

# AUTOMATIC EVALUATION OF HUMAN OOCYTE DEVELOPMENTAL POTENTIAL FROM MICROSCOPY IMAGES

Denis Baručić\*    Jan Kybic\*    Olga Teplá†    Zinovij Topurko†    Irena Kratochvílová‡

\* Department of Cybernetics, Faculty of Electrical Engineering, Czech Technical University in Prague, Czech Republic

† Department of Obstetrics and Gynecology, The First Faculty of Medicine and General Teaching Hospital, Czech Republic

‡ Institute of Physics of the Czech Academy of Sciences, Czech Republic

## ABSTRACT

Infertility is becoming an issue for an increasing number of couples. The most common solution, in vitro fertilization, requires embryologists to carefully examine light microscopy images of human oocytes to determine their developmental potential. We propose an automatic system to improve the speed, repeatability, and accuracy of this process. We first localize individual oocytes and identify their principal components using CNN (U-Net) segmentation. We calculate several descriptors based on geometry and texture. The final step is an SVM classifier. Both the segmentation and classification training are based on expert annotations. The presented approach leads to the classification accuracy of 70%.

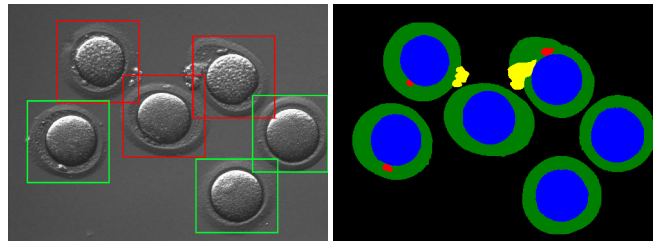
**Index Terms**— human oocytes, fertilization, microscopy, classification, segmentation

## 1. INTRODUCTION

Infertility has been an issue for several years and is expected to further grow. Nowadays, the most common solution for an infertile couple is in vitro fertilization (IVF). One of the important steps is choosing the best oocytes to be fertilized, since, for practical, ethical, and legal reasons, it is not feasible to fertilize more than a few of them — and even fewer can be implanted.

The situation is easier when a patient’s own oocytes are used and are, therefore, more readily available. In this case, the embryologists attempt to fertilize almost all collected eggs that are not apparently damaged. The knowledge gained from these attempts can be used later to determine the apriori developmental potential of newly collected oocytes, reducing both the cost and the failure rate of the IVF. This is especially important when the oocytes (eggs) come from a donor. In this case, each oocyte is very valuable, and it is important to reliably determine its developmental potential. Currently, this is performed by an expert using a microscope.

In this work, we aim to replace this subjective process by an automatic evaluation of the developmental ability of oocytes from digitized light microscopy images to improve



**Fig. 1.** Example of an input image and the corresponding expert segmentation. Classes *background*, *cytoplasm*, *zona pellucida*, *polar body*, and *cumulus cells* are denoted in black, blue, green, red, and yellow, respectively. The green and red frames denote viable and nonviable oocytes, respectively.

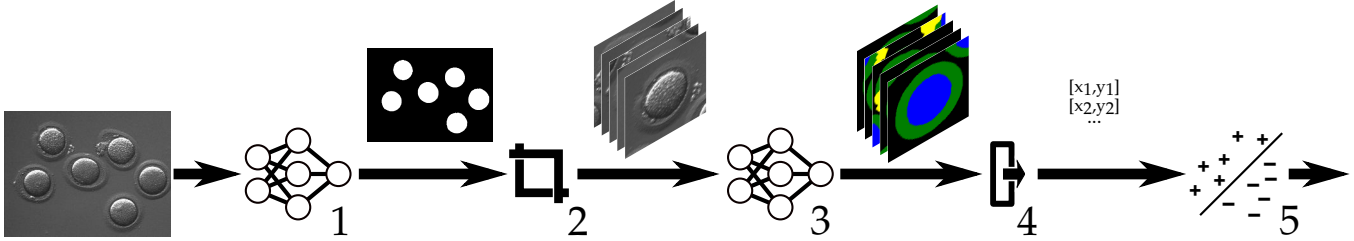
its speed, repeatability, and accuracy. We present a proof-of-concept solution on a small dataset, showing the viability of this approach.

### 1.1. Previous work

To the best of our knowledge, the task of fully automatic oocyte developmental potential assessment has not been addressed before. Manna et al. [1] attempt to predict whether an oocyte or embryo (i.e., fertilized oocyte) will lead to birth. The method uses LBP texture features and an ensemble of shallow neural networks for the classification process but requires manual segmentation. Moreover, our target classification is different, since not all well-developing embryos lead to birth.

