Paper Title	Objective	ML Techniques	Max success rate	Observations
Machine Learning-Based Analysis of Sperm Videos and Participant Data for Male Fertility Prediction	Predict sperm motility (3 classifications)	Linear Regression Random Forests Gaussian Process Sequential Minimal SMOreg Elastic Net Random Trees CNN pipeline 3-fold cross validation	MAE: 7.260 to 13.220	LIRE(java) WEKA(library) Multimodal Participant data: age, BMI, and days of sexual abstinence. Tamura features
Classification of Mouse Sperm Motility Patterns Using an Automated Multiclass Support Vector Machines Model1	Predict sperm motility (5 classifications)	Multiclass SVM model: “CASAnova”	Accuracy: 52.1 to 94.7 Total: 88.2%	Matlab LibSVM Analysis with GraphPad Prism 5 CASA parameters
CASAnova: a multiclass support vector machine model for the classification of human sperm motility patterns	Predict sperm motility (5 classifications)	This study uses the same model and technologies as the previous but focuses on human sperm.	Precision Sensitivity Specificity Accuracy: 89.92%	GraphPad Prism 6
Faster region convolutional neural network and semen tracking algorithm for sperm analysis	Analyzing the sperm motility and morphology	Linear Regression CNN vgg16 FRCNN	Accuracy: 97.37% Execution time: 1.12s	(weird english) Very detailed on the making of the images for analysis & the CNN workflow THMA: Tail to Head Movement Algorithm
Deep Convolutional and Recurrent Neural Networks for Cell Motility Discrimination and Prediction	Predict different cell types by motility	CNN RNN Long Short-Term Memory Random Forest	Up to 95% Accuracy	Good detailed explanation of RNN architecture used
Detection of DNA base modifications by deep recurrent neural network on Oxford Nanopore sequencing data	Detect DNA base modifications	RNN Long Short-Term Memory 5-fold Cross Validation 90/10 train test split DeepMod	F1: 0.983 Accuracy: 98.8%	Nanopore Sequencing based on electrical signals to identify altered DNA bases Bidirectional RNN Accuracy of 57 features == to 7 features

Machine-learning algorithm incorporating capacitated sperm intracellular pH predicts conventional in vitro fertilization success in normospermic patients	To measure sperm intracellular pH and predict successful conventional IVF. Success if fertilization ratio greater than 0.66	Gradient Boosting 75/25 train-test split Recursive Feature Elimination for feature selection	Accuracy: 72% AUC: 0.81 Sensitivity: 0.65 Specificity: 0.8	Bootstrapping models Very small dataset, focused on specific analysis. Not a lot of information on ML model
The use of machine learning methods to predict sperm quality in Holstein bulls	Predict total sperm motility, morphological abnormalities and sperm output.	Lasso Group Lasso Gradient Boosting Cross Validation Binary classification for motility and morphology 85/15 train-test split	Accuracy: 80+%	Bull sperm, not human sperm. Interesting analysis on flexible variable selection for best predictions vs overfitting Caution with MAPE and MAE
Quantification of human sperm concentration using machine learning-based spectrophotometry	Quantification of human sperm concentration	Artificial Neural Network Two-layer Backpropagation 70/15/15 train-test-val split 13-fold Cross Validation	Accuracy: 93+%	Spectrophotometry The metric for the ML model is the absorption response of sperm to different light. Sperm concentration seems to be a good metric to qualify samples
motilitAI: A machine learning framework for automatic prediction of human sperm motility	Predict sperm motility Predict percentage of each classification	Neural Network (MLP, CNN & RNN) Support Vector Regression Bag of Words	R2 score for each class: 74% 26% 66%	2 sperm tracking methods and 2 feature extraction methods that might be relevant

Datasets:

[Simula Datasets - VISEM]
[Human Sperm Head Morphology dataset (HuSHeM) - Mendeley Data]
[soroushj/mhsma-dataset: MHSMA: The Modified Human Sperm Morphology Analysis Dataset (github.com)]
[Sperm Morphological Quality | Kaggle]
[Demozsj/Detection-Sperm (github.com)]