International Journal of Computer Vision and Signal Processing, 8(1), 12-20(2018)

Ovarian Follicle Classification using Convolutional Neural Networks from Ultrasound Scanning Images

Manabu Nii*, Yusuke Kato, Masakazu Morimoto, Syoji Kobashi, Naotake Kamiura

Graduate School of Engineering, University of Hyogo, Himeji, Hyogo, Japan

Yutaka Hata

Graduate School of Simulation Studies, University of Hyogo, Kobe, Hyogo, Japan

Seturo Imawaki

Ishikawa Hospital, Hyogo, Japan

Tomomoto Ishikawa, Hidehiko Matsubayashi

Reproduction Clinic Osaka, Osaka, Japan

Abstract

In this paper, a new approach to classify ovarian follicles into two classes is proposed. A smoothing filter is applied for filtering ovarian follicle images. The smoothing filter is designed to consider speckle patterns under the resolution of the ultrasound devices Then, for extracting features from the filtered ovarian follicle images, two types of convolutional neural networks are utilized. One is the convolutional autoencoder, and the other is the layered convolutional neural network. Finally, both features extracted by the CNN-AE or the CNN from the filtered ovarian follicle images and numerical features defined by our previous works are used for classification. Several types of classifiers are examined in our experiments. From experimental results, we show the effectiveness of our proposed method. Especially, when image features extracted by the CNN and numerical features are both used, we have better classification performance than the other cases.

Keywords: ultrasound scanning, B-mode images, ovarian follicle classification, infertility, convolutional neural networks, machine learning

IJCVSP
International Journal of Computer Vision and Signal Processing

ISSN: 2186-1390 (Online)
http://cennser.org/IJCVSP

Article History:
Received: 24 October 2017
Revised: 17 August 2018
Accepted: 26 August 2018
Published Online: 28 August 2018

1. INTRODUCTION

Lower birth rate and population aging are serious problems in Japan as well as other countries [1]. Especially, the lower birth rate is a pressing issue which should be tackled. One of the reasons for the lower birth rate is the higher average marriage age. The late marriage may be a risk for infertility. In [2], the authors have reported that there were about 60% of causes of infertility (about 40 % of cases are due to female infertility and about 20 % of cases are due to both male and female infertility) on the female side. Since it is hard to improve the late marriage, therefore, we need to tackle the infertility problem to improve the birth rate in Japan.

In vitro fertilization is a typical medical treatment for infertility. First, in the treatment of in vitro fertilization, ova are extracted from the mother's ovaries and then extracted ova are fertilized outside of the mother's body. An embryo which has the best cell division will be returned to the mother's womb [3]. Therefore, in order to perform in vitro fertilization, ova have to be extracted from the mother's ovary. Presently, ultrasound scanning devices are used by medical doctors to find ovarian follicles from ovaries. Since ultrasound scanning is noninvasive, ultrasound devices are major tools in medical examinations.

*Corresponding author
Email addresses: nii@eng.u-hyogo.ac.jp (Manabu Nii), ei17h008@steng.u-hyogo.ac.jp (Yusuke Kato), morimoto@eng.u-hyogo.ac.jp (Masakazu Morimoto), kobashi@eng.u-hyogo.ac.jp (Syoji Kobashi), kamiura@eng.u-hyogo.ac.jp (Naotake Kamiura), hata@sim.u-hyogo.ac.jp (Yutaka Hata), imawaki@ishikawa-hp.or.jp (Seturo Imawaki), iskwtmmt@hotmail.com (Tomomoto Ishikawa), matsubayashi@me.com (Hidehiko Matsubayashi)

In ovarian follicle paracentesis operation, all ovarian follicles observed by using ultrasound scanning devices are extracted by medical doctors. Medical doctors can not distinguish between ovarian follicles with ova and without ova (called "vacuole") on the ultrasound B-mode image because ova do not appear in ultrasound B-mode image directly. Thus, all observed follicles are extracted at paracentesis operations. However, since paracentesis operations are invasive, subject persons are accompanied by pain from the operation. If medical doctors can find ovarian follicles with ova from ultrasound B-mode images, then the medical doctors are able to decide to do the paracentesis operation. Otherwise, unnecessary paracentesis operations need not do. Therefore, our purpose is to classify ovarian follicles into two classes (i.e., ovarian follicles with ova or vacuole) from the B-mode images. When ovarian follicles are able to be classified before paracentesis operations using a computer-aided system, only ovarian follicles with ova are extracted by paracentesis operations. For example, let us consider if the patient has many ovarian follicles but those follicles are vacuoles. Currently, all ovarian follicles are extracted by paracentesis operations because medical doctors could not distinguish ovarian follicles with ova and vacuole, however, they are unnecessary operations. On the other hand, if medical doctors can find ovarian follicles with ova before operations, the above mentioned unnecessary operations are able to be reduced.

To tackle the problem of infertility, several types of research have been reported [4, 5, 6, 7]. In [4, 5], stenosis detection method has been proposed for tubal infertility. By using an ultrasound measurement device, new features have been defined based on the difference between the minimum and maximum amplitude values. In [6], a visualization method has been proposed for measuring the diameter of the columnar soft matter. In [7], an identification method of ovarian follicles has been proposed. The proposed method is based on the brightness level relation between ovarian follicles with ova and vacuoles. This research shows the possibility of the ovarian follicle classification from ultrasound B-mode images.

