



This MICCAI paper is the Open Access version, provided by the MICCAI Society. It is identical to the accepted version, except for the format and this watermark; the final published version is available on SpringerLink.

CS3: Cascade SAM for Sperm Segmentation

Yi Shi^{1,2(*)}, Xu-Peng Tian^{1,2(*)}, Yun-Kai Wang^{1,2}, Tie-Yi Zhang^{1,2}, Bing Yao^{3,4}, Hui Wang^{3,4}, Yong Shao^{3,4}, Cen-Cen Wang^{3,4}, Rong Zeng^{3,4}, De-Chuan Zhan^{1,2(✉)}

¹ School of Artificial Intelligence, Nanjing University, Nanjing, China

² State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing, China

³ Jinling Hospital, Affiliated Hospital of Medical School, Nanjing University, Nanjing, China

⁴ Jiangsu Provincial Medical Key Discipline Cultivation Unit, Nanjing, China

zhandc@nju.edu.cn

Abstract. Automated sperm morphology analysis plays a crucial role in the assessment of male fertility, yet its efficacy is often compromised by the challenges in accurately segmenting sperm images. Existing segmentation techniques, including the Segment Anything Model (SAM), are notably inadequate in addressing the complex issue of sperm overlap—a frequent occurrence in clinical samples. Our exploratory studies reveal that modifying image characteristics by removing sperm heads and easily segmentable areas, alongside enhancing the visibility of overlapping regions, markedly enhances SAM’s efficiency in segmenting intricate sperm structures. Motivated by these findings, we present the Cascade SAM for Sperm Segmentation (CS3), an unsupervised approach specifically designed to tackle the issue of sperm overlap. This method employs a cascade application of SAM to segment sperm heads, simple tails, and complex tails in stages. Subsequently, these segmented masks are meticulously matched and joined to construct complete sperm masks. In collaboration with leading medical institutions, we have compiled a dataset comprising approximately 2,000 unlabeled sperm images to fine-tune our method, and secured expert annotations for an additional 240 images to facilitate comprehensive model assessment. Experimental results demonstrate superior performance of CS3 compared to existing methods.

Keywords: Cascade SAM · Sperm image segmentation · Segmentation for overlapping structures

1 Introduction

Sperm morphology analysis is a crucial clinical technique for evaluating male fertility [6,4,21]. Currently, this analysis predominantly depends on the subjective evaluations of andrology specialists using microscopy, a process that is both time-consuming and labor-intensive, failing to satisfy the demands for large-scale, efficient assessments. Furthermore, the assessment of sperm morphology is highly subjective and lacks standardized criteria, presenting a significant challenge to reproductive medicine [26,8,19].

(*) means equal contribution to this work.

Consequently, there is a pressing need for an automated system to assess sperm morphology in andrology clinical diagnostics, with the primary challenge being the precise segmentation of individual sperm within images. Sperm microscopic images frequently exhibit issues such as overlapping, indistinct boundaries, and various disruptive elements, which complicate the task for existing image segmentation models to accurately isolate sperm. This complexity further hampers the manual annotation of sperm image data, leading to a critical shortage of labeled sperm image datasets.

Image segmentation technology has seen widespread application across various fields in recent years [23,14,7,20,17], with its use in medical imaging becoming increasingly prevalent [28,32,30,31,11]. In the specific context of sperm image segmentation, existing models can be categorized into three types. The first type [16,25] focuses solely on segmenting the sperm head, neglecting the tail. The second type [5,9] is capable of segmenting only non-overlapping sperms, failing to address instances where sperms overlap. The third type [13,11] treats overlapping sperms as a single entity, without distinguishing between individual sperms. These approaches all rely on labeled data for training, yet the acute scarcity of labeled sperm image datasets has led to a simplification in their segmentation targets. However, sperm overlap is a common occurrence in clinical settings, rendering these methods inadequate for practical clinical applications. The ability to effectively segment overlapping sperms, generating independent and complete masks for each sperm without relying on labeled training data, is crucial for automated sperm morphology analysis.

