# Enhancing Polycystic Ovary Syndrome (PCOS) Diagnosis Using Machine Learning Techniques

**Dr. Baraa Mohammed Abed[1], Rasha Amer Kadhim[2], Ihsan Salman[3], Hussam Shaker Hussein[5], Dr. Abdullah Farhan Mahdi[6]**

[1.] Ministry of Education, Directorate of Education in AL-Anbar, Iraq, Artificial Intelligence, burasoft@gmail.com
[2.] College of Engineering for Artificial Intelligence Technology, University of Diyala,
[3.] rashakadhim@uodiyala.edu.iq
[4.] College of Engineering For Artificial Intelligence Technology, Artificial, Intelligence Department, Diyala University, Diyal, Iraq, drihsan@uodiyala.edu.iq
[5.] College of Education for Pure Sciences, University of Diyala, hussam.shakir@uodiyala.edu.iq
[6.] College of Engineering For Artificial Intelligence Technology, Artificial Intelligence Department, Diyala University, Diyala, Iraq, abdullahmahdi@uodiyala.edu.iq

**Abstract:** *Polycystic Ovary Syndrome (PCOS) is a prevalent medical condition affecting women and is associated with hormonal imbalances, infertility, and irregular menstruation. Accurate and early diagnosis of PCOS is critical for effective treatment and management. This study explores the application of a Siamese Neural Network in diagnosing PCOS using ovarian ultrasound images. By comparing extracted features from healthy and abnormal ovarian images, the model provides a robust method to identify PCOS-specific abnormalities, such as cystic formations, thickened ovarian tissues, and disrupted vascular patterns. Advanced image enhancement techniques are employed to preprocess the data, while the Siamese Network analyzes structural and histological features to determine the presence of PCOS. The results demonstrate the model's ability to detect significant variations between healthy and abnormal ovaries, including hormonal patterns, blood vessel distributions, and tissue abnormalities and the proposed model achieves reliable classification accuracy, offering a potential tool for aiding clinicians in early PCOS detection and this research highlights the transformative role of machine learning, particularly Siamese Networks, in medical imaging and diagnosis, paving the way for transparent and efficient healthcare solutions. Further advancements are recommended to incorporate additional data sources and predictive treatment tools for enhanced patient outcomes.*
**Keywords:** Polycystic Ovarian Syndrome (PCOS) - Medical Diagnosis - Machine Learning Techniques - Predictive Models - Biological Analysis.

## 1. Introduction.

Polycystic Ovary Syndrome (PCOS) is one of the most common conditions among women of reproductive age, affecting approximately 5 to 10 percent of women between the ages of 15 to 45 years. It is primarily characterized by hormonal imbalances and metabolic complications that significantly impact reproductive health. PCOS disrupts the normal production of estrogen and androgens, hormones critical for ovulation and other physiological processes. This disruption often results in the formation of cysts on the ovaries, reducing ovulation frequency and consequently lowering fertility potential [1]. Furthermore, PCOS is associated with an increased risk of comorbidities such as type 2 diabetes, hypertension, coronary artery disease, endometrial hyperplasia, and carcinoma. Psychological issues, including depression, anxiety, and sleep disorders, are also prevalent among patients [2]. Genetic predisposition and environmental factors contribute to the pathogenesis of PCOS. Early and accurate diagnosis, along with tailored interventions and lifestyle modifications, is essential to mitigate long-term complications associated with the syndrome [3].

Artificial intelligence (AI) has revolutionized medical diagnostics, offering promising solutions for conditions like PCOS. By leveraging machine learning (ML) and deep learning (DL) techniques, AI systems are capable of processing complex datasets, such as hormone levels, ultrasound images, and genetic markers, to identify patterns indicative of PCOS and these