Sensors & Diagnostics

PAPER

View Article Online
View Journal | View Issue

Check for updates

Cite this: *Sens. Diagn.*, 2025, **4**, 336

Leveraging synthetic imagery and YOLOv8 for a novel colorimetric approach to paper-based point-of-care male fertility testing†

Olgac Özarslan,‡[a] Begum Kubra Tokyay, ‡[ai] Cansu Soylemez,[a] Mehmet Tugrul Birtek, [a] Zihni Onur Uygun, [d] İpek Keles, [e] Begum Aydogan Mathyk, [f] Cihan Halicigil [g] and Savas Tasoglu *[bchij]

The development of paper-based systems has revolutionized point-of-care (POC) applications by enabling rapid, robust, accurate and sensitive biochemical analysis, infectious disease diagnosis, and fertility monitoring, in particular, in male fertility monitoring, offering portable, cost-effective solutions compared to traditional methods. This innovation addresses high costs and limited accessibility of male fertility testing in resource-poor settings. Male infertility, a significant issue globally, often faces stigma, hindering men from seeking care. This study introduces a novel approach to male fertility testing using colorimetric analysis of paper-based assays, enhanced by synthetic imagery and the YOLOv8 (You Only Look Once) object detection algorithm. Synthetic imagery was employed to train and fine-tune YOLOv8, enhancing its capability to accurately detect color changes in paper-based tests. This colorimetric detection leverages smartphone imaging, making it both accessible and scalable. Initial experiments demonstrate that YOLOv8's precision and efficiency, when combined with synthetic data, significantly enhance the system's ability to recognize and analyze colorimetric signals, positioning it as a promising tool for male fertility POC diagnostics. In our study, we evaluated 39 semen samples for pH and sperm count using standard clinical tests, comparing these results with a novel paper-based semen analysis kit. This kit utilizes reaction zones that exhibit color changes when exposed to semen samples, with images captured using a smartphone under varied lighting conditions. Despite a limited number of images, our synthetically trained YOLOv8 model achieved an accuracy of 0.86, highlighting its potential to improve the reliability of colorimetric analysis for both home and clinical use.

Received 14th November 2024,
Accepted 16th January 2025

DOI: 10.1039/d4sd00348a

rsc.li/sensors

Open Access Article. Published on 31 January 2025. Downloaded on 4/21/2026 9:00:16 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Introduction

Male infertility is a critical global health issue, affecting millions of couples and posing a challenge in contexts where access to laboratory-based fertility testing is limited. Infertility affects approximately 15% of couples worldwide, with male factors contributing to about half of these cases.[1,2] In fact, male factor infertility has been identified as the sole cause of 20–30% of infertility cases. Suboptimal sperm parameters account for 40–50% of infertility cases in males, and sperm analysis can enable the assessment of several factors that can lead to an infertility diagnosis, including sperm concentration, morphology, motility, pH, and DNA analysis.[3] Traditional semen analysis relies on skilled healthcare professionals and expensive, complex equipment. This dependency on specialized resources limits access to semen analysis in resource-poor areas and can discourage men from seeking testing due to cultural norms or privacy concerns. Furthermore, traditional methods often involve high medical costs and slow processing times. Among semen parameters, pH plays a critical role, as it influences the stability of

[a] *Graduate School of Sciences & Engineering, Koc University, Istanbul, 34450, Turkiye*
[b] *Koç University Is Bank Artificial Intelligence Lab (KUIS AI Lab), Koç University, Sariyer, Istanbul 34450, Turkiye. E-mail: stasoglu@ku.edu.tr*
[c] *Mechanical Engineering Department, School of Engineering, Koç University, Istanbul, 34450 Turkiye*
[d] *Department of Medical Biochemistry, Faculty of Medicine, Kafkas University, Kars 36100, Turkiye*
[e] *Koc University Hospital, Assisted Reproduction Unit, Istanbul, Turkiye*
[f] *Department of Obstetrics and Gynecology, Morsani College of Medicine, University of South Florida, Tampa, Florida, 33606, USA*
[g] *Yale University School of Medicine, Dept. Obstetrics, Gynecology & Reproductive Sciences, Yale University, Connecticut 06520, USA*
[h] *Koç University Arçelik Research Center for Creative Industries (KUAR), Koc University, Istanbul, 34450 Turkiye*
[i] *Koç University Translational Medicine Research Center (KUTTAM), Koc University, Istanbul, 34450 Turkiye*
[j] *Bogazici Institute of Biomedical Engineering, Bogazici University, Istanbul, 34684 Turkiye*

† Electronic supplementary information (ESI) available. See DOI: **https://doi.org/10.1039/d4sd00348a**

‡ These authors contributed equally to this work.

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 31 January 2025. Downloaded on 4/21/2026 9:00:16 PM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

biomolecules and supports sperm functionality essential for fertilization. According to World Health Organization (WHO) guidelines (2010), a healthy semen pH should be between 7.2 and 8.0.[4] Alongside pH, sperm count, or the concentration of sperm cells in a semen sample, serves as a vital indicator of male fertility. It directly reflects fertility health by helping to diagnose potential issues in sperm production, which may arise from hormonal imbalances, genetic factors, or lifestyle influences.[5] A sperm count of 15 million per milliliter or more is considered to be within the normal range based on the WHO laboratory manual.[4] Accordingly, evaluating both the pH and sperm count is essential for a comprehensive fertility assessment, as low sperm counts combined with an imbalanced pH can significantly hinder conception potential.