The sperm image segmentation task is essentially an instance segmentation task where the goal is to acquire distinct masks for each sperm without necessarily differentiating their semantic attributes. Recent advancements, such as CutLER [27] and U2Seg [22], have demonstrated promising results in unsupervised instance segmentation within common scenarios. However, these techniques often fall short in more specialized medical contexts, such as distinguishing sperms. The Segment Anything Model (SAM) [12] represents a significant leap forward in image segmentation technology. Operating in everything mode, SAM can segment images it has never seen before without any annotations or prompts. This capability has led to successful applications across various domains [15,33,1,2,29,27]. Nonetheless, our experiments indicate that SAM tends to segment only those tails that are distinct and non-overlapping, neglecting or inaccurately grouping overlapping tails without further segmentation. This limitation highlights the need for tailored approaches to address the unique challenges of sperm image segmentation.

To tackle the challenges identified in sperm image segmentation, particularly with the SAM in everything mode, we begin with preliminary experiments that yield three pivotal insights. Firstly, We find that SAM will give priority to segmentation by color, and will mainly consider geometric features when there is no obvious difference in color. Secondly, we discover that excluding easily segmentable regions from the original image prompts SAM to target more complex areas. Lastly, we find that enlarging and thickening the lines of overlapping sperm tails render these previously indistinct areas separable. Leveraging these insights, we introduce our method CS3, a cascade SAM algorithm for the end-to-end segmentation of sperm images. This unsupervised approach effectively mitigates the issues posed by overlapping sperms, enabling precise

differentiation and segmentation of individual sperms, including both heads and tails, to generate independent and complete masks for each. CS3 begins with an initial segmentation using SAM on pre-processed images, followed by the application of color filters to isolate sperm head masks. These masks are then saved and removed from the image, leaving an image with only sperm tails. Subsequently, the cascade process focuses on sperm tails, each time isolating and preserving masks that adhere to the morphological criteria of individual tails, thus allowing for a progressive segmentation from simpler to more complex forms until no further changes are observed between two successive rounds. At this stage, while most tails are independently segmented, a few remain intertwined. For these, we apply an enlargement and line-thickening technique before subjecting them to further segmentation with SAM, with the segmented results then resized to their original dimensions. Finally, CS3 matches the obtained head and tail masks based on criteria such as distance and angle to assemble complete and independent sperm masks. To our knowledge, this cascade approach using SAM for segmenting intersecting elongated structures, such as those found in sperm tails, is novel. Moreover, CS3 offers a fresh perspective for segmenting similar structures in other domains, such as vascular and neural imaging.

In collaboration with several leading hospitals, we collect approximately 2,000 unlabeled sperm images to refine the parameters of our proposed method. Additionally, we engage relevant experts to meticulously annotate 240 sperm images, serving as a benchmark for model evaluation. The empirical evidence suggests that our CS3 surpasses previous methodologies in segmentation efficacy, particularly in segmenting overlapping sperms. This advancement paves the way for leveraging artificial intelligence to fully automate sperm morphology analysis.

Our contributions are articulated as follows: (1) We identify three limitations of SAM in sperm segmentation and provide actionable solutions. (2) We develop an unsupervised method named CS3 for sperm image segmentation, employing cascade applications of SAM. (3) Experiments demonstrate that CS3 outperforms existing methods in segmentation accuracy, especially in resolving overlapping sperm instances, marking a significant advancement in the field.

2 Method

In this section, we present a detailed exposition of our CS3, a novel algorithm designed for the unsupervised segmentation of sperm images using cascade applications of SAM. Initially, we highlight three critical experimental observations identified during the application of SAM in everything mode for the task of sperm image segmentation. These observations serve as foundational insights that inspire the development and comprehensive delineation of the CS3 process.

2.1 Preliminary Investigations into SAM

In our exploratory studies, we employ SAM in everything mode to perform segmentation on clinical sperm images. Raw images are initially pre-processed for better performance of SAM. Our results show that SAM effectively segments sperm heads but

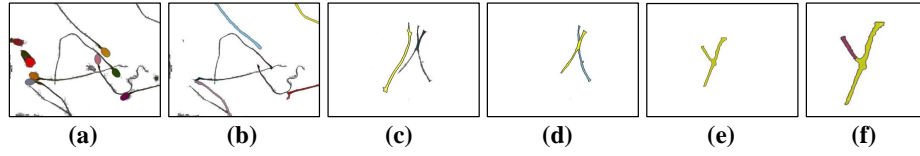


Fig. 1. Preliminary Investigations into SAM on pre-processed sperm images. (a)-(b) After removing sperm heads, SAM begins to segment remaining parts from a geometric perspective. (c)-(d) After removing simple parts, SAM begins to focus on complex areas. (e)-(f) After enlarging and bolding, overlapping parts that are difficult to split become separable.

