

Review

Not peer-reviewed version

Exploration of the Clinical Application of Multimodal Artificial Intelligence in Sperm Screening: Identifying Opportunities and Addressing Challenges

[Jungi Cui](#) *

Posted Date: 25 March 2025

doi: 10.20944/preprints202503.1877.v1

Keywords: multimodal data; artificial intelligence; assisted reproductive technology; machine learning; deep learning; sperm screening



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Review

Exploration of the Clinical Application of Multimodal Artificial Intelligence in Sperm Screening: Identifying Opportunities and Addressing Challenges

Junqi Cui

School of Health Sciences, University of New South Wales, Kensington, NSW 2033 Australia
807692046@qq.com

Abstract: One limitation of Assisted Reproductive Technology (ART) is its relatively low pregnancy success rate per cycle. A key contributing factor to this issue is the lack of precise and efficient sperm selection methods. The emergence of multimodal artificial intelligence (AI), which integrates diverse data types—including imaging, genomics, and clinical parameters—represents a transformative advancement in reproductive medicine. This system enhances the accuracy and comprehensiveness of sperm screening through dynamic functional assessments, molecular feature recognition, and advanced data integration strategies, thereby enabling the identification of sperm with high genomic integrity. This review explores recent innovations in multimodal AI for sperm screening, highlighting its potential to overcome the inherent limitations of traditional static morphological assessments. Furthermore, this paper addresses significant challenges, including data heterogeneity, model interpretability, and barriers to clinical translation. By systematically integrating machine learning, deep learning, and explainable AI techniques, multimodal AI offers promising strategies to improve ART outcomes through precise and comprehensive sperm analysis.

Keywords: multimodal data; artificial intelligence; assisted reproductive technology; machine learning; deep learning; sperm screening

1. Introduction

Nearly half of all cases of infertility are caused by male infertility [1], male infertility is idiopathic in approximately 30-40% of cases [2]. Pregnancy rates per embryo transfer using ART remain relatively low [3]. One of the key challenges is the current lack of effective methods to isolate this specific sperm subpopulation for use in ART [4].

Artificial intelligence is playing a pivotal role in driving transformative changes across the medical technology field. The application of AI in medicine has advanced from single-modal analysis to sophisticated multi-dimensional intelligent systems, with its technical framework encompassing machine learning (ML), deep learning (DL), multi-modal fusion, and intelligent robotics. This evolution has significantly improved the efficiency of disease diagnosis, treatment planning, and medical resource allocation [5–7].

The emergence of multimodal AI has further transcended the limitations of single-source data, with its primary aim being to overcome the constraints of traditional empirical medicine through data-driven intelligent systems [8]. As one of the most intricate clinical application scenarios in reproductive medicine, ART encompasses multi-dimensional decision-making processes such as follicular monitoring, dynamic embryo assessment, and IVF outcome prediction. There is an urgent need for intelligent analysis systems capable of integrating multimodal heterogeneous data, including time-series embryo images, genomic maps, clinical biochemical indicators, and patient electronic health records [9,10]. In recent years, artificial intelligence systems have progressively

transformed the clinical decision-making paradigm in assisted reproductive technology by leveraging multimodal data integration [11].

This article focuses on the clinical translation of multimodal artificial intelligence technology in sperm screening for assisted reproductive technology. By combining dynamic functional analysis with molecular feature recognition, it systematically investigates the innovative approaches of intelligent algorithms to enhance the accuracy of sperm motility assessment and optimize metabolic function evaluation, thereby offering multi-dimensional and precise strategies for clinical evaluation and decision-making in ART.

2. The Evolution of Artificial Intelligence

The term artificial intelligence was first formally introduced by John McCarthy at the Dartmouth Conference in 1956. McCarthy defined AI as “the science and engineering of making intelligent machines,” with the core objective of replicating human cognitive abilities such as reasoning, learning, perception, and language understanding through algorithms and computational models [12,13]. The introduction of this concept not only inaugurated a new domain within computer science but also marked a pivotal shift in humanity’s exploration of intelligence, transitioning from philosophical speculation to practical technological implementation [14,15].

In the field of medical imaging diagnosis, AI has emerged as a robust image analysis tool, aiding radiologists in early disease detection and reducing misdiagnosis rates [16]. For instance, AI has shown promising results in the early identification of various diseases, including breast cancer, skin cancer, ophthalmic conditions, and pneumonia [16,17]. Additionally, AI can be utilized for predicting disease risk and managing health outcomes. Studies have utilized machine learning models to predict the likelihood of diabetes onset, with an enhanced decision tree model demonstrating superior performance in forecasting diabetes-related variables [18,19]. In drug development, AI also exhibits significant potential. AI utilizing bioinformatics and chemoinformatic can significantly reduce the time and cost associated with drug discovery [16,20]. Overall, AI is driving innovation in healthcare, encompassing areas such as medical imaging, disease prediction, and drug discovery [21].

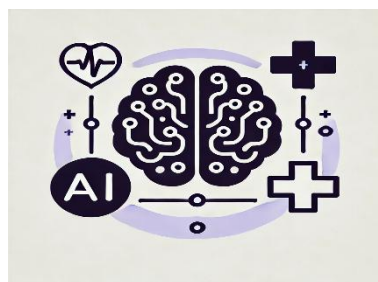


Figure 1. AI + Human Brain > Brain: A Minimalist Vision of Intelligent Medicine.

3. The Paradigmatic Innovation and Multi-dimensional Empowerment of Medical AI Approaches in ART

AI methods are typically categorized into supervised, semi-supervised, and unsupervised approaches.

Supervised learning, grounded in meticulous annotations, enables precise analysis of germ cells in microscopic images [22,23]. Deep convolutional neural networks (CNNs) utilizing high-resolution microscopic images can accurately detect subcellular structural variations in sperm—such as acrosome integrity and mitochondrial sheath defects—achieving discrimination accuracy that surpasses traditional manual microscopic examination standards [24–26]. Cross-modal supervised models further integrate spectral data and dynamic parameters to enable non-invasive quantitative assessment of the DNA fragmentation index, thereby significantly reducing the clinical risks associated with invasive testing [27,28].

Semi-supervised, guided by clinical endpoints, addresses the practical challenge of limited annotation resources by leveraging coarse or indirect labels [29]. In analysing testicular tissue pathology sections, researchers employ a multi-instance learning framework to locate spermatogenic functional units by associating surgical outcome labels with specific tissue features, providing interpretable decision-making support for predicting sperm detection rates [30,31]. This approach effectively correlates overall surgical outcomes with local tissue features, enhancing both the interpretability and predictive performance of the model [32]. Cross-modal contrastive learning, through unsupervised alignment of motion trajectories and metabolic features, uncovers the potential regulatory mechanisms of oxidative stress levels on flagellar movement patterns, paving new pathways for diagnosing and treating male infertility [24,33,34].

Unsupervised learning, utilizing advanced algorithms, explores biological principles beyond empirical observations [35,36]. Self-supervised contrastive clustering techniques have identified clinically significant novel kinetic subgroups from large-scale motion parameters, with helical propulsion patterns showing a significant positive correlation with fertilization success rates [37–39]. The generative model synthesizes a sperm image library with controllable pathological characteristics, effectively addressing the generalization limitations of algorithms caused by the scarcity of data on rare sperm morphological abnormalities [40,41].

Table 1. Artificial intelligence and machine learning algorithms used to evaluate sperm morphology.

