# Automated Detection of PCOS using Follicle Recognition Techniques

## (Polycystic Ovarian Syndrome Detection)

Zuha Azmi[1]
Dept. of Computer Science & Engineering
GSSS Institute of Engineering and Technology for Women, VTU Belagavi, India

Vishwesh J[2]
Dept. of Computer Science & Engineering
GSSS Institute of Engineering and Technology for Women, VTU Belagavi, India

Madhu M Nayak[3]
Dept. of Computer Science & Engineering
GSSS Institute of Engineering and Technology for Women, VTU Belagavi, India

Gururaj K S[4]
Dept. of Computer Science & Engineering
GSSS Institute of Engineering and Technology for Women, VTU Belagavi, India

***Abstract*--Polycystic ovaries cause infertility in women because the development of follicles is inhibited, resulting in a large number of follicles. Follicle identification is still configured by a gynaecologist and quantity and size of follicles in the ovaries are also counted, which takes a lengthy schedule and requires a peak level of precision. PCO can be discovered via stereology calculations or extraction and classification of features in follicles. PCOS is diagnosed in this study utilizing a follicle count retrieved from ultra sound pictures using the K-Means clustering technique. A decision tree classifier with greater than 90% accuracy is used to perform the classification.**

***Keywords--Polycystic ovarian syndrome, k-means clustering, SVM algorithm, follicle detection, ultrasound images, reduced levels of FSH and LH, Decision Tree Classifier.***

## 1. INTRODUCTION

The alteration of digital images in order to extract more information than is evident on the original image is known as image processing. Image processing is the technique of altering photos using digital computers. Its popularity has risen dramatically in recent decades. It has a wide range of applications, including medical and entertainment, as well as remotely sensed and geotechnical analysis.

Multifunctional peripherals, which are one of the foundations of today's information society, rely heavily on digital image processing. Picture enhancement, image restoration, image analysis, and image compression are some of the types of digital image processing.

An image which is digital is a 2-D matrix of different-valued pixels that define the image's color or grey level. The number of pixels in an image increases as the resolution increases. Each pixel represents a different shade of grey or color. The resolution of grey levels is measured in bits; an 8-bit image has 256 grey levels. Filters are used in image processing techniques to enhance an image. Their primary functions are to alter an image's contrast, brightness, resolution, and noise level. Image processing functions include contouring, image sharpening, blurring, embossing, and edge recognition.

## 2. LITERATURE REVIEW

The existing research involved in the detection PCOS is presented as follows:

Amsy Denny et al. [1] proposed Machine Learning Techniques for identifying Polycystic Ovary Syndrome (PCOS) also includes detection and prediction. It basically segment the images and then classification results will be applied accordingly. Various algorithms has been used to compare the performance of each algorithm.

Subrato Bharati et al. [2]. The best features that can predict PCOS are found using a univariate feature selection technique or algorithm. To segregate training from test data, cross-validation methods are used on the dataset. It also separates trained and test sets.

Vishal Vishwakarma et al. [3]. ANN was used to diagnose whether a patient had PCOS, achieving an accuracy of 87.96%. The classification results will help to classify the disease by severity: high, moderate or weak.

R. Boomidevi et al. [4] from Department of CSE Engineering Kongu Engineering paper, 2020, it uses the technique of neural networks measurements were taken every other day to determine the serial follicular score to monitor ovulation induction. Methods and mathematical assertions designed to devise solutions for enhancing the prediction accuracy of infertility in women.

Jojo Cheng et al. [5] used the method of Performance evaluation of rule-based classifier and gradient tree model for automatic PCOS feature extraction and classification.

E. Setiawati et al. [6]. Using particle swarm optimization (PSO) and a new modified asymmetric fitness function, a new picture clustering method for follicle segmentation is proposed.

C. Gopalakrishnan et al. [7] . it used the method of proposed feature extraction GISTMDR model. Efficient automated PCOS detection and classification system suggested from ultrasound images by analyzing affected and unaffected cases.

Ranjitha S. et al. [8] uses an efficient automated follicular identification system. Methods for Detecting Ovarian

Folles When It consists of two main phases including: - Treatment phase, the identification of follicles based on the growth of the object.

O. Rabiu1 et al. [9] explained Computer-aided ways for detecting and diagnosing Polycystic Ovarian Syndrome (PCOS) cysts offer great benefits such as stool Quickly analyze ultrasound images in the shortest possible time, minimizing diagnostic errors.

Noraishikin et al. [10] from Ultrasound Imaging generative Medicine, 2020. The vesicle score was established considering the summation of points obtained once measurement the mean diameter of every follicle.

