal association of detected entities across consecutive frames. This association task commences with the evaluation of spatial overlaps among the bounding boxes enclosing detected objects. Specifically, the overlap between bounding boxes is quantified, and a correspondence is established when the overlap exceeds an 80% threshold. This criterion is predicated on the assumption that a substantial overlap between bounding boxes in successive frames likely indicates the persistence of the same sperm cell.

After the initial matching phase, attention is directed towards sperm cells that could not be associated with counterparts in the preceding frame through bounding box overlap. Such instances of unsuccessful matching may arise from several scenarios, including the non-detection of the sperm cell in a previous frame or the occurrence of collisions. Collisions, especially with stationary sperm cells or extraneous objects, can result in an expansion of the bounding box. This expansion, in turn, diminishes the likelihood of surpassing the overlap threshold, thereby complicating the task of accurate sperm cell tracking.

This methodological approach underscores the complexities inherent in tracking the dynamic and often unpredictable movement of sperm cells through video microscopy. The challenges posed by detection inconsistencies and the potential for collisions necessitate a nuanced strategy for maintaining continuity in sperm cell trajectories across video frames.

The scenario where a sperm cell is not being identified in a sequential frame can be attributed to one of three principal scenarios. Initially, the optimal scenario is the introduction of a new sperm cell into the field of view, marking its initial detection. Alternatively, the sperm cell's visibility might have been insufficient in the antecedent frame, rendering it undetectable due to the processes of background subtraction or noise reduction. The final scenario entails the convergence of trajectories from multiple sperm cells, culminating in their placement within a singular bounding box. The current implementation of the analysis algorithm specifies that each detected bounding box from a preceding frame correlates uniquely with a single bounding box in the subsequent frame. Such convergence events are notably frequent in video samples characterized by a high density of sperm cells.

3.3 The Kalman Filter Implementation

To mitigate the challenges posed by these scenarios, particularly in maintaining the continuity of sperm cell tracking, the application of a Kalman filter is proposed. This filter functions by extrapolating the future position of each sperm cell based on its previously observed location. In instances where direct matching through bounding box overlap proves infeasible, the Kalman filter's predictive capability allows for alternative matching based on the spatial proximity between the observed position of a sperm cell and the forecasted position

from the preceding frame. This approach enhances the robustness of the tracking process, ensuring more accurate and continuous monitoring of sperm cell movement across frames, even in densely populated video sequences.

The Kalman filter is a recursive algorithm utilized for the estimation and prediction of the state of a linear dynamic system from a series of noisy measurements. It enhances path prediction accuracy by filtering out the inherent noise found in all measurements, which are invariably imperfect. The application of the Kalman filter spans numerous fields due to its ability to optimally estimate system states under the assumption of process and measurement noises being Gaussian and white.

The dynamic system under consideration is represented by the state-space model. State Equation:

$$\mathbf{x}_k = \mathbf{F}_{k-1}\mathbf{x}_{k-1} + \mathbf{B}_{k-1}\mathbf{u}_{k-1} + \mathbf{w}_{k-1} \tag{1}$$

Measurement Equation:

$$\mathbf{z}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k \tag{2}$$

where: $\mathbf{x}_k$ is the state vector at time step k, $\mathbf{F}_{k-1}$ is the state transition model, $\mathbf{B}_{k-1}$ is the control-input model, $\mathbf{u}_{k-1}$ is the control vector, $\mathbf{w}_{k-1}$ represents the process noise, $\mathbf{z}_k$ is the measurement vector, $\mathbf{H}_k$ is the measurement model, $\mathbf{v}_k$ denotes the measurement noise.

Both process noise $\mathbf{w}_k$ and measurement noise $\mathbf{v}_k$ are assumed to be Gaussian with mean zero and covariance matrices $\mathbf{Q}_k$ and $\mathbf{R}_k$, respectively. The Kalman filter operates in two main phases: the prediction phase and the update phase.