Model Name	Subtype	Application
Supervised Learning	Cross-modal models, Deep CNNs	Acrosome integrity detection, DNA fragmentation index prediction
Semi-Supervised Learning	Multi-instance learning, Attention-based models	Testicular tissue analysis, sperm detection prediction
Unsupervised Learning	Self-supervised clustering, GANs	Kinetic subgrouping, sperm motion pattern discovery
Generative Models	DCGAN, VAE	Synthetic sperm image generation for rare cases
Transformer Models	Biomedical Transformers	Multimodal fusion, sperm DNA prediction
Contrastive Learning	Cross-modal alignment, Self-supervised contrastive clustering	Motion-metabolic pattern alignment, sperm phenotype matching
Explainable AI (XAI)	Grad-CAM, Visual heatmaps	Interpretable sperm quality prediction
Meta-learning & Causal Reasoning	Causal representation learning	Generalizable and interpretable sperm screening

4. The Technological Evolution of Multimodal Data Fusion and the Paradigm Shift in Medical Interpretation

Multimodal data fusion is reshaping the analytical dimension of complex biological systems by leveraging the complementarity and synergy of heterogeneous data sources [6]. Supervised learning establishes precise cross-modal mapping benchmarks; semi-supervised learning mitigates the limitations posed by scarce annotated data; and unsupervised learning uncovers intrinsic data associations. [42,43] These three approaches form a closed-loop optimization through techniques such as adversarial training, contrastive learning, and generative models.

Supervised learning establishes robust associations between modalities using annotated data [44]. In medical image diagnosis, the XLIP framework employs cross-modal attention masking

strategies to interactively exchange multi-modal features between annotated medical images and pathological reports, reconstructing both the image features and textual descriptions of lesion areas, thereby enhancing the learning of pathological characteristics [45]. In drug development, hierarchical fusion architectures guided by supervised signals (e.g., Stacking frameworks) optimize the allocation of multi-modal feature weights, integrating molecular structure data with clinical trial results, which significantly improves the cross-modal interpretability of drug efficacy predictions [6,46]. Additionally, joint training strategies promote complementary information exchange between modalities through supervised loss functions [47]. Supervised transfer learning combines single-modal pre-trained models (e.g., ImageNet-pretrained CNNs) with newly introduced supervised data from other modalities [48,49], enabling rapid adaptation to multi-modal tasks [50].

semi-supervised facilitates semantic alignment between modalities using sparsely labelled data [51]. In pathological analysis, only malignant regions of some tissue sections are labelled, and a contrastive learning framework correlates pathological image features with unlabelled proteomics data across scales, revealing morphological-molecular interaction patterns [52]. Additionally, semi-supervised fusion models incorporating attention mechanisms, such as multi-instance learning, address discrepancies in annotation granularity between modalities [53]. Adversarial training frameworks enhance the robustness of semi-supervised fusion models [54]. Furthermore, pre-training Transformer architectures using unlabelled medical images and electronic health records captures latent cross-modal patterns of disease progression [6]. This approach offers an efficient and interpretable multimodal fusion pathway for medical image analysis, drug development, and other domains by reducing labelling costs and mining cross-modal relationships.

Unsupervised learning highlights the superior performance of contrastive learning frameworks in unsupervised multi-modal alignment [55]. Through modal negative sampling strategies, it enhances feature discriminability, and adversarial generative networks (GANs) achieve cross-modal data distribution matching [56]. The Conditional Alignment Time Diffusion (CATD) framework proposed by Yao et al. aligns electroencephalogram (EEG) signals with functional magnetic resonance imaging (fMRI) images in a latent space, uncovering cross-modal dynamic mechanisms of neural activities [57]. Variational autoencoders (VAEs) model the joint distribution of multi-modal data, enabling the generation of virtual compounds that satisfy both structural features and activity predictions in drug design [58,59]. Additionally, constructing cross-modal associations of organ anatomical structures through unsupervised contrastive learning, followed by fine-tuning attention mechanisms with weakly labelled lesion annotations, optimizes diagnostic decision-making under full supervision [60,61]. This hierarchical supervision mechanism achieves an optimal balance between data labelling costs and model performance [62]. In mixed supervision learning models, meta-learning frameworks adaptively adjust the weights of different supervision signals, enhancing model generalization [63]. Incorporating causal reasoning modules distinguishes correlations from causations among variables, thereby improving the interpretability of decisions made by mixed supervision models [64]. The combination of meta-learning and causal reasoning not only boosts model performance but also enhances adaptability and reliability in complex tasks.

The essence of multimodal data fusion lies in constructing alignment and mapping mechanisms across domain feature spaces to fully exploit the complementarity of different modalities [65]. Spatio-temporal convolution architectures extract dynamic patterns of organ function evolution from dynamic image sequences and form spatio-temporally coupled joint representations with genomic features [66]. Self-supervised contrastive learning frameworks overcome the semantic limitations of traditional single-modal representations, achieving cross-scale semantic alignment of pathological image texture features and proteomics molecular markers [49,67], revealing the dynamic interaction network of morphology, function, and molecules during disease progression. Generative models mitigate distribution shifts arising from multimodal data heterogeneity by modelling joint distributions in latent space [68]. In the medical field, research on cross-modal intervention counterfactual models is emerging. For instance, Hu et al. proposed the interpretable multimodal fusion network gCAM-CCL, which performs automatic diagnosis and result interpretation

simultaneously. The gCAM-CCL model generates interpretable activation maps by combining intermediate feature maps with gradient-based weights, thereby quantifying pixel-level contributions of input features [69]. These multimodal data technologies offer cross-dimensional evidence chains for clinical decision-making, signifying a shift in medical artificial intelligence from fragmented single-modal analyses to system-level integrated diagnostics.

5. The Limitations of Traditional Sperm Screening Methods and Potential Directions for Breakthrough

In ART, traditional sperm screening predominantly depends on static morphological parameters, such as head shape, acrosome integrity, and flagellar structure. However, the inherent limitations of this approach have become increasingly evident, particularly in its inability to assess dynamic functionalities and genetic integrity. The traditional two-dimensional morphological assessment system is fundamentally grounded in empirical thresholds and thus limited to reflecting the phenotypic characteristics of sperm at a single time point, failing to capture critical biological information such as dynamic functionality and genetic integrity [24,70,71]. A mechanistic study by Komatsu et al. provides compelling evidence supporting this limitation. Although sperm selected based on traditional morphological criteria can form high-quality embryos, their mitochondrial membrane potential ($\Delta\Psi_m$) may consistently remain below the threshold observed in embryos derived from *in vivo* fertilization. Additionally, significant mitochondrial ultrastructural abnormalities may be detected. Consequently, the failure to incorporate dynamic evaluations of mitochondrial function may increase the risk of embryo arrest [72].

At the molecular level, Zhao et al. demonstrated that sperm with reduced expression of the *Mfn2* gene maintained normal morphology but had diminished embryonic developmental potential due to compromised mitochondrial fusion [73]. Collectively, these findings highlight a fundamental disconnect between traditional morphological assessments and the functional status of sperm mitochondria. Therefore, overcoming the limitations of static morphological assessments and developing a multimodal framework that integrates dynamic functionality, molecular integrity, and three-dimensional morphology is a critical scientific challenge to improve the success rates of ART.

6. The Application of Multimodal Data Integration Strategies in Sperm Screening

Theoretical breakthroughs in artificial intelligence for sperm screening have fundamentally redefined the evaluation criteria for “high-quality sperm” in reproductive medicine. This advancement addresses the limitations of traditional static morphological assessments by integrating dynamic functional data and genomic information through cross-scale modelling, thereby enabling the simultaneous evaluation of sperm morphology, motility, and genetic integrity.

