ORIGINAL ARTICLE

Reproductive Medicine and Biology    WILEY

# Development of an automated two pronuclei detection system on time-lapse embryo images using deep learning techniques

Noritaka Fukunaga[1,2] (iD)    |    Sho Sanami[3]    |    Hiroya Kitasaka[1] (iD)    |    Yuji Tsuzuki[3]    |
Hiroyuki Watanabe[1] (iD)    |    Yuta Kida[1]    |    Seiji Takeda[3]    |    Yoshimasa Asada[1,2] (iD)

[1]Asada Ladies Clinic, Nagoya, Japan

[2]Asada Institute for Reproductive Medicine, Kasugai, Japan

[3]Research & Development Center, Dai Nippon Printing Co., Ltd, Tokyo, Japan

**Correspondence**
Noritaka Fukunaga, Asada Ladies Clinic, Nagoya Building 5F, 4-6-17 Meieki, Nakamura-ku, Nagoya, Aichi 450-0002, Japan.
Email: n_fukunaga@ivf-asada.jp

## Abstract

**Purpose:** To establish an automated pronuclei determination system by analysis using deep learning technology which is able to effectively learn with limited amount of supervised data.

**Methods:** An algorithm was developed by explicitly incorporating human observation where the outline around pronuclei is being observed in determining the number of pronuclei. Supervised data were selected from the time-lapse images of 300 pronuclear stage embryos per class (total 900 embryos) clearly classified by embryologists as 0PN, 1PN, and 2PN. One-hundred embryos per class (a total of 300 embryos) were used for verification data. The verification data were evaluated for the performance of detection in the number of pronuclei by regarding the results consistent with the judgment of the embryologists as correct answers.

**Results:** The sensitivity rates of 0PN, 1PN, and 2PN were 99%, 82%, and 99%, respectively, and the overlapping 2PN being difficult to determine by microscopic observation alone could also be appropriately assessed.

**Conclusions:** This study enabled the establishment of the automated pronuclei determination system with the precision almost equivalent to highly skilled embryologists.

**KEYWORDS**

computer-aided diagnosis, deep learning technology, pronuclei determination, supervised learning, time-lapse images

## 1 | INTRODUCTION

We investigated the establishment of a complete continuous blastocyst culture system without exchange of culture medium by using a time-lapse device and a single-step culture medium in 2014. Consequently, the chi-square test showed that the rates of blastocyst development and good blastocyst development were both significantly higher ($P \leq .05$) in the complete continuous culture on day 5 of culture than in the case of direct microscopic observation and the medium replacement on days 3 and 5 of culture using a sequential culture medium; we confirmed the superiority of this complete continuous blastocyst culture system without observation and the exchange of culture medium, where no embryo was casually exposed to atmospheric oxygen during the culture period.[1]

As described above, the greatest advantages of a time-lapse incubator in combination with the time-lapse device and the single-step

medium are as follows: (a) the embryonic development state can be checked non-invasively at any time without taking out the embryo outside the incubator, and (b) the improvement in blastocyst development rate is expected, since the embryo can be observed without opening or closing the incubator and thus the inside of incubator can be kept in a specific condition (hypoxic environments are maintained); in addition, a good embryo can be selected without missing, because the embryo development process may be observed and/or recorded over times via the videos and the images recorded at regular intervals. On the other hand, there are problems: (a) it is difficult to apply this system to all patients uniformly due to high costs of installation and operation, and (b) since annotation of a large amount of the images by physicians or embryologists is beyond their capacity, all image data obtained might not be clinically utilized; as such, how to process the huge image data collected has been an emerging challenge.

In a "Report of a social gathering on promotion of AI utilization in the field of health care" issued by the Japanese Ministry of Health, Labour and Welfare on June 27, 2017, the results of discussion on the areas where AI should be utilized and on the securing of efficacy and safety in the field of health care are shown. This report shows that "individualized medicine" by utilizing the AI for genome analysis, etc is expected to be realized, because "machine learning" has dramatically evolved by deep learning technology since 2012. At the same time, as a specific example of utilization of the AI, the report refers to the possibility of reducing the workload of physicians when the AI interprets thousands to tens of thousands of images obtained by capsule endoscopy.[2]

After the establishment of the complete continuous blastocyst culture system in 2014, we aimed to develop an algorithm to automatically detect the number of pronuclei by analysis using deep learning technology based on images from an embryo from immediately after the insemination to the disappearance of the pronuclei and attempted to establish an automated pronuclei determination system that can withstand actual clinical use in infertility treatment setting.

## 2 | MATERIALS AND METHODS

### 2.1 | System concept

Deep learning is an excellent machine learning technology that leads to breakthroughs in various fields, which is also highly expected as an approach to solve problems that have ever been difficult to achieve in image recognition. As an example showing the strength of deep learning, there may be a description that the technology can automatically create a solution from the data.[3] However, such an example implies an investigative meaning to understand human intelligence and expand the applicable range in the future; therefore, a huge amount of data, large networks, and complex learning processes must be used in order to realize learning that creates a solution method, which requires considerable costs. We adopted

a system that explicitly incorporates human observation into deep learning as a concept to realize and continuously improve the detection of the number of pronuclei at a realistic cost while utilizing the strong performance of deep learning.