Classification of embryo quality was targeted in [2]. The authors worked with an extensive database of 50 000 embryo images and proposed a method based on deep learning to classify the embryos into three quality classes. Time-lapse analysis predicting the number of cells in embryos was proposed in [3].

Classification of swine cumulus oocyte complexes (i.e., before stripping away the cumulus cells, unlike in our data) was considered in [4]. The authors examined the number of cumulus cell layers and the homogeneity of the cytoplasm and used random forests for the automatic classification of



**Fig. 2.** The classification pipeline. (1) Binary segmentation. (2) ROI extraction. (3) Five class segmentation. (4) Feature extraction. (5) Classification.

oocytes. For segmentation, a semi-automatic (snake) method was used. Pure texture analysis has appeared to be useful for cytoplasm clustering [5] or cytoplasm segmentation [6].

## 1.2. Proposed approach

This paper describes a fully automatic approach to classify oocytes in light microscopy images (see Fig. 1) into two categories, viable and nonviable, where viable oocytes have a good potential of becoming well-developed embryos. We learn from subjective expert annotations of individual oocytes. Our approach, described in Sec. 3, consists of five consecutive stages — localization, patch extraction, segmentation, feature extraction, and classification.

## 2. DATA

Our anonymized dataset contains 34 greyscale images of groups of oocytes after cumulus cells denudation. The images of  $1392 \times 1040$  pixels were acquired using Nikon Diaphot 300 inverted microscope, Eppendorf (Hamburg, Germany) micro-manipulation system equipped with a thermoplate (Tokai Hit, Japan).

The ground truth (GT) segmentations (see Fig. 1) were created using the GIMP software. We considered four classes: background (*bg*), cytoplasm (*ct*), zona pellucida (*zp*), polar body (*pb*), and cumulus cells (*cc*). Furthermore, the embryology expert (OT) subjectively assessed the individual oocytes’ quality and classified them as viable or nonviable.

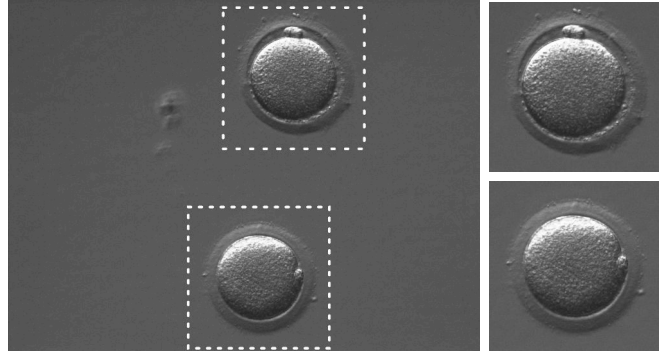
Each image contained 1 ~ 7 oocytes. Only oocytes fully within the image were considered, yielding 103 oocytes — 50 viable and 53 nonviable.

## 3. METHOD

The proposed pipeline is depicted in Fig. 2.

### 3.1. Oocyte localization

We first perform a binary segmentation of the cytoplasm (*cc*) versus the remaining classes because the cytoplasm is clearly distinguishable in the images. We use a U-Net [7] CNN with



**Fig. 3.** Example of an input image (left) and two ROIs extracted from the image.

MobileNetV2 [8] as the encoder and a pixel-wise softmax final layer. MobileNetV2 is a fast architecture with a relatively low number of parameters. The network was trained for 500 epochs. We used the Dice loss in combination with the ADAM optimizer (learning rate  $10^{-4}$ ). To prevent overfitting, multiple data augmentation methods (shifting and rotation, contrast and brightness adjustments, and blurring) were applied to the training images during training.

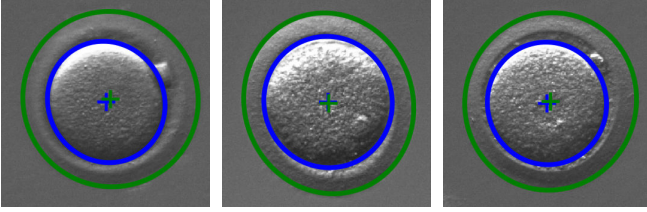
Connected foreground components smaller than 10 000 pixels are suppressed. Regions of interest (ROI) of size  $416 \times 416$  are extracted from around the centers of gravity of the remaining components (see Fig. 3).

### 3.2. Oocyte segmentation

Once the ROIs are extracted, they are segmented into the five classes (described in Sec. 2) using another CNN. Since the *pb* or *cc* classes are more challenging, we use the U-Net with the powerful ResNet50 architecture [9] as the encoder, trained for 600 epochs. The rest of the procedure was identical to Sec. 3.1. Example segmentations are shown in Fig. 5.