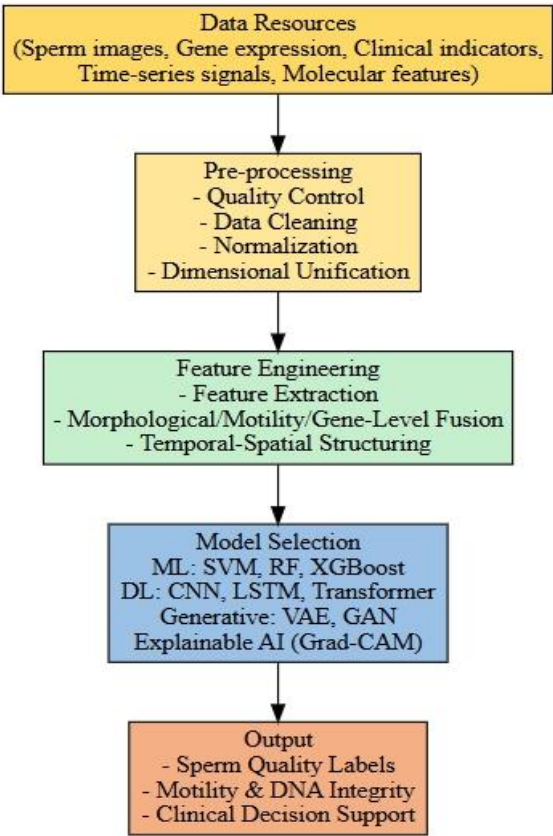


Figure 2. An integrated pipeline of multimodal artificial intelligence for sperm screening.

6.1. Strategies for Sperm Morphology Assessment

Assessing sperm morphology is crucial for the success of assisted reproductive technologies. Normal sperm morphology indicates proper development, essential for fertilization, and correlates positively with successful pregnancy outcomes. The World Health Organization (WHO) has established morphological criteria, defining normal semen samples as those containing at least 4% morphologically normal sperm [74]. However, compared to other parameters such as sperm concentration and motility, standardizing sperm morphology assessment has proven particularly challenging [75]. Traditional methods primarily rely on staining techniques to evaluate static structural features, including head shape, acrosome integrity, and midpiece dimensions [76]. These approaches are limited by subjectivity, potential staining effects on sperm motility, and lack of dynamic physiological data [24].

Integrating multimodal data strategies offers innovative solutions for assessing sperm morphology. By combining diverse data sources, including microscopic imaging, molecular markers, and computer vision, a more comprehensive and objective evaluation of sperm morphology can be achieved. Yoav N. Nygate et al. developed HoloStain, a deep learning method that integrates holographic microscopy with virtual staining technology to analyse individual sperm cells without chemical staining. This approach uses quantitative phase imaging for high-resolution morphological features and employs a deep convolutional generative adversarial network (DCGAN) to generate virtual staining images. Consequently, HoloStain provides an accurate, artifact-free assessment of sperm morphology by eliminating potential distortions caused by traditional staining methods [77]. Kamieniczna et al. utilized digital holographic microscopy (DHM) for the morphological evaluation of live sperm, overcoming limitations of conventional techniques. DHM, a non-contact, label-free imaging technology, avoids compromising sperm activity and provides three-dimensional morphological parameters, including detailed head and tail characteristics [25].

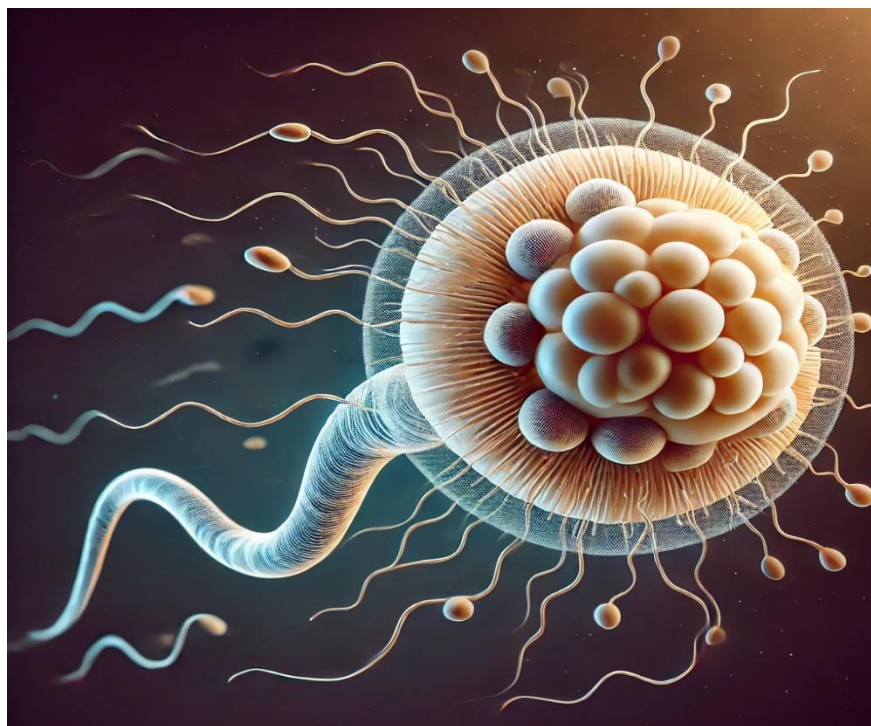


Figure 3. The three-dimensional structure, flagellar movement and biological characteristics of sperm cells.

Zou et al. designed TOD-CNN, a convolutional neural network for detecting sperm cells in microscopic videos. The model was trained on an extensive dataset of 111 high-quality sperm videos, encompassing over 278,000 annotated instances, and achieved an impressive average precision (AP_{50}) of 85.60% in real-time sperm detection tasks. Integrating TOD-CNN into multimodal data strategies is anticipated to significantly enhance the comprehensiveness and precision of sperm morphology evaluations [78]. Dardikman-Yoffe et al. introduced an innovative high-resolution, stain-free imaging technique for optical computed tomography (CT) of freely swimming sperm. This approach reconstructs the three-dimensional trajectories and structures of individual sperm cells, capturing intricate details of internal organelles and flagellar dynamics. Employing a high-speed off-axis holographic system, this technique images live sperm without the need for cell staining or mechanical stabilization. The reconstruction process relies solely on the sperm's natural movement and advanced computational algorithms, enabling precise four-dimensional reconstructions [79]. By leveraging these extensive datasets, artificial intelligence can construct dynamic models to enhance sperm morphology assessments [80].

6.2. Evaluate the Motility Characteristics of Individual Spermatozoa with Precision

In the assessment of individual sperm motility, traditional methods have predominantly relied on microscopic examination and computer-assisted sperm analysis (CASA). While conventional microscopic examination requires trained laboratory personnel to evaluate sperm motility through visual observation of movement patterns, this methodology remains vulnerable to observer subjectivity, thereby limiting both the reproducibility and diagnostic accuracy of results [81]. The CASA system addresses these limitations by offering standardized quantitative assessment of sperm motility characteristics and morphological parameters via automated tracking algorithms coupled with video microscopy [82]. Nevertheless, traditional CASA platforms demonstrate reduced diagnostic reliability when analysing semen specimens with extreme concentrations (hyper-concentrated or hypoconcentrated), particularly in scenarios involving cellular contaminants such as non-gamete cells and particulate debris [83]. These limitations underscore the need for methodological refinements in conventional semen analysis techniques.

The integration of multimodal artificial intelligence technology, which combines high-resolution imaging of sperm trajectories, physiological signals, and molecular-level information, offers more objective, accurate, and efficient solutions for sperm motility assessment. Klumpp et al. developed the Syntalos software, enabling synchronized operation across multiple devices and facilitating the detection of complex correlations between behaviour and single-cell activities. By integrating high-resolution imaging of sperm trajectories with synchronous physiological signal recording, this multimodal data synchronization function allows simultaneous tracking of sperm motility behaviour and associated physiological parameters. This integration provides a more comprehensive analytical framework for assessing sperm motility. Compared to traditional sperm motility assessment methods, which predominantly rely on single motility parameters or physiological indicators, Syntalos excels in its capacity to integrate multimodal data [84]. Additionally, Syntalos supports the synchronized recording of high-resolution imaging and physiological signals, capturing subtle changes in sperm at different time points and revealing the dynamic processes underlying sperm functional characteristics [84].