### 2.2 | Framework

The entire system framework is shown in Figure 1. The processing is roughly divided into four steps.

1. Preprocessing: Detection of an area of a fertilized oocyte
2. Main processing 1: Detection of an outline around pronuclei
3. Main processing 2: Determination of the number of pronuclei
4. Postprocessing: Integration of time-series information

Each process simulates a procedure which humans seem to almost automatically perform in observing the number of pronuclei; by dividing the processing steps, the contents of each function will be clarified and realized more easily. In addition, classic imaging techniques are used, including the circular Hough transform[4] for the step "1. Detection of an area of a fertilized oocyte" and the hidden Markov model[5] for the step "4. Integration of time-series information." These classic methods are easy to reflect human observation in setting parameters and conditions, which are suitable for use in steps that do not require advanced processing. The neural network learning through deep learning is used in the steps of "2. Detection of an outline around pronuclei" and "3. Determination of the number of pronuclei." There may also be a framework realized as one neural network that determines the number of pronuclei from the images without dividing steps; however, this framework is divided into two neural networks and each network learns independently in order to improve the performance by explicitly incorporating human observation where the outline around pronuclei is being observed in the determination of the number of pronuclei.

The deep learning software we used was Microsoft Cognitive Toolkit (CNTK), an open-source deep learning toolkit with brain script. We built the framework on the personal workstation (i7-3.4GHz, RAM64GB, GeForce GTX1080) by using development software including deep learning library (Matlab2016, CNTK version 2.3.). The neural networks are described by brain script of CNTK.

### 2.2.1 | Preprocessing: detection of an area of a fertilized oocyte

Humans focus only on the inside of a fertilized oocyte when observing the number of pronuclei. This is because there is no necessary information to determine the number of pronuclei on the outside of the fertilized oocyte. Similarly, it is easier to improve the final performance in deep learning if the unnecessary information is not included. In this system, the minimum range of information necessary for determining the number of pronuclei is defined as the inside of
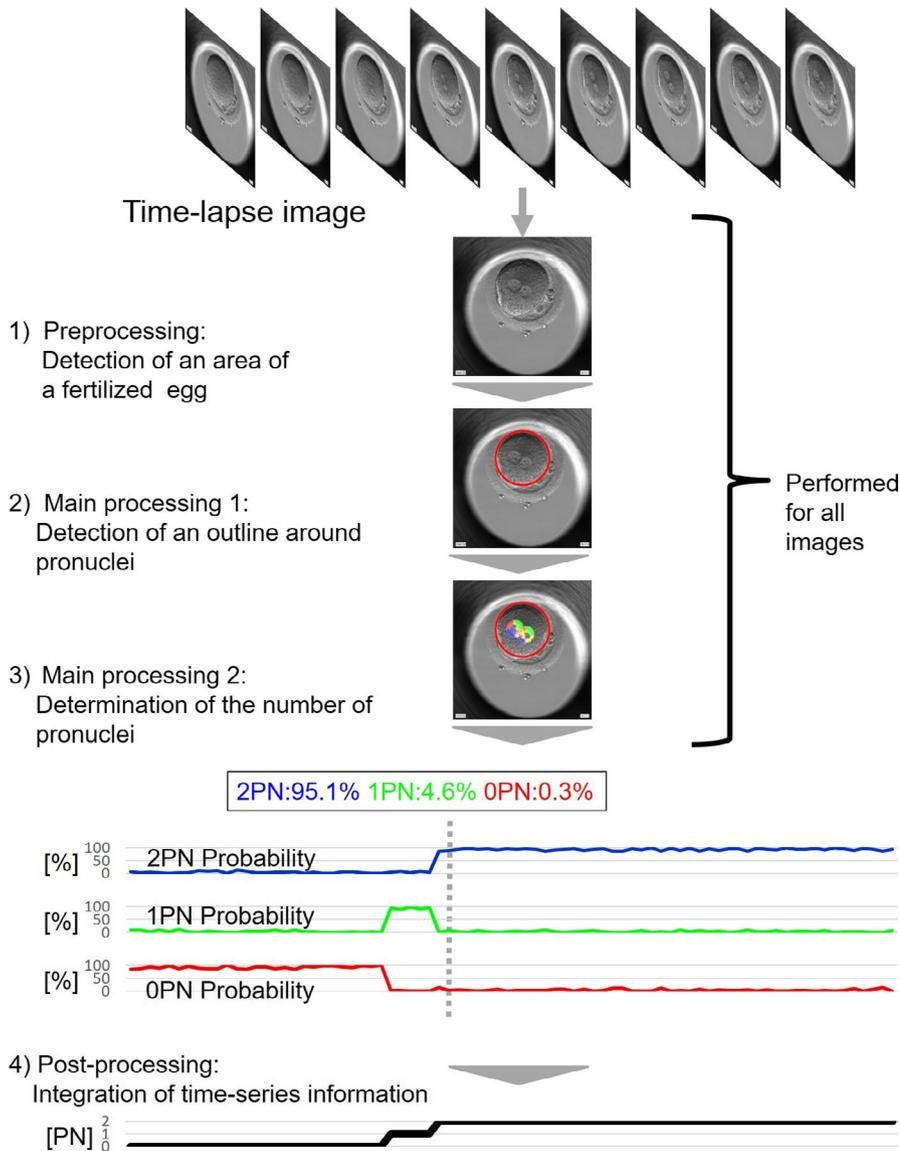
the zona pellucida of the fertilized oocyte, and the range is detected in the first processing. The fertilized oocyte is detected using the circular Hough transform, a technique to detect a circle because the shape of the oocyte in the images can be approximated to a true circle in many cases.