### 3.3. Feature extraction

Using the segmentation from the previous section, for each oocyte (ROI), we compute the 24 features described below.



**Fig. 4.** Fitted ellipses and their centers for  $ct$  (blue) and  $zp$  (green).

First, to handle the case where the ROI contains parts of several oocytes, we keep only the largest  $ct$  and  $zp$  components. We also suppress  $pb$  components smaller than 500 pixels (for a bad oocyte, it is possible to have multiple  $pb$  components, so we cannot just keep the largest).

Ellipses are fitted to the boundary of the  $ct$  class and to the outer boundary of the  $zp$  class by least squares fitting (see Fig. 4). We calculate the following features based on the cytoplasm ellipse:

- mean axis  $\mu_c = \frac{a_c + b_c}{2}$ ,
- eccentricity  $e_c = \sqrt{1 - \frac{a_c^2}{b_c^2}}$ ,  $a_c \geq b_c$ ,
- compactness  $\gamma_c = \frac{a_c b_c \pi}{S_c}$ ,

where  $a_c, b_c$  are the estimated semi-axes, and  $S_c$  is the area of the  $ct$  component. The features  $\mu_z, e_z, \gamma_z$  are calculated similarly for the  $zp$  class. We also define the misalignment  $m = \|\mathbf{c}_c - \mathbf{c}_z\|$ , the Euclidean distance between the ellipse centers, and the ratio of the cytoplasm and zona pellucida areas,  $r = \frac{S_c}{S_z}$ , inspired by [5].

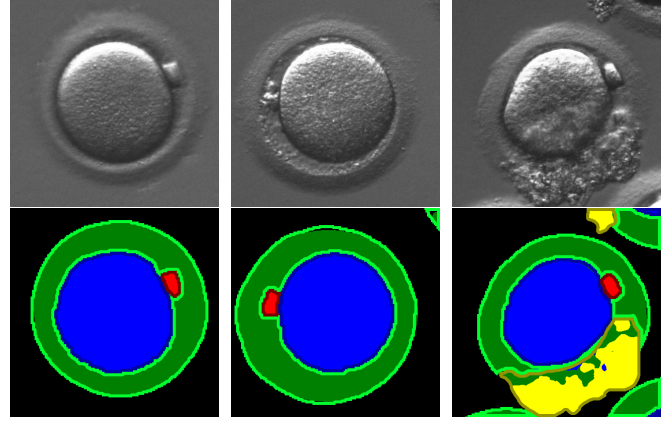
Regarding polar bodies, two features are used: the number of connected components,  $n_{pb}$ , and the total area,  $S_{pb}$ . The presence of cumulus cells is not related to the oocyte quality but may influence the other features, hence we also calculate the total area of cumulus cells  $S_{cc}$ , which completes the 11 geometrical features.

The remaining 13 features describe the texture of the cytoplasm, which appears to be important [1]. The first 10 texture features are calculated from a three-level undecimated wavelet transform and correspond to the energies in the low pass channel and 9 high frequency channels [10]. The remaining three features are the mean and variance of the pixel intensities and the entropy of the cytoplasm histogram.

### 3.4. Oocyte classification

For each oocyte, the extracted feature vector is normalized and fed into a binary classifier to produce the binary label, *viable* or *nonviable*. Several classifiers were tried with similar results. For the sake of space, only kernel SVM results are reported.

The RBF kernel with  $\gamma = 10^{-2}$  and a cost parameter  $C = 0.1$  were selected using cross-validation and grid search on a training dataset of 83 randomly sampled ROIs, yielding



**Fig. 5.** Three extracted ROIs in the top row and the corresponding predicted and expert segmentations in the bottom row, denoted as colored areas and outlines, respectively. The colors encode the same classes as in Fig. 1.

a validation accuracy  $A_{\text{val}} \approx 74\%$ . The remaining 20 ROIs were used for testing (Sec. 4.3).

## 4. RESULTS

We evaluate the three stages of the pipeline separately.

### 4.1. Oocyte localization

An 8-fold cross-validation was performed to evaluate the oocyte localization described in Sec. 3.1 using the 34 training images. On our data, the method works perfectly. The number of detected oocytes was always equal to the ground truth and the localization error between the centers of gravity of the  $ct$  class and the ground truth cytoplasm segmentation was inferior to 10 pixels in 98% of cases.