In the prediction phase, the filter predicts the current state and error covariance based on the previous estimates.
Predicted state estimate:

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{F}_{k-1}\hat{\mathbf{x}}_{k-1|k-1} + \mathbf{B}_{k-1}\mathbf{u}_{k-1} \tag{3}$$

Predicted error covariance:

$$\mathbf{P}_{k|k-1} = \mathbf{F}_{k-1}\mathbf{P}_{k-1|k-1}\mathbf{F}_{k-1}^T + \mathbf{Q}_{k-1} \tag{4}$$

In the update phase, upon receiving a new measurement, the filter updates the state and error covariance estimates.
Kalman gain:

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{H}_k^T(\mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}_k^T + \mathbf{R}_k)^{-1} \tag{5}$$

Updated state estimate:

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{z}_k - \mathbf{H}_k\hat{\mathbf{x}}_{k|k-1}) \tag{6}$$

Updated error covariance:

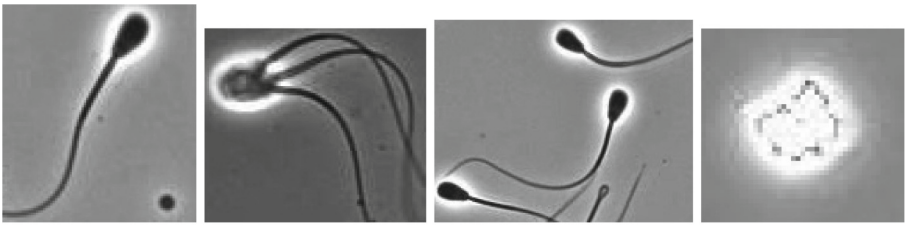
$$\mathbf{P}_{k|k} = (I - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_{k|k-1} \tag{7}$$

where I denotes the identity matrix of appropriate dimensions.

In this work, we apply the Kalman filter using a constant velocity model for motion in 2D space, where the position is predicted from the current position and velocity. This model, which simplifies the system by treating acceleration as noise, guides our estimations of the object's trajectory. The Kalman filter, provided with this model, refines these predictions by incorporating measurements, thus enhancing the accuracy of future estimations.

3.4 Bounding Box Classification

To address bounding box classification, we employ machine learning techniques to categorize the entities encapsulated within the bounding boxes detected in our video recordings. Classification framework distinctly identifies four categories: *a single sperm cell* Fig. 3a, *a bundle of sperm cells* Fig. 3b, *a group of nearby cells* Fig. 3c, and *others* Fig. 3d. Visual examples of these categories are elucidated in Fig. 3.



(a) single sperm cell (b) bundle of sperm cells (c) group of nearby cells (d) others

Fig. 3. Examples of each classified class

The primary objective of this identification process is to differentiate between single sperm cells and sperm cell bundles. On the one hand to be able to compare the motility capabilities of those two and on the other hand to help with detecting bundle formation. A focal point of our analysis is to elucidate the dynamical transitions that occur from *single sperm cells* to *a group of cells*, and ultimately to *a sperm cell bundle*. The secondary reason for bounding box identification is to filter out irrelevant background constituents present in the fluid medium that have not been filtered out during the sperm cell detection process. The outcomes of our classification process are subsequently presented to researchers, enabling them to efficiently discern the recordings of interest from those that may be deemed non-essential to their investigative pursuits. To recognize individual classes, we used the ResNet18 architecture [11] pretrained on the ImageNet dataset.

The classification process relies on the content within the bounding boxes. These bounding boxes are extracted from the frames and fed into a ResNet model. Given that ResNet requires input images to have uniform dimensions, the bounding boxes are resized to 224×224 pixels.

To train the algorithm, a custom dataset was created from manually annotated bounding boxes. In total, 10,568 bounding boxes were collected for training and validation purposes. Figure 3 illustrates examples from all four categories.

The dataset exhibits a high degree of class imbalance. The most frequent category is *single sperm cells*, whereas sperm cell bundles are comparatively rare. To develop a robust model, random undersampling was employed to reduce the prevalence of the majority class. Concurrently, the minority classes were oversampled by duplicating images. To counterbalance the oversampling and mitigate overfitting, all images were randomly cropped to introduce variability among the duplicated samples.

3.5 Final Analysis

In the final analysis, we combined the results from previous steps to evaluate the defined rules and calculate several metrics: VSL, VAP, VCL, path straightness (STR), and path linearity (LIN), highlighting significant events in the recordings.