Many ultrasound computer-aided diagnostic (CAD) systems have been proposed in the medical field. In recent years, the deep learning technology proposed by G. E. Hinton and R. R. Salakhutidinov [8] have been widely used in the image processing, speech recognition, and text processing. As a survey reported by Q. Huang, et al [9], several ultrasound CAD systems with deep learning technology have been successfully utilized in the major medical application which includes the breast lesion, the liver lesion, the fetal ultrasound standard plane detection, and the thyroid nodule, and also the carotid ultrasound image classification. In the ultrasound CAD systems with deep learning technology, the deep learning method is utilized to extract features from ultrasound images. While man-made features are used for the conventional ultrasound CAD systems, features extracted by the deep learning techniques are used for the newest ultrasound CAD systems. Such features extracted by the deep learning are sometimes more effective than the man-made features [9].

We have proposed a classification system which uses some kinds of ovarian follicle deformation information by contacting the ovarian collection needle in [10]. Six kinds of features were defined in [10]. Using these six kinds of features, a k-means clustering method had two clusters. One cluster consisted of eight ovarian follicles with ova and two vacuoles. The other cluster consisted of one ovarian follicle with ovum and five vacuoles. Eventually, 13 of 16 ovarian follicles in B-mode images from three patients were correctly classified. Although these six features are effective to classify ovarian follicles with ova or vacuoles, we can not obtain the feature values before paracentesis operations because these features were based on the deformation information by contacting the needle. Accordingly, features which were not based on the deformation information by the ovarian collection needle have been proposed in [11]. Eight features were proposed which are defined by using some frames of a B-mode moving image. To extract values for these eight features, 30 frames until just before the needle contacting were used. Therefore, since the ovarian collection needle does not touch to an ovarian follicle, these eight features were not based on the deformation information directly.

In this paper, we propose a novel method to classify ovarian follicles on ultrasound B-mode moving images. The proposed method utilizes both some features defined in our previous works and image features which are vectors extracted by convolutional neural networks (CNNs). In the proposed method, some ovarian follicle B-mode images are presented to a CNN for extracting image features. We also examine the effectiveness of some kinds of filters which are applied to B-mode images input to the CNN. The smoothing filter is designed to consider speckle patterns under the resolution of the ultrasound devices. Feature vectors related to image features from the trained CNNs and feature vectors which were defined in [11] are both used for classification. This paper is organized as follows. In the section 2, the previous work [10, 11] are explained. Then the proposed method is described in section 3. In section 4, we compare some results of the proposed method with the conventional one.

2. PREVIOUS WORKS

We define our problem as a two-class classification problem. Ovarian follicles on ultrasound B-mode images are classified into two classes; (1) ovarian follicles with ova, and (2) ovarian follicles without ova called "vacuoles". B-mode moving image sets obtained by medical ultrasound scanning devices are available for classifying ovarian follicles. Currently, the B-mode moving image sets are obtained at the time of paracentesis operations. To make the training data sets, medical doctors classified ovarian follicles into two classes whenever an ovarian follicle is extracted from

Table 1: Six features for identifying an ovarian follicle with an ovum.

No.	Name	Description
1	C	Circularity of outline of ovarian follicles.
2	$\theta_{1/\sqrt{2}}$	Angle formed between the two lines defined by the two points which have the same $D_{1/\sqrt{2}}$.
3	$\theta_{0.01}$	Angle formed between the two lines defined by the two points which have the same $D = 0.01$.
4	$D^{\mathrm{diff}}_{\mathrm{max}}$	the difference between the maximum distances in both FOP and FOC.
5	$\theta^{\mathrm{diff}}_{\theta_{1/\sqrt{2}}}$	difference between the $\theta^{\mathrm{FOP}}_{1/\sqrt{2}}$ and $\theta^{\mathrm{FOC}}_{1/\sqrt{2}}$ in both FOP and FOC.
6	$\theta^{\mathrm{diff}}_{\theta_{0.01}}$	difference between the $\theta^{\mathrm{FOP}}_{0.01}$ and $\theta^{\mathrm{FOC}}_{0.01}$ in both FOP and FOC.

ovaries. Actually, medical doctors extract several ovarian follicles in a single puncturing at paracentesis operation because of minimizing patient's pain by the paracentesis operation. Since medical doctors classified extracted ovarian follicles one by one, the paracentesis operations for making training data sets took a long time.

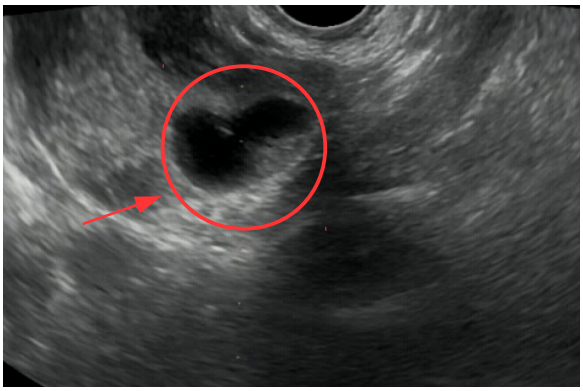
Our purpose is to classify ovarian follicles which appeared in the B-mode moving images into two classes. We have proposed several kinds of features to classify ovarian follicles in our previous works. First, the six features which are defined according to the deformation information by the ovarian collection needle have been proposed. In [10], we assumed that there are some differences in the amount of deformation between ovarian follicles with ova and vacuoles. Therefore, the frames, which are from the point of the ovarian collection needle contacted to the ovarian follicle until the point of the collection needle pierced the follicle, are considered. The six features were defined based on the above-mentioned frames of the B-mode moving images. That is, the six features were related to the deformation information by contacting an ovarian follicle collection needle. We defined the two kinds of frames as;

Frame of piercing(FOP); the frame contains the scene of the ovarian collection needle is piercing the follicle (see Fig. 1a).