Paper-based sensor systems offer a solution to many existing challenges by enabling user-friendly sperm testing in the comfort and privacy of a patient's home.[6] Additionally, paper-based microfluidics integrated into POC tests can provide low-cost, disposable, rapid, and sensitive sperm analysis[7] to: (i) assist clinicians in initial infertility screening without imposing significant financial burdens on patients,[8] (ii) facilitate male infertility diagnosis in resource-limited settings, and (iii) help alleviate concerns and stigma associated with male infertility testing.[9–11] Colorimetric detection using a mobile phone camera is challenging due to variations in lighting, angle, and camera quality, which can lead to inconsistent color readings and reduced accuracy in quantitative analysis. Additionally, differences in color calibration across devices and ambient light interference can complicate the reliable detection and interpretation of color changes, especially for subtle variations.[12] YOLO is a high-speed, accurate object detection model capable of recognizing and locating items in images, making it ideal for real-time applications. For colorimetric detection in home-based paper tests, YOLO can be fine-tuned to detect and analyze specific color changes or intensities on test strips, enabling accurate, automated interpretation of results from smartphone-captured images, even in variable lighting and environments.[13] However, the number of images required to fine-tune a YOLO model effectively can vary depending on the complexity of the task, the variation within the dataset, and the accuracy desired. Generally, a minimum of 500–1000 labeled images per class is recommended for initial fine-tuning, but more images (often in the range of 5000–10 000) can significantly improve performance, especially for challenging or nuanced detection tasks like colorimetric analysis.[14]

To create a YOLO model capable of accurately capturing custom sensing regions and assigning correct labels, the dataset should include high-resolution images of the paper-based sensors with bounding box annotations around each region of interest (ROI), labeled according to the analyte level or test result. Additionally, a variety of lighting conditions, color variations, augmented samples, and synthetic images (if necessary) will help the model adapt to real-world variations and detect subtle color changes effectively. Obtaining a dataset for a colorimetric semen test kit is challenging due to the need for precise color variations that represent different fertility levels, which may not be naturally abundant or easy to replicate. Additionally, capturing consistent, high-quality images under diverse real-world conditions, while also labeling and annotating each sensing region accurately, requires extensive time, specialized equipment, and rigorous quality control to ensure the model's effectiveness in various environments.

Unreal Engine and Unity are popular, powerful game engines widely used for creating 3D environments, simulations, and interactive experiences due to their sophisticated rendering and lighting capabilities, physics engines, and flexible development tools. They enable developers to design realistic scenes by combining advanced shaders, global illumination, and ray tracing techniques to simulate natural light behaviors. Through these tools, artists and developers can fine-tune lighting and textures, allowing for highly accurate visual renderings that are essential for realism in games, simulations, and even architectural visualization. We utilized Unity to generate synthetic images that closely mimic actual semen test kits by implementing custom shaders that procedurally generate varying sensing regions based on the sensing regions of actual semen test images under varying lighting conditions.

Here, we present a paper-based colorimetric semen analysis sensor to accurately measure the sperm count and pH along with a mobile application that includes an ML-enabled colorimetric image analysis system which can overcome several drawbacks associated with conventional semen testing and provide a low-cost, accurate, and user-friendly male fertility monitoring strategy. A laser cutter was used to fabricate multiple channels and reaction zones on Whatman filter paper. Reaction zones were chemically modified to allow color changes based on the sperm count and pH values of patient semen samples and the resulting color changes captured using a smartphone. The captured images are standardized by two pre-processing steps. YOLOv8 by Ultralytics employed to detect and quantify color changes and map the corresponding labels, thereby minimizing inter-user variability in result interpretation, and the pipeline of this study is shown in Fig. 1. Following the development of paper-based tests, samples with known pH and sperm count values (established using conventional clinical laboratory tests) were applied onto test strips. Images of test strips were then captured using a smartphone at various orientations and light conditions. The sensing regions of those images are then used for procedural image generation to fine tune the YOLOv8 model. Although the model achieved 0.86 accuracy, it is significant to acknowledge the success of the classification approach under highly varying disturbances of the smart phone imaging with a scarce dataset. Leveraging computer graphics algorithms, synthetic images with the fine-tuned powerful YOLO-based model prove to be a highly promising tool for addressing paper-based colorimetry, since it eliminates troublesome data gathering processes. This

Open Access Article. Published on 31 January 2025. Downloaded on 4/21/2026 9:00:16 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