The defined rules specify the required percentage of classifications per frame for an object to be classified into a particular category. For instance, to classify an object as a bundle, 40% of the classifications must indicate it as such. Additional rules detect patterns suggesting unusual behavior, such as the disappearance of sperm cells within a frame or classification patterns indicating bundle formation.

The final part of the analysis involves calculating velocities, as depicted in Fig. 1. Using this approach, we calculate VSL, VAP, and VCL. From these velocities, we derive STR, indicating the linearity of the average path, and LIN, indicating the linearity of the curvilinear path [13,19].

4 Results

The detection accuracy has been tested over a random sample of videos, the path recovery has been tested using four selected path scenarios, and the neural network classification accuracy has been evaluated using a random test sample of 59 videos.

First, we present the sperm cell detection accuracy. For detection accuracy, a random subset of 13 videos was selected, and all moving sperm cells were annotated for each frame. The expected number of detections for each sperm cell was then compared to the number of times this sperm cell was detected using our approach described in Sect. 3.1. The detection accuracy was 95.99%, and all the missed detections were due to the noise filtration by size rule.

The efficacy of the path reconstruction methodology was evaluated through the annotation of three distinct sperm cell trajectories, each exemplifying one of the three predominant scenarios encountered in our analysis as illustrated in Fig. 4. Path 1 delineates an ideal trajectory scenario, wherein the sperm cell is consistently detected across all video frames. Path 2 outlines a scenario in which a sperm cell's path intersects with another, rendering its measurements temporarily unattainable. During such instances, the sperm cell's location is approximated

using the Kalman filter, effectively demonstrating the filter's capability to predict the trajectory amidst measurement loss. Path 3 describes the trajectory of a sperm cell that amalgamates with another to form a bundle, leading to a cessation of measurement (red 'x') acquisition from the juncture of bundle formation. This comprehensive evaluation underscores the robustness and versatility of the proposed path reconstruction approach in accommodating various complex motion patterns exhibited by sperm cells.

To assess the accuracy of the reconstructed sperm cell paths relative to their actual trajectories, the Optimal Sub-Pattern Assignment (OSPA) metric was employed. This metric quantitatively evaluates the discrepancy between the true path of a sperm cell and its counterpart estimated through the application of the Kalman filter and recorded measurements [3]. For a comprehensive analysis, our findings were compared with accuracies derived from a prior study [27], which examined and compared the effectiveness of four distinct algorithms in tracking sperm cell movements. This comparative evaluation, illustrated in Fig. 4, facilitates a nuanced understanding of our methodology's performance in the context of existing tracking techniques.

We refined and calibrated this program using 366 videos of sperm cell samples, recorded under various scenarios and in differing qualities. Across all videos, we identified 983 sperm cells, 455 bundles, and 312 groups of sperm cells. Additionally, 257 objects were classified as 'other' and excluded from the analysis. In 153 videos, the tracking system failed to follow some objects. Overall, 117 videos were flagged for potential bundle formation.

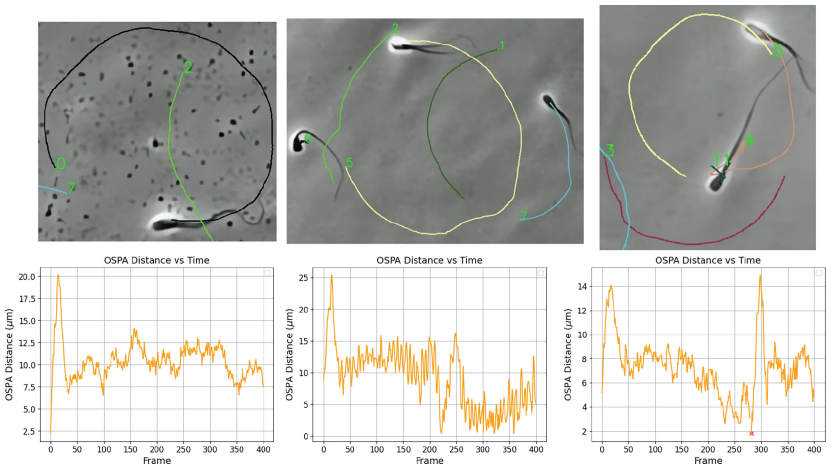


Fig. 4. Reconstruction of three different paths and their distance in μm from reality using OSPA metric. Path 1 an ideal path where measurements are not lost at all of the sperm cell 0 is shown in black (left). Path 2 a path of a sperm cell 5 (shown in yellow) that collided with other sperm cells leading to loss of measurements for several frames (middle). Path 3 a path of two sperm cells that formed a bundle (right). The red 'x' signals the point of bundle formation. (Color figure online)

