A Classification of Polycystic Ovary Syndrome Based on Follicle Detection of Ultrasound Images

Bedy Purnama1, Untari Novia Wisesti2, Adiwijaya1, Fhira Nhita4, Andini Gayatri3, Titik Mutiah6
School of Computing
Telkom University
Bandung, Indonesia
1bedypurnama@telkomuniversity.ac.id, 2untarinw@telkomuniversity.ac.id, 3adiwijaya@telkomuniversity.ac.id, 4fhiranhita@telkomuniversity.ac.id, 5andhinigyt@gmail.com, 6titik.mutiah@gmail.com

Abstract—Polycystic Ovary Syndrome (PCOS) is an endocrine abnormality that occurred in female reproductive cycle. This paper designed an application to classify Polycystic Ovary Syndrome based on follicle detection using USG images. The first stage of this classification is preprocessing, which employs low pass filter, equalization histogram, binarization, and morphological processes to obtain binary follicle images. The next stage is segmentation with edge detection, labeling, and cropping the follicle images. The following stage is feature extraction using Gabor wavelet. The cropped follicle images are categorized into two groups of texture features: (1) Mean, (2) Mean, Entropy, Kurtosis, Skewness, and Variance. This result in 2 datasets prepared for classification process, i.e. (1) data set A has 40 images that consist of 26 normal images and 14 PCOS-indicated images. It counted by Mean texture feature and obtained 275 follicle images. (2) Dataset B has 40 images consist of 34 normal images and 6 PCOS-indicated images. It counted by Mean, Entropy, Kurtosis, Skewness, and Variance texture features then obtained 339 follicle images. The last stage is classification. It identifies the features of PCO and non-PCO follicles based on the feature vectors resulted from feature extraction. Here, three classification scenarios are designed: (1) Neural Network-Learning Vector Quantization (LVQ) method, (2) KNN - euclidean distance, and (3) Support Vector Machine (SVM) – RBF Kernel. The best accuracy gained from SVM - RBF Kernel on C=40. It shows that dataset A reach 82.55% while dataset B that obtained from KNN-euclidean distance classification on K=5 reach 78.81%.

Key words: follicle detection, polycystic ovary syndrome, ultrasonography images, Gabor wavelet.

1. INTRODUCTION

Currently, digital medical images have been widely used in making diagnosis for data mobility and authentication [1]. One of the digital medical images is USG images. The diagnostic criteria on USG image is analyzing manually the ovary USG image to detect the number and size of ovary’s follicle. According to the criteria of National Institutes of Health (NIH), 6% - 10% female has PCOS while Rotterdam can widely define as 15% prevalency of PCOS [11]. Infertility is a failed ovulation, i.e. a process when an egg is released from the ovary. There are a lot of factors causing infertility; one of which is the abnormal number and size of follicle growth in ovulation phase. This abnormality is the initial symptom of Polycystic Ovary Syndrome (PCOS).

Ultrasonography (USG) diagnosis is an effective method to monitor internal organs, including to detect infertility in women. Infertility can be identified from the problem of uterus caused by PCOS. According to Rotterdam Consensus Criteria [6] in 2003, PCOS may be diagnosed if 2 of the 3 following criteria are identified: (1) oligo/amenorrhea, (2) clinical or biochemical symptom of excess androgen activity, and (3) PCO follicles according to USG result.

USG is an ultrasound-based diagnosis at 2-15 MHz. Sound waves are sent to the object and reflected in the form of electrical pulses. The processed pulses are displayed in a monitor in the form of grayscale image. The low contrast and high noise characteristics of USG images require better accuracy to detect PCO follicles, which are spherical (resembling a necklace) and clustered.