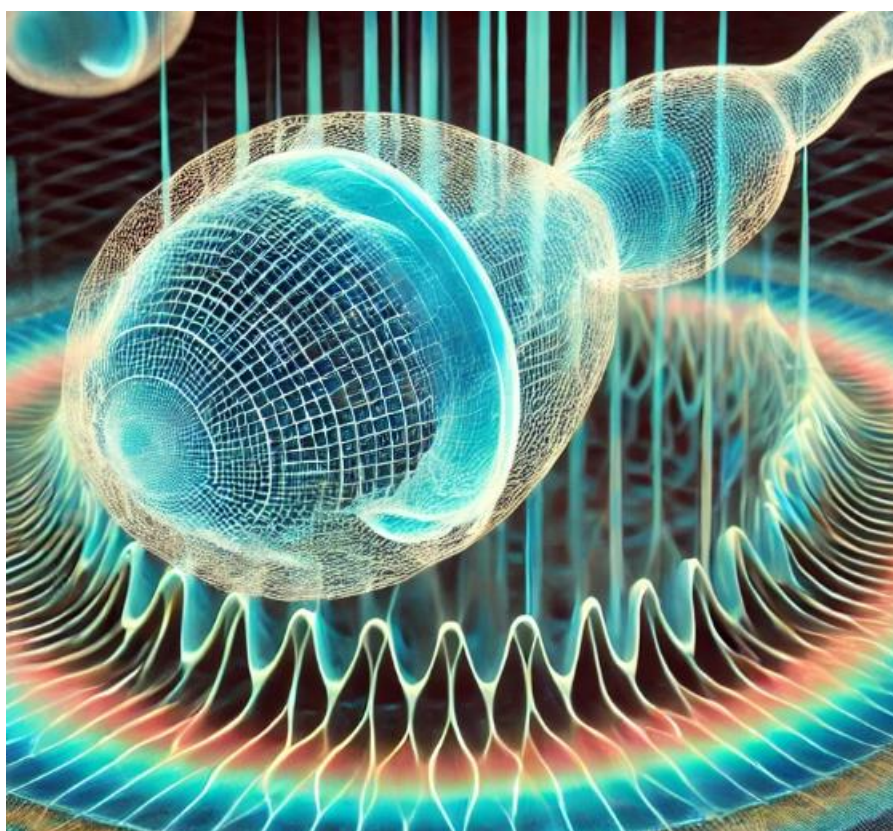


Figure 4. 4D Imaging of Sperm Cell Movement.

Pinto et al.'s research highlights the critical role of the CatSper channel, particularly CATSPER1, in regulating calcium ion influx within sperm cells—a process essential for sperm motility and male fertility. Traditional sperm motility assessment methods focus primarily on basic motility parameters and may fail to fully capture the comprehensive functional state of sperm. Incorporating the molecular understanding of CatSper channel function into the multimodal data integration strategy for sperm screening is anticipated to significantly enhance the accuracy of assessments [85]. By linking molecular-level information on CATSPER1 expression with dynamic imaging data (e.g., flagellar oscillation patterns), researchers can develop predictive models that correlate specific gene expression with functional sperm characteristics [86]. Dardikman-Yoffe et al. introduced a high-speed off-axis holographic system capable of performing high-resolution four-dimensional reconstruction of freely swimming human sperm cells without staining or mechanical components. Combining this 4D imaging data with complementary modalities such as Raman spectroscopy or

metabolomics-based molecular profiles enables the development of cross-scale predictive models. These models connect molecular-level gene expression information with sperm functional traits, offering a deeper understanding of sperm motility [79]. This multimodal data integration strategy, coupled with advancements in CASA systems, overcomes the limitations of conventional sperm motility assessment techniques [88].

6.3. Evaluating the Integrity and Damage of Sperm DNA

In traditional sperm DNA integrity assessment, clinical practice primarily depends on endpoint detection techniques such as chromatin structure analysis, in situ DNA break detection, and comet assays [88–90]. While these methods indicate the extent of DNA damage at a single time point, they have limitations, such as procedural complexity, subjective result interpretation, and standardization challenges [81,91]. Moreover, these methods involve significant sample destruction, are time-consuming, and cannot dynamically monitor damage progression [92,93].

The innovative integration of multimodal data is expected to overcome these limitations. With the emergence of deep learning, unprocessed sperm images can be utilized to train predictive models, potentially uncovering important features overlooked by humans and improving prediction accuracy [94]. In a study by McCallum et al., a deep convolutional neural network was trained using approximately 1,000 human sperm cells with known DNA quality. This approach avoids destructive sperm processing and enables non-invasive DNA integrity assessment. The results demonstrated a moderate correlation (bivariate correlation coefficient of approximately 0.43) between sperm morphology images and DNA quality and could identify sperm cells with DNA integrity above the median [95]. Popova et al.'s research revealed that structured illumination microscopy (SIM) can capture high-resolution images of live sperm cells at high resolution, particularly in the mitochondrial region of the midpiece, achieving spatial resolutions up to 100 nm. Non-invasive quantitative phase microscopy (QPM), combined with machine learning techniques, elucidated the association between oxidative stress-induced morphological changes in the sperm head and reduced motility. By applying deep neural networks to QPM images, the study achieved 85.6% accuracy in distinguishing healthy and damaged sperm. Furthermore, a systematic review indicated that sperm mitochondrial DNA copy number correlates with DNA quality. Collectively, these findings suggest that advanced imaging techniques such as SIM and QPM, combined with analyses of mitochondrial DNA copy number, can serve as valuable tools for assessing sperm DNA integrity [96].

Micro-Raman spectroscopy, a non-invasive method, has been employed to detect sperm DNA damage [97]. Du et al.'s research integrated micro-Raman spectroscopy with image analysis to simultaneously provide dynamic morphological and biochemical information. They developed a rapid Raman spectroscopy-based method that uses fine glass pipettes to adsorb samples onto a metal substrate, measuring Raman spectra at the pipette tip to evaluate DNA damage [97]. Madan et al. reviewed the development and application of Transformer architectures in analysing diverse biomedical datasets, such as text data, protein sequences, structured longitudinal data, images, and graphs. Incorporating Transformer models into multimodal data strategies for sperm screening can effectively capture complex patterns across different data types, significantly enhancing analytical capabilities. Using the Transformer architecture to integrate genomic data (e.g., CATSPER1 gene expression) with dynamic imaging data (e.g., flagellar oscillation patterns) can more accurately predict sperm motility, DNA fragmentation index (DFI), and aneuploidy risk [98].

6.4. Explainable Artificial Intelligence Technology

The “black box” problem is a central challenge in applying deep learning technology to medical contexts [99]. Its core issue stems from the inherent conflict between the irreducible complexity of a model's internal decision-making logic and the necessity for clinical verifiability [100]. Within traditional deep learning frameworks, models abstract and integrate features from input data through multiple layers of nonlinear transformations, ultimately producing diagnostic outputs [27]. However, this end-to-end mapping process lacks transparency, leading to two critical concerns: first,

the discriminative features on which the model relies may diverge from established biological mechanisms [101]. For instance, in sperm quality assessment, algorithms might erroneously classify microscopic imaging artifacts, such as slide scratches, as abnormalities in sperm head structure [105]. Second, clinicians are unable to trace the decision-making rationale, making it challenging to verify whether the model's judgments are based on appropriate biomarkers [106]. In sperm multimodal screening, the "black box" problem is further compounded by data heterogeneity. When multi-source data, such as digital holographic microscopy and Raman spectroscopy, are simultaneously input into the model, the feature fusion process within hidden layers can lead to unintuitive weight distributions [103].