### 2.2.2 | Main processing

There are precedent cases in the study of counting objects using deep learning,[6,7] where it is common to perform the procedure of counting objects after their individual recognition. The pronuclei counted in this system have characteristics that several adjacent and overlapping pronuclei can be very often observed, as well as the pronuclei are transparent so that the other pronucleus can also be seen even if overlapping. "Individual identification of transparent and overlapping objects" is more challenging than general object recognition and may require a huge cost to determine the number of pronuclei with high precision in the traditional counting framework.

However, humans are possible to intuitively grasp the number without the procedure of clearly counting as many as five objects when counting them. As with the intuitive grasp of the number by humans, this system will determine the number of pronuclei without counting. Thus, the difficult challenge of the procedure of "individual identification of transparent and overlapping objects" is no longer needed to be performed, which allows us to grasp the number of pronuclei at a lower cost. Additionally, humans may intuitively determine the number of pronuclei based on the appearance such as the connection and shapes of the outline. This system also initially detects an outline around pronuclei in the same way and the number of pronuclei is determined from the outline.
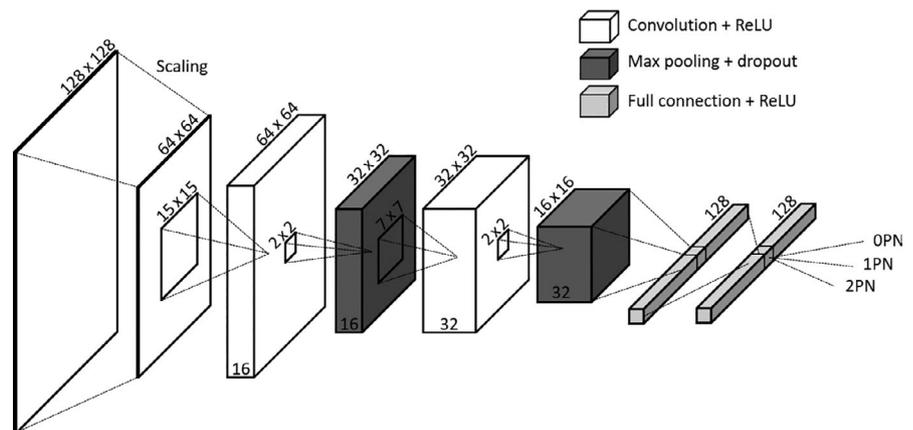
#### Detection of an outline around pronuclei

The neural network learning through deep learning is used in the detection of the outline. The model of the neural network consists of two convolution layers and two full-connection layers (Figure 2.) Classic imaging techniques that address similar tasks include edge detection filter and template matching. However, the granular

**FIGURE 2** Neural network for detecting outline of pronuclei. This software has two convolution layers and two full-connection layers with ReLU function as activation function
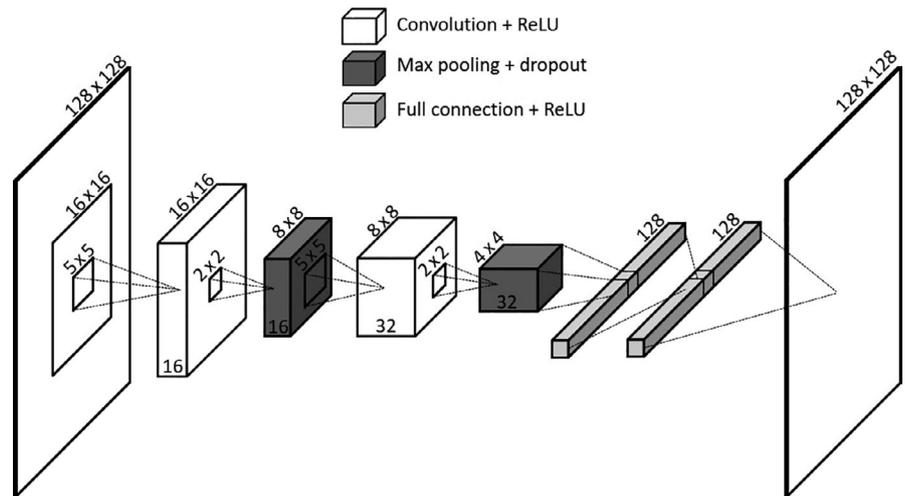