In terms of morphological features [9], ovaries consist of 2 types: normal and polycystic (multiple small cysts). In normal ovulation phase, mature follicles grow in various sizes. Follicles under 18 mm are called antral and those in 18-28 mm are categorized as dominant. Normal ovary is characterized by 5-9 follicles sized 2-28 mm [4] (as validated by medical experts). In the polycystic ovary, follicles are mature enough to ovulate when their size reaches 18-28 mm. However, in PCO, follicles stop growing in ovulation phase when they reach 5-7 mm in size. Immature follicles keep...
producing estrogen, a hormone that regulates mucous membrane growth in uterine cavity.

In a long period, the overproduced estrogen causes a thick mucous membrane. When endometrium cannot accommodate blood circulation, bleeding occurs in the form of spotting or even excessive bleeding. It is a symptom caused by PCOS.

According to a research conducted by Akkasaligar and Malagavi [4], PCOS ovary contains more than 10 PCO follicles. On the other hand, medical experts assisting in this research validate that PCOS ovary contains 10 or more PCO follicles. In PCO, the numerous growth of cyst-structured antral follicles causes the ovary to possess polycystic characteristic (multiple small cysts) [4]. The main difference from normal ovary is that antral follicles in polycystic ovary grows in number but not in size (immature), resulting in failed ovulation.

II. RELATED WORK

In general, the applicable approaches to detect PCO follicles are (1) stereology and (2) feature extraction and classification. In stereology, two-dimensional images are viewed as projections of three-dimensional objects. Stereology relates three-dimensional parameters of structures to two-dimensional measurements that are obtained from 2D slices through the structures. A variety of geometric attributes of follicles can be calculated using stereology, such as the follicle count, distribution of follicles within the ovary, and follicle size [10]. In a paper entitled Follicle Detection on the USG Images to Support Determination Polycystic Ovary Syndrome [2], stereology is used to measure the follicle diameter and euclidean distance is used on quantification of follicle. On the other hand, in a paper titled Ovarian Follicle Detection for Polycystic Ovary Syndrome using Fuzzy C-Means Clustering [3], feature extraction and classification is in use.

III. PROPOSED SCHEME

A. Problem Definition

Generally, obstetrician can directly identify PCO follicles. However, extensive time is needed to conduct a series of analysis. In this research, an application is designed to help physicians to detect PCO follicles. The proposed method is preprocessing and feature extraction phases. In the classification phase, KNN, LVQ and SVM are used to find the most optimal one.

B. General System Design

Figure 1 shows process blocks designed to detect PCO follicles, consist of 5 stages.

1) Medical Ultrasound Images

Original ultrasound image of the ovary is considered as input. There are 80 images, consist of 60 normal ovary images and 20 PCO ovary images. All of the data have been validated by physicians. The input images are shown in grayscale.

2) Preprocessing

Preprocessing stage is essential to improve input image quality in order to easily acquire important information to be processed in feature extraction stage. The first step in preprocessing is to obtain ROI. Then, Gaussian-based noise filtering 5x5 windows is used to eliminate noise. Afterward, contrast level is increased to distinguish the background from the foreground using equalization histogram. To facilitate follicle labeling process, the resulted images of equalization histogram are
inverted with negative transform. Therefore, the follicles are detected as white objects.

The information needed is only the background and foreground. Therefore, binarization is performed, resulting in black background and white foreground.

Binary images are sharpened and cleared to gain the desired follicle data through morphology process within erosion and dilation. Output of the preprocessing stage is binary images containing follicles detected as white objects. Figure 2 illustrates preprocessing diagram, while Figure 3 shows the image before and after preprocessing.

Figure 2. Block Diagram of Preprocessing

Figure 3. Before preprocessing (left), after preprocessing (right)

3) Segmentation

Figure 4 is a Block Diagram of segmentation. This stage is needed to distinguish the background of the image from the desired object for easier analysis. The initial step of this stage is edge detection using Canny method. The detected edge follicles are labeled as shown in the left image of Figure 5.

Each labeled follicle in the image is cropped in bounding box. Each follicle becomes a new image to be processed in the next stage, as displayed in the right side of Figure 5.