struggles with tail segmentation. We speculate that SAM will give priority to segmentation based on color, so that the segmentation effect on sperm heads with obviously different colors is better, which is also consistent with the findings in [10,1,3,17]. When we try to remove sperm heads from the image so that the remaining parts have similar colors, SAM intends to divide it from a geometric perspective. At this time, SAM can segment some simple sperm tails, as shown in Figure 1(a)-(b). Notably, SAM’s capability is confined to the identification of distinct tails, showing limitations in recognizing overlapping sperm tails or in delineating them into separate entities. To enhance SAM’s segmentation proficiency for sperm imagery, modifications are attempted on the tail images. Figure 1(c)-(d) demonstrates that omitting the readily segmentable tails from the imagery prompts SAM to shift focus towards the intricate tails, thereby initiating their segmentation. This discovery inspires us to design a cascade structure to segment the entire image step by step. Besides, Figure 1(e)-(f) shows that in some cases SAM faces challenges in segmenting overlapping slender tails into distinct entities. We find that enlarging these regions and bolding the slender structures enable SAM to recognize and further partition these areas. All these experimental observations are instrumental in the conceptualization and formulation of the CS3 algorithm.

2.2 Cascade SAM for Sperm Segmentation (CS3)

The architecture of our CS3 is shown in Figure 2. Initially, a series of preprocessing steps are applied to the raw images to enhance quality for segmentation. These steps include adjustments to brightness, contrast, and saturation, along with background whitening, aimed at reducing noise and emphasizing the primary features of the sperm. Then, CS3 uses a sequence of SAMs, denoted as S_1, S_2, \dots, S_n , for stepwise segmentation of pre-processed images. Inspired by our first observation, S_1 is used to segment sperm heads, simplifying the remaining parts’ colors for easier geometric division by subsequent SAM applications. By intersecting the obtained masks from S_1 with the purple regions of the raw image and applying a threshold filter based on the intersection’s proportion, we efficiently isolate all head masks. These masks are kept and then eliminated from the original image, resulting in an image that only contains sperm tails.

The process from S_2 to S_n takes full advantage of our second preliminary observation. After each round of SAM, the newly acquired masks undergo a filtration process.

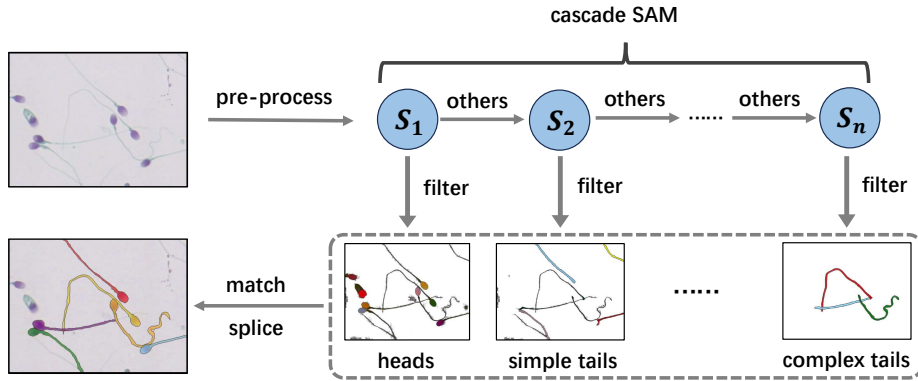


Fig. 2. The architecture of our method CS3. Through the cascade application of SAM, sperm heads, simple tails, and complex tails are segmented step by step. All masks are matched and spliced to generate complete masks. S_1, S_2, \dots, S_n denotes a series of SAM in everything mode.

This step involves preserving and discarding single tail masks, thereby focusing on overlapping tails and those not yet detected by SAM. To accurately identify single tail masks, obtained masks are first skeletonized into one-pixel-wide lines. These lines are then subjected to two critical filtration criteria: (1) the presence of a single connected segment, and (2) the line terminating in *exactly* two endpoints. The first criterion ensures the exclusion of masks representing multiple aggregated tails, while the second confirms the mask outlines a solitary tail. Masks conforming to these specifications are deemed single tail masks, saved, and then removed from the original image. Following this extraction, the image undergoes a denoising process to enhance clarity and reduce potential segmentation errors. This cascade process persists until SAM’s segmentation outputs remain consistent across two successive rounds.