**FIGURE 3** Neural network for determining number of pronuclei. This software has two convolution layers and two full-connection layers with ReLU function as activation function



brightness changes are very intense within a fertilized oocyte, and the "edgy" structures are included such as polar bodies and vesicles (circles) in addition to the pronuclei; therefore, it is extremely difficult to detect the outline around pronuclei with high precision by combining simple processing such as the edge detection filter and the template matching. On the other hand, deep learning enables processing including cognitive functions, and thus, even in the "edgy" structures, a function to distinctively detect whether they are pronuclei or polar bodies can be realized. Furthermore, the connection of outline may be important information for the recognition of an overlapping outline. We can explicitly handle the information on how to connect the outline in the subsequent processing, "the determination of the number of pronuclei," by obtaining the information on which part of the outline around pronuclei is detected as well. In the initial analysis, the outline was divided into four parts in our system. However, as the greater the number of parts, the more precise the detection, the system was changed to detect the outline divided into 8 parts, considering that too many divisions require huge number of supervised data for neural network learning.

*Determination of the number of pronuclei*
While it is easy for humans to intuitively determine the number of pronuclei from the outline, it is difficult to clearly regularize the specific determination methods. However, the task of classification of images intuitively performed by humans has been actively studied as application of deep learning in image recognition, and many results have shown that the task is very compatible with deep learning.[8-11] In this system, the outline images are handled as a task classifying them into three classes of 0PN, 1PN, and 2PN, from which the number of pronuclei is determined. The actual outputs are the probability that the input outline images belong to each class, with three values corresponding to each class such as 0PN, 1%; 1PN, 4%; and 2PN, 95%. We built the model of the neural network with two convolution layers and two full-connection layers (Figure 3.)

### 2.2.3 | Postprocessing: integration of time-series information

In the visual observation by embryologists, the number of pronuclei may not be accurately determined only by the image at a certain time due to the development or disappearance of pronuclei over time and the overlapping of pronuclei. In this case, the decision is suspended at that time and made with the image suitable for the judgment at another time. However, the task of how much the image at a given time is suitable for determining the number of pronuclei for the image analysis

is much more challenging than that of determining the number of pronuclei, even if it is easy for humans. Thus, in this system, "the determination of the number of pronuclei" was carried out for all images of the time lapse, and the change over time in the number of pronuclei was to be determined from the probabilities of each class at all times.

We used 20 hours of culture as an endpoint for analysis, and the conditions for the change over time in the number of pronuclei were defined as follows:

1. 0PN (prior to pronuclear development) immediately after the start of culture.
2. The number of pronuclei observed up to 20 hours after the start of image acquisition with the time-lapse device was classified by each pronuclear class (0PN, 1PN, and 2PN). The number of pronuclei may increase during the culture, but will not decrease.
3. If the pronuclei disappear before 20 hours of culture, the disappearance period is very short.

In this system, the change over time in the optimal number of pronuclei meeting the above conditions is determined by the hidden Markov model, based on the analysis of the number of pronuclei at each time. Learning by incorporating the above explicit conditions is difficult for deep learning; however, since the hidden Markov model can be set to an arbitrary probability of the state change, the conditions may be expressed in the form of 0PN of 100% immediately after the start of culture and the change from 2PN to 1PN of 0% during the culture. The probability is defined as shown in Table 1. The transition probabilities to the next pronuclei stage were set to 0.01 (1/100) during the time-lapse observation. Since the number of pronuclei will not decrease over time, we set such fields to 0, and the remaining fields are set to make the sum of the numerical values of each row 1. These values showed the best accordance in the number of pronuclei evaluated by the embryologist with the one calculated by the computer.

The number of pronuclei at the final time obtained by the hidden Markov model is considered as the final number of pronuclei in the entire time lapse by finding arbitrary transition sequence by using the Viterbi algorithm.

## 2.3 | Supervised data set

For supervised data, an image processing engineer selected clearly classified images from the time-lapse images of 300 embryos per class (a total of 900 embryos) judged by embryologists as 0PN, 1PN, and 2PN from the pronuclear stage embryos generated by infertility treatment at the Asada Ladies Clinic.

In the neural net outputting the outline around pronuclei, the results manually annotated by the engineer to the outline features for the images with the clear outline around pronuclei were prepared as supervised data by reference to the number of pronuclei in the time lapse determined by the embryologists for the abovementioned 900 pronuclear stage embryos.