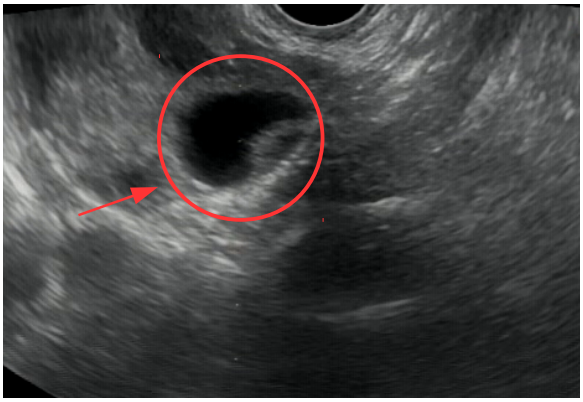
Frame of before contacting(FOC); the frame contains the scene of the ovarian collection needle located 1mm in front of the target ovarian follicle (see Fig. 1b).

Based on these two frames, six features shown in Table 1 were depicted in [10]. Since the six features were based on the deformation information by external force given by the ovarian collection needle, we could not avoid invasion on patients.

In [11], we have proposed eight features which were not based on the deformation information. The proposed eight features were extracted from frames which were before the ovarian collection needle is touching to the ovarian follicle. 30 frames until just before the ovarian collection needle contacting to the ovarian follicle were extracted. These



(a) Frame of piercing(FOP)



(b) Frame of before contacting(FOC)

Figure 1: Two kinds of frames of ultrasound images of an ovarian follicle.

frames have no direct influence from the ovarian collection needle. The target ovarian follicle is extracted from the 30 frames. The area S_i, circularity C_i, average of brightness B_{ave}, and variance of brightness B_{var} are calculated from the extracted ovarian follicle. Then, the following differences are calculated.

$$S^{\mathrm{diff}} = \max\{S_i\} - \min\{S_i\}, \quad i = 1, \ldots, 30, \tag{1}$$

$$C^{\mathrm{diff}} = \max\{C_i\} - \min\{C_i\}, \quad i = 1, \ldots, 30, \tag{2}$$

$$B^{\mathrm{diff}}_{\mathrm{ave}} = \max\{B_{\mathrm{ave},i}\} - \min\{B_{\mathrm{ave},i}\}, \quad i = 1, \ldots, 30, \tag{3}$$

$$B^{\mathrm{diff}}_{\mathrm{var}} = \max\{B_{\mathrm{var},i}\} - \min\{B_{\mathrm{var},i}\}, \quad i = 1, \ldots, 30, \tag{4}$$

where $i = 1$ means the frame of just before the ovarian collection needle contacting on the target ovarian follicle, $i = 30$ means the frame of one second before the needle contacting. Moreover differences of the area S_i, circularity C_i, average of brightness B_{ave}, and variance of brightness B_{var} for every adjacent frame are calculated. Therefore, the following features for every adjacent frame are defined. These eight features are summarized in Table 2.

$$S^{\mathrm{diff}}_{\mathrm{adj}} = \max\{S_{\mathrm{adj}}\} - \min\{S_{\mathrm{adj}}\}, \tag{5}$$

$$C^{\mathrm{diff}}_{\mathrm{adj}} = \max\{C_{\mathrm{adj}}\} - \min\{C_{\mathrm{adj}}\}, \tag{6}$$

$$B^{\mathrm{diff}}_{\mathrm{ave,adj}} = \max\{B_{\mathrm{adj,ave}}\} - \min\{B_{\mathrm{adj,ave}}\}, \tag{7}$$

Table 2: Proposed eight features for identifying an ovarian follicle with an ovum.

No.	Name	Description
1	S^{diff}	The area difference between $S^{\max}$ and $S^{\min}$.
2	$S^{\mathrm{diff}}_{\mathrm{adj}}$	The area difference between $S^{\max}_{\mathrm{adj}}$ and $S^{\min}_{\mathrm{adj}}$.
3	C^{diff}	The circularity difference between $C^{\max}$ and $C^{\min}$.
4	$C^{\mathrm{diff}}_{\mathrm{adj}}$	The circularity difference between $C^{\max}_{\mathrm{adj}}$ and $C^{\min}_{\mathrm{adj}}$.
5	$B^{\mathrm{diff}}_{\mathrm{ave}}$	Difference of average of brightness between $B^{\max}_{\mathrm{ave}}$ and $B^{\min}_{\mathrm{ave}}$.
6	$B^{\mathrm{diff}}_{\mathrm{ave,adj}}$	Difference of average of brightness between $B^{\max}_{\mathrm{adj,ave}}$ and $B^{\min}_{\mathrm{adj,ave}}$.
7	$B^{\mathrm{diff}}_{\mathrm{var}}$	Difference of variance of brightness between $B^{\max}_{\mathrm{var}}$ and $B^{\min}_{\mathrm{var}}$.
8	$B^{\mathrm{diff}}_{\mathrm{var,adj}}$	Difference of variance of brightness between $B^{\max}_{\mathrm{adj,var}}$ and $B^{\min}_{\mathrm{adj,var}}$.

$$B^{\mathrm{diff}}_{\mathrm{var,adj}} = \max\{B_{\mathrm{adj,var}}\} - \min\{B_{\mathrm{adj,var}}\}. \tag{8}$$

3. PROPOSED METHOD

In this paper, we consider classifying ovarian follicles into two classes by using the image of the target ovarian follicle in ultrasound B-mode images.

There are some ovarian follicles which appeared in each frame of the B-mode moving image. First, we need to extract the target ovarian follicle. Since we focus on to classify ovarian follicles, we extract ovarian follicle images by hand. Each extracted ovarian follicle image is a 256 × 256 pixel size.

In the extracted image, an exclusive part of the target ovarian follicle is considered as the background. Every pixel value on the background is set to 255, that is, the background color is considered as white).