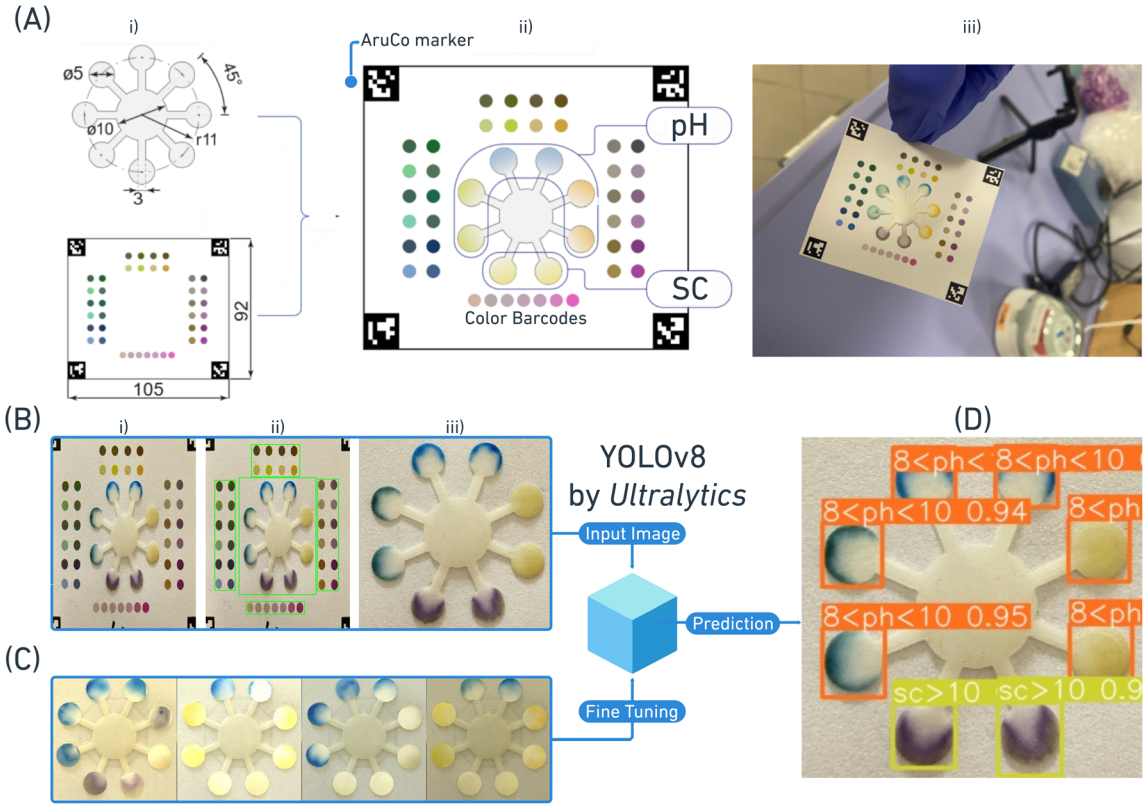

Fig. 1 Proposed instrument and workflow for enabling colorimetric paper-biosensor sensing with scarce semen samples. (A) Proposed instrument. i) Dimensions of the reference color paper and cellulose paper of the assay. All dimensions are given in millimeters. *r*: radius; Ø: diameter. ii) Key elements of the instrument. AruCo markers for edge detection. 6 pH-sensing regions and 2 sperm count-sensing regions. Color barcodes to experiment with color formations under varying lighting conditions. iii) Actual image of the designed instrument. (B) Preprocessing steps after imaging interfaced paper using a mobile phone. i) AruCo markers are *detected,* and perspective warp is applied to the image by stretching the detected corner coordinates to the edge of the image. ii) Pattern matching is applied to separate a flower-shaped sensing region. iii) Obtained flower shape resized to a 256 × 256 image. (C) Due to the challenging nature of obtaining a well quantified, representative semen analysis dataset, a procedural image generator is designed to produce synthetic images to fine tune the YOLOv8 by Ultralytics to obtain bounding boxes and the corresponding labels. (D) Fine-tuned model prediction.

system can revolutionize male fertility tracking particularly in areas with limited access to healthcare resources and simultaneously support clinicians in screening for male infertility.

Materials and methods

The detailed information is provided in the ESI.†

Results and discussion

Effectiveness of laser cutting

The laser cutting method significantly improved channel precision and eliminated leakage issues observed with traditional hydrophobic marker plotting. Unlike ink-based barriers, which were disrupted by the surfactant molecules in semen samples, laser-cut channels and double-sided tape provided robust hydrophobic barriers, ensuring consistent test performance. Details and illustrations of the design are provided in the ESI† (Fig. S1).

pH analysis

In this study, we examined the effects of pH variations on semen samples using a microfluidic testing platform. One of the semen samples was precisely adjusted across a pH range of 5 to 10 using 1 M NaOH and 1 M HCl, with each pH condition tested in duplicate for robustness. Following sample application to the tests, the tests were incubated for 10 minutes at room temperature to permit chemical interactions to be apparent. The results indicated a specific response correlating with pH adjustments, aligning closely with theoretical expectations. Notably, the platform we developed demonstrated robust pH measurement capability, which, combined with the use of synthetic images, enhances the potential for improved results and insights (Fig. 2B).

MTT assay and the calibration curve

The MTT assay was utilized to evaluate the sperm concentration by assessing the conversion of yellow tetrazolium dye to purple formazan, catalyzed by the flavoprotein enzyme diaphorase present in metabolically active human sperm (Fig. S2†). This

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 31 January 2025. Downloaded on 4/21/2026 9:00:16 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