The cascade SAM methodology delineated herein proves effective in resolving the majority of instances involving overlapping sperm. However, a marginal subset of these overlaps presents a notable challenge, resisting separation through cascade processing. To address these particularly resilient cases, CS3 uses the enlargement and bold method inspired by our third preliminary observation to solve the problem in the last round of SAM. Specifically, we first use the following two rules to filter out those parts that still overlap: (1) the presence of a single connected segment, and (2) the line terminating in *more than* two endpoints. Each selected mask is then emphasized by enlarging its presence within the image and augmenting the outline’s width. This process involves isolating the mask as the foreground, cropping the background to enhance the foreground’s relative size, and applying edge detection to delineate and thicken the outline. Subsequent smoothing of the enlarged outline mitigates any resultant jaggedness. These tailored steps facilitate a further round of SAM segmentation to these obstinate overlapping areas, with the segmented masks thereafter resized and repositioned to their original context within the image.

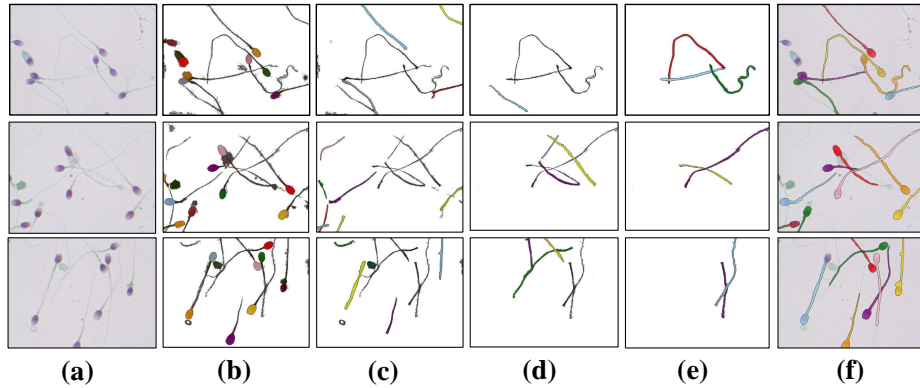


Fig. 3. Segmentation results of CS3. (a) The raw images. (b)-(e) The results of cascade SAM gradually segmenting different parts of sperm from simple to complex. (f) The results after mask matching and splicing. More results can be found in the appendix.

The culmination of this process entails the matching and splicing of the corresponding head and tail masks to construct comprehensive sperm masks. To achieve this, an elliptical fitting technique is applied to delineate each sperm head, leveraging the ellipse’s major axis to closely approximate the head’s geometry. Tail structures are represented by their skeletonized lines. Subsequently, the endpoints of these skeletonized tails and the major axes of the heads are identified, with particular emphasis on the terminal slopes at these endpoints. Our matching process is grounded in the hypothesis that a minimal distance between endpoints and a similarity in the terminal slopes indicate a higher likelihood of the endpoints belonging to the same sperm entity. Therefore, we establish a distance threshold λ_{dis} and an angle threshold λ_{angle} for filtering. In cases where multiple endpoints meet these criteria for a single anchor endpoint, preference is given to the one closest in angle. Successful matching yields the assembly of complete sperm masks, effectively bridging the head and tail components. All results in the process of CS3 are shown in Figure 3.

3 Experiment

Dataset. In a collaborative effort with leading hospitals, we collect a dataset of approximately 2,000 unlabeled sperm images directly from clinical practice, vital for our model’s parameter optimization. In addition, relevant andrology experts also annotate an additional 240 sperm images, and the annotation results of each image are checked and approved by at least three senior experts for model evaluation. These original images are all 720*540 pixels in size. For those comparing methods that require labeled samples for training, we divide all 240 existing labeled images according to 3:1, that is, 180 images are used for their training, and 60 images for their performance evaluation.