3.1. Smoothing filter

To classify ovarian follicle images directly, a smoothing filter is applied to every B-mode image of the ovarian follicle. Speckles in ultrasound B-mode scanning images have been reported in [12]. Speckle patterns in some scanning motions (i.e, linear scan or sector scan) were investigated theoretically in [12]. In [12], the author described that the speckle in B-mode images obtained from the ultrasound scanning can be treated in a similar manner as laser speckle. As a case of utilizing speckle patterns in B-mode images in the medical area, speckle patterns were used to find parts of pathological change on the liver [13]. Speckles appeared in B-mode images do not show any structure of the target organ directly, however, speckles are created from the interference of echo signals by scatterers under the resolution of the ultrasound devices. Some papers have reported that, for example, speckles are almost distributed uniformly in the B-mode images from the normal liver. On the other hand, speckles are distributed disproportionally on the abnormal liver. In [13], the above-mentioned characteristics of the speckles have been used as a smoothing filter of the B-mode scanning images for the liver medical examination.

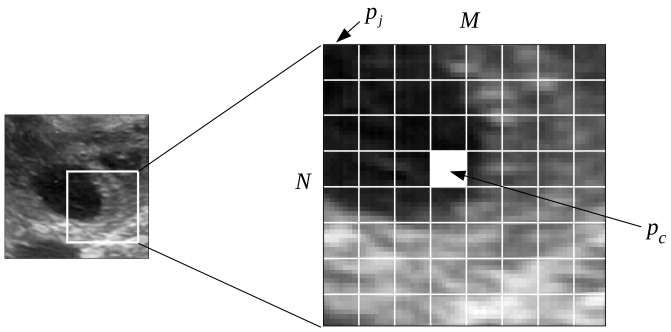


Figure 2: A region of interest (ROI).

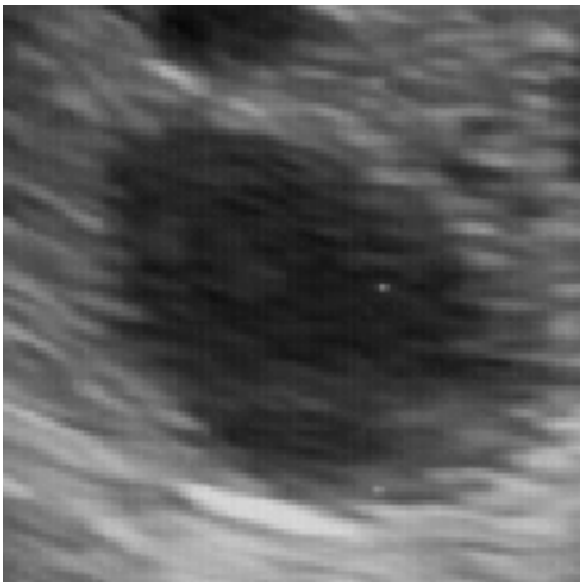
Although there is the difference between liver and ovaries, we consider that the speckle pattern analysis can be applied to our ovarian follicle classification problem. In our ovarian follicle classification, the difference between ovarian follicles with ova and vacuoles is the existence of ovum in the follicle.

An ovum is enough small, that is, its size is under the resolution of the medical ultrasound scanning devices. Therefore, we can not find it in the B-mode image. And also several parameters of the medical ultrasound scanning devices are optimized to find the shape of ovarian follicles. If speckle patterns are different between ovarian follicles with ova and vacuoles, we can consider that the images obtained by filtering have some differences.

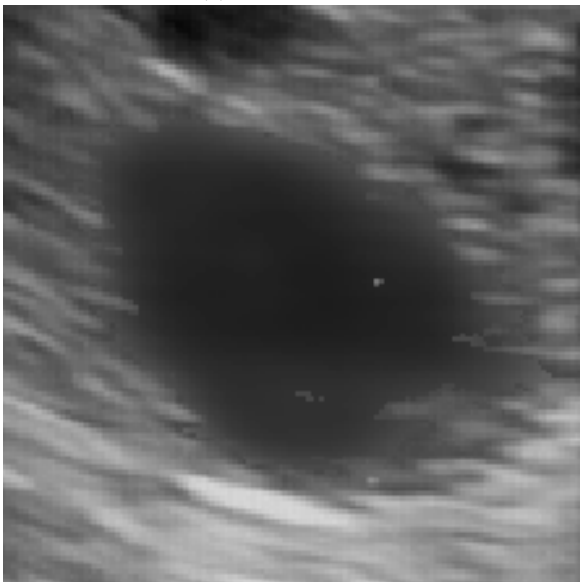
As reported in [13], the amplitude of echo signals by scatterers under the resolution of the ultrasound scanning devices is approximately represented as the Rayleigh distribution. Let us consider a $M \times N$ area in a B-mode image as the region of interest (ROI) as shown in Fig.2. p_{C} is the center of the ROI. For all pixels in the ROI, weights w_j $j = 1, 2, \ldots, M \times N$ are calculated as the following;

$$\begin{aligned} w_j &= \left(1 - \left(\frac{A_j - A_{\mathrm{C}}}{\alpha\sigma}\right)^2\right)^2, \\ &\text{if } 1 - \left(\frac{A_j - A_{\mathrm{C}}}{\alpha\sigma}\right)^2 < 0, \quad \text{then } w_j = 0, \end{aligned} \tag{9}$$

where A_{C} and A_j are amplitude of p_{C} and p_j, respectivly.



(a) Before filtering



(b) After filtering

Figure 3: Example images of before and after filtering.