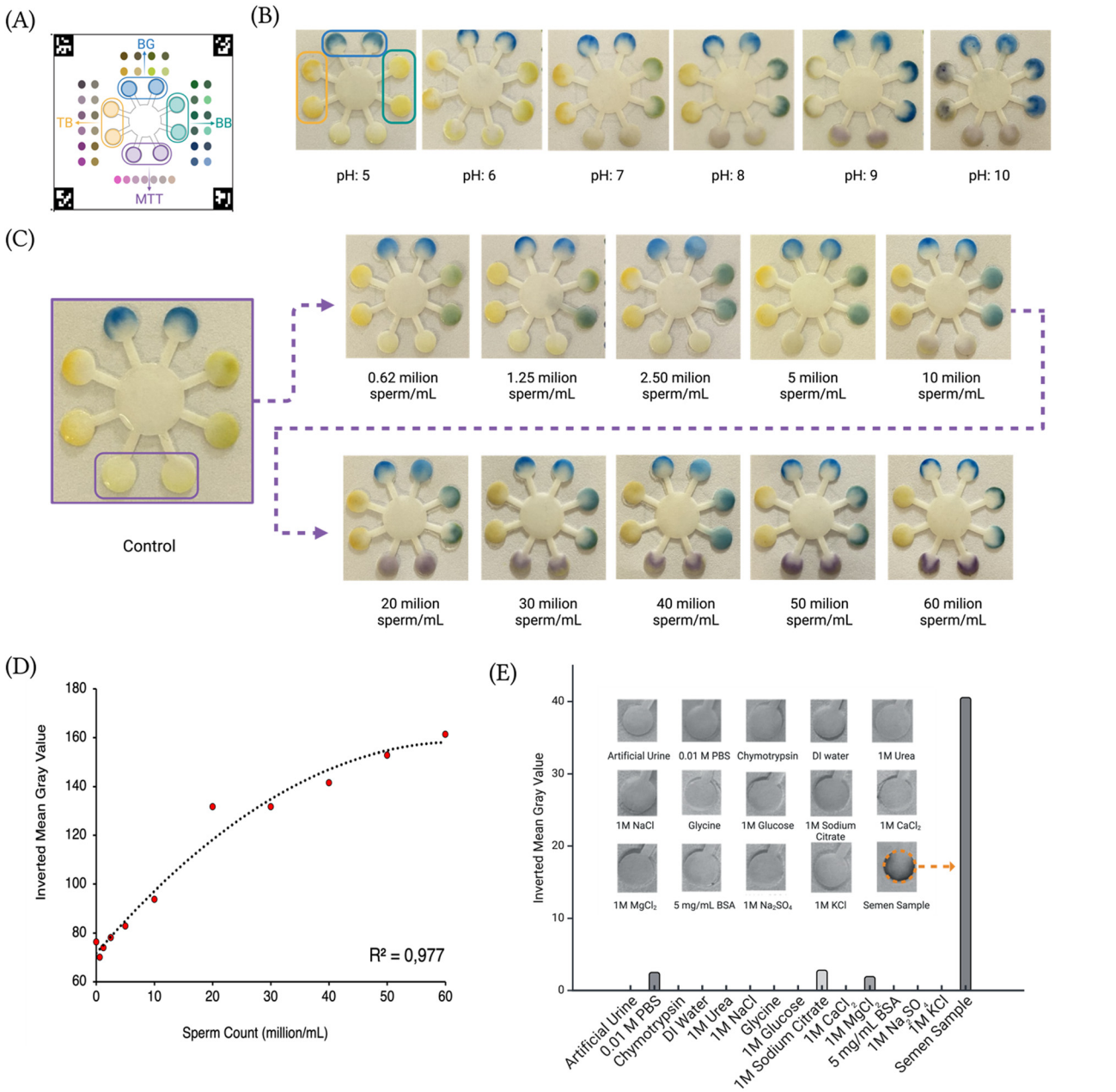

Fig. 2 Effect of key parameters on the test signal response. (A) A schematic representation of the sensor layout, indicating the design arrangement and specific detection function of each indicator dye. The indicators are labeled as follows: TB (thymol blue), BG (bromocresol green), BB (bromothymol blue), and MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide); (B) sequential image results captured on the device, showing the response of a single semen sample tested under varying pH conditions. The dynamic color changes of the indicators illustrate their sensitivity to pH variations; (C) results obtained by applying serially diluted semen samples to the device with control group distilled water. (D) and (C) The graph demonstrates a strong linear correlation (R^2 = 0.977) between the sperm concentration and the mean gray value, with inverted mean gray values calculated by subtracting the mean value from 255. (E) Selectivity assay for the MTT-indicator test, revealing a significant increase in the semen sample compared to control substances, indicating specific binding and response.

conversion was directly proportional to the sperm count, with higher concentrations resulting in increased formazan production. Absorbance measurements of formazan at 570–590 nm allowed for a linear correlation between sperm concentration and color intensity.[15] Through this assay, the sperm count in each sample was quantified, offering valuable insights into the sperm concentration across samples.

To further investigate the relationship between sperm concentration and MTT assay results, a sperm dilution series was prepared. Initial sperm concentrations were determined using a hemocytometer,[16] and semen samples were subsequently diluted to a series of concentrations. Each diluted sample was applied to a paper-based semen assay developed for this study, as illustrated in Fig. 2C. A clear

Open Access Article. Published on 31 January 2025. Downloaded on 4/21/2026 9:00:16 PM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

correlation was observed between formazan intensity and sperm concentration; higher sperm counts consistently produced more formazan, indicating greater metabolic activity.

Images captured with a smartphone were analyzed and validated using ImageJ software, and a calibration curve was constructed based on these results (Fig. 2D). From this result, the limit of detection (LOD) and limit of quantification (LOQ) for sperm concentration were determined to be 8.27 million sperm per mL and 25.3 million sperm per mL, respectively, underscoring the assay's sensitivity and quantitative capability.

Selectivity assay

The selectivity assay for the semen sensor aimed to assess the sensor's response to a variety of analytes commonly found in biological fluids, emphasizing its specificity for the metabolically active sperm count. Unlike pH, which does not require selectivity, we focused on assessing selectivity specifically for sperm count using the MTT indicator assay. Tested substances included artificial urine, PBS, chymotrypsin, deionized water, urea, NaCl, glycine, glucose, sodium citrate, $CaCl_2$, $MgCl_2$, BSA, Na_2SO_4, and KCl. Each test was carefully conducted to evaluate the sensor's selectivity towards individual analytes under controlled conditions, providing insights into its ability to detect specific biomarkers in complex biological samples.

The assay results revealed distinct responses of the sensor to the range of tested chemicals. Only PBS, sodium citrate, and $MgCl_2$ produced measurable signals, suggesting potential interaction with the sensor. Notably, these substances displayed only modest responses relative to the baseline, implying a level of specificity in the sperm sensor's detection capabilities for sperm count. Conversely, chymotrypsin, urea, KCl, glycine, glucose, $CaCl_2$, BSA, Na_2SO_4, and NaCl exhibited minimal or no response, indicating limited or no interaction (Fig. 2E). This response highlights the sensor's selectivity towards metabolically active sperm count like we expected.

Deionized water served as a blank to create baseline responses in Fig. 2E, also, with the unadjusted results shown in ESI† Fig. S5, providing a complete view of the responses to each analyte tested. This comprehensive dataset underscores the sensor's ability to selectively detect metabolically active sperm within complex biological matrices.

