Follicle Detection on the USG Images to Support Determination of Polycystic Ovary Syndrome

2015 J. Phys.: Conf. Ser. 622 012027

(http://iopscience.iop.org/1742-6596/622/1/012027)

Follicle Detection on the USG Images to Support Determination of Polycystic Ovary Syndrome

Adiwijaya[1], B. Purnama[2], A. Hasyim[3], M. D. Septiani[4], U. N.Wisesty[5], W. Astuti[6]

School of Computing Telkom University, Bandung

[1]adiwijaya@telkomuniversity.ac.id, [2]bedy@telkomuniversity.ac.id,
[5]untarinw@telkomuniversity.ac.id, [6]astutiwidi@telkomuniversity.ac.id

Abstract. Polycystic Ovary Syndrome (PCOS) is the most common endocrine disorders affected to female in their reproductive cycle. This has gained the attention from married couple which affected by infertility. One of the diagnostic criteria considereded by the doctor is analysing manually the ovary USG image to detect the number and size of ovary's follicle. This analysis may affect low varibilites, reproducibility, and efficiency. To overcome this problems, automatic scheme is suggested to detect the follicle on USG image in supporting PCOS diagnosis. The first scheme is determining the initial homogeneous region which will be segmented into real follicle form The next scheme is selecting the appropriate regions to follicle criteria, then measuring the segmented region attribute as the follicle. The measurement remains the number and size that aimed at categorizing the image into the PCOS or non-PCOS. The method used is region growing which includes region-based and seed-based. To measure the follicle diameter, there will be the different method including stereology and euclidean distance. The most optimum system plan to detect PCO is by using region growing and by using euclidean distance on quantification of follicle.

Keywords : polycystic ovary syndrome, region growing, stereology, euclidean distance, follicle segmentation, follicle detection.

1. Introduction

Polycystic Ovary Syndrome (PCOS) is the most common endocrine disorders affected to female in their reproductive cycle. According to the criteria of National Institutes of Health (NIH), 6% - 10% female who has PCOS and Rotterdam can widely define as 15% prevalency of PCOS. PCOS has gained the attention from married couple when they acquire infertility. Yet, there have been 4 conferences in terms of identifying the definition and criteria of PCOS dignosis. The last conference was sponsored by the experts from National Institutes of Health (NIH) which suggested the existence of diagnostic criteria produced by Rotterdam conference [2]. One of the Rotterdam criteria in

diagnosing PCOS remains the ovary polycystic (PCO) [3]. Female can be said that they have PCO syndrome when there has been 12 or more a 2-9mm follicle or the increasing in ovary volume to more than 10 cm3 [3]. For now, to recognize that female is affected by polycystic ovary syndrome, doctor can manually calculate number of follicle on ovary and identifies its size on ultrasonographic (USG) image. In this research, doctor needs the thoroughness and high accuracy so there will not be low variability, reproducibility, and efficiency. Thus, there will be the system built automatically which can detect PCOS. Based on the ultrasonographic images as the input, this research can idetify PCOS considering to follicle morphology on ovary using NIH criteria so it is expected to assist the doctor in analyzing PCO accurately on PCOS patient using USG image.

The method used in identifying follicle object on USG image is algorithm approach of Region Growing. This aims at segmenting both structural characteristic follicle and geometric follicle comprehensively [4]. Moreover, Region Growing is known as follicle segmentation algorithm which has the highest follicle recognition about 78% and has a better accuracy than its competitor [4]. After segmentation process of follicle structure and existence, there will be stereology method and euclidean distance to identify geometric structures quantitatively to the follicle on ovary [5], so it can be used to detect PCO on ultrasonographic image. These approaches will be used on medical images for mobility and authentication[1]. In this paper, detection of a follicle is done automatically. Hence, it will be faster then conventional ones when the measurement carried out offline.