## 2.4 | Verification data set

From the pronuclear stage embryos produced by infertility treatment at the Asada Ladies Clinic, 100 embryos per class (a total of 300 embryos) judged as 0PN, 1PN, and 2PN by the embryologists were selected, and time-lapse images (approximately 70 images per embryo) taken from each embryo from immediately after the insemination to the disappearance of the pronuclei were used.

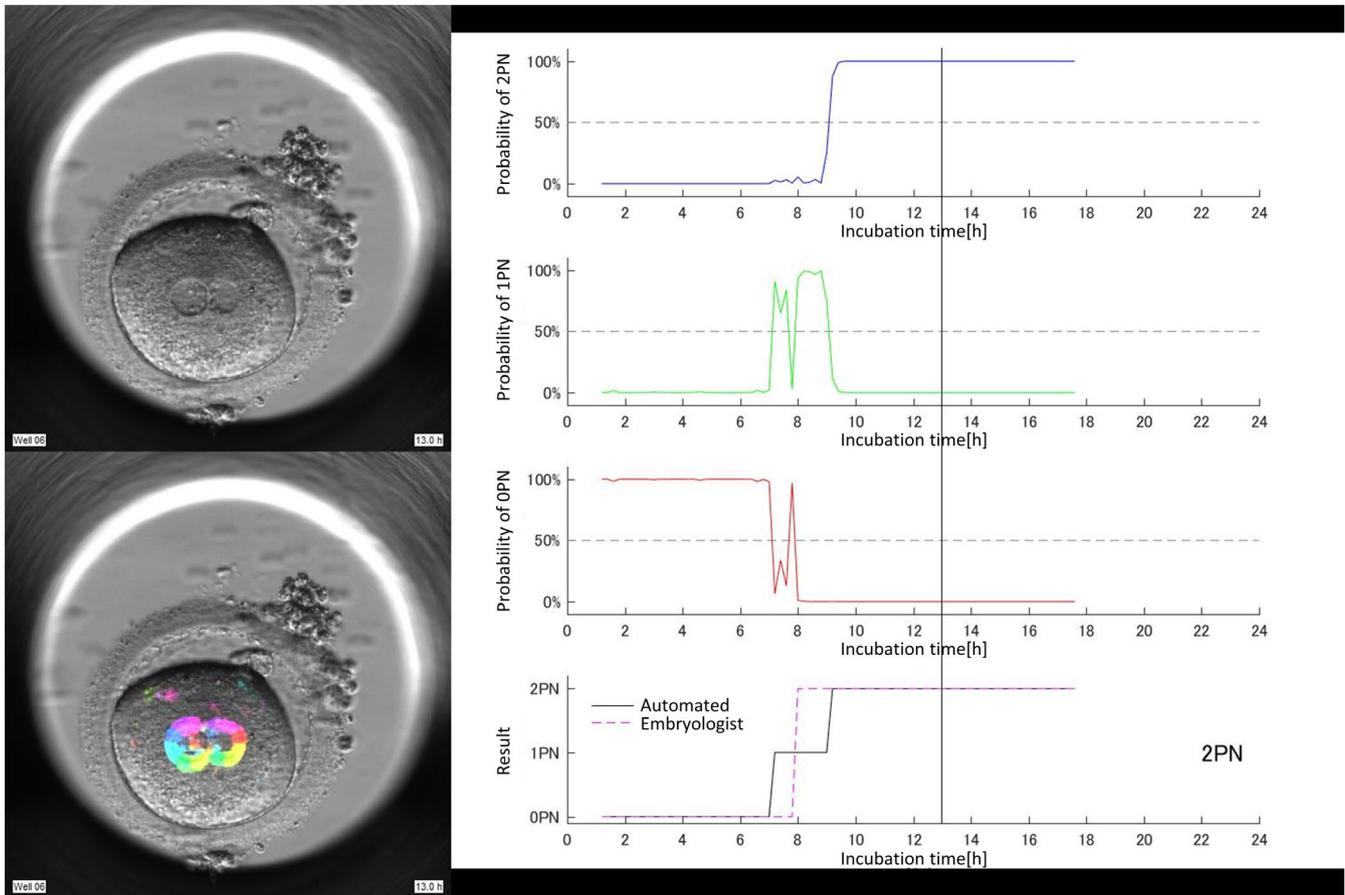## 2.5 | Applying the neural net to time-lapse images and evaluating the system

Approximately 70 time-lapse images of each embryo were input into the neural net after learning to determine the number of pronuclei at each time point and the maximum number of pronuclei detected from immediately after the insemination to the disappearance of the pronuclei, as well as to evaluate the performance of detection for the number of pronuclei by regarding the results consistent with the judgment of the embryologists as correct answers.

## 3 | RESULTS

A case of the automated pronuclei determination and the visual judgment by the embryologists over time is shown in Figure 4. The detected outline of pronuclei is visualized by colors (Figure 5.)

By combining the neural net detecting the outline around pronuclei with that detecting the number of pronuclei, the sensitivity rates of 0PN, 1PN, and 2PN in this system were 78%, 68%, and 97%, respectively (Table 2A).