Evaluation metric. Drawing on previous studies [5,9,16], this research adopts two prevalent evaluation metrics in the domain of instance segmentation: the mean Intersection over Union (mIOU) and the mean Dice coefficient (mDice). Here, “mean” signifies the computation of average metric values across all instances. mIOU and mDice assess a model’s accuracy by measuring the extent of overlap between the actual areas (denoted as $A_i, i = 1, \dots, N$) and the predicted areas (denoted as $B_i, i = 1, \dots, N$). A greater overlap corresponds to a higher metric value, indicating superior model performance. The specific definitions are as follows:

$$\text{mIoU} = \frac{1}{N} \sum_{i=1}^N \left(\frac{|A_i \cap B_i|}{|A_i \cup B_i|} \right), \quad \text{mDice} = \frac{1}{N} \sum_{i=1}^N \left(\frac{2 \times |A_i \cap B_i|}{|A_i| + |B_i|} \right).$$

Before calculating mIOU and mDice, we ascertain the optimal pairing relationship by comparing the IOU ratios between each instance in the Ground Truth and the predicted results. It is important to note that the segmentation outcomes of Ground Truth and some methods may not correspond on a one-to-one basis, potentially resulting in disparate quantities. In such cases, we employ a filtering criterion predicated on the logic that pairs with an optimal pairing IoU ratio are selected for calculation.

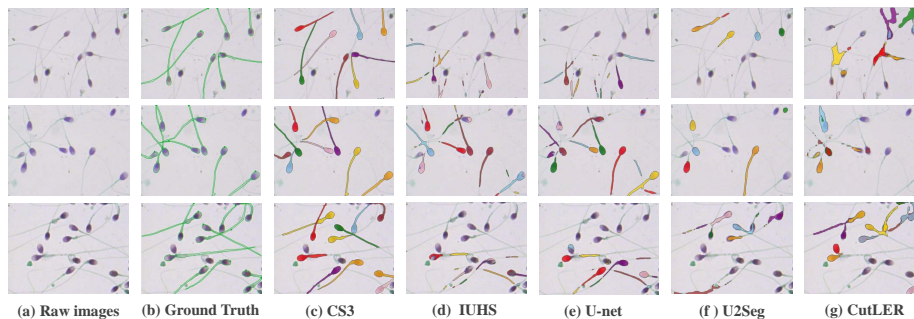
Comparison methods. This study employs two types of principal methodologies for comparison. The first type centers on supervised learning strategies derived from U-net [24], extensively applied on sperm segmentation, such as IUHS [16] and CN2UA [18]. Owing to the unavailability of source codes, this research undertake independent replication of IUHS and CN2UA, guided by the detailed descriptions provided in these papers. The second type encompasses unsupervised instance segmentation, such as SAM [12], U2Seg [22], and CutLER [27]. For these approaches, the deployment of models are facilitated by utilizing source codes provided by the original authors.

Implementation details. Through statistical analysis of unlabeled images, we identify the HSV (Hue, Saturation, Value) range for detecting purple regions as (100.0 – 180.0, 20.0 – 255.0, 20.0 – 255.0). This range is crucial for the filtering process in generating head masks. Furthermore, we employ a 3×3 convolution kernel to determine line endpoints, defining them as pixels connected to only one neighboring pixel. The distance threshold λ_{dis} and the angle threshold λ_{angle} are set to 20 pixels and 60 degrees, respectively. The enlarging and bolding as well as the corresponding restoration steps are all completed automatically by Python code without manual intervention. Our analysis indicates that images in our test set require an average of 4.8 rounds of SAM for optimal processing, with some needing up to 7 rounds. Codes of CS3 are available at <https://github.com/shiy19/CS3>.

Performance evaluation and visualization. The experimental findings on our sperm dataset are illustrated in Table 1, where both the comparative methods and our CS3 are applied to pre-processed images. Supervised techniques such as U-Net, IUHS, and CN2UA necessitate extensive labeled datasets for training, rendering them less effective on sperm datasets that lack substantial labeled data. These methods often oversimplify the complex task of sperm segmentation, leading to their ineffectiveness on our real-

Table 1. mIOU and mDice of different methods on our collected sperm image dataset. The best results are in bold.

Method	U-net	IUHS	CN2UA	SAM	U2Seg	CutLER	CS3
mIOU	49.21	47.56	19.26	34.17	16.44	31.21	72.50
mDice	62.17	58.46	32.94	46.75	28.86	44.17	80.26