2. PCO detection system

The primary four processes in creating PCO detection system are identification of homogeneous regions, region growing, follicle extraction, and follicle quantification. The process can be described as follows:

1. The identification process of homogeneous region is begun when the user selects template 1 and template 2. The cropping size on template 1 is 330x389 in which the whole ovary is filled up. The template 2 has the cropped image which remains variety of size. The next process goes to determining threshold value (T1) from the cropped image by calculating the average value of greylevel in each pixel. This T1 threshold value will be the algorithm limitation of adaptive neighborhood median filter in filtering process. Lower intensity pixel of threshold T1 will be filtered using a 11x11 neighborhood, while higher one will be using 5x5 neighborhood. The 11x11 neighborhood is selected because a 50-60 pixel on USG image is approximately 10mm in a real.
   The twice-filtered images by adaptive neighborhood filter will determine threshold T2 value by calculating the average pixel intensity on smoothed image then it is subtructed by its pixel deviation standard. In threshold T2 value will be used by binary segmentation to create binary image in which homogeneous regions remain. It is conducted by determining lower pixel intensity of T2 as the white pixel while higher one as black pixel. The white pixel or the pixel which has white neighborhood is stated as the region.
   The previous procedure has produced number of unexpected homogeneous region since it is too small to become follicle region. This region will be deleted by considering to the criteria where the area of the region is less than threshold T3. It is stated by 50 which is predicted as well as a quarter of the detected minimum wide-region of follicle [4]. T1, T2, and T3 value are obtained from the earlier research in refference [4]. Furthermore, the region passed through the limit on the image, will be deleted considering to the artificial region in the edge of image. Therefore, this process can produce some homogeneous region.
2. The identified homogeneous region of the previous step can represent the prediction of initial follicle region and can determine the seed of region. Region growing step in this paper aims to compare both the identified seed and homogeneous region to enlarge its region for reaching the real

follicle limits. Each of homogeneous region will enlarge its region repeatedly using region's pixel outer limits.

a. Regarding to reffrence [4], there will be some criteria used to the region growing based on the region as follows:
   1) Grey-level criteria, by equation:

$$\left|I_k(p)-\mu(R_0)\right| \geq \alpha\sigma R_0 \quad (1)$$

   2) Edge criteria and texture information, by equation:

$$grad(p_i) = \left\|\nabla I_k(p)\right\|\left(\frac{c}{e^{tex(p)}-1}\right) \quad (2)$$

   3) Grey-level criteria using weighted gradient, by equation:

$$\left|grad(p)-\mu(grad(R_0))\right| \leq \left|grad(p)-\mu(grad(R_0))\right| \leq \alpha\sigma\mu(grad(R_0)) \quad (3)$$

   The refference [4] describes that the pixel's outer limits of region filled up the two criteria (i and iii) is marked as the potential candidate which may be gathered with its region. It will be conducted either when there has been four neighboring pixels of connected environment in their region which fulfilled those criteria or the pixel marked as potential candidate. The connected environment is observed using 2 type including 4-neighborhood and 8-neighborhood. However, if the pixel is not marked by potential candidate or does not comply the integrated criteria, so it can not enlarge its region. The repeated process from the outer pixels of region will be stopped since there is no one which fills up the criteria of integrated pixel and region. Therefore, it will be added to segmented region image.

b. Seed-based region growing uses six criteria to enlarge homogeneous region as follows:

K1: *if* $S_{q(x)}+i>0$ *and* $S_{q(x)}+i \leq Row$ (4)

K2: *if* $S_{q(y)}+j>0$ *and* $S_{q0}+j \leq Col$ (5)

K3:*if* $\sqrt{S_{q(x)}+i-(S_{q(x)})^2} < MaxDist$ (6)

K4:*if* $\sqrt{S_{q(y)}+j-(S_{q(y)})^2} < MaxDist$ (7)

K5: *if* $I(S_{q(x)}+i, S_{q(y)}+j) \leq (\mathrm{Re}\,gVal + ThresVal)$ (8)

K6: *if* $I(S_{q(x)}+i, S_{q(y)}+j) \geq (\mathrm{Re}\,gVal - ThresVal)$ (9)

Based on the reference [7], Row is the number of the original image lines, Col is the number of the original image coloums, MaxDist is the maximum distance to the initial *seed* with infinite value, RegVal is the value of graylevel pixel which is the initial *seed (centroid),* and an empirical value that will be observed. If it meets those six criteria, the pixel will be marked as follicle region.