The system incorporating the pronuclei outlining step obtained a high 2PN detection power of 97%, but with the aim of further improving performance, we changed the selection strategy for supervised data and re-performed deep learning using division numbers of the outline changed from 4 to 8 (improved system). As a result, the sensitivity rates of 0PN, 1PN, and 2PN increased to 99%, 82%, and 99%, respectively (Table 2B). The overlapping 2PN being difficult to determine by microscopic observation alone could also be appropriately assessed using this improved determination system (Figure 6). There is no relationship between the overlapping rate of pronuclei and determination

**TABLE 1** Transition probabilities

| | Next | | |
| --- | --- | --- | --- |
| | 0PN | 1PN | 2PN |
| Current | | | |
| 0PN | 0.98 | 0.01 | 0.01 |
| 1PN | 0.00 | 0.99 | 0.01 |
| 2PN | 0.00 | 0.00 | 1.00 |

**FIGURE 4** A case of the automated determination and the visual judgment by the embryologists over time. The probability of 0PN had been almost 100% (up to 7 h of incubation) before the pronuclear formation occurred, the probability of 1PN increased immediately after the start of the pronuclear formation (from 7 to 9 h of incubation), and the probability of 2PN increased to almost 100% after the pronuclei were clearly confirmed (after 9 h of incubation), indicating that the system could correctly detect the pronuclear formation at almost the same timing as the embryologists
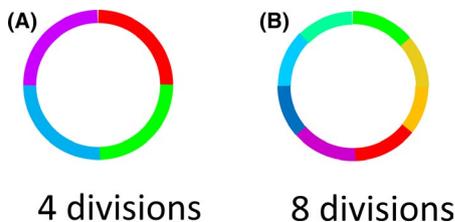


**FIGURE 5** Colors for outline parts. Each color indicates which part of the outline around pronuclei. Outline of pronuclei is divided into 4 or 8 parts and shown by specific colors

of number of pronuclei on our system (Figure 7.) The overlapping rate is defined as Equation 1, which $S_A$ and $S_B$ are areas of one of pronuclei.

$$\text{overlapping rate} = \frac{S_A \cap S_B}{S_A \cup S_B} \qquad (1)$$

The false-negative results are caused by the conditions of acquiring images in which the pronuclei are not observed clearly, such as fragment, out focusing, and embryo touching the wall.

**TABLE 2** Detection results of pronuclei

| A. Evaluating the contour as 4 divisions | | | | B. Evaluating the contour as 8 divisions | | | |
|---|---|---|---|---|---|---|---|
| | **Ground truth** | | | | **Ground truth** | | |
| | **0PN** | **1PN** | **2PN** | | **0PN** | **1PN** | **2PN** |
| Results from image analysis | | | | Results from image analysis | | | |
| 0PN | 78 | 12 | 0 | 0PN | 99 | 15 | 0 |
| 1PN | 18 | 68 | 3 | 1PN | 0 | 82 | 1 |
| 2PN | 4 | 20 | 97 | 2PN | 1 | 3 | 99 |