## 3. PRESENT SYSTEM

Researchers looked at potential environmental factors throughout prenatal and postnatal periods because early symptoms of PCOS usually appear in teens. Excessive intrauterine exposure to androgens or glucocorticoids during specific critical phases of fetal development contributes to the development of PCOS symptoms, according to evidence from experimental research and determines phenotypic expression of PCOS in adulthood.

It suggests that there is a possibility of doing so. Intrauterine growth retardation (IUGR) is thought to increase prenatal exposure to androgens / glucocorticoids and can induce fetal programming in PCOS.

Studies about this evidence for this possible pathway is mixed, with studies pointing to a relationship between gestational age and PCOS, and other data showing no such link. Although there is a lack of additional evidence that postnatal exposure to environmental risk factors is connected with PCOS, based on human investigations of putative prenatal environmental factors.

## 4. PROBLEM STATEMENT

According to the Rotterdam Conference, patients may have PCOS if they have two of the three symptoms: (1) Ovulation inability (2) High levels of androgen hormone, or (3). Existence of polycystic ovary. Morphologically, the presence of polycystic ovaries indicates 12 or follicles many with a range or measure of 29 mm, or the volume of the ovaries has increased by 10 cm3 or more. Ultrasound images are manually checked by a doctor by counting the number and size of ovarian follicles.

However, the test is time consuming and requires high accuracy to determine if it is polycystic ovary syndrome. In PCOS-affected ovaries, follicles cannot grow and mature due to Fertilization and Luteinizing hormones are low, while prolactin levels are high.

## 5. OBJECTIVES

The objectives of the following PCOS work is mentioned below:

1. Collection of Ultrasound PCOS images to create dataset.
2. Data preprocessing using SVM algorithm.
3. Classification of PCOS images as positive or negative PCOS.

## 6. METHODOLOGY

Figure.1 shows the present research work steps to be followed. The block diagram consists of steps performed while evaluating PCOS.

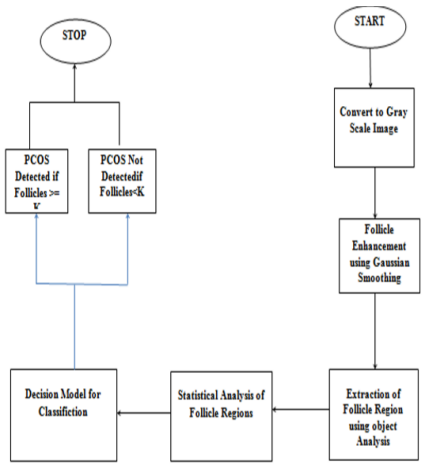


Figure 1: Methodology

Figure.1 shows the following steps of Methodology to detect PCOS:

1. Start the procedure.
2. Collect ultra sound images.
3. Convert the image to gray scale images.
4. Follicle is enhanced using gaussian smoothing.
5. Extract the follicle region using object analysis.
6. Follicle region is statistically analyzed.
7. Decision model is designed for classification.
8. Classifies as positive or negative PCOS.
9. Stop the procedure.

### *6.1 Image to grey scale*

As the name implies, RGB images are made up of three levels: R, G, and B. This is a three-dimensional matrix, for example, three consecutive pages of a book. Here, the grayscale image has only two dimensions and the range of values is 0 to 255 (8-bit unsigned integer).

Therefore, some algorithms only apply to 2D images, not 3D, so convert RGB images to grayscale images. B. Black-and-white conversion of images, image convolution, etc.

### *6.2 Follicle enhancement using Gaussian Smoothing*

Effective ovarian ultrasound imaging enhancements provide better results. The purpose of this module is to further improve the contrast of the image while preserving the information. Low contrast ultrasound images must be preprocessed before applying the pseudo-staining procedure. Stretching of contrast can increase contrast, while selecting a zone might draw attention to a specific spot.

In ovarian ultrasound imaging, a combination of image processing techniques such as area selection, contrast extending seed-based region development, and pseudo-staining procedures can improve contrast and offer additional information.

*6.3 Follicle region extraction using connected component analysis*

Seed Based Region Growing is a technique of segmentation for segmenting a region of interest (ROI). In seed-based region growth, the pixels in the region are used as the starting spot, and then all indistinguishable pixels (spatial near pixels that share the similar function, that is, adjacent pixels) are collected in the same region. The area expansion method that defines the ROI is based on the difference in grayscale.

*6.4 Compute follicle count*

After binarization and edge detection, a number of features are recognized. According to the literature reviewed, follicle sizes range from 29 mm in range or measure for highly PCOS pretended ovaries to approximately 20 mm in normal ovaries. Taking into consideration that the follicle parts are circular, each area is about 480 mm for PCOS ovaries and about 314 mm for common follicles.

The follicles are assumed to be round, so the queerness is approximately equal to 1. All these estimates of the entry are taken into account to calculate the area of each area or feature detected to determine the number of important follicles detected.