3. Follicle extraction stage is the stage of final decision taken on each *region* which has been expanding its territory, whether it is follicle or not. Each *region* in the segmented image has area, which is the number of pixels bounded by a line that forms a homogeneous space. *Region* that has area of less than 220 pixels will be deleted. The *threshold* is obtained from a minimum size of follicle (2-3mm dimeter) [4]. Each *region* will calculate its ratio between the width of *the bounding box* and the area. *Region* that has a ratio of more than 0.5 will be marked as a *region* in accordance with the follicle, but if the ratio is less than 0.5, the *region* will be deleted and not marked as the actual follicle [5].

4. In the process of quantifying the quantification of follicles will be conducted to obtain the feature vector from the image that has been recognized as the follicles in the previous stage. The information will be used to classify follicles on ovary whether it is PCOS or not. Information of structural and geometric attributes taken from the follicles in the image is the number and the size of follicles. In this paper, the process of follicle quantification uses two approaches, namely stereology and *euclidean distance.* Both of these methods will later be compared to find out which one is better in measuring the diameter of the detected follicles so later it can be determined whether a follicle belongs to PCO or not. In process of stereology, each recognized follicle in the image will be determined its pixel area. Then, the recognized pixel area will be calculated in order to determine the number and size of each follicle in the image using the equation (10):

$$\frac{\sqrt{4\times\frac{Area}{\pi}}}{treshold} \tag{10}$$

where *Area* is the number of pixels that forms the follicle region *(region)*. *Threshold* is obtained from the size of pixels per millimeter, usually 50-60 pixels in the image of USG ovary has a distance of 1cm in its original size. It means that every millimeter has 5 to 6 pixels, so the value of *threshold* that will be used in order to measure the follicles per millimeter is 5,5piksel / mm [4].

Meanwhile, *euclidean distance* method is used to measure the diameter of follicles. *Euclidean distance* is a calculation method that is most often used to calculate the similarity of two vectors [6]. *Euclidean distance* calculating the squared root of the difference of two vectors. In this research, a vector is a follicle coordinate. The formula for calculating the diameter of follicles and *euclidean distance* uses the equation (11).

$$D=\sqrt{\sum_{t=1}^{n}(x_{it}-x_{jt})^2+(y_{it}-y_{jt})^2} \tag{11}$$

where,

D = diameter of the follicle (pixel),
x_i, x_j = coordinate of axis *x* follicle,
y_i, y_j = coordinate of axis *y* follicle,
t = the number of points.

In this test, the coordinate of axis *x* follicle and axis *y* follicle is farthest pixel of the axis x and y of a follicle. The farthest pixel can be found by using the function of MajorAxisLength and MinorAxisLength which are already contained in the Matlab program.

3. Testing and Analysis

In this testing, a total of 12 images of USG Ovary is obtained from Clinic of Permata Bunda Syari'ah, Cirebon (Indonesia). Each image contains two images, the left ovary and the right ovary. We use a total of 12 images of USG Ovary on the testing process. It was felt sufficient the amount of data for testing in this research, because this research is focused on the size and number of follicles were detected, not a person identification PCO syndrome or not. The number and size of follicles of all images have been validated by Dr. Zulkifli Ahmad, SpOG., K.FER., M.Kes as Female Fertility Consultant in Clinic of Permata Bunda Syari'ah. Testing the system is conducted by performing simulations on some pre-defined scenarios. Testing the success of estimating the size and number of

follicles is conducted by calculating the accurate success rate of follicle number detection (RR), the error rate of follicle size (MR), and the error rate of detection (FR).

3.1. Scenario Parameter Values of Alpha and Thresval on Region Growing

For area-based region growing, it is used the empirical values of variable alpha in equation (1), (2) and (3). Experiment A is conducted by changing the value of variable (alpha = 1, or alpha = 1.5, or alpha = 2). The value of observed threshold is assumed that the distribution of pixel intensity and weighted gradient in the image is classified as normal distribution. Therefore, the observed value of alpha is derived from the fixed normal distribution curve by selecting the interval 66% - 95% of the data.

Meanwhile, for the seed-based region growing it is used the empirical value on variable ThresVal in equation (8) and (9). Experiment B is conducted by changing the value of variable (Thresval = 0.025, or Thresval = 0.050, or Thresval = 0075). Empirical values of alpha and ThresVal provide different performance measured from the average success of detection (RR).