σ is the standard deviation of the population, and α is an arbitrary defined coefficient. The new pixel value of p_{C} is calculated as the following;

$$p_{\mathrm{C}} = \frac{\sum w_j p_j}{\sum w_j}. \tag{10}$$

Figure 3 shows an example in before and after filtering for an ovarian follicle.

3.2. Two types of neural networks

We consider two types of convolutional neural networks [14] to extract image features from the filtered ultrasound B-mode images. One is a convolutional autoencoder and the other is a layered convolutional neural network.

3.2.1. Convolutional autoencoder

In order to extract image features, a convolutional autoencoder is utilized. The convolutional autoencoder is a kind of unsupervised learning neural networks. Autoencoders transform every input into lower dimensional representation by encoding layers. Then, the lower dimensional representation is reconstructed for the input by decoding layers. The training of autoencoders performs to adjust weights for realizing the above mentioned encode and decode processes. The convolutional autoencoder includes convolutional layers in encoding and decoding layers. Therefore, by using convolutional autoencoders, we aim to extract image features. The image features extracted by the convolutional autoencoders are features of ovarian follicles in ultrasound B-mode images, which are not related to ovarian follicles with ova or vacuoles.

3.2.2. Layered convolutional neural networks

While the convolutional autoencoders extract image features unrelated to the classes, a layered convolutional neural network is used for extracting features related to the difference between ovarian follicles with ova and vacuoles. The convolutional neural network is a supervised learning method. In this paper, the last layer (i.e., classification layer) is designed to classify ovarian follicles with ova or vacuoles. Therefore, the training data set consists of two classes. The convolutional neural network includes some convolutional layers and pooling layers. A softmax layer which is the last layer of the convolutional neural networks classifies feature vectors extracted by convolutional and pooling layers into two classes. As the training of the convolutional neural network is performed to classify ovarian follicles into two classes, the extracted features are expected to realize image features related to the ovarian follicles with ova or vacuoles.

3.3. Classification procedure

After every B-mode image is filtered by the above-mentioned filter, each filtered image is input to a convolutional autoencoder (CNN-AE) or a convolutional neural network (CNN). The CNN-AE or CNN extracts features from input filtered images to distinguish ovarian follicles with ova and vacuoles. Our CNN-AE and CNN architectures are shown in Fig.4 and Fig.5. Feature vectors at the last layer of the trained CNN-AE or CNN are able to be used as input vectors to many other machine learning classification techniques.

In this paper, the two-dimensional features which are proposed in our previous works (i.e., the area S_i and circularity C_i) of our previous method in [11] and the feature vectors from the trained CNN-AE or CNN are used as input vectors for several types of machine learning classification techniques.

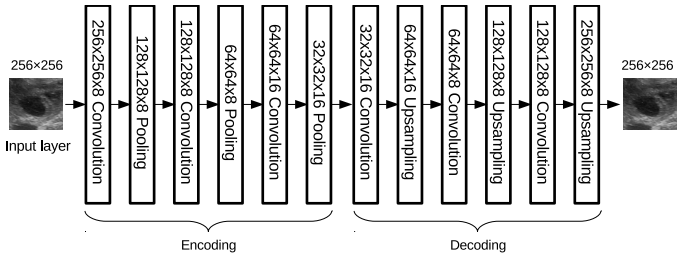


Figure 4: Our CNN-AE architecture.

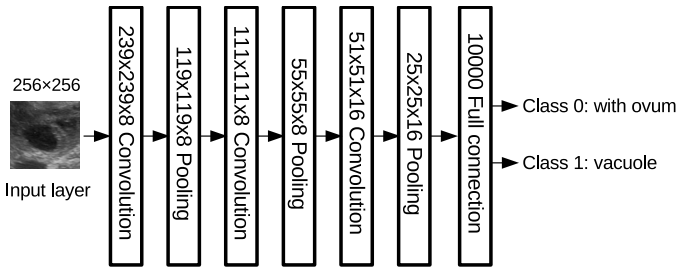


Figure 5: Our CNN architecture.

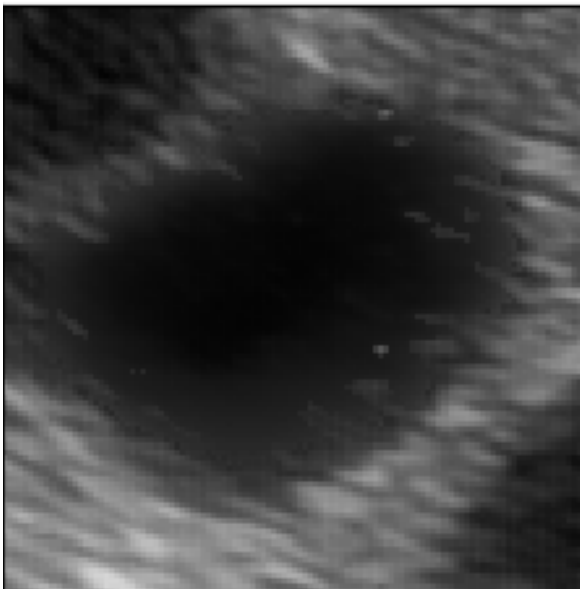
Table 3: Details of B-mode moving images used in this paper.