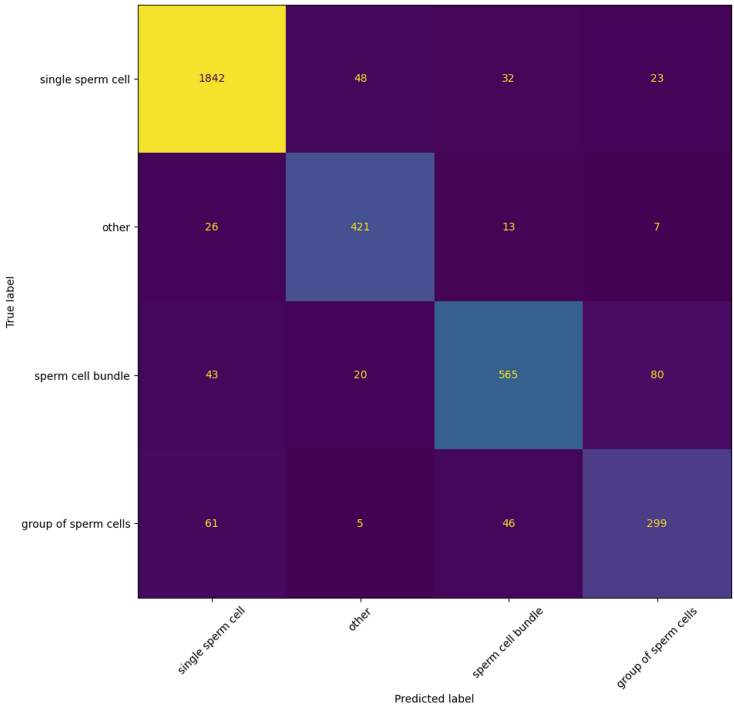


Fig. 5. Object classification confusion matrix

5 Conclusion

In this paper, we introduce a novel, multistep methodology employing machine learning techniques to track sperm cells within video recordings. Our approach demonstrates the capability to accurately identify and track individual sperm cells, as well as to distinguish these single entities from conglomerations of sperm cells, referred to here as "bundles." Furthermore, this paper highlights our method's proficiency in recognizing significant events within the recordings, including the instances of sperm cells coalescing into bundles. Additionally, our system offers the functionality to exclude recordings deemed devoid of noteworthy activity, thereby optimizing the efficiency of research efforts. Through extensive experimentation and analysis, our results affirm the high precision and reliability of our machine learning-based tracking solution in the context of sperm cell observation.

Presented within the results section of this manuscript, our findings elucidate that the proposed methodology for individual sperm cell detection achieves a detection accuracy of 95.99% for live sperm cells. This technique effectively diminishes the interference caused by stationary background elements and non-viable cells, thereby enhancing both accuracy and performance metrics. Utilizing a combination of bounding box intersections and the application of the Kalman

filter, our approach mitigates inaccuracies inherent in bounding box detection methodologies. The empirical evidence demonstrates that, under optimal conditions where measurement data remains unlost, the deviation between the path identified by our system and a path annotated manually is quantified at a distance of 5378 units. This identified path forms the basis for subsequent estimations of key motility parameters, namely the Straight Line Velocity (VSL), Average Path Velocity (VAP) and curvilinear velocity (VCL), offering insights into the dynamic behavior of sperm cells.