Figure 4. Block Diagram of Segmentation

4) Feature Extraction

In this stage, the essential features of segmented follicle images are extracted. Gabor wavelet method is applied to produce filter that is adjustable to the configuration of frequency and orientation utilized. The more combinations used, the more feature vectors produced. However, only the important vectors are in use, i.e. vectors that represent PCO and non-PCO follicles, for easier identification.

Gabor 2D Filter is Gaussian ellipse kernel function modulated with complex sinusoidal wave consist of real and imaginary parts. It is formulated in the following general equation [5]:

\[
G(x, y) = \frac{1}{2\pi} e^{-\alpha^2\left(\frac{x^2+y^2}{2}\right)} e^{i\pi\theta x \cos \theta + y \sin \theta}
\]  

2D Gabor wavelet filtering is a convolution operation between image matrix, for example I(x,y), and filter bank. In this research, image matrix I(x,y) is the detected follicle image with filter bank contains convolved filter mask matrix (kernel) in real and imaginary parts, where both matrices have the same dimension as image matrix I(x,y). The resulted real and imaginary filter depends on the combination of frequency j and orientation \( \theta \).

To obtain the extracted image features using Gabor wavelet [7], the next step is to convolve image matrix I(x,y) with filter bank using the following formula :

\[
O_{j,\theta}(x, y) = I(x, y) * G(x, y)
\]  

The convolved matrix is transformed into magnitude form to acquire matrix containing absolute value of the image. Then, feature vectors are normalized by dividing those using elements with maximum value. The result is then changed
into line vector and finally combined with feature vector. This feature vector is the result of the convolution output.

To determine the best texture features of Gabor wavelet, 2 groups of texture features are used: (1) Mean, and (2) Mean, Entropy, Kurtosis, Skewness, and Variance. As a result, there are 2 groups of dataset: (1) data A, 40 images consist of 26 normal images and 14 PCOS-indicated images, with Mean texture feature, obtained 275 follicle images, (2) data B, 40 images consist of 34 normal images and 6 PCOS-indicated images, with Mean, Entropy, Kurtosis, Skewness, and Variance texture features, obtained 339 follicle images.

5) Classification
This stage identifies the features of PCO and non-PCO follicles based on feature vectors resulted from feature extraction.

Three classification scenarios are designed. First, Neural Network-Learning Vector Quantization (LVQ) method. In training process, training data is processed and undergo the learning process using LVQ method. The competitive layer of this method will automatically learn to classify the given input vector to easily determine the target class based on similarity level. Target output is determined by user's decision/input. The result of this learning process is the change in weight value of hidden neuron. In testing process, test data is treated the same as training data. However, follicle feature recognition refers to trained network model generated in training process. The output identifies whether the follicle images are classified as PCO or non-PCO follicle class. To maximize the accuracy, the configuration of Neural Network-LVQ parameters is designed, including input neuron, hidden neuron, learning rate, and number of iterations. The best combination of those parameters value influences the success of the system in classifying the target and detecting PCO follicles.

Second, Support Vector Machine (SVM) method. There are 2 steps in this method: Training and Testing. The type of data distribution is non-linear; therefore, kernel function is needed, i.e. RBF, and Polynomial, using C parameter for data tolerance. This process discovers the best hyperline function. In training step, two training datasets are prepared, i.e. positive-labeled dataset (1) and negative-labeled dataset (-1). This enables testing step using max-voting method for each voting value hyperline function, and add one to data point x. As a result, prediction can be made to define which class data x belongs to, based on the highest voting value.