Explainable artificial intelligence technology is expected to solve this problem. Suara et al. conducted a comprehensive study on interpretable deep learning principles, focusing particularly on the application of Gradient-weighted Class Activation Mapping (Grad-CAM) in medical imaging. They demonstrated how Grad-CAM enhances the interpretability of deep learning models by highlighting critical regions within medical images that significantly influence model decision-making. By backpropagating gradients through the convolutional neural network, Grad-CAM generates spatial heat maps that correlate strongly with the model's decisions, transforming the 'input-output' black box into a visually interpretable 'feature-decision' association [104]. The integration of Grad-CAM into multimodal data strategies for sperm screening can substantially enhance the interpretability of predictive models [105]. Through the application of Grad-CAM, clinicians gain a clearer understanding of which morphological features and molecular markers most significantly influence the model's predictions regarding sperm quality parameters [106]. Such an intuitively interpretable model enables direct validation of the artificial intelligence decision-making process, fostering greater trust and acceptance of AI-assisted diagnosis in reproductive medicine [107] [108].

7. Challenges and Breakthrough Paths in the Medical Application of Multimodal Data Fusion

The clinical application of multimodal data fusion faces significant challenges due to the interplay between technical limitations and biological complexity [109–111]. The primary hurdle originates from the spatiotemporal asynchrony and semantic gap inherent in heterogeneous data sources [112,113]. Specifically, the millisecond temporal resolution of dynamic imaging contrasts sharply with the static nature of genomics data, creating a parsing scale discontinuity. Moreover, the nonlinear relationships between microscopic phenotypic features and metabolic molecular markers far exceed the characterization capabilities of traditional feature engineering approaches [114–116]. Additionally, the challenge of balancing information redundancy and complementarity across modalities limits model generalization. For example, in neurological disease diagnosis, the brain network topology derived from functional magnetic resonance imaging (fMRI) and electrophysiological oscillation signals exhibit both synergistic verification and potential interference [117,118]. A deeper challenge lies in overcoming the barrier to biological interpretability. Current black-box fusion mechanisms struggle to trace the causal chains underlying cross-modal associations, raising concerns about the credibility of clinical decision support systems [119–121].

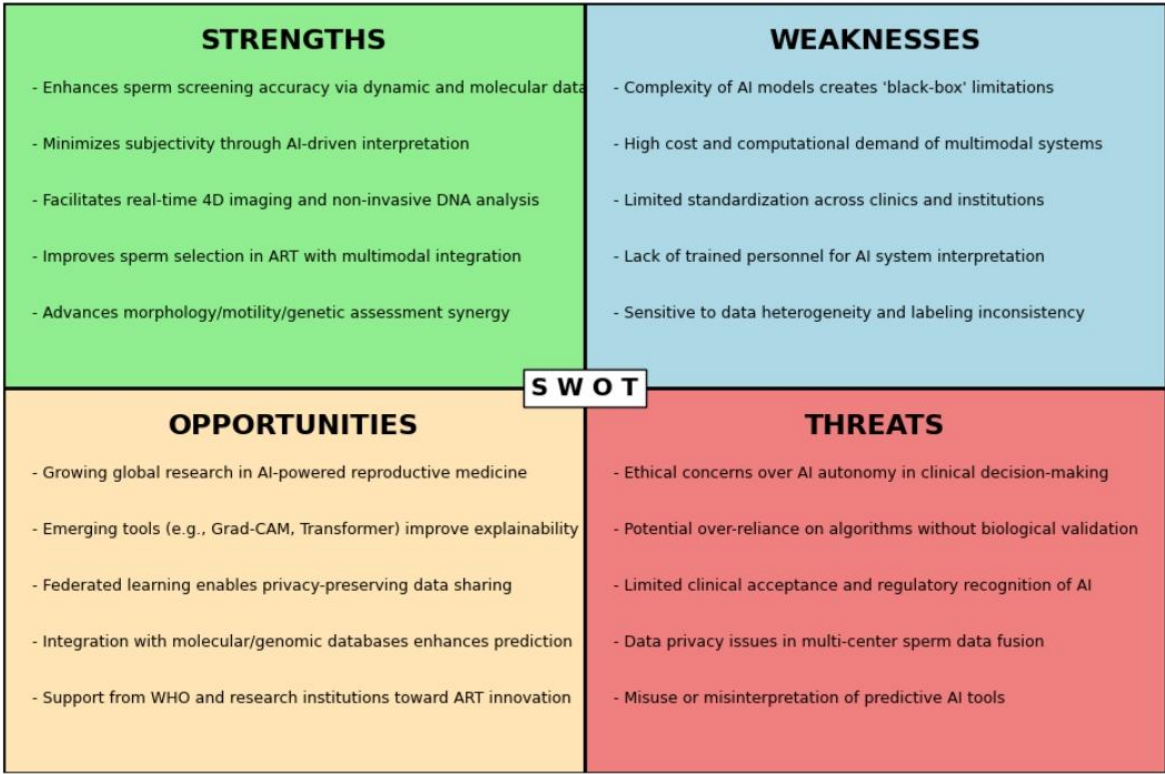


Figure 5. SWOT analysis of multimodal artificial intelligence applications in sperm screening and assisted reproductive technologies.

Frontier research is exploring multiple pathways to address these challenges. Adaptive hierarchical fusion architectures leverage gating mechanisms to dynamically weight the contributions of different modalities, suppressing noise propagation while preserving modality-specific characteristics [122,123]. Knowledge distillation techniques compress large multimodal models into lightweight clinical reasoning engines, mitigating the tension between computational constraints and real-time performance requirements [124,125]. Distributed multi-center data collaboration under the federated learning framework aims to resolve the trade-off between privacy protection and model generalization [126]. Notably, the introduction of causal representation learning offers a promising approach by simulating counterfactual interventions to uncover the driving relationships between imaging feature variations and molecular pathway abnormalities, providing a novel paradigm for constructing interpretable fusion decision systems [127,128]. The co-evolution of these technologies is propelling multimodal fusion toward a substantive transition from laboratory validation to routine clinical application [129].

8. Conclusions

In the provided document, the application of multimodal AI in sperm screening within ART is comprehensively discussed. The study emphasizes overcoming the limitations of traditional sperm selection methods, which primarily rely on static morphological parameters and subjective assessments. Multimodal AI incorporates diverse data, including dynamic imaging, genomic profiles, biochemical markers, and clinical parameters, enhancing the accuracy of sperm motility assessments and optimizing metabolic function evaluations.

Key advancements highlighted include the integration of sophisticated AI models—such as machine learning, deep learning, and multimodal fusion techniques—to systematically assess sperm quality in multiple dimensions. These innovative approaches enable precise evaluations of sperm morphology, dynamic motility, and genetic integrity simultaneously, significantly improving clinical outcomes in ART.

However, the clinical application of multimodal AI faces several challenges. Issues such as data heterogeneity, model interpretability, biological complexity, and clinical translation barriers remain critical hurdles. Addressing these concerns requires advanced data fusion strategies, adaptive hierarchical modelling, and explainable artificial intelligence methods to facilitate clinical validation and acceptance. Overall, the study underscores the transformative potential of multimodal AI for reproductive medicine, providing a clear path forward by highlighting current technological opportunities and outlining strategic approaches to address ongoing challenges.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: I am grateful to the teachers from the School of Health Science of the University of New South Wales for their guidance, which enabled me to complete the interdisciplinary research.