Concurrently with the process of path identification, our study also focuses on determining the nature of the object enclosed within each bounding box. For this purpose, we employ the ResNet architecture, a convolutional neural network renowned for its efficacy in image recognition tasks. The implementation of ResNet in our methodology is underpinned by a foundation of manually annotated data, ensuring the training phase is robust and tailored to the specific nuances of our domain. The F1 scores for each category are as follows:

- Sperm: 0.940
- Other: 0.876
- Bundle: 0.828
- More: 0.729

The confusion matrix is depicted in Fig. 5. Even though from the confusion matrix it can be seen that the model tends to confuse "bundle" with "sperm" in single shots, sequential tracking works reasonably well to detect potential bundle formation.

In the concluding phase of our analysis, the classification of objects on a per-frame basis is aggregated to facilitate the classification of entire paths. This aggregation process also involves the identification of pivotal points of transition within the paths. The synthesized classifications and identified points of transition are meticulously highlighted for the researchers' review. This strategic delineation enables researchers to prioritize their focus on segments of interest within the recordings and to discern which recordings may be extraneous to their investigative pursuits, thereby streamlining the research process.

Regarding the enhancement of path identification capabilities, we assume that transitioning to a non-linear modeling approach would more accurately encapsulate the dynamics of sperm cell motion. In parallel, the adoption of the Unscented Kalman Filter, or the exploration of more sophisticated methodologies such as the Joint Probabilistic Data Association Filter, which facilitates the simultaneous tracking of multiple targets, could offer substantial improvements. These proposed enhancements are predicated on existing literature and methodologies that have demonstrated efficacy in similar contexts [2,6,27,29].

Acknowledgments. This work was supported by the Student Summer Research Program 2023 of FIT CTU in Prague

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Amann, R.P., Waberski, D.: Computer-assisted sperm analysis (casa): capabilities and potential developments. Theriogenology **81**(1), 5-17.e3 (2014). https://doi.org/10.1016/j.theriogenology.2013.09.004. https://www.sciencedirect.com/science/article/pii/S0093691X13003555
2. Bar-Shalom, Y., Daum, F., Huang, J.: The probabilistic data association filter. IEEE Control Syst. Mag. **29**(6), 82–100 (2009). https://doi.org/10.1109/MCS.2009.934469
3. Beard, M., Vo, B.T., Vo, B.N.: Ospa(2): using the ospa metric to evaluate multi-target tracking performance. In: 2017 International Conference on Control, Automation and Information Sciences (ICCAIS), pp. 86–91 (2017). https://doi.org/10.1109/ICCAIS.2017.8217598
4. Brownscombe, J.W., et al.: Application of machine learning algorithms to identify cryptic reproductive habitats using diverse information sources. Oecologia **194**(1), 283–298 (2020)
5. Burnett, W.E., Heinze, J.: Sperm bundles in the seminal vesicles of sexually mature lasius ant males. PLOS ONE **9**(3), 1–4 (2014). https://doi.org/10.1371/journal.pone.0093383
6. Choi, J.w., Alkhoury, L., Urbano, L.F., Masson, P., VerMilyea, M., Kam, M.: An assessment tool for computer-assisted semen analysis (casa) algorithms. Sci. Rep. **12**(1), 16830 (2022). https://doi.org/10.1038/s41598-022-20943-9
7. Dorado, J., Rijsselaere, T., Muñoz-Serrano, A., Hidalgo, M.: Influence of sampling factors on canine sperm motility parameters measured by the sperm class analyzer. Syst. Biol. Reprod. Med. **57**(6), 318–325 (2011)
8. Fisher, H.S., Giomi, L., Hoekstra, H.E., Mahadevan, L.: The dynamics of sperm cooperation in a competitive environment. Proc. Roy. Soc. B: Biol. Sci. **281**(1790), 20140296 (2014). https://doi.org/10.1098/rspb.2014.0296
9. Flores, A., Wiff, R., Donovan, C.R., Gálvez, P.: Applying machine learning to predict reproductive condition in fish. Eco. Inf. **80**, 102481 (2024)
10. Haralick, R.M., Sternberg, S.R., Zhuang, X.: Image analysis using mathematical morphology. IEEE Trans. Pattern Anal. Mach. Intell. **9**(4), 532–550 (1987). https://doi.org/10.1109/TPAMI.1987.4767941
11. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition (2016)
12. Herrero, S., Bescós, J.: Background subtraction techniques: systematic evaluation and comparative analysis. In: Blanc-Talon, J., Philips, W., Popescu, D., Scheunders, P. (eds.) ACIVS 2009. LNCS, vol. 5807, pp. 33–42. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-04697-1_4
13. Hidayatullah, P., Awaludin, I., Kusumo, R.D., Nuriyadi, M.: Automatic sperm motility measurement. In: 2015 International Conference on Information Technology Systems and Innovation (ICITSI), pp. 1–5 (2015). https://doi.org/10.1109/ICITSI.2015.7437674
14. Johnston, S.D., Smith, B., Pyne, M., Stenzel, D., Holt, W.V.: One-sided ejaculation of echidna sperm bundles. Am. Nat. **170**(6), E162–E164 (2007)
15. Khalil, I.S., Dijkslag, H.C., Abelmann, L., Misra, S.: Magnetosperm: a microrobot that navigates using weak magnetic fields. Appl. Phys. Lett. **104**(22), 223701 (2014)
16. Magdanz, V., et al.: Ironsperm: sperm-templated soft magnetic microrobots. Sci. Adv. **6**(28), eaba5855 (2020)