**Fig. 4.** Segmentation results of different methods. More results can be found in the appendix.

world clinical sperm dataset. On the other hand, unsupervised instance segmentation techniques like SAM, U2Seg, and CutLER depend on recognizing texture and structural patterns within the images. However, sperm images significantly diverge from conventional natural images due to their high contrast, scant texture, and the interference caused by dye blocks. These distinct characteristics hinder the learning process from such images. While some methods utilize advanced data augmentation strategies to introduce variability, these attempts can inadvertently remove essential features of sperm imagery, such as tail position and morphology, thereby degrading the model’s performance. Additionally, the common occurrence of numerous elongated tails and their uneven distribution complicates the task of accurately characterizing these structures. The comparative methods generally struggle with identifying and processing these complex, overlapping structures. In contrast, our CS3 adeptly segments sperm images by progressively handling simpler to more challenging regions through the cascade application of SAM, particularly excelling in segmenting overlapping parts. By effectively assembling the head and tail components, CS3 achieves independent and complete segmentation of individual sperm without relying on labeled data for training. Therefore, our CS3 results in superior performance across various evaluated metrics. Visual results achieved by different methods are displayed in Figure 4.

4 Conclusion

We experimentally explore the difficulties of SAM in sperm image segmentation and propose effective solutions. Based on these experimental observations, we develop an

unsupervised algorithm named CS3 for the automated segmentation of sperm morphology images, specifically addressing the challenge of overlapping. Through the cascade application of SAM, CS3 can gradually segment different parts of the sperm image from simple to difficult, and then combine these segmentation results into complete sperm masks. Our experimental results on a clinical sperm dataset collected in-house demonstrate superior segmentation performance of CS3 compared to previous methods. Future work will focus on researching automated sperm morphological classification methods based on the segmentation results from CS3.

Limitations. CS3 faces challenges in processing images with excessively complex tail overlaps, such as scenarios where more than ten sperm are entangled, a situation that even human experts find difficult to discern. Therefore, adequate dilution of semen samples is still necessary in clinical practice to ensure manageable analysis conditions.

Acknowledgments. This research was supported by NSFC (61921006) and Clinical Technology Foundation of Jinling Hospital (22LCZLXJS68).

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

References

1. Archit, A., Nair, S., Khalid, N., Hilt, P., Rajashekar, V., Freitag, M., Gupta, S., Dengel, A., Ahmed, S., Pape, C.: Segment anything for microscopy. *bioRxiv* pp. 2023–08 (2023)
2. Cheng, J., Ye, J., Deng, Z., Chen, J., Li, T., Wang, H., Su, Y., Huang, Z., Chen, J., Jiang, L., et al.: Sam-med2d. *CoRR* **abs/2308.16184** (2023)
3. Deng, R., Cui, C., Liu, Q., Yao, T., Remedios, L.W., Bao, S., Landman, B.A., Wheless, L.E., Coburn, L.A., Wilson, K.T., et al.: Segment anything model (sam) for digital pathology: Assess zero-shot segmentation on whole slide imaging. *CoRR* **abs/2304.04155** (2023)
4. Eliasson, R.: Semen analysis with regard to sperm number, sperm morphology and functional aspects. *Asian journal of andrology* **12**(1), 26 (2010)
5. Fraczek, A., Karwowska, G., Miler, M., Lis, J., Jezierska, A., Mazur-Milecka, M.: Sperm segmentation and abnormalities detection during the ICSI procedure using machine learning algorithms. In: *HSI*. pp. 1–6 (2022)
6. Gatimel, N., Moreau, J., Parinaud, J., Léandri, R.: Sperm morphology: assessment, pathophysiology, clinical relevance, and state of the art in 2017. *Andrology* **5**(5), 845–862 (2017)
7. Ghosh, S., Das, N., Das, I., Maulik, U.: Understanding deep learning techniques for image segmentation. *ACM computing surveys* **52**(4), 1–35 (2019)
8. van den Hoven, L., Hendriks, J.C., Verbeet, J.G., Westphal, J.R., Wetzels, A.M.: Status of sperm morphology assessment: an evaluation of methodology and clinical value. *Fertility and sterility* **103**(1), 53–58 (2015)
9. Ilhan, H.O., Sigirci, I.O., Serbes, G., Aydin, N.: A fully automated hybrid human sperm detection and classification system based on mobile-net and the performance comparison with conventional methods. *Medical biological engineering computing* **58**(5), 1047–1068 (2020)
10. Ji, W., Li, J., Bi, Q., Li, W., Cheng, L.: Segment anything is not always perfect: An investigation of sam on different real-world applications. *CoRR* **abs/2304.05750** (2023)