Third, K-Nearest Neighbor (K-NN) method. Classification is conducted using distance measurement and distance type parameters to measure the distance among objects. In this method, learning data is projected to multidimensional room, in which each dimension presents its features. The room is divided into several parts based on learning data classification. To determine which category a point belongs to, it is marked with k number of neighbors most frequently appearing. There are four ways to measure the similarity level, i.e. based on Euclidian, correlation, cosine, and cityblock. In this research we only use Euclidian. K-nearest neighbor analyzes the influence of similarity measurement and k value on the accuracy and failed system in classifying follicles, whether they are PCO or non-PCO follicles. K values used in this research are only the odd ones, i.e. 1, 3, 5, 7, and 9, for easier voting process in its category selection.

IV. RESULTS AND DISCUSSION
These are the testing scenarios designed to test the system:

A. Segmentation Validation

Figure 6. Follicles generated from system detection segmentation (left), Follicles generated by physician validation (right)

Figure 6 shows an example of image comparing follicles generated from system detection segmentation (left) with follicles generated by physician validation (right). The figure shows 2 segmentation mistakes, i.e. no. 6 and no. 13. Each
follcle is given target output as validated by the physicians. Classification process is then performed using various methods to match the features with the training follicle bank categorized into PCO and non-PCO classes. In this matching process, system classification output is obtained. Output is correct if it matches the physician's target, and vice versa.

B. Neural Network – LVQ test

In NN-LVQ test, the parameter combination is changed in terms of the number of hidden neuron (HN), learning rate (LR), and epoch. Data A and data B are compared. In addition, Hidden Neuron is set at 32, 256, 512 and 1024. The contrast in numbers is intended to distinguish the difference more clearly. Learning Rate is set at 0.01 and 0.5. 0.01 indicates slow rate, and 0.5 defines fast rate. The iteration is set at 100, 300, 500 and 1000. The classification rates of each classifier are determined using k-fold cross validation methodology. Dataset A and B are randomly divided into k = 10 folds for evaluation using k-fold cross validation technique and the holdout cross validation is carried out 10 times using nine folds as the training set and the remaining fold as the test set.

Table I shows that dataset B produces better accuracy than dataset A in NN-LVQ classification, in which dataset A and dataset B reach 72.36% and 74.63% respectively. A greater amount of hidden neuron influence both datasets accuracy, which means the more the hidden neuron is, the more increase the accuracy. The difference between learning rate and iteration amount do not affect the number of hidden neuron but influence a few of them. The maximum accuracy on NN-LVQ classification is on HN=1024, LR=0.01, and Epoch=1000 in which dataset A and dataset B reach 72.36% and 74.63 respectively.  The high number of texture features in data B causes NN-LVQ learning machine to build better knowledge.

C. KNN - Euclidean Distance Test

In KNN - Euclidean distance test, the parameter combination is changed in terms of k value. Only the odd k values are selected, which are 1, 3, 5, 7 and 9, because the number of groups that will be classified are even. It enables the system to easily detect new groups.

Table II shows that dataset A generates higher accuracy than dataset B in KNN - euclidean distance classification. The increase of value K started from 1 up to 5 improves the accuracy while that from 7 until 9 reach the fluctuated value by maximum K=5. The KNN-euclidean distance classification has reached its maximum accuracy on
K=5 in which the dataset A gains 80.73% and dataset B gets 78.81%. Therefore, this classification shows that a few texture characteristic may produce better accuracy compared to a great one.

D. SVM Test

In SVM test, the parameter combination is changed in terms of C Parameter. It is set at 10 folds, ranging from 10 to 200. The applied kernels are RBF and Polynomial.