Conflicts of Interest: the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Inhorn, Marcia C., and Pasquale Patrizio. "Infertility around the globe: new thinking on gender, reproductive technologies and global movements in the 21st century." *Human reproduction update* 21.4 (2015): 411-426.
2. Zegers-Hochschild, Fernando, et al. "The international glossary on infertility and fertility care, 2017." *Human reproduction* 32.9 (2017): 1786-1801.
3. Oseguera-López, Iván, et al. "Novel techniques of sperm selection for improving IVF and ICSI outcomes." *Frontiers in cell and developmental biology* 7 (2019): 298.
4. Yániz, J. L., C. Soler, and P. Santolaria. "Computer assisted sperm morphometry in mammals: a review." *Animal reproduction science* 156 (2015): 1-12.
5. Chen, Xieling, et al. "Artificial intelligence and multimodal data fusion for smart healthcare: topic modeling and bibliometrics." *Artificial Intelligence Review* 57.4 (2024): 91.
6. Duan, Junwei, et al. "Deep learning based multimodal biomedical data fusion: An overview and comparative review." *Information Fusion* (2024): 102536.
7. Mohsen, Farida, et al. "Artificial intelligence-based methods for fusion of electronic health records and imaging data." *Scientific Reports* 12.1 (2022): 17981.
8. AlSaad, Rawan, et al. "Multimodal large language models in health care: applications, challenges, and future outlook." *Journal of medical Internet research* 26 (2024): e59505.
9. Liu, Junwei, et al. "Challenges in AI-driven Biomedical Multimodal Data Fusion and Analysis." *Genomics, Proteomics & Bioinformatics* (2025): qzaf011.
10. Zhang, Qing, Xiaowen Liang, and Zhiyi Chen. "A review of artificial intelligence applications in in vitro fertilization." *Journal of Assisted Reproduction and Genetics* 42.1 (2025): 3-14.
11. Medenica, Sanja, et al. "The future is coming: artificial intelligence in the treatment of infertility could improve assisted reproduction outcomes—the value of regulatory frameworks." *Diagnostics* 12.12 (2022): 2979.
12. Rajaraman, Vaidyeswaran. "John McCarthy—Father of artificial intelligence." *Resonance* 19 (2014): 198-207.
13. Negnevitsky, Michael. "The History Of Artificial Intelligence Or From The." *WIT Transactions on Information and Communication Technologies* 19 (2024).
14. Rai, Dilli Hang. "Artificial Intelligence Through Time: A Comprehensive Historical Review." (2024).
15. Anurag, A. S. "Early beginnings of AI: the field of research in computer science." *Cases on AI Ethics in Business*. IGI Global, 2024. 1-20.

16. Al Kuwaiti, Ahmed, et al. "A review of the role of artificial intelligence in healthcare." *Journal of personalized medicine* 13.6 (2023): 951.
17. Shaikh, Khalid, Sabitha Krishnan, and Rohit M. Thanki. *Artificial intelligence in breast cancer early detection and diagnosis*. Cham: Springer, 2021.
18. Xie, Yi, et al. "Integration of artificial intelligence, blockchain, and wearable technology for chronic disease management: a new paradigm in smart healthcare." *Current medical science* 41.6 (2021): 1123-1133.
19. Ghaffar Nia, Nafiseh, Erkan Kaplanoglu, and Ahad Nasab. "Evaluation of artificial intelligence techniques in disease diagnosis and prediction." *Discover Artificial Intelligence* 3.1 (2023): 5.
20. Sarkar, Chayna, et al. "Artificial intelligence and machine learning technology driven modern drug discovery and development." *International Journal of Molecular Sciences* 24.3 (2023): 2026.
21. Pinto-Coelho, Luís. "How artificial intelligence is shaping medical imaging technology: a survey of innovations and applications." *Bioengineering* 10.12 (2023): 1435.
22. Gedefaw, Lealem, et al. "Artificial intelligence-assisted diagnostic cytology and genomic testing for hematologic disorders." *Cells* 12.13 (2023): 1755.
23. Liu, Jiazheng, et al. "A robust transformer-based pipeline of 3D cell alignment, denoise and instance segmentation on electron microscopy sequence images." *Journal of Plant Physiology* 297 (2024): 154236.
24. Dai, Changsheng, et al. "Advances in sperm analysis: techniques, discoveries and applications." *Nature Reviews Urology* 18.8 (2021): 447-467.
25. Kamieniczna, Marzena, et al. "Human live spermatozoa morphology assessment using digital holographic microscopy." *Scientific reports* 12.1 (2022): 4846.
26. Kang, Mi-Sun, et al. "Accuracy improvement of quantification information using super-resolution with convolutional neural network for microscopy images." *Biomedical Signal Processing and Control* 58 (2020): 101846.
27. Duan, Junwei, et al. "Deep learning based multimodal biomedical data fusion: An overview and comparative review." *Information Fusion* (2024): 102536.
28. Hussain, Dildar, et al. "Revolutionizing tumor detection and classification in multimodality imaging based on deep learning approaches: Methods, applications and limitations." *Journal of X-Ray Science and Technology* 32.4 (2024): 857-911.
29. Gou, Fangfang, et al. "Research on artificial-intelligence-assisted medicine: a survey on medical artificial intelligence." *Diagnostics* 14.14 (2024): 1472.
30. Dai, Wei, et al. "Automated Non-Invasive Analysis of Motile Sperms Using Sperm Feature-Related Network." *IEEE Transactions on Automation Science and Engineering* (2024).
31. Lee, Ryan, et al. "Automated rare sperm identification from low-magnification microscopy images of dissociated microsurgical testicular sperm extraction samples using deep learning." *Fertility and Sterility* 118.1 (2022): 90-99.
32. Abou Ghayda, Ramy, et al. "Artificial intelligence in andrology: from semen analysis to image diagnostics." *The World Journal of Men's Health* 42.1 (2023): 39.
33. Ignatieva, Elena V., et al. "A catalog of human genes associated with pathozoospermia and functional characteristics of these genes." *Frontiers in genetics* 12 (2021): 662770.
34. Botezatu, A., et al. "Advanced molecular approaches in male infertility diagnosis." *Biology of Reproduction* 107.3 (2022): 684-704.
35. Sharma, Saurabh, Praveen Kumar Mannepalli, and Harshita Chourasia. "Exploring Cutting-Edge Deep Learning Techniques in Image Processing." 2024 International Conference on Artificial Intelligence and Quantum Computation-Based Sensor Application (ICAIQSA). IEEE, 2024.
36. Swamy, Samatha R., and KS Nandini Prasad. "Revolutionizing healthcare intelligence multisensory data fusion with cutting-edge machine learning and deep learning for patients' cognitive knowledge." 2024 International Conference on Knowledge Engineering and Communication Systems (ICKECS). Vol. 1. IEEE, 2024.
37. Wang, Xu-Wen, Tong Wang, and Yang-Yu Liu. "Artificial Intelligence for Microbiology and Microbiome Research." *arXiv preprint arXiv:2411.01098* (2024).