17. Middelhoek, K.I., Magdanz, V., Abelmann, L., Khalil, I.S.: Drug-loaded ironsperm clusters: modeling, wireless actuation, and ultrasound imaging. Biomed. Mater. **17**(6), 065001 (2022)
18. Mortimer, S.T., Van der Horst, G., Mortimer, D.: The future of computer-aided sperm analysis. Asian J. Androl. **17**(4), 545 (2015)
19. Okumuş, F., Kocamaz, F., Özgür, M.E.: Using polynomial modeling for calculation of quality parameters in computer assisted sperm analysis. Comput. Sci. **6**(3), 152–165 (2021)
20. Otsu, N.: A threshold selection method from gray-level histograms. IEEE Trans. Syst. Man Cybern. **9**(1), 62–66 (1979)
21. Raef, B., Ferdousi, R.: A review of machine learning approaches in assisted reproductive technologies. Acta Informatica Medica **27**(3), 205 (2019)
22. Rijsselaere, T., Van Soom, A., Maes, D., de Kruif, A.: Effect of technical settings on canine semen motility parameters measured by the Hamilton-Thorne analyzer. Theriogenology **60**(8), 1553–1568 (2003)
23. ching S. Cheung, S., Kamath, C.: Robust techniques for background subtraction in urban traffic video. In: Proceedings of SPIE, vol. 5308, pp. 881–892 (2004). https://doi.org/10.1117/12.526886
24. Schoeller, S.F., Holt, W.V., Keaveny, E.E.: Collective dynamics of sperm cells. Philos. Trans. R. Soc. B **375**(1807), 20190384 (2020)
25. Morcillo i Soler, P., et al.: Bundle formation of sperm: influence of environmental factors. Front. Endocrinol. **13** (2022). https://doi.org/10.3389/fendo.2022.957684. https://www.frontiersin.org/articles/10.3389/fendo.2022.957684
26. Surmacz, P., Niwinska, A., Kautz, E., Gizinski, S., Faundez, R.: Comparison of two staining techniques on the manual and automated canine sperm morphology analysis. Reprod. Domest. Anim. **57**(6), 678–684 (2022)
27. Urbano, L.F., Masson, P., VerMilyea, M., Kam, M.: Automatic tracking and motility analysis of human sperm in time-lapse images. IEEE Trans. Med. Imaging **36**(3), 792–801 (2017). https://doi.org/10.1109/TMI.2016.2630720
28. Virkki, N.: Sperm bundles and phylogenesis. Z. Zellforsch. Mikrosk. Anat. **101**(1), 13–27 (1969)
29. Wan, E.A., Merwe, R.V.D.: The unscented kalman filter for nonlinear estimation. In: Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No.00EX373), pp. 153–158 (2000). https://doi.org/10.1109/ASSPCC.2000.882463
30. Zhang, K., Klingner, A., Le Gars, Y., Misra, S., Magdanz, V., Khalil, I.S.M.: Locomotion of bovine spermatozoa during the transition from individual cells to bundles. Proc. Natl. Acad. Sci. **120**(3), e2211911120 (2024). https://doi.org/10.1073/pnas.2211911120