	Subject 1	Subject 2	Subject 3	Subject 4
Pixels	510×360			
Frame rate	30fps			
# of frams	3163	10384	2445	47522
Length	106 sec.	353 sec.	81 sec.	1584 sec.
# of extracted	3	17	2	53

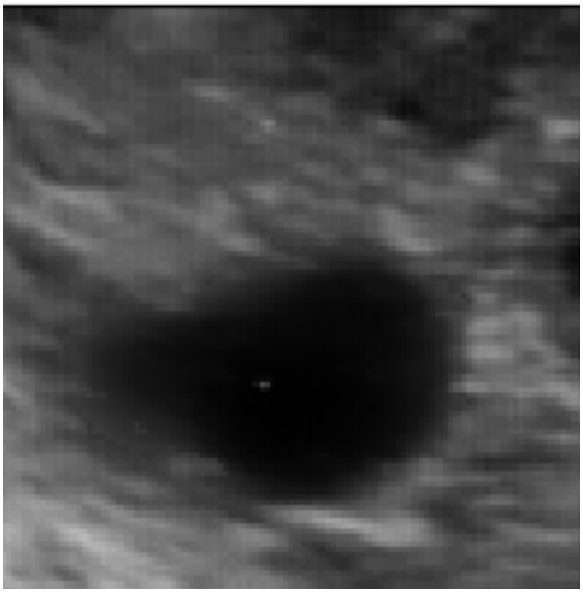
4. EXPERIMENTAL RESULTS

In this section, some experimental results are shown. The data set was provided from "Reproduction Clinic Osaka". Totally, 75 ovarian follicles were extracted from four patients. 32 of 75 ovarian follicles had ova. The other 43 ovarian follicles had no ovum, that is, they were vacuoles. The details of the data set are shown in Table 3. The presence of the ovum in every extracted ovarian follicle was examined by medical doctors under a microscope. Ultrasound moving images were obtained at the paracentesis operation for the four patients. The length of each ultrasound moving image is different because the number of extracted ovarian follicles was different between the four patients. Three of the 75 ovarian follicles in the ultrasound moving images were not enough quality to extract features. For example, the outline of the ovarian follicle was not clear. Or it appeared on just a few frames in the moving image. Therefore, 72 ovarian follicles were used in our experiments.

Figure 6 shows images which were filtered using the above-mentioned smoothing filter. These figures (both Fig.6 (a) and (b)) are smoother than the original B-mode images. As shown in Fig. 6, however, we can not detect the clear



(a) Ovarian follicle with an ovum



(b) Vacuole

Figure 6: Example images of after filtering.

difference between images of ovarian follicles with ova and vacuole by our eyes. Therefore, we need to extract some features from the B-mode images by CNN-AEs or CNNs, and then classify these B-mode images using extracted features. Our proposed method employs input vectors which are clipped from ultrasound B-mode images and filtered by a speckle-based smoothing filter. Moreover, the CNN-AEs and CNNs used in this paper are not so deep architecture. Therefore, our proposed method can be considered to be designed for the small number of data.

The leaving-one-out method was used for training in the following experiments. First, we examined the effectiveness of the smoothing filter which considers speckle patterns in the ultrasound B-mode images. Tables 4–6 show

Table 4: Classification results of CNN with filtered image.

	Predict		
	with an ovum	vacuole	Percent Correct
with an ovum	**28**	12	70.0%
vacuole	17	**15**	46.9%
Overall Percent	62.2%	55.6%	**59.7%**

Table 5: Classification results of CNN with original image.

	Predict		
	with an ovum	vacuole	Percent Correct
with an ovum	**29**	11	72.5%
vacuole	20	**12**	37.5%
Overall Percent	59.2%	52.2%	**56.9%**

Table 6: Classification results of our previous work (eight features).

	Predict		
	with an ovum	vacuole	Percent Correct
with an ovum	**30**	10	75.0%
vacuole	18	**14**	43.8%
Overall Percent	62.5%	58.3%	**61.1%**

the classification results. Table 4 shows CNN results with filtered images. Table 5 shows CNN results with original images. And Table 6 shows previous eight feature vectors [11]. From the Table 4, the classification result of CNN when filtered images were used is better than the case of using original images (Table 5). Unfortunately, however, both results (Tables 4 and 5) which used images only did not improve the classification performance of our previous works. The difference of the classification results is quite small between CNN results with filtered images (Table 4) and CNN results with original images (Table 5).

Both extracted features at the last layer of the CNN and the output vector at the encoding layer of CNN-AE can be used with two features (i.e., the area S_i and circularity C_i) of our previous method in [11] simultaneously. Since the other six features of the eight features in our previous work are redundant when images are used as features, we omitted these six features. We tried to classify using several classification algorithms; (1) AdaBoost, (2) decision trees, (3) k-nearest neighbors, and (4) random forests. The results are shown in Tables 7–14. In these tables, sensitivity, specificity, and accuracy are used as indicators of the effectiveness of our method. The best result was obtained by random forest algorithms with 70 trees and feature vectors by CNN-based method using the filtered image feature and the additional two numerical features (see Table 14). From

Table 7: Classification results of the CNN-AE in using non-filtered images.

Classifier	Sensitivity	Specificity	Accuracy
AdaBoost	0.625	0.344	0.500
Decision tree	0.500	0.375	0.444
k-NN ($k = 1$)	0.625	0.500	0.569
k-NN ($k = 3$)	0.600	0.594	**0.597**
k-NN ($k = 5$)	0.575	0.563	0.569
k-NN ($k = 7$)	0.575	0.500	0.542
k-NN ($k = 9$)	0.575	0.531	0.556
Rain forest (70 trees)	0.675	0.438	0.569
Rain forest (80 trees)	0.700	0.344	0.542
Rain forest (90 trees)	0.700	0.344	0.542
Rain forest (100 trees)	0.700	0.375	0.556

Table 8: Classification results of the CNN-AE in using conbined features (i.e., two numerical and non-filtered image features).