38. Wang, Guangyu, et al. "A generalized AI system for human embryo selection covering the entire IVF cycle via multi-modal contrastive learning." *Patterns* 5.7 (2024).
39. Liu, Mark, et al. "WISE: whole-scenario embryo identification using self-supervised learning encoder in IVF." *Journal of Assisted Reproduction and Genetics* 41.4 (2024): 967-978.
40. Zhang, Chongming, et al. "Sperm YOLOv8E-TrackEVD: A Novel Approach for Sperm Detection and Tracking." *Sensors* 24.11 (2024): 3493.
41. Kresch, Eliyahu, et al. "Novel methods to enhance surgical sperm retrieval: a systematic review." *Arab Journal of Urology* 19.3 (2021): 227-237.
42. Fan, Keqiang. *Machine learning techniques for medical image analysis with data scarcity*. Diss. University of Southampton, 2025.
43. Galić, Irena, et al. "Machine learning empowering personalized medicine: A comprehensive review of medical image analysis methods." *Electronics* 12.21 (2023): 4411.
44. Baltrušaitis, Tadas, Chaitanya Ahuja, and Louis-Philippe Morency. "Multimodal machine learning: A survey and taxonomy." *IEEE transactions on pattern analysis and machine intelligence* 41.2 (2018): 423-443.
45. Wu, Biao, et al. "Xlip: Cross-modal attention masked modelling for medical language-image pre-training." *arXiv preprint arXiv:2407.19546* (2024).
46. Chen, Jintai, et al. "Trialbench: Multi-modal artificial intelligence-ready clinical trial datasets." *arXiv preprint arXiv:2407.00631* (2024).
47. Yang, Jucheng, and Fushun Ren. "Drug–Target Affinity Prediction Based on Cross-Modal Fusion of Text and Graph." *Applied Sciences* 15.6 (2025): 2901.
48. Lu, Haoyu, et al. "Uniadapter: Unified parameter-efficient transfer learning for cross-modal modeling." *arXiv preprint arXiv:2302.06605* (2023).
49. Ye, Yiwen, et al. "Continual self-supervised learning: Towards universal multi-modal medical data representation learning." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2024.
50. Liu, Qijiong, et al. "Multimodal pretraining, adaptation, and generation for recommendation: A survey." *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 2024.
51. Chen, Cheng, et al. "Unsupervised bidirectional cross-modality adaptation via deeply synergistic image and feature alignment for medical image segmentation." *IEEE transactions on medical imaging* 39.7 (2020): 2494-2505.
52. Li, Xiangyun, et al. "Machine Learning-Based Pathomics Model to Predict the Prognosis in Clear Cell Renal Cell Carcinoma." *Technology in Cancer Research & Treatment* 23 (2024): 15330338241307686.
53. Lai, Qi, et al. "Hybrid multiple instance learning network for weakly supervised medical image classification and localization." *Expert Systems with Applications* 260 (2025): 125362.
54. Ren, Zeyu, Shuihua Wang, and Yudong Zhang. "Weakly supervised machine learning." *CAAI Transactions on Intelligence Technology* 8.3 (2023): 549-580.
55. Koteluk, Oliwia, et al. "How do machines learn? artificial intelligence as a new era in medicine." *Journal of Personalized Medicine* 11.1 (2021): 32.
56. Wang, Yang. "Survey on deep multi-modal data analytics: Collaboration, rivalry, and fusion." *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* 17.1s (2021): 1-25.
57. Yao, Weiheng, and Shuqiang Wang. "CATD: Unified Representation Learning for EEG-to-fMRI Cross-Modal Generation." *arXiv preprint arXiv:2408.00777* (2024).
58. Zeng, Xiangxiang, et al. "Deep generative molecular design reshapes drug discovery." *Cell Reports Medicine* 3.12 (2022).
59. Born, Jannis, and Matteo Manica. "Trends in deep learning for property-driven drug design." *Current medicinal chemistry* 28.38 (2021): 7862-7886.
60. Weng, Wei-Hung. *Learning Representations for Limited and Heterogeneous Medical Data*. Diss. Massachusetts Institute of Technology, 2022.
61. Sun, Kai, et al. "Medical Multimodal Foundation Models in Clinical Diagnosis and Treatment: Applications, Challenges, and Future Directions." *arXiv preprint arXiv:2412.02621* (2024).

62. Chen, Yanbei, et al. "Semi-supervised and unsupervised deep visual learning: A survey." *IEEE transactions on pattern analysis and machine intelligence* 46.3 (2022): 1327-1347.
63. Sun, Ya, Sijie Mai, and Haifeng Hu. "Learning to learn better unimodal representations via adaptive multimodal meta-learning." *IEEE Transactions on Affective Computing* 14.3 (2022): 2209-2223.
64. Chudasama, Yashrajsinh, et al. "Towards Interpretable Hybrid AI: Integrating Knowledge Graphs and Symbolic Reasoning in Medicine." *IEEE Access* (2025).
65. Zou, Xingchen, et al. "Deep learning for cross-domain data fusion in urban computing: Taxonomy, advances, and outlook." *Information Fusion* 113 (2025): 102606.
66. Gao, Bin, et al. "An Explainable Unified Framework of Spatio-Temporal Coupling Learning with Application to Dynamic Brain Functional Connectivity Analysis." *IEEE Transactions on Medical Imaging* (2024).
67. Deldari, Shohreh, et al. "Beyond just vision: A review on self-supervised representation learning on multimodal and temporal data." *arXiv preprint arXiv:2206.02353* (2022).
68. Dimitri, Giovanna Maria, et al. "Multimodal and multicontrast image fusion via deep generative models." *Information Fusion* 88 (2022): 146-160.
69. Hu, Wenxing, et al. "Interpretable multimodal fusion networks reveal mechanisms of brain cognition." *IEEE transactions on medical imaging* 40.5 (2021): 1474-1483.
70. Fernández-López, Pol, et al. "Predicting fertility from sperm motility landscapes." *Communications biology* 5.1 (2022): 1027.
71. Tanga, Bereket Molla, et al. "Semen evaluation: Methodological advancements in sperm quality-specific fertility assessment—A review." *Animal bioscience* 34.8 (2021): 1253.
72. Komatsu, Kouji, et al. "Mitochondrial membrane potential in 2-cell stage embryos correlates with the success of preimplantation development." *Reproduction* 147.5 (2014): 627-638.
73. Zhao, Na, et al. "Mfn2 affects embryo development via mitochondrial dysfunction and apoptosis." *PloS one* 10.5 (2015): e0125680.
74. Zenoaga-Barbăroșie, Cătălina, and Marlon Martinez. "Sperm Morphology." *Human Semen Analysis: From the WHO Manual to the Clinical Management of Infertile Men* (2024): 135.
75. Gatimel, N., et al. "Sperm morphology: assessment, pathophysiology, clinical relevance, and state of the art in 2017." *Andrology* 5.5 (2017): 845-862.
76. Milewska, Anna Justyna, et al. "The Influence of Cryopreservation on Sperm Morphology and Its Implications in Terms of Fractions of Higher-Quality Sperm." *Journal of Clinical Medicine* 13.24 (2024): 7562.
77. Nygate, Yoav N., et al. "Holographic virtual staining of individual biological cells." *Proceedings of the National Academy of Sciences* 117.17 (2020): 9223-9231.
78. Zou, Shuojia, et al. "TOD-CNN: An effective convolutional neural network for tiny object detection in sperm videos." *Computers in Biology and Medicine* 146 (2022): 105543.
79. Dardikman-Yoffe, Gili, et al. "High-resolution 4-D acquisition of freely swimming human sperm cells without staining." *Science advances* 6.15 (2020): eaay7619.
80. Gongora, A. "A New Perspective on Sperm Analysis Through Artificial Intelligence: The Path Toward Personalized Reproductive Medicine." *Ann Clin Med Case Rep* 14.11 (2025): 1-10.
81. You, Jae Bem, et al. "Machine learning for sperm selection." *Nature Reviews Urology* 18.7 (2021): 387-403.
82. O'meara, Ciara, et al. "The effect of adjusting settings within a Computer-Assisted Sperm Analysis (CASA) system on bovine sperm motility and morphology results." *Animal Reproduction* 19.1 (2022): e20210077.
83. Dias, Tania R., Chak-Lam Cho, and Ashok Agarwal. "Sperm assessment: Traditional approaches and their indicative value." *In vitro fertilization: A textbook of current and emerging methods and devices* (2019): 249-263.
84. Klumpp, Matthias, et al. "Syntalos: a software for precise synchronization of simultaneous multi-modal data acquisition and closed-loop interventions." *Nature Communications* 16.1 (2025): 708.
85. Pinto, Francisco M., et al. "The role of sperm membrane potential and ion channels in regulating sperm function." *International Journal of Molecular Sciences* 24.8 (2023): 6995.