## 4 | DISCUSSION

### 4.1 | Preprocessing

As a nature of the circular Hough transform used for detection of an area of a fertilized oocyte, it is possible to detect the oocyte with some deformation or defect from a true circle, and it may exert high performance if the size and number of circles are known. Since the time-lapse imaging device can take images of one fertilized
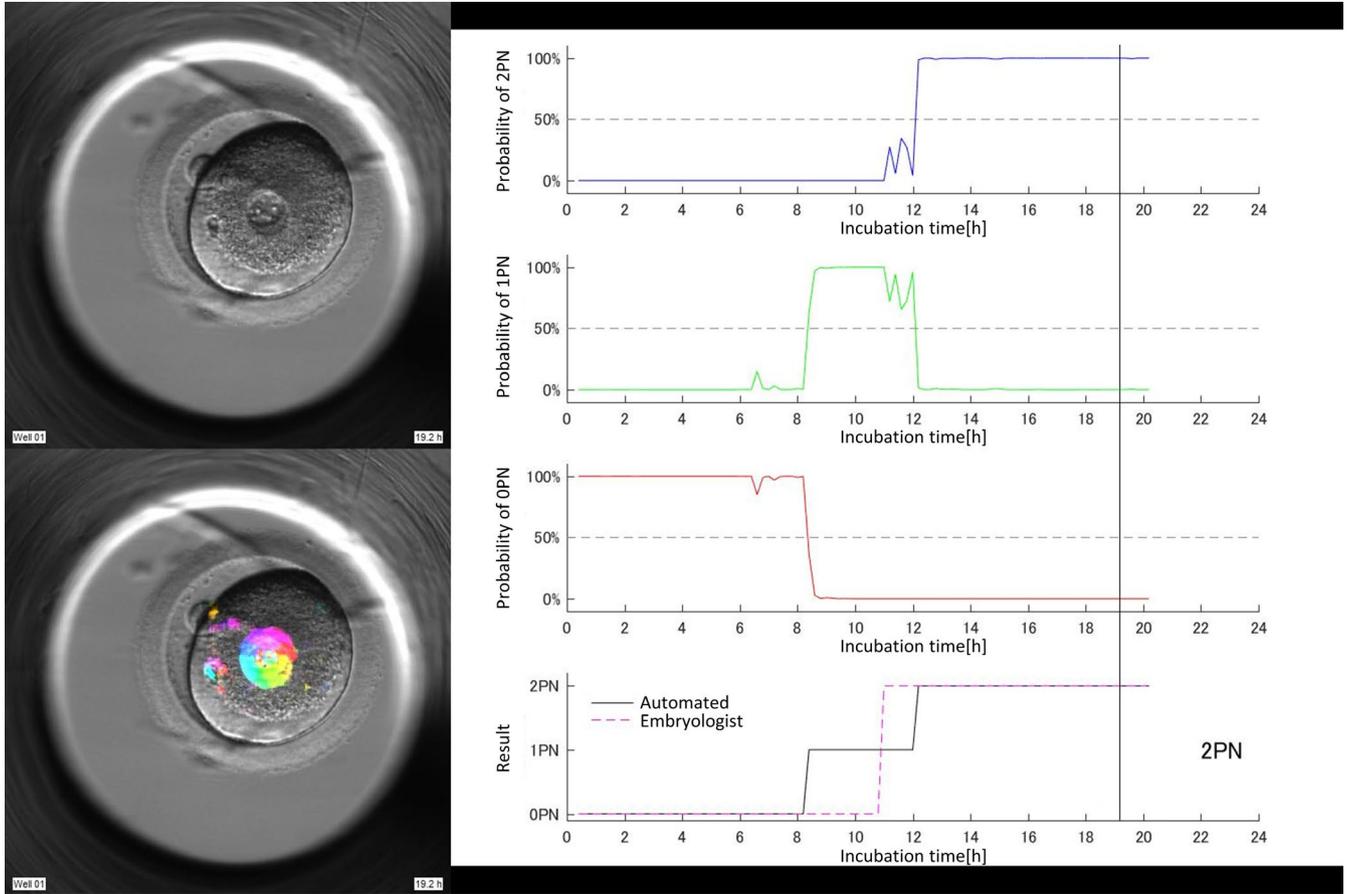
**FIGURE 6** A case of automated determination of overlapping 2PN with improved system and the visual judgment by the embryologists over time. Although most part of the pronuclei overlapped each other at 19 h of incubation, the probability of 2PN was almost 100%, indicating that the system could correctly determine them as 2PN
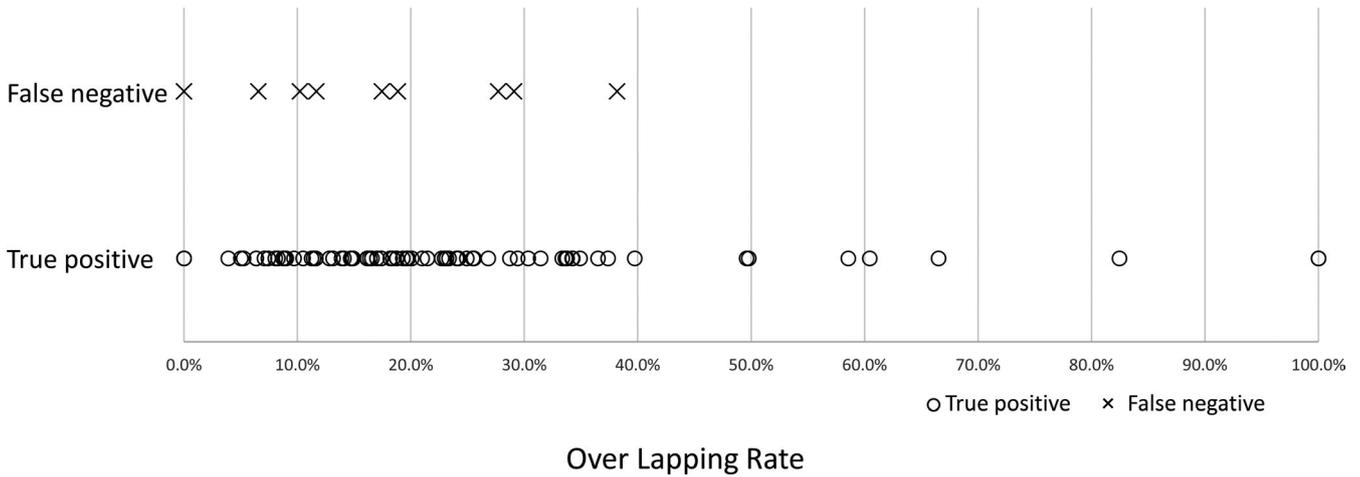


**Over Lapping Rate**

**FIGURE 7** Overlapping rate for 2PN case. Relationship between overlapping rate and accuracy (false negative or true positive) of 100 of 2PN embryos. This is four classification model of 0, 1, 2, and ≥3, which differs from the one we used in this paper

oocyte at a given magnification, any images can be detected as fertilized oocyte and detection with sufficient performance can be realized.