<table>
<thead>
<tr>
<th>C</th>
<th>Accuracy of Dataset A</th>
<th>Accuracy of Dataset B</th>
<th>Accuracy of Dataset A</th>
<th>Accuracy of Dataset B</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>80.73%</td>
<td>75.22%</td>
<td>80.36%</td>
<td>78.17%</td>
</tr>
<tr>
<td>20</td>
<td>81.45%</td>
<td>74.63%</td>
<td>80.36%</td>
<td>76.99%</td>
</tr>
<tr>
<td>30</td>
<td>81.82%</td>
<td>75.81%</td>
<td>80.73%</td>
<td>76.99%</td>
</tr>
<tr>
<td>40</td>
<td>82.55%</td>
<td>76.40%</td>
<td>80.73%</td>
<td>76.7%</td>
</tr>
<tr>
<td>50</td>
<td>81.82%</td>
<td>76.12%</td>
<td>80.73%</td>
<td>76.7%</td>
</tr>
<tr>
<td>60</td>
<td>81.45%</td>
<td>76.12%</td>
<td>80.73%</td>
<td>76.99%</td>
</tr>
<tr>
<td>70</td>
<td>81.82%</td>
<td>76.12%</td>
<td>80.73%</td>
<td>76.99%</td>
</tr>
<tr>
<td>80</td>
<td>81.09%</td>
<td>76.12%</td>
<td>80.73%</td>
<td>77.29%</td>
</tr>
<tr>
<td>90</td>
<td>81.82%</td>
<td>76.40%</td>
<td>80.73%</td>
<td>77.29%</td>
</tr>
<tr>
<td>100</td>
<td>81.09%</td>
<td>76.7%</td>
<td>80.73%</td>
<td>77.29%</td>
</tr>
<tr>
<td>110</td>
<td>81.45%</td>
<td>76.7%</td>
<td>80.73%</td>
<td>77.29%</td>
</tr>
<tr>
<td>120</td>
<td>81.45%</td>
<td>76.7%</td>
<td>80.73%</td>
<td>77.29%</td>
</tr>
<tr>
<td>130</td>
<td>81.45%</td>
<td>76.7%</td>
<td>80.73%</td>
<td>77.29%</td>
</tr>
<tr>
<td>140</td>
<td>81.45%</td>
<td>77.29%</td>
<td>80.73%</td>
<td>77.29%</td>
</tr>
<tr>
<td>150</td>
<td>81.09%</td>
<td>77.58%</td>
<td>80.73%</td>
<td>77.29%</td>
</tr>
<tr>
<td>160</td>
<td>81.09%</td>
<td>76.99%</td>
<td>80.73%</td>
<td>77.29%</td>
</tr>
<tr>
<td>170</td>
<td>81.45%</td>
<td>76.99%</td>
<td>80.36%</td>
<td>77.29%</td>
</tr>
<tr>
<td>180</td>
<td>81.45%</td>
<td>76.99%</td>
<td>80.36%</td>
<td>77.29%</td>
</tr>
<tr>
<td>190</td>
<td>81.09%</td>
<td>76.99%</td>
<td>80.36%</td>
<td>77.29%</td>
</tr>
<tr>
<td>200</td>
<td>81.09%</td>
<td>76.99%</td>
<td>80.36%</td>
<td>77.29%</td>
</tr>
</tbody>
</table>

Table III shows that the overall process of using both RBF Kernel and Polynomial Kernel on SVM classification have made dataset A better than dataset B. The maximum accuracy on SVM classification with RBF kernel C=40 remains dataset A reaching 82.55%. In addition, dataset B with polynomial kernel C=10 has reached 78.17%. Therefore, this classification shows that a few texture characteristic may produce better accuracy compared to a great one.

V. CONCLUSION

Based on the analysis towards the test on the PCO detection system of ultrasonography image, it can be concluded that the proposed scheme can be implemented for detecting PCO. The best accuracy gained from SVM - RBF Kernel on C=40. It shows that dataset A reach 82.55% while dataset B that obtained from KNN-euclidean distance classification on K=5 reach 78.81%. It means a few texture feature may produce better accuracy compared to a great one.

ACKNOWLEDGMENT

The authors would like to thank Telkom University for financial support on this research. Also, the authors would like to thank Dr. Zulkifli Ahmad, SpOG., K.FER., M.Kes as Female Fertility Consultant in Permata Bunda Syari'ah Clinic for the data and consultation.

REFERENCES