Classifier	Sensitivity	Specificity	Accuracy
AdaBoost	0.600	0.281	0.458
Decision tree	0.550	0.438	0.500
k-NN ($k = 1$)	0.450	0.563	0.500
k-NN ($k = 3$)	0.525	0.594	0.556
k-NN ($k = 5$)	0.500	0.563	0.528
k-NN ($k = 7$)	0.550	0.594	**0.569**
k-NN ($k = 9$)	0.550	0.563	0.556
Rain forest (70 trees)	0.625	0.313	0.486
Rain forest (80 trees)	0.600	0.281	0.458
Rain forest (90 trees)	0.575	0.281	0.444
Rain forest (100 trees)	0.600	0.281	0.458

Tables 7–10, unfortunately, the results from the CNN-AE based method are almost the same as our previous work. That is, we had only one case whose accuracy is the same as our previous work.

5. CONCLUSIONS

In this paper, we have proposed a new approach to classify ovarian follicles into two classes. We examine two types of convolutional neural networks; (1) convolutional autoencoders (CNN-AEs), and (2) layered convolutional neural networks (CNNs). The CNN-AEs were used to extract ovarian follicle image features from B-mode images. On the other hand, CNNs were used to extract image features related to the difference between ovarian follicles with ovum and vacuoles. For both CNN-AEs and CNNs, the smoothing filter which is designed to consider speckle patterns under the resolution of the ultrasound devices was used for filtering ovarian follicle B-mode images. Then, CNN-AEs and CNNs were used for extracting features from the filtered ovarian follicle images. Finally, we

Table 9: Classification results of the CNN-AE in using filtered images.

Classifier	Sensitivity	Specificity	Accuracy
AdaBoost	0.625	0.375	0.514
Decision tree	0.625	0.500	0.569
k-NN ($k = 1$)	0.575	0.438	0.514
k-NN ($k = 3$)	0.600	0.438	0.528
k-NN ($k = 5$)	0.600	0.406	0.514
k-NN ($k = 7$)	0.625	0.406	0.528
k-NN ($k = 9$)	0.625	0.438	0.542
Rain forest (70 trees)	0.750	0.406	0.597
Rain forest (80 trees)	0.775	0.344	0.583
Rain forest (90 trees)	0.775	0.406	**0.611**
Rain forest (100 trees)	0.775	0.375	0.597

Table 10: Classification results of the CNN-AE in using conbined features (i.e., two numerical and filtered image features).

Classifier	Sensitivity	Specificity	Accuracy
AdaBoost	0.675	0.406	0.556
Decision tree	0.600	0.594	**0.597**
k-NN ($k = 1$)	0.625	0.500	0.569
k-NN ($k = 3$)	0.625	0.563	**0.597**
k-NN ($k = 5$)	0.550	0.500	0.528
k-NN ($k = 7$)	0.575	0.563	0.569
k-NN ($k = 9$)	0.600	0.531	0.569
Rain forest (70 trees)	0.675	0.406	0.556
Rain forest (80 trees)	0.675	0.406	0.556
Rain forest (90 trees)	0.675	0.406	0.556
Rain forest (100 trees)	0.700	0.406	0.569

Table 11: Classification results of the CNN in using non-filtered images.

Classifier	Sensitivity	Specificity	Accuracy
AdaBoost	0.650	0.563	0.556
Decision tree	0.600	0.625	0.500
k-NN ($k = 1$)	0.650	0.438	0.611
k-NN ($k = 3$)	0.700	0.406	0.653
k-NN ($k = 5$)	0.725	0.438	0.653
k-NN ($k = 7$)	0.775	0.469	**0.667**
k-NN ($k = 9$)	0.750	0.438	**0.667**
Rain forest (70 trees)	0.825	0.594	0.639
Rain forest (80 trees)	0.825	0.594	0.639
Rain forest (90 trees)	0.825	0.594	0.639
Rain forest (100 trees)	0.825	0.594	0.639

Table 12: Classification results of the CNN in using conbined features (i.e., two numerical and non-filtered image features).

Classifier	Sensitivity	Specificity	Accuracy
AdaBoost	0.725	0.563	0.597
Decision tree	0.625	0.625	0.514
k-NN ($k = 1$)	0.650	0.438	0.611
k-NN ($k = 3$)	0.700	0.406	0.653
k-NN ($k = 5$)	0.725	0.438	0.653
k-NN ($k = 7$)	0.775	0.469	**0.667**
k-NN ($k = 9$)	0.750	0.438	**0.667**
Rain forest (70 trees)	0.800	0.563	0.639
Rain forest (80 trees)	0.850	0.563	**0.667**
Rain forest (90 trees)	0.850	0.563	**0.667**
Rain forest (100 trees)	0.825	0.563	0.653

proposed a classification method which used both features extracted by the CNN-AEs or the CNNs from the filtered ovarian follicle images and numerical features defined by our previous works (i.e., the area S_i and circularity C_i). From the experimental results, the extracted features by CNNs are better than those by CNN-AEs. We obtained better classification results when we used both features extracted by CNNs from the filtered ovarian follicle images and numerical features defined by our previous works than the other results. As shown in the experimental results, we obtained the best accuracy when two numerical and image features extracted by the CNN-based method from the filtered ovarian follicle images were used as inputs.

One of our future works is to examine the parameters of smoothing filters. The parameters were originally proposed for the liver diagnosis in [13]. We need to optimize them for ovarian follicles classification.

Currently, medical doctors cannot distinguish between ovarian follicles with ova and vacuoles from ultrasound CAD systems. On the other hand, about 68% of ovarian follicles were correctly classified from ultrasound B-mode images by our proposed classification method. Unfortunately, although the classification accuracy is not enough to use our classification system in the clinical situation, we believe that our proposed method has a possibility of classifying ovarian follicles at paracentesis operations.

Acknowledgment

This work was supported in part by a Grant-in-Aid from Hyogo COE Program Promotion Project and Terumo Foundation for Life Sciences and Arts.