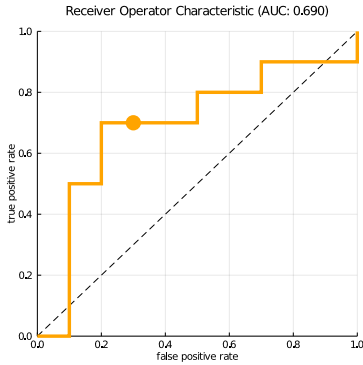
### 4.2. Oocyte segmentation

The five-class oocyte segmentation (Sec. 3.2) was evaluated using a 10-fold cross-validation on the 103 ROIs, each approximately centered on one oocyte.

The Intersection over Union (IoU) metric computed over the folds for the  $ct$ ,  $zp$ ,  $pb$ , and  $cc$  classes was 95.48%, 89.72%, 42.85%, and 60.29%, respectively. While  $ct$  and  $zp$  are segmented very well, the  $pb$  class suffers from false detections and is more difficult to segment due to its small size. Although the  $cc$  segmentation performance is also far from perfect, it does not seem to influence the overall classification performance.

### 4.3. Oocyte classification

The performance of the SVM from Sec. 3.4 was evaluated using the 20 testing ROIs that were omitted during the training.



**Fig. 6.** ROC curve obtained for the SVM classifier on the testing data. The dot represents the operating point we used.

The classifier obtained the testing accuracy  $A_{\text{test}} = 70\%$ , sensitivity  $S_{e_{\text{test}}} = 70\%$ , specificity  $S_{p_{\text{test}}} = 70\%$ , precision  $P_{\text{test}} = 70\%$ , and the area under the ROC curve was  $AUC_{\text{test}} = 0.69$  (see Fig. 6).

## 5. DISCUSSION AND CONCLUSIONS

In this paper, a proof-of-concept solution was proposed for automatic detection of oocytes that have a good developmental potential and are therefore viable for fertilization. When interpreting the results, it is important to realize that the expert assessment of the oocyte quality from a single image is difficult and rather subjective. For example, the polar body, which is an important indicator of oocyte viability, may not be visible in a given image. Our performance is nevertheless comparable to the accuracy achieved by Manna et al. [1] ( $AUC=0.68$ ) who worked with the more reliable information of whether an oocyte leads to birth.

We expect our accuracy to improve when more data is available, including objective information about the fertilization success. This should enable us to employ deep learning techniques instead of hand-crafted features, boosting the performance.

## 6. REFERENCES

- [1] C. Manna, L. Nanni, A. Lumini, and S. Pappalardo, “Artificial intelligence techniques for embryo and oocyte classification,” *Reproductive biomedicine online*, vol. 26, no. 1, pp. 42–49, 2013.
- [2] P. Khosravi, E. Kazemi, Q. Zhan, et al., “Deep learning enables robust assessment and selection of human blastocysts after in vitro fertilization,” *npj Digital Medicine*, vol. 2, no. 1, pp. 1–9, 2019.
- [3] V. Raudonis, A. Paulauskaite-Taraseviciene, K. Sutiene, and D. Jonaitis, “Towards the automation of early-stage

human embryo development detection,” *Biomedical Engineering OnLine*, vol. 18, no. 1, pp. 1–20, 2019.

- [4] P. S. Viswanath, T. Weiser, et al., “Grading of mammalian cumulus oocyte complexes using machine learning for in vitro embryo culture,” in *2016 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)*. IEEE, 2016, pp. 172–175.
- [5] T. M. A. Basile et al., “A texture-based image processing approach for the description of human oocyte cytoplasm,” *IEEE Transactions on Instrumentation and Measurement*, vol. 59, no. 10, pp. 2591–2601, 2010.
- [6] L. Caponetti, G. Castellano, et al., “Cytoplasm image segmentation by spatial fuzzy clustering,” in *International Workshop on Fuzzy Logic and Applications*. Springer, 2011, pp. 253–260.
- [7] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional networks for biomedical image segmentation,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2015, pp. 234–241.
- [8] M. Sandler, A. Howard, M. Zhu, et al., “MobileNetV2: Inverted residuals and linear bottlenecks,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 4510–4520.
- [9] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778.
- [10] M. Unser, “Texture classification and segmentation using wavelet frames,” *IEEE Transactions on Image Processing*, vol. 4, no. 11, pp. 1549–1560, 1995.

## Compliance with ethical standards

This research study was conducted retrospectively using anonymized data collected as part of a standard treatment procedure. The data acquisition was approved by the Ethics Committee of the General University Hospital, Prague, Czech Republic, on October 12, 2018, reference number 79/17.

## Acknowledgments

The authors acknowledge the support of the OP VVV funded project “CZ.02.1.01/0.0/0.0/16\_019/0000765 Research Center for Informatics” and the Grant Agency of the Czech Technical University in Prague, grant No. SGS20/170/OHK3/3T/13. The authors have no relevant financial or nonfinancial interests to disclose.