Procedurally generated images

A significant challenge in biological sample-based research is often the quantity and quality of samples. In our studies, after successfully demonstrating the sensor's functionality and selectivity through initial experiments, we encountered limitations while obtaining representative semen samples. To overcome this problem and enable highly accurate colorimetric detection, we implemented a strategy involving the generation of synthetic images. This approach allowed us to expand our dataset and refine our analysis, mitigating the impact of limited biological samples on our research.

YOLOv8 is an object detection model that operates in a single-stage process, directly predicting bounding boxes and class probabilities for objects in an image. It efficiently extracts features from the input image using convolutional layers and then generates bounding box predictions and class probabilities using a prediction head where non-maximum suppression is applied to filter out redundant detections, resulting in a set of bounding boxes with associated class labels and confidence scores.[17] YOLOv8 is known for its speed, accuracy, flexibility, and scalability, making it suitable for various real-time object detection applications.

To enable capturing class probabilities that the model hasn't seen before, a representative collection of images containing the target in a compatible format (image and its corresponding .txt file that contains class ID, bounding box coordinates, and the bounding box size between 0 and 1) and training the model on the dataset using a pre-trained model is required. Then, the model's performance on validation and testing sets are evaluated to assess accuracy. Therefore, detecting custom classes relies on the protocol-compliant dataset and data augmentation techniques to make sure the dataset effectively captures the desired features and the model learns them effectively.[18] In our case, obtaining a dataset with similar class distribution is a challenging task since regular semen pH clusters within the 7.2–8 range, and sperm count clusters within 15–200 million,[19] but images with pH and sperm count values that do not fall into these ranges are also needed in a similar amount. Besides the challenging nature of obtaining well-distributed class instances, trained operators, a significant number of donors, time-consuming lab procedures and expensive lab equipment to obtain gold standard semen analysis data are strict requirements. In addition, varying lighting conditions also affect the bitwise spatial information of the captured image. Therefore, sample images need to be populated under varying lighting conditions to enable accurate prediction.

Procedural image generation can be a valuable tool for augmenting or creating training datasets and improving the generalization capabilities of YOLOv8 models. By generating synthetic images with diverse variations of the target objects, one can increase the model's exposure to different scenarios and reduce the risk of overfitting. Techniques like GANs (generative adversarial networks) can be used to create realistic-looking images that mimic the appearance and variations of real-world objects.[20] This can be particularly useful when dealing with limited or biased training data, as it allows us to artificially expand the dataset and introduce new variations that might not be present in the original images. By incorporating procedurally generated images into the training process, YOLOv8's ability to detect and localize target objects in a wider range of conditions can be enhanced.[21]

To procedurally generate images that preserve the disturbances and dynamics of the real images and applicable to

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 31 January 2025. Downloaded on 4/21/2026 9:00:16 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

the YOLOv8 convention, we organized our dataset with 3 labels for pH ($5 < pH < 7$, $7 < pH < 8$, and $8 < pH < 10$) and 2 labels for sperm count (sperm count <10 million and sperm count >10 million); we manually extracted the region of interests for each label and combined cropped sensing regions falling into label intervals in the same group (Fig. 3A). Texture transformation function is employed to manipulate texture coordinates with different mapping algorithms (Fig. S9†), resulting in varying flow patterns and color distribution, mimicking spatial and bitwise information of the real sample (Fig. 3B). These transformed textures are then inserted into the corresponding locations on top of the 3D paper body mesh. A virtual camera is placed to obtain the top-down view of the generated 3D sample and saves the rendered results as an image file to create the desired dataset (Fig. 3C). Game engines like Unity and Unreal Engine employ a variety of techniques to simulate realistic lighting in virtual worlds. These techniques include ray tracing for highly accurate and realistic effects, hybrid rendering for a balance of quality and performance, rasterization for speed and efficiency, screen space ambient occlusion for depth and realism, lightmaps for pre-calculated lighting, and deferred shading for efficient handling of multiple light sources. By carefully selecting and combining these methods, game engines can create visually stunning and immersive lighting effects that enhance the overall player experience. Therefore, with a bare-minimum virtual scene and custom designed algorithms using Unity, we manage to control various parameters of the paper sensor such as the texture array index (spread pattern type), shear strength, shear rotation inclination (left–right-up-down) and lighting conditions such as the incident light angle, intensity, hue, and saturation. As a result, we obtained 2500 images capturing the target domain by