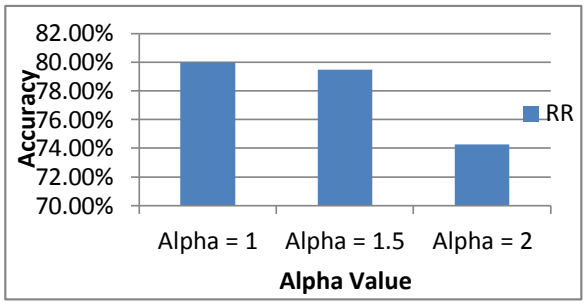

Figure 1. Success Rate Detection Accuracy Testing Follicles (RR) in Experiment A

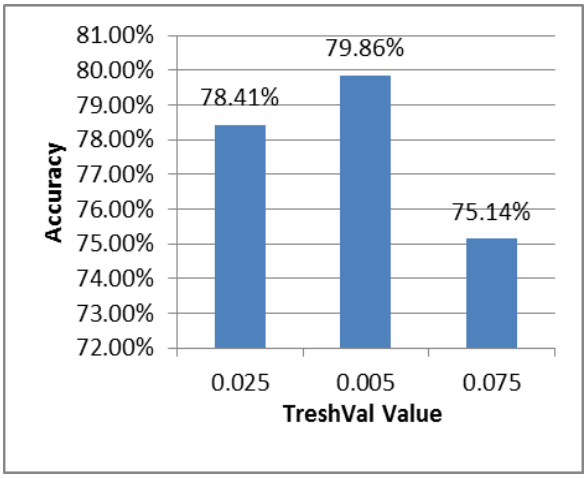

Figure 2. Succes Rate Detection Accuracy Testing Follicles (RR) in Experiment B

Figure 1 shows the average success rate of the detection on alpha equals 1 is 80%, alpha equals 1.5 is 79.5%, and alpha equals 2 is 74.25%. While in Figure 2, it can be seen that the average success rate of detection on thresVal Value equals 0.025 is 78.41%, thresVal Value equals 0050 is 79.86%, and thresVal Value equals 0075 is 75.14%.

3.2. Scenario Size of Template Image

In this scenario, the test is conducted with the configuration on the size of the template image that is used. Test A is combined with the region growing parameter (alpha = 1, or alpha = 1.5, or alpha = 2 and connectivity pixel 4- neighborhood or 8 -neighborhood). However, test B is combined with the parameter region growing (ThresVal = 0.025, or ThresVal = 0.050, or ThresVal = 0.075 and connectivity pixel 4- neighborhood or 8 -neighborhood).

Testing in this scenario aims to determine the initial estimation of the region of optimal follicles that can affect the region growing method in determining the actual follicles. Based on the testing that has been conducted, the performance in this scenario is affected by the size of the template image that is used. The size of the template used is the template 1 and template 2 that give different performance measured from the average error rate of detection (FR).

Based on Figure 3, it can be seen that the error level rate on template 1 is 17.67% and template 2 is 6.17%. Meanwhile, from figure 4, it can be seen that the error rate of template 1 is 16.36% and template 2 is 15.93%.

