

Successful classification of normal and abnormal ram sperm by a machine learning model

Katherine R Seymour¹, Jessica P Rickard¹, Simon P de Graaf¹

1. The University of Sydney, Faculty of Science, School of Life and Environmental Sciences, Australia

Sperm morphology is a key in vitro predictor of ram fertility. A regular feature of semen assessment, sperm morphology assists in the identification of infertile or sub-fertile rams or semen samples, thus ensuring maximum reproductive efficiency of natural and artificial breeding programs. However, the efficacy of morphology assessment is limited by high levels of variability between assessors and labs and it's difficult to standardise. Indeed, preliminary work by our group indicated industry experts (n=4) using a simplified yes/no system of normal/abnormal agreed on only 73% (585/800) of sperm, which fell to 27% (212/800) of sperm in a comprehensive 30 category labelling system. One alternative to subjective morphological assessment of sperm is the application of machine learning. Machine learning has a multitude of applications and has already been applied to object recognition and labelling in other reproductive assessment contexts such as embryo quality. As such, this study aimed to utilise machine learning, specifically a convolutional neural network, to create an objective tool for morphological classification of ram sperm.

Ram sperm obtained from 73 sires were fixed (5% formalin in PBS) and mounted on glass slides. Images of 100-200 sperm per sire were captured using differential interference contrast microscopy (400x magnification; numerical aperture = 0.75), on a Olympus DP23 camera, and cropped to 512x512 px resolution using a novel machine learning tool. Sperm were labelled as normal or abnormal by experienced morphologists. Sperm (N=2323) that had 100% classification consensus between the morphologists were randomly divided into training, validation and test data sets for the development of the machine learning model, at a ratio of 60:20:20, respectively. The model was constructed using a Resnet 50 scaffold with two dense layers of 512 then 272 units and an optimal learning rate was decided upon running multiple hyperparameters. The data points were put through the model at a 224x224 px resolution and kept as colour (RGB) images. TensorFlow was utilised to monitor for overfitting and to automatically stop the model learning before it reached that point.

The model was found to have a 95% accuracy for classifying sperm as normal or abnormal on the unseen test data set and was able to achieve a 97% accuracy on the training data.

This study demonstrated that machine learning can be used to successfully classify ram sperm as normal or abnormal. Future research into this area will focus on expanding the number of morphological categories the model is able to classify in order to facilitate objective assessment of sperm in clinical and research situations that require more complex morphological classification systems.