References

[1] Ministry of Health, Labor and Welfare, "Annual Health, Labour and Welfare Report 2013-2014", `http://www.mhlw.go.jp/english/wp/wp-hw8/index.html`, accessed Apr. 13, 2016.

[2] F. Zegers-Hochschild, G. D. Adamson, J. de Mouzon, O. Ishihara, R .Mansour, K. Nygren, E. Sullivan, S. van der Poel, "The International Committee for Monitoring Assisted Reproductive Technology (ICMART) and the World Health Organization (WHO) Revised Glossary on ART Terminology," Human Reproduction Vol. 24, No. 11, pp. 2683–2687, 2009.

Table 13: Classification results of the CNN in using filtered images.

Classifier	Sensitivity	Specificity	Accuracy
AdaBoost	0.650	0.594	0.542
Decision tree	0.650	0.531	0.569
k-NN ($k = 1$)	0.625	0.375	0.625
k-NN ($k = 3$)	0.650	0.438	0.611
k-NN ($k = 5$)	0.650	0.438	0.611
k-NN ($k = 7$)	0.650	0.438	0.611
k-NN ($k = 9$)	0.650	0.469	0.597
Rain forest (70 trees)	0.800	0.563	0.639
Rain forest (80 trees)	0.800	0.531	**0.653**
Rain forest (90 trees)	0.800	0.563	0.639
Rain forest (100 trees)	0.800	0.563	0.639

Table 14: Classification results of the CNN in using conbined features (i.e., two numerical and filtered image features).

Classifier	Sensitivity	Specificity	Accuracy
AdaBoost	0.650	0.625	0.528
Decision tree	0.600	0.500	0.556
k-NN ($k = 1$)	0.625	0.375	0.625
k-NN ($k = 3$)	0.650	0.438	0.611
k-NN ($k = 5$)	0.650	0.438	0.611
k-NN ($k = 7$)	0.650	0.438	0.611
k-NN ($k = 9$)	0.650	0.469	0.597
Rain forest (70 trees)	0.875	0.563	**0.681**
Rain forest (80 trees)	0.825	0.594	0.639
Rain forest (90 trees)	0.850	0.563	0.667
Rain forest (100 trees)	0.850	0.563	0.667

[3] T. Hamatani, Y. Yoshimura, "Diagnosis, Treatment and Management of Gynecologic Diseases", Acta Obstetrica et Gynaecologica Japonica, Vol. 60, No. 12, pp. 495–504, 2008. (in Japanese)

[4] N. Kamiura, A. Emura, T. Yumoto, T. Isokawa, Y. Hata, T. Ishikawa, H. Matsubayashi, "On ultrasound measurement system estimating diameters of fallopian tube models", 2017 6th International Conference on Informatics, Electronics and Vision & 2017 7th International Symposium in Computational Medical and Health Technology (ICIEV-ISCMHT), pp. 1–6, doi: 10.1109/ICIEV.2017.8338604, 2017.

[5] A. Emura and N. Kamiura, "Stenosis detection in fallopian tubal model using ultrasonic measurement", 2017 6th International Conference on Informatics, Electronics and Vision & 2017 7th International Symposium in Computational Medical and Health Technology (ICIEV-ISCMHT), pp. 1–1, doi: 10.1109/ICIEV.2017.8338526, 2017.

[6] Y. Hata, Y. Takashima, K. Tsukuda, S. Kikuchi, T. Ishikawa, "Ultrasonic-Frequency-Based Visualization of Columnar Soft Matter Beyond Wavelength Limitations", IEEE Transactions on Systems, Man, and Cybernetics: Systems, Vol.48, No.2, pp. 224–231, doi: 10.1109/TSMC.2016.2598294, 2018.

[7] Y. Koya, T. Ishikawa, H. Matsubayashi, Y. Hata, "Identification of Ovarian Follicles for Infertility Treatment", Proceedings of the ISCIE International Symposium on Stochastic Systems Theory and its Applications, pp. 93–97, doi: 10.5687/sss.2017.93, 2017.

[8] G. E. Hinton and R. R. Salakhutdinov, "Reducing the Dimensionality of Data with Neural Networks", Science, 2006 Jul 28;313(5786):504-7.

[9] Q. Huang, F. Zhang, X. Li, "Machine Learning in Ultrasound Computer-Aided Diagnostic Systems: A Survey," BioMed Research International, Vol. 2018, Article ID 5137904, 10 pages, 2018.

[10] M. Nii, H. Kozakai, M. Morimoto, S. Kobashi, N. Kamiura, Y. Hata, S. Imawaki, T. Ishikawa, H. Matsubayashi, "Identification of ovarian follicle with ovum from ultrasonic images," Proc. of 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 1514–1518, 2016.

[11] M. Nii, R. Kashiwaki, M. Morimoto, S. Kobashi, N. Kamiura, Y. Hata, S. Imawaki, T. Ishikawa, H. Matsubayashi, "Ovarian Follicle Classification without Direct Deformation Information from Ultrasonic Images," International Journal of Biomedical Soft Computing and Human Science, Vol. 22, No. 1, pp. 19–28, 2017.

[12] C. B. Burckhardt, "Speckle in Ultrasound B-Mode Scans," IEEE Trans. on Sonics and Ultrasonics, Vol. SU-25, No. 1, pp. 1–6, 1978.

[13] N. Kamiyama, T. Yamaguchi, H. Hachiya, "Tissue Chracterization Using Statistical Information from Ultrasound Echo Signals," Medical Imaging Technology, Vol. 21, No. 2, pp. 112–116, 2003. (in Japanese)

[14] I. Goodfellow, Y. Bengio, A. Courville, "Deep Learning", MIT Press, `http://www.deeplearningbook.org`, 2016.