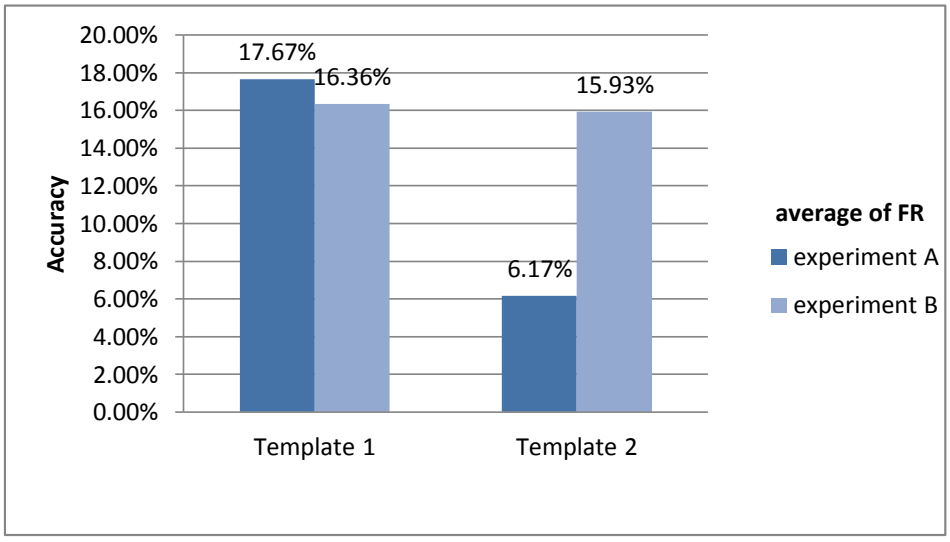

Figure 3. Detection Error Rate (FR) based on Template of Experiment A and B

Based on the experiment A and B, it can be seen that the error rate (FR) of template 2 is lower than FR of template 1. It is because template 2 only takes the most essential image part. The most essential image part is the part that contains follicles reffered by specialists that makes speckle noise at template 2 lesser than template 1 that captures the complete image of a single ovarium that results in some findings of the other structures at ovarium, such as blood vessel, lymph node and dendrite.

3.3. Scenario of Pixel Connectivity Type

This scenario is conducted by comparing the region of pixel connectivity type of 4-neighborhood and 8-neighborhood. The pixel used 4-neighborhood connectivity is classified as a part of an object if the neighborhood along the horizontal and vertical line fits with the criteria, while 8-neighborhood adds neighborhood on the diagonal line. Observation is conducted by changing the pixel connectivity type. On this scenario, the observation is conducted towards the most appropriate pixel connectivity size on the region growing method as to extend the region. Type of connectivity is proportional to the

similarity between pixels in each template. But inversely proportional to the long process of computation.

Experiment A is combined with the parameter of region growing (alpha =1, or alpha=1.5, or alpha =2, and 4-neighborhood or 8-neighborhood pixel connectivity) and the parameter of image template size (template 1 or template 2). Meanwhile, experiment B is combined with the parameter of region growing (ThresVal =0.025, or ThresVal =0.050, or ThresVal =0.075 and pixel connectivity of 4-neighborhood or 8-neighborhood). Each experiment is conducted by using two quantification methods, where the measuring methods used are stereologi and euclidean distance.

The experiment in this scenario is aimed for determining the pixel connectivity region in the most optimum region growing method. The determination of pixel connectivity region can influence the region growing method used to determine the outest region of pixel, whether the pixel is connected to the region or not so that it can influence the stereology method in the measurement of the follicle region. The second aim of the scenario of experiment is to find the follicle quantification method that has the lowest error level of follicle measurement.

Based on the experiment, the performance in the scenario is influenced by the pixel connectivity type and the quantification method used. The pixel connectivity types used are 4-neighborhood and 8 neighborhood and the follicle quantification methods used are stereology and euclidean distance that give different performance measured from the error level of follicle detection (MR).

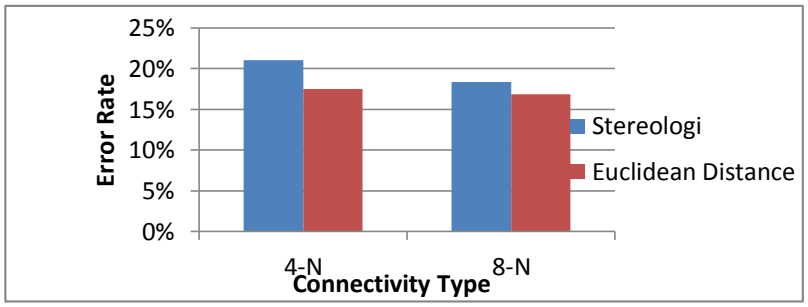

Figure 4. Follicle Size Error Rate Testing (MR) Based on the Type of Pixels Connectivity using Stereologi Method and Euclidean Distance in Experiment A

Based on figure 4, it can be seen that the error rate of follicle detection (MR) of the 4-neighborhood connectity using stereology is 21% and the one of 8-neighborhood is 18.33%, meanwhile that of the 4-neighborhood connectivity using euclidean distance is 17.50% and that of 8-neighborhood is 16.83%.