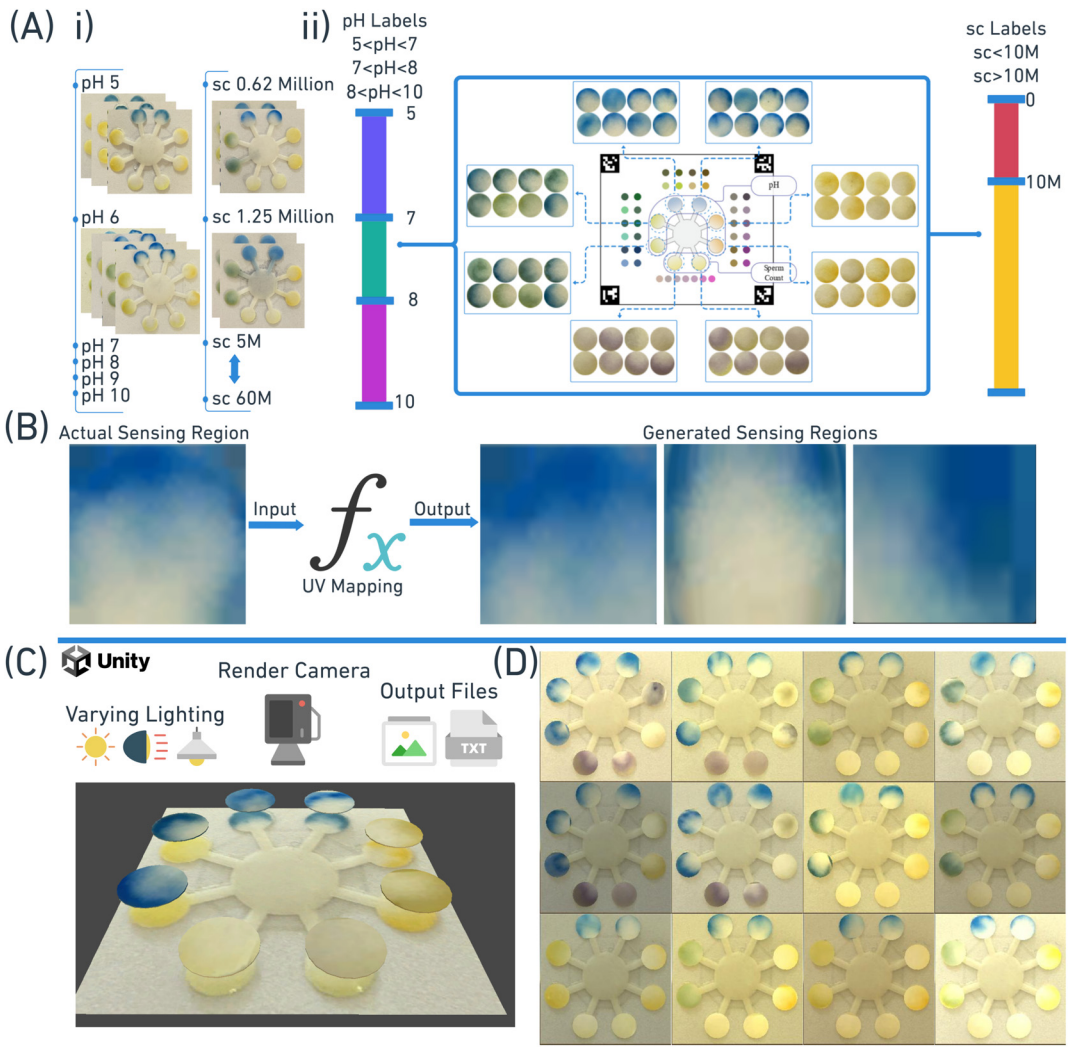

Fig. 3 Proposed procedural image generation for achieving a comprehensive dataset to fine-tune the YOLOv8 model to detect sensing regions with their labels. A) Prepared dataset and decided labels. i) Quantified, scarce semen samples are divided into 5 groups ($5 < pH < 7$, $7 < pH < 8$, $8 < pH < 10$, sc < 10 M, and sc > 10 M). ii) Sensing regions falling into the intervals as labels indicated are manually cropped and grouped (only $7 < pH < 8$ is shown). (B) In Unity, UV mapping is applied to cropped sensing regions to algorithmically manipulate the cropped sensing regions. (C) Output texture obtained from the UV mapping function is placed on top of the actual image while the camera renders from the top view. (D) The result of the procedural image generator, synthetic images to fine-tune the YOLO model.

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 31 January 2025. Downloaded on 4/21/2026 9:00:16 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

slightly mimicking the lateral flow patterns and real-life disturbances applicable to the YOLO training format (Fig. 3D).

YOLOv8 assessment

We utilized a lightweight pretrained model "yolov8s.pt" to transfer the weights that we are interested in with their corresponding labels. With a small YOLOv8s model, we achieved a high mean average precision (mAP) of 0.987 on the validation dataset indicating that it learned the ROIs and corresponding labels effectively from our synthetic dataset.

When test images were used as input, the results revealed that the model achieved 0.71 accuracy for $5 < pH < 7$, 0.875 for $7 < pH < 8$, and 0.75 for $8 < pH < 10$; the model predicted sperm count labels without any error for protocol-compliant dataset while it struggled to assign the correct labels to the $7 < pH < 8$ label, achieving 0.75 for $5 < pH < 7$ and 0.24 for $7 < pH < 8$. (Fig. 4A). In addition, the model accurately detected regions of interest (ROIs) and produced well-bounded boxes (Fig. 4B and C). The results highlight the significance of the testing conditions and the limitations of using synthetic data for training. The YOLOv8s model, trained solely on synthetic images, demonstrated promising accuracy when the waiting time and protocols were strictly followed, as observed in the protocol-compliant dataset test. However, the challenging dataset introduced additional complexity. This dataset used artificial solutions with varying pH values, interfaced with the instrument without adhering to a standardized protocol. Moreover, the image quality varied significantly—images were captured in less than 3 minutes after interfacing, often while the solutions were still wet (Fig. 4C). These factors, combined with inconsistent lighting and environmental conditions, contributed to the model's poor accuracy, particularly for the $7 < pH < 8$ range, underscoring the importance of controlled conditions.