*6.5 Classification using decision tree classifier*

First the ultra sound image is given as input and tested. Later, converted to gray scale image. Gray scale image of follicles is enhanced using Gaussian smoothing. The follicle region is extracted using a connected component. After this procedure the follicle is computed.

Table.1 shows the follicles are classified using decision tree algorithm and then finally it is classified as PCOS or Non PCOS.

Table 1: Value of Parameters for classifying Normal and PCOS Affected Patients

| Parameters | Normal | PCOS Affected |
|---|---|---|
| No. of Ovarian Follicles | 1 Dominant Follicle | >8-10 |
| BMI (kg/m$^2$) | <=24 | >24 |
| Cycle Length(Days) | 28-32 | >32 |
| LH (IU/L) | 2-8 | >8 |
| FSH (IU/L) | 2-5 | >5 |

## 7. SYSTEM IMPLEMENTATION

In order to design a system, first step is to collect the system requirements, functional and non-functional requirements, constraints from the user.

Second step is designing the system in an abstract manner, this step provides outline of all major components that is required for designing system architecture.

Third step is detecting and addressing bottlenecks generated in the abstract or high level design due to violation of some constraints specified by the user.

*7.1 K-Means Clustering*

K-means clustering is one of the best and famous unsupervised gadget mastering algorithms. “The goal of K-means is simple: organization comparable records factors collectively and find out underlying patterns. To acquire this goal, K-means appears for a hard and fast number (k) of clusters in a dataset.”

A focal point is the real or imaginary place that is the center of the group. Each data point is assigned to each of the groups by reducing the sum of the squares of the groups.

*7.1.1 K-Means Algorithm:*

Let X = {x1,x2,x3,……..,xn} be the set of data points and V = {v1,v2,…….,vc} be the set of centers.

1. Pick the cluster C center at random.
2. Determine the distance between each data points and the center of the cluster.
3. Assign the data point to the cluster center.
4. Re compute the new cluster center.
5. Re compute the interspace between each data point and the newly obtained cluster center.
6. Quit and restart from step 3 while no data points were reassigned.

*7.2 Decision Tree Classifier*

Figure.2 depicts the decision tree used for classification. One of the simplest categorization techniques to comprehend and analyze is the decision tree. It can be used to solve problems involving classification and regression.

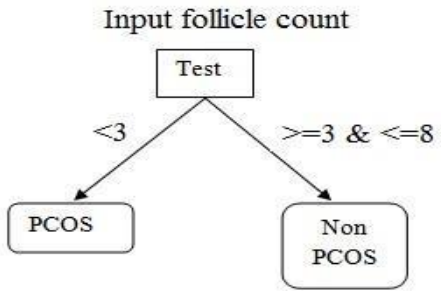


Figure 2: Proposed decision tree classifier for PCOS classification

*7.2.1 Decision Tree classifier Algorithm*

Figure.3 shows the working of following decision tree classifier as a activity diagram:

1. Make S the root node of the original set.
2. Revisit through the unused attribute of the set S and compute the entropy at each enumeration of the procedure (H).
3. Next, choose the property that has the lowest information obtained or the greatest information obtained.
4. A subset of the data is generated by dividing the set S by the selected attribute.
5. The algorithm repeats itself on each subset.

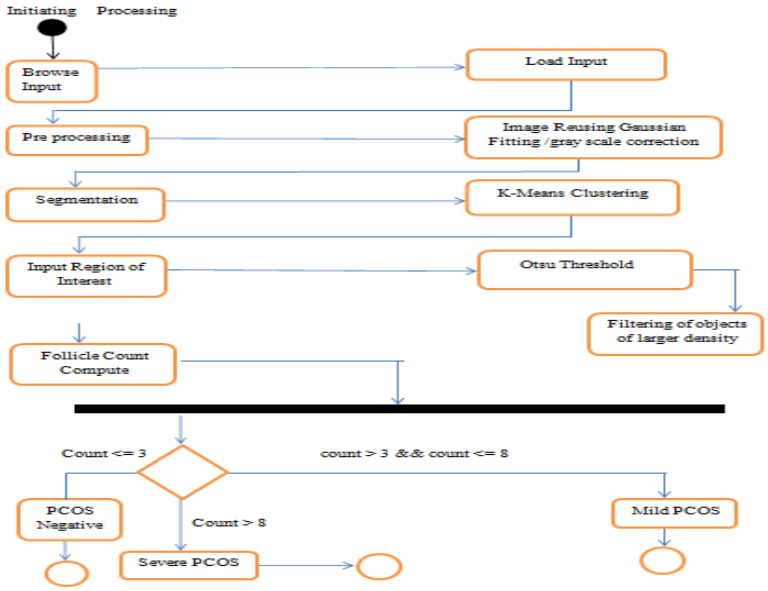


Figure 3: Decision Tree Classifier as Activity Diagram

## 8. EXPECTED OUTCOME

From literature it is observed that, most of the works focus on classification and segmentation of follicle region using morphological operations and or either applied classifier directly on ultrasound images without segmentation. It is also noticed that accuracy of classification is lagging down to 78%. Therefore proposing an efficient technique for segmentation of follicle regions is very significant. Thus in the proposed approach, automated method for diagnosis of PCOS is devised.