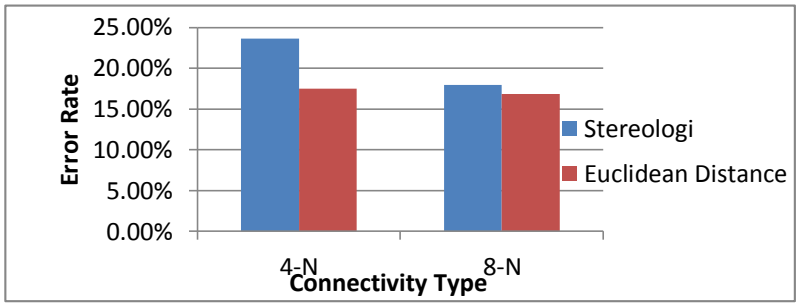

Figure 5. Follicle Size error Rate Testing (MR) Based on the Type of Pixels Connectivity using Stereologi Method and Euclidean Distance in Experiment B

Figure 5 provide the error rate of follicle detection (MR) on the 4-neighborhood connectivity using stereology is 23.62% and the one of 8-neighborhood is 17.59%; while that of 4-neighborhood connectivity using euclidean distance method is 17.50% and that of 8-neighborhood is 16.83%. Based on test A and B, it can be seen that the rate of MR shows that 8-neighborhood connectivity is the one that has lower error level on follicle detection compared to stereology. It is because 8-neighborhood has more neighbors that touch each pixels, so that it influences the number of pixels found on the follicle border that causes the follicle region wider resembling the real shape of the follicle. Euclidean distance method has a better performance compared to the stereology since euclidean distance conducts the follicle quantification referred to the farthest diameter. While the stereology conducts the follicle quantification through diameter equivalen. Euclidean distance method is in accordance with the validation conducted by specialists so that the performance of euclidean distance is better than that of stereology.

4. Conclusion

Based on the analysis towards the test on the PCO detection system of ultrasonographic image, it can be concluded that the region growing and euclidean distance methods can be implemented for detecting PCO. In the experiment A, the system configuration scenario that generates the most optimum performance is resulted from template 2, 8-neighborhood connectivity, and the threshold of alpha of 2 while the most optimum quantification method is euclidean distance. The testing scenario, when the performance is compared to the validation of specialists, results in success level of follicle detection (RR) of 78%, error level of follicle size (MR) of 17%, and error level of follicle detection (FR) of 12%. In the experiment B, the system configuration scenario that generates the most optimum performance is resulted from template size 2, 8-neighborhood connectivity, and the threshold of thresVal of 0.050 while the most optimum quantification method is euclidean distance. The testing scenario, when the performance is compared to the validation of specialists, results in success level of follicle detection (RR) of 76%, error level of follicle size (MR) of 20%, and error level of follicle detection 9FR) of 9%. Based on the second experiment, it can be concluded that the most optimum system to detect PCO is by using region growing and by using euclidean distance on quantification of follicle. It can be seen from the success level of follicle detection (RR), the error level of follicle size 9MR) and error level of follicle detection (FR).

References

[1] Adiwijaya, T.A.B. Wirayuda, S.D. Winanjuar, U. Muslimah, *The Multiple Watermarking on Digital Medical Image for Mobility and Authenticity*, Operations Research Proceedings 2012, 457-462

[2] National Institutes of Health, *Evidence-based Methodology*, Final report of Workshop on Polycystic Ovary Syndrome. 2012 December 3–5.

[3] M.J. Lawrence, M.G. Eramian, R.A. Pierson, E. Neufeld, *Computer Assisted Detection of Polycystic Ovarian Morphology in Ultrasound Images*, Proceedings of the 4th Canadian Conference on Computer and Robot Vision. 2007, IEEE Computer Society.

[4] Rotterdam ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group. *Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome*. Fertil Steril. 2004 Jan; 81(1) : 19-25.

[5] B. Potoˇcnik and D. Zazula. *Automated analysis of a sequence of ovarian ultrasound images. PartI : segmentation of single 2D images*. Image and Vision Computing. 2002, 20(3) : 217–225.

[6] E. Weibel, *Stereological Methods. Vol. 1: Practical Methods for Biological Morphometry.* 1979, Academic Press, London, U.K.