Our approach, utilizing a YOLOv8 model trained with synthetic images, effectively demonstrates the potential of this methodology for colorimetric sperm analysis. The results suggested that by employing a more sophisticated virtual scene setup, higher classification accuracy might be achievable. In our

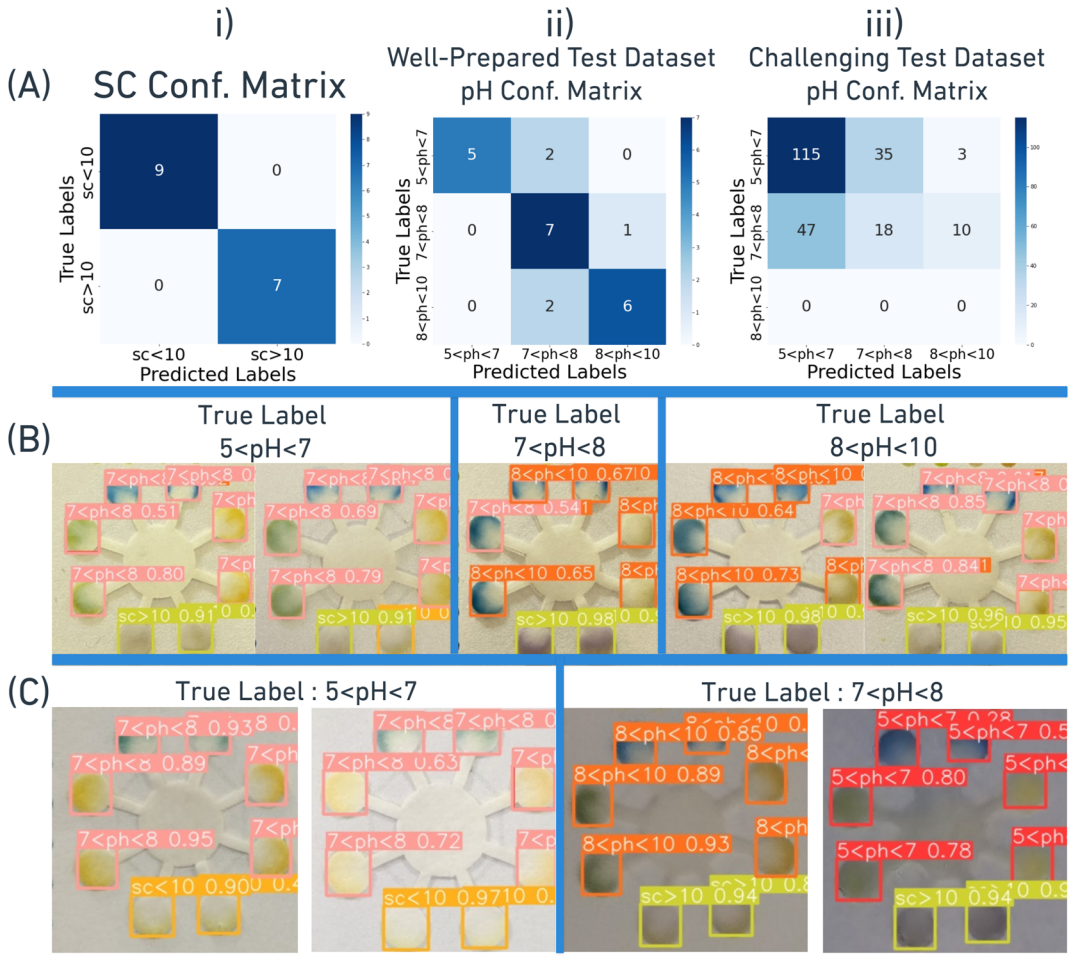

Fig. 4 YOLOv8 model evaluation. (A) Ground truth – prediction comparison. i) Sperm count classification accuracy on the protocol-compliant dataset. ii) pH classification confusion matrix on the protocol-compliant dataset. iii) pH classification confusion matrix on a challenging dataset. (B) All the misclassification results of the model over protocol-compliant input images indicated with their correct label on top. (C) Example images from the challenging dataset obtained from YOLOv8 prediction with their corresponding label on top.

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 31 January 2025. Downloaded on 4/21/2026 9:00:16 PM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

case, the difficulty of obtaining gold-standard sperm image data in large quantities is mitigated by algorithmically generating sample images by preserving the significant features in actual images. Procedural image generation emerged as a powerful tool for creating a representative dataset, became an essential tool for us to tackle various challenges related to colorimetric sperm analysis using mobile phone cameras. We explored a single-step solution using paper-based colorimetric sensing and it holds significant promise with the potential to significantly reduce the workload associated with time-consuming procedures for obtaining ground truth samples, while simplifying the real-life applications, eventually leading to the more adopted continuous health-monitoring platforms.[22]

Conclusion

Male infertility is a widespread problem that affects millions of families worldwide, and because males are typically reluctant to request clinical tests in conventional healthcare settings, early diagnosis and treatments are sometimes hampered.[23,24] This study illustrates the potential of paper-based colorimetric microfluidic test kits in combination with ML-based image analysis as a strong POC diagnostic tool for male infertility in a variety of situations, including resource-limited regions. Our custom-designed paper-based semen analysis kit provides numerous advantages, including cost effectiveness, simplicity in manufacturing, portability, disposability, and the capacity to achieve POC testing in the comfort of a person's home. This easy kit can greatly increase access to male infertility diagnoses by removing the requirement for external power sources and knowledgeable clinicians.

Our study emphasizes the significance of integrating cutting-edge ML methodologies with breakthrough diagnostic technology to address urgent healthcare concerns. We further demonstrated an effective approach for enabling colorimetric detection that requires a minimal number of precious samples and expensive laboratory hardware with trained personnel. By relying on a YOLO model, fine-tuned with synthetic data, generated by bare-minimum rendering settings and a virtual scene setup, we achieved 0.86 when the test protocol was strictly followed. We aim to increase the early detection of male infertility and ultimately give millions of couples who are trying to start families hope by bridging the accessibility and accuracy gap. These paper-based diagnostic kits can be enhanced and made more comprehensive through additional research and development, making them more potent weapons in the global war against infertility.