## 9. CONCLUSION

In this work, cyst identification and classification using SVM to try to build an automated technique for detecting polycystic ovarian syndrome (PCOS). To improve image quality, preprocessing of image such as: contrast enhancement and filtering is utilised. The glossy and shiny transform and a multi-scale morphometry technique were used to extract features.

Using the canny facet detection approach, the ultrasound picture is binarized and segmented. Different locations are used to isolate important follicles.

The data is classified into: clinical, biochemical and imaging, it is executed with the use of Support Vector Machine (SVM) algorithm.

The output of manual classification and physicians are compared with the results. 95% is the accuracy of this method. Therefore, this algorithm can be effectively used to automatically detect PCOS patients.

## REFERENCES

[1] Amsy Denny, Anita Raj, C Maneesh Ram “Detection And Prediction System For Polycystic Ovary Syndrome (PCOS) Using Machine Learning Techniques”, 2019.

[2] Subrato Bharati; Prajoy Podder; M. Rubaiyat Hossain Mondal “Diagnosis of Polycystic Ovary Syndrome Using Machine Learning Algorithms”@ IEEE Transaction on medical imaging, 2020.

[3] Vishwakarma, S. Chethan, Mohana Tejasri Datla, Mohammed Mukkaram Aqib, Subhajeet Roy, T. Thasni. “Prediction of Severity of Polycystic Ovarian Syndrome”. @ Artificial Neural Networks, 2021.

[4] R. Boomidevi, S. Usha “Polycystıc Ovary Syndrome (PCOS) Detectıon System Usıng Neural Network Approach” @ from Department of CSE Engineering Kongu Engineering paper, 2020.

[5] Jojo Cheng, Shruthi Mahalingaiah” Data mining and classification of polycystic ovaries in ultrasound reports.” @ published in Fertility Research and Practice. 2021.

[6] E. Setiawati; Adiwijaya; A.B.W. Tjokorda” Particle Swarm Optimization on follicles segmentation to support PCOS detection.”International conference on Information and Communication Technology (ICoICT). 2020.

[7] C. Gopalakrishnan & M. Iyapparaja” ultrasound images using machine learning” @ International Journal of System Assurance Engineering and Management. 2021.

[8] Ranjitha Sitheswaran S. Malarkhodi” [8] An effective automated system in follicle identification for Polycystic Ovary Syndrome using ultrasound images” @ International Conference on Electronics and Communication Systems (ICECS) 2019

[9] O. Rabiu1 , A. D. Usman2 , A. M. S. Tekanyi,” Computer Assisted Follicle Detection Techniques and Polycystic Ovarian Syndrome (PCOS) Diagnostic Systems.” @ Department of Computer Engineering, Federal University of Technology, Minna, 2019.

[10] A A Nazarudin, Noraishikin Zulkarnain, A. Hussain, S. S. Mokri, I. N. A. M. Nord “Review on automated follicle identification for polycystic ovarian syndrome.” @ Ultrasound Imaging Reproductive Medicine, 2020.

[11] Moran LJ, Noakes M, Clifton PM, Wittert GA, Williams G, Norman RJ: Short-term meal replacements followed by dietary macronutrient restriction enhance weight loss in polycystic ovary syndrome, 2006.

[12] Herriot AM, Whitcroft S, Jeanes Y: An retrospective audit of patients with polycystic ovary syndrome: the effects of a reduced glycaemic load diet. J Hum Nutr Diet, 2008.

[13] Meyer C, McGrath BP, Teede HJ: Effects of medical therapy on insulin resistance and the cardiovascular system in polycystic ovary syndrome. Diabetes Care, 2007.

[14] Tang T, Lord JM, Norman RJ, Yasmin E, Balen AH: Insulin-sensitising drugs (metformin, rosiglitazone, pioglitazone, D-chiro- inositol) for women with polycystic ovary syndrome, oligo amenorrhoea and subfertility. Cochrane Database Syst Rev, 2009.

[15] Teede H, Zoungas S, Deeks A, Farrell E, Moran L, Vollenhoven B: Check: independent learning program for GPs. Polycystic ovary syndrome. Edited by: Royal Australian College of General Practitioners, 2008.