86. Amini Mahabadi, Javad, et al. "The Use of Machine Learning for Human Sperm and Oocyte Selection and Success Rate in IVF Methods." *Andrologia* 2024.1 (2024): 8165541.
87. Abou Ghayda, Ramy, et al. "Artificial intelligence in andrology: from semen analysis to image diagnostics." *The World Journal of Men's Health* 42.1 (2023): 39.
88. Esteves, Sandro C., et al. "Sperm DNA fragmentation testing: Summary evidence and clinical practice recommendations." *Andrologia* 53.2 (2021): e13874.
89. Agarwal, Ashok, and Rakesh Sharma. "Sperm chromatin assessment." *Textbook of assisted reproductive techniques*. CRC Press, 2023. 70-93.
90. Marinaro, Jessica A., and Peter N. Schlegel. "Sperm DNA damage and its relevance in fertility treatment: a review of recent literature and current practice guidelines." *International journal of molecular sciences* 24.2 (2023): 1446.
91. Tvrdá, Eva, et al. "Dynamic assessment of human sperm DNA damage III: The effect of sperm freezing techniques." *Cell and Tissue Banking* 22 (2021): 379-387.
92. Diaz, Parris, et al. "Future of male infertility evaluation and treatment: brief review of emerging technology." *Urology* 169 (2022): 9-16.
93. Li, Philip S., and Ranjith Ramasamy. "Future of Male Infertility Evaluation and Treatment: Brief Review of Emerging Technology." (2022).
94. Esteva, Andre, et al. "A guide to deep learning in healthcare." *Nature medicine* 25.1 (2019): 24-29.
95. McCallum, Christopher, et al. "Deep learning-based selection of human sperm with high DNA integrity." *Communications biology* 2.1 (2019): 250.
96. Popova, Daria. "Advanced methods in reproductive medicine: Application of optical nanoscopy, artificial intelligence-assisted quantitative phase microscopy and mitochondrial DNA copy numbers to assess human sperm cells." (2021).
97. Du, Shengrong, et al. "Micro-Raman analysis of sperm cells on glass slide: potential label-free assessment of sperm DNA toward clinical applications." *Biosensors* 12.11 (2022): 1051.
98. Madan, Sumit, et al. "Transformer models in biomedicine." *BMC Medical Informatics and Decision Making* 24.1 (2024): 214.
99. Murad, Nafeesa Yousuf, et al. "Unraveling the black box: A review of explainable deep learning healthcare techniques." *IEEE Access* (2024).
100. Ali, Sajid, et al. "Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence." *Information fusion* 99 (2023): 101805.
101. Perry, Joseph. Streamlining the image capture and analysis for high throughput unbiased quantitation and categorization of nuclear morphology. Diss. University of Essex, 2024.
102. Datta Burton, Saheli, et al. "Clinical translation of computational brain models: Understanding the salience of trust in clinician–researcher relationships." *Interdisciplinary Science Reviews* 46.1-2 (2021): 138-157.
103. Ruppli, Camille. Methods and frameworks of annotation cost optimization for deep learning algorithms applied to medical imaging. Diss. Institut Polytechnique de Paris, 2023.
104. Suara, Subhashis, et al. "Is grad-cam explainable in medical images?." *International Conference on Computer Vision and Image Processing*. Cham: Springer Nature Switzerland, 2023.
105. Nazir, Asifa, et al. "A novel approach in cancer diagnosis: integrating holography microscopic medical imaging and deep learning techniques—challenges and future trends." *Biomedical Physics & Engineering Express* 11.2 (2025): 022002.
106. Isa, Iza Sazanita, Umi Kalsom Yusof, and Murizah Mohd Zain. "Image Processing Approach for Grading IVF Blastocyst: A State-of-the-Art Review and Future Perspective of Deep Learning-Based Models." *Applied Sciences* 13.2 (2023): 1195.
107. Bagheri, Maryam, et al. "AI-driven decision-making in healthcare information systems: a comprehensive review." *Preprints* (2024).
108. Luong, Thi-My-Trang, and Nguyen Quoc Khanh Le. "Artificial intelligence in time-lapse system: advances, applications, and future perspectives in reproductive medicine." *Journal of assisted reproduction and genetics* 41.2 (2024): 239-252.

109. Lilhore, Umesh Kumar, et al., eds. *Multimodal Data Fusion for Bioinformatics Artificial Intelligence*. John Wiley & Sons, 2025.
110. Kaur, Simranjit, et al. "Precision medicine with data-driven approaches: A framework for clinical translation." *AIJMR-Advanced International Journal of Multidisciplinary Research* 2.5 (2024).
111. Bai, Long, et al. "AI-enabled organoids: construction, analysis, and application." *Bioactive materials* 31 (2024): 525-548.
112. Putrama, I. Made, and Péter Martinek. "Heterogeneous data integration: Challenges and opportunities." *Data in Brief* (2024): 110853.
113. Li, Jian, et al. "Image Encoding and Fusion of Multi-modal Data Enhance Depression Diagnosis in Parkinson's Disease Patients." *IEEE Transactions on Affective Computing* (2024).
114. Lürig, Moritz D., et al. "Computer vision, machine learning, and the promise of phenomics in ecology and evolutionary biology." *Frontiers in Ecology and Evolution* 9 (2021): 642774.
115. Chakraborty, Sohini, et al. "Multi-OMICS approaches in cancer biology: New era in cancer therapy." *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease* 1870.5 (2024): 167120.
116. Huan, Changxiang, et al. "Spatially Resolved Multiomics: Data Analysis from Monoomics to Multiomics." *BME frontiers* 6 (2025): 0084.
117. Liu, Jinyuan, et al. "Coconet: Coupled contrastive learning network with multi-level feature ensemble for multi-modality image fusion." *International Journal of Computer Vision* 132.5 (2024): 1748-1775.
118. Tan, Bo, et al. "Aberrant whole-brain resting-state functional connectivity architecture in obsessive-compulsive disorder: an EEG study." *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 30 (2022): 1887-1897.
119. Wysocka, Magdalena, et al. "A systematic review of biologically-informed deep learning models for cancer: fundamental trends for encoding and interpreting oncology data." *BMC bioinformatics* 24.1 (2023): 198.
120. Sun, Shilin, et al. "A review of multimodal explainable artificial intelligence: Past, present and future." *arXiv preprint arXiv:2412.14056* (2024).
121. Joshi, Gargi, Rahee Walambe, and Ketan Kotecha. "A review on explainability in multimodal deep neural nets." *IEEE Access* 9 (2021): 59800-59821.
122. Dong, Aimei, et al. "Co-Enhancement of Multi-Modality Image Fusion and Object Detection via Feature Adaptation." *IEEE Transactions on Circuits and Systems for Video Technology* (2024).
123. Zhang, Qingyang, et al. "Multimodal fusion on low-quality data: A comprehensive survey." *arXiv preprint arXiv:2404.18947* (2024).
124. Friha, Othmane, et al. "Llm-based edge intelligence: A comprehensive survey on architectures, applications, security and trustworthiness." *IEEE Open Journal of the Communications Society* (2024).
125. Bai, Guangji, et al. "Beyond efficiency: A systematic survey of resource-efficient large language models." *arXiv preprint arXiv:2401.00625* (2024).
126. Rauniyar, Ashish, et al. "Federated learning for medical applications: A taxonomy, current trends, challenges, and future research directions." *IEEE Internet of Things Journal* 11.5 (2023): 7374-7398.
127. Jiao, Licheng, et al. "Causal inference meets deep learning: A comprehensive survey." *Research* 7 (2024): 0467.
128. Jing, Yankang. *Machine/Deep Learning Causal Pharmacology-Analytics for Preclinical System Pharmacology Modeling and Clinical Outcomes Analytics*. Diss. University of Pittsburgh, 2021.
129. Xu, Xi, et al. "A comprehensive review on synergy of multi-modal data and ai technologies in medical diagnosis." *Bioengineering* 11.3 (2024): 219.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.