Data availability

The data supporting this study's findings are available from the corresponding author upon reasonable request.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

S. T. acknowledges the TUBITAK 2232 International Fellowship for the Outstanding Researchers Award (118C391), Alexander von Humboldt Research Fellowship for Experienced Researchers, Marie Skłodowska-Curie Individual Fellowship (101003361), and Royal Academy Newton-Katip Çelebi Transforming Systems Through Partnership award (120N019) for financial support of this research. We thank Sara Asghari Dilmani for her help in the experiment and insightful feedback. Opinions, interpretations, conclusions, and recommendations are those of the author and are not necessarily endorsed by the TUBİTAK. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

References

1 A. Agarwal, A. Mulgund, A. Hamada and M. R. Chyatte, A unique view on male infertility around the globe, *Reprod. Biol. Endocrinol.*, 2015, **13**, 37.
2 S. S. Dhumal, P. Naik, S. Dakshinamurthy and K. Sullia, Semen pH and its correlation with motility and count - A study in subfertile men, *JBRA Assist. Reprod.*, 2021, **25**(2), 172–175.
3 N. Kumar and A. K. Singh, Trends of male factor infertility, an important cause of infertility: A review of literature, *J. Hum. Reprod. Sci.*, 2015, **8**(4), 191–196.
4 WHO, *WHO Laboratory Manual For The Examination And Processing Of Human Semen*, World Health Organization, 2021.
5 S. E. M. Lewis, Is sperm evaluation useful in predicting human fertility?, *Reproduction*, 2007, **134**(1), 31–40.
6 S. Tasoglu, Toilet-based continuous health monitoring using urine, *Nat. Rev. Urol.*, 2022, **19**(4), 219–230.
7 S. Tasoglu, H. Safaee, X. Zhang, J. L. Kingsley, P. N. Catalano and U. A. Gurkan, *et al.*, Exhaustion of Racing Sperm in Nature-Mimicking Microfluidic Channels During Sorting, *Small*, 2013, **9**(20), 3374–3384.
8 M. R. Sarabi, D. Yigci, M. M. Alseed, B. A. Mathyk, B. Ata and C. Halicigil, *et al.*, Disposable paper-based microfluidics for fertility testing, *iScience*, 2022, **25**(9), 104986.
9 T. Wischmann and P. Thorn, (Male) infertility: what does it mean to men? New evidence from quantitative and qualitative studies, *Reprod. BioMed. Online*, 2013, **27**(3), 236–243.
10 S. Akbari Nakhjavani, B. K. Tokyay, C. Soylemez, M. R. Sarabi, A. K. Yetisen and S. Tasoglu, Biosensors for prostate cancer detection, *Trends Biotechnol.*, 2023, **41**(10), 1248–1267.
11 S. M. Knowlton, M. Sadasivam and S. Tasoglu, Microfluidics for sperm research, *Trends Biotechnol.*, 2015, **33**(4), 221–229.
12 L. Shen, J. A. Hagen and I. Papautsky, Point-of-care colorimetric detection with a smartphone, *Lab Chip*, 2012, **12**(21), 4240–4243.

Open Access Article. Published on 31 January 2025. Downloaded on 4/21/2026 9:00:16 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

13 Large-scale object detection of images from network cameras in variable ambient lighting conditions, *2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*, ed. C. Tung, M. R. Kelleher, R. J. Schlueter, B. Xu, Y.-H. Lu and G. K. Thiruvathukal, *et al.*, IEEE, 2019.
14 T. Diwan, G. Anirudh and J. V. Tembhurne, Object detection using YOLO: Challenges, architectural successors, datasets and applications, *Multimed. Tools Appl.*, 2023, **82**(6), 9243–9275.
15 K. Buranaamnuay, The MTT assay application to measure the viability of spermatozoa: A variety of the assay protocols, *Open Vet. J.*, 2021, **11**(2), 251–269.
16 K. Lemon, *Determining the Concentration of Sperm with a Hemocytometer*, 2017.
17 J. Terven, D.-M. Córdova-Esparza and J.-A. Romero-González, A comprehensive review of yolo architectures in computer vision: From yolov 1 to yolov 8 and yolo-nas, *Mach. Learn. Knowl. Extr.*, 2023, **5**(4), 1680–1716.
18 J. Ding, B. Chen, H. Liu and M. Huang, Convolutional neural network with data augmentation for SAR target recognition, *IEEE Geosci. Remote Sens. Lett.*, 2016, **13**(3), 364–368.
19 E. Levitas, E. Lunenfeld, N. Weisz, M. Friger and G. Potashnik, Relationship between age and semen parameters in men with normal sperm concentration: analysis of 6022 semen samples, *Andrologia*, 2007, **39**(2), 45–50.
20 GAN-based synthetic brain MR image generation, *2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018)*, ed. C. Han, H. Hayashi, L. Rundo, R. Araki, W. Shimoda and S. Muramatsu, *et al.*, IEEE, 2018.
21 T. Haavisto, Impact of synthetic dataset on the accuracy of YOLO object detection neural network, *Master thesis*, Turku University of Applied Sciences, 2024.
22 M. Vagnoli, R. Remenyte-Prescott and J. Andrews, Railway bridge structural health monitoring and fault detection: State-of-the-art methods and future challenges, *Struct. Health Monit.*, 2018, **17**(4), 971–1007.
23 S. W. Leslie, T. L. Soon-Sutton and M. A. B. Khan, *Male Infertility*, Disclosure: Taylor Soon-Sutton declares no relevant financial relationships with ineligible companies, Disclosure: Moien AB Khan declares no relevant financial relationships with ineligible companies, Stat Pearls Publishing Copyright © 2023, Stat Pearls Publishing LLC., Stat Pearls, Treasure Island (FL), ineligible companies, 2023.
24 R. Nosrati, P. J. Graham, B. Zhang, J. Riordon, A. Lagunov and T. G. Hannam, *et al.*, Microfluidics for sperm analysis and selection, *Nat. Rev. Urol.*, 2017, **14**(12), 707–730.

© 2025 The Author(s). Published by the Royal Society of Chemistry