It seems also possible to detect the pronuclei by the circular Hough transform because they can be approximated to a true circle; however, this technique is not suitable for the detection of

the pronuclei since it is difficult to detect the correct position and number in case of several circles in proximity.

## 4.2 | Main processing

Deep learning focuses on learning to retrieve useful information to distinguish classes because of its nature. Without learning to explicitly extract the outline around pronuclei, the equivalent learning is expected to be automatically performed since the information of the pronuclei is surely valuable for determining their number. However, learning should find statistically valuable information for classification in deep learning. Consequently, a statistically sufficient amount of data is required for adequate learning. Furthermore, the smaller the constraints in a task, the larger the amount of data needed. The CIFAR-10,[12] which is currently used as a standard benchmark for classification tasks, is intended to learn from 5000 images per class. If the similar number of data were needed for learning in the determination of the number of pronuclei without any constraints, the data for 5000 embryos would have to be collected even for 1PN with low clinical incidence rates. This learning approach might be impossible to implement at a realistic cost because of enormous effort and time. However, learning with fewer data is possible for tasks to detect the outline and to determine the number of pronuclei from the outline, because conditions are added to each task and thus the constraints become greater. In this paper, although deep learning data of 300 embryos per class is small as supervised data, they indicate that appropriate learning is possible both in the detection of the outline and the determination of pronuclei from the outline.

## 4.3 | Postprocessing

Images of 2PN with overlapping pronuclei appear close to 1PN and may be output with a higher probability of 1PN than 2PN on the "the determination of the number of pronuclei." However, some images have high probability of 2PN because of changes in the position relationship between the pronuclei over the time-series observation. Due to the integration of multiple times in this system, the correct results can be determined even if the determination is temporarily inaccurate. In addition, it is possible to make an appropriate determination even if the overlapping of pronuclei occurs during the second half of culture since the condition that the number of pronuclei will not decrease is added to the constraints at the integration of time-series information.

## 4.4 | Evaluation of the system

While the embryo evaluation by embryologists is essentially a technique acquired by accumulated experience, the precision of machine learning by deep learning is close to that of the embryologists in a short period of time.

In this system, the combination of the neural net detecting the outline around pronuclei from the time-lapse images of an embryo with that detecting the number of pronuclei from the outline enabled the establishment of the automated pronuclei determination system with the precision almost equivalent to highly skilled embryologists. Reducing the oversight of the expression of pronuclei as low as possible would allow for the better evaluation of quality of the embryo itself beyond the conventional framework of morphologic evaluation based on images of an embryo. We will continue investigation to improve the precision of detection of pronuclei.

## ORCID

*Noritaka Fukunaga* https://orcid.org/0000-0003-2161-3319
*Hiroya Kitasaka* https://orcid.org/0000-0001-5850-2773
*Hiroyuki Watanabe* https://orcid.org/0000-0001-8744-9799
*Yoshimasa Asada* https://orcid.org/0000-0002-0448-9953

## REFERENCES

1. Fukunaga N, Kitasaka H, Yoshimura T, et al. Establishing a continuous blastocyst culture system without direct observation or exchange of culture medium using a live-embryo imaging system (Embryo Scope(TM)) (Japanese). *J Fertil Implant*. 2014;31(2):176-180.
2. Report of a social gathering on promotion of AI utilization in the field of health care. https://www.mhlw.go.jp/file/05-Shingikai-10601000-Daijinkanboukouseikagakuka-Kouseikagakuka/0000169230.pdf. Accessed October 30, 2019.
3. Le QV, Ranzato M, Monga R, et al.Building high-level features using large scale unsupervised learning. Proceedings of the 29th International Conference on Machine Learning (ICML-2012).
4. Ballard DH. Generalizing the Hough transform to detect arbitrary shapes. *Pattern Recogn*. 1981;13(2):111-122.
5. Viterbi A. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Trans Inf Theory*. 1967;13(2):260-269.
6. Segue S, Pujol O, Vitria J.Learning to count with deep object features. 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW).
7. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015;234-241.
8. Krizhevsky A, Sutskever I, HintonIlya GE.ImageNet classification with deep convolutional neural networks. Advances in Neural

Information Processing Systems 25 (NIPS 2012). https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks. Accessed October 30, 2019.

9. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. The 3rd International Conference on Learning Representations (ICLR 2015).

10. Szegedy C, Liu W, Jia Y, et al.Going deeper with convolutions. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

11. He K, Zhang X, Ren S, Sun J.Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

12. CIFAR-10 and CIFAR-100 datasets. https://www.cs.toronto.edu/~kriz/cifar.html. Accessed October 30, 2019.