11. Kheirkhah, F.M., Mohammadi, H.R.S., Shahverdi, A.: Efficient and robust segmentation and tracking of sperm cells in microscopic image sequences. *IET Computer Vision* **13**(5), 489–499 (2019)
12. Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W., Dollár, P., Girshick, R.B.: Segment anything. In: *ICCV*. pp. 3992–4003 (2023)
13. Lewandowska, E., Wesierski, D., Mazur-Milecka, M., Lis, J., Jezierska, A.: Ensembling noisy segmentation masks of blurred sperm images. *Computers in Biology and Medicine* **166**, 107520 (2023)
14. Liu, X., Deng, Z., Yang, Y.: Recent progress in semantic image segmentation. *Artificial Intelligence Review* **52**, 1089–1106 (2019)
15. Liu, Y., Zhu, M., Li, H., Chen, H., Wang, X., Shen, C.: Matcher: Segment anything with one shot using all-purpose feature matching. *CoRR* **abs/2305.13310** (2023)
16. Lv, Q., Yuan, X., Qian, J., Li, X., Zhang, H., Zhan, S.: An improved u-net for human sperm head segmentation. *Neural Processing Letters* **54**(1), 537–557 (2022)
17. Ma, J., He, Y., Li, F., Han, L., You, C., Wang, B.: Segment anything in medical images. *Nature Communications* **15**(1), 654 (2024)
18. Marín, R., Chang, V.: Impact of transfer learning for human sperm segmentation using deep learning. *Computers in Biology and Medicine* **136**, 104687 (2021)
19. Menkveld, R., Holleboom, C.A., Rhemrev, J.P.: Measurement and significance of sperm morphology. *Asian journal of andrology* **13**(1), 59 (2011)
20. Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N., Terzopoulos, D.: Image segmentation using deep learning: A survey. *IEEE transactions on pattern analysis and machine intelligence* **44**(7), 3523–3542 (2021)
21. Moruzzi, J.F., Wyrobek, A.J., Mayall, B.H., Gledhill, B.L.: Quantification and classification of human sperm morphology by computer-assisted image analysis. *Fertility and sterility* **50**(1), 142–152 (1988)
22. Niu, D., Wang, X., Han, X., Lian, L., Herzig, R., Darrell, T.: Unsupervised universal image segmentation. *CoRR* **abs/2312.17243** (2023)
23. Patil, D.D., Deore, S.G.: Medical image segmentation: a review. *International Journal of Computer Science and Mobile Computing* **2**(1), 22–27 (2013)
24. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: *MICCAI*. pp. 234–241. Springer (2015)
25. Shahzad, S., Ilyas, M., Lali, M.I.U., Rauf, H.T., Kadry, S., Nasr, E.A.: Sperm abnormality detection using sequential deep neural network. *Mathematics* **11**(3), 515 (2023)
26. Sun, F., Ko, E., Martin, R.H.: Is there a relationship between sperm chromosome abnormalities and sperm morphology? *Reproductive Biology and Endocrinology* **4**(1), 1–5 (2006)
27. Wang, X., Girdhar, R., Yu, S.X., Misra, I.: Cut and learn for unsupervised object detection and instance segmentation. In: *CVPR*. pp. 3124–3134 (2023)
28. Wei, X.Y., Yang, Z.Q., Zhang, X.L., Liao, G., Sheng, A.L., Zhou, S.K., Wu, Y., Du, L.: Deep collocative learning for immunofixation electrophoresis image analysis. *IEEE Transactions on Medical Imaging* **40**(7), 1898–1910 (2021)
29. Wong, H.E., Rakic, M., Guttag, J., Dalca, A.V.: Scribbleprompt: Fast and flexible interactive segmentation for any medical image. *CoRR* **abs/2312.07381** (2023)
30. Xie, Y., Liao, H., Zhang, D., Chen, F.: Uncertainty-aware cascade network for ultrasound image segmentation with ambiguous boundary. In: *MICCAI*. pp. 268–278 (2022)
31. Yu, X., Pang, W., Xu, Q., Liang, M.: Mammographic image classification with deep fusion learning. *Scientific Reports* **10**(1), 14361 (2020)
32. Zeng, B., Lin, Y., Wang, Y., Chen, Y., Dong, J., Li, X., Zhang, Y.: Semi-supervised pr virtual staining for breast histopathological images. In: *MICCAI*. pp. 232–241 (2022)

33. Zou, X., Yang, J., Zhang, H., Li, F., Li, L., Wang, J., Wang, L., Gao, J., Lee, Y.J.: Segment everything everywhere all at once. In: NeurIPS (2023)