AUTOMATED EMBRYO GRADING BASED ON ARTIFICIAL INTELLIGENCE

Ali Akbar Septiandri[1], Ade Jamal[1], Oki Riayati[2], Pritta Ameilia Iffanolida[2], Budi Wiweko[2,3,4]

[1]Faculty of Science and Technology, Universitas Al-Azhar Indonesia

[2]Human Reproductive, Infertility, and Family Planning Research Center, Indonesian Medical Education and Research Institute (IMERI), Faculty of Medicine Universitas Indonesia, Jakarta, 10430, Indonesia

[3]Division of Reproductive Endocrinology and Infertility, Department of Obstetrics and Gynecology Faculty of Medicine, Universitas Indonesia, Jakarta, 10430, Indonesia

[4]Yasmin IVF Clinic, Dr. Cipto Mangunkusumo General Hospital, Jakarta, 10430, Indonesia

INTRODUCTION

BACKGROUND: Visual morphology assessment for embryo grading is routinely used for evaluating of embryo selection for transfer after in vitro fertilization (IVF). However, the assessment produces different results between embryologist and as a result, the success rate of IVF remains low. To overcome uncertainties in embryo quality, multiple embryos are often implanted resulting in undesired multiple pregnancies and complication. Unlike in other imaging fields, human embryology and IVF have not yet develop artificial intelligence (AI) for unbiased, automated embryo assessment. The aim of this study was to develop the automated embryo assessment based on AI.

MATERIALS AND METHODS: This study included a total of 1084 images from 1226 embryos of 246 IVF cycles at Yasmin IVF Clinic, Jakarta, Indonesia. The images were capture by inverted microscope at day 3 after fertilization. Using a fine-tuned ResNet50, a pre-trained deep learning model trained on the ImageNet dataset, we predict the labels of embryo images. The predicted grading results were compared with the grading results from trained embryologists to evaluate the model performance.

RESULTS

The images were labeled based on Veeck Criteria that differentiate embryo to grade 1 until 3 based on size of the blastomere and grade of fragmentation. We divided the data into training and test sets with 3:1 ratio.

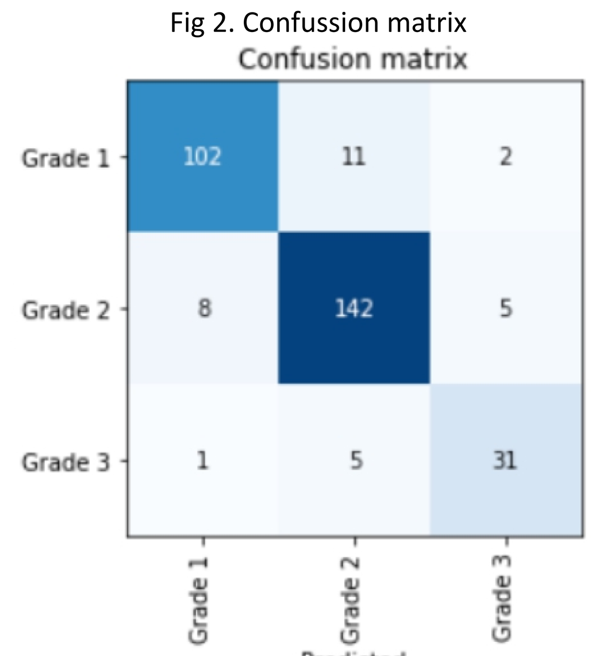
Fig. 1. Embryo images after the preprocessing steps

We found that the best learning rates are around 5 × 10−3 from 8 epochs. The results from all models can be seen in Table I. We can see that the ResNet50 got the highest accuracy and the lowest cross-entropy loss of all models. While increasing the depth of the ResNet model from 18, 34, to 50 increased the accuracy, it stopped increasing afterwards. We argue that this is because the dataset is relatively simpler compared to ImageNet where we have different objects in different colours and sizes. This might also be the case why DenseNet models failed to achieve better performance while being more complex.

TABLE I
MODEL COMPARISON

model	accuracy	loss
ResNet18	$89.38\% \pm 0.75\%$	0.3312 ± 0.0330
ResNet34	$89.97\% \pm 1.27\%$	0.3495 ± 0.0343
ResNet50	$\mathbf{91.79\% \pm 0.48\%}$	$\mathbf{0.3114 \pm 0.0253}$
ResNet101	$91.07\% \pm 1.00\%$	0.3749 ± 0.0623
DenseNet121	$89.97\% \pm 0.27\%$	0.3567 ± 0.0365
DenseNet169	$91.14\% \pm 0.54\%$	0.3472 ± 0.0366
Xception	$88.86\% \pm 0.96\%$	0.3209 ± 0.0206
MobileNetV2	$91.14\% \pm 0.84\%$	0.3442 ± 0.0258

Fig 2. Confussion matrix



CONCLUSION

Although the results are only preliminary, it showed an interesting classification performance. The AI presented is an automated, non-invasive, objective, and highly recommendation for embryo assessment.

We have shown in this study that we can grade day 3 embryo images automatically with the best accuracy of 93.16% by fine-tuning a ResNet50 model. We found that more complex models failed to achieve better accuracy compared to the ResNet50.

Since we are still manually cropping the embryos from the original images, we can extend this work to automate this task, e.g. using an image segmentation model like YOLOv3 [23] or U-Net [24]. In the future, we hope that this model can be developed as an embedded system for point-of-care diagnostics such as found in

REFERENCES

[1] J. Cummins, T. Breen, K. Harrison, J. Shaw, L. Wilson, and J. Hennessey, “A formula for scoring human embryo growth rates in in vitro fertilization: its value in predicting pregnancy and in comparison with visual estimates of embryo quality,” Journal of In Vitro Fertilization and Embryo Transfer, vol. 3, no. 5, pp. 284–295, 1986.

[2] N. Nasiri and P. Eftekhari-Yazdi, “An overview of the available methods for morphological scoring of pre-implantation embryos in in vitro fertilization,” Cell Journal (Yakhteh), vol. 16, no. 4, p. 392, 2015.

[3] A. E. B. Bendus, J. F. Mayer, S. K. Shipley, and W. H. Catherino, “Interobserver and intraobserver variation in day 3 embryo grading,” Fertility and sterility, vol. 86, no. 6, pp. 1608–1615, 2006.

[4] L. L. Veeck, An atlas of human gametes and conceptuses: an illustrated reference for assisted reproductive technology. CRC Press, 1999.

[5] P. Khosravi, E. Kazemi, Q. Zhan, J. E. Malmsten, M. Toschi, P. Zisimopoulos, A. Sigaras, S. Lavery, L. A. Cooper, C. Hickman et al., “Deep learning enables robust assessment and selection of human blastocysts after in vitro fertilization,” npj Digital Medicine, vol. 2, no. 1, p. 21, 2019.

[6] Z. Akkus, A. Galimzianova, A. Hoogi, D. L. Rubin, and B. J. Erickson, “Deep learning for brain mri segmentation: state of the art and future directions,” Journal of digital imaging, vol. 30, no. 4, pp. 449–459, 2017.

[7] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. Van Der Laak, B. Van Ginneken, and C. I. Sanchez, ´ “A survey on deep learning in medical image analysis,” Medical image analysis, vol. 42, pp. 60–88, 2017.

[8] J. A. Quinn, R. Nakasi, P. K. Mugagga, P. Byanyima, W. Lubega, and A. Andama, “Deep convolutional neural networks for microscopybased point of care diagnostics,” in Machine Learning for Healthcare Conference, 2016, pp. 271–281.

[9] N. Zaninovic, O. Elemento, and Z. Rosenwaks, “Artificial intelligence: its applications in reproductive medicine and the assisted reproductive technologies,” Fertility and sterility, vol. 112, no. 1, pp. 28–30, 2019.

[10] M. F. Kragh, J. Rimestad, J. Berntsen, and H. Karstoft, “Automatic grading of human blastocysts from time-lapse imaging,” Computers in Biology and Medicine, p. 103494, 2019.

CONTACT

E-mail: ina.repromed@gmail.com

The 10th Congress of the Asia Pacific Initiative on Reproduction

BRIDGING THE GAP: FERTILITY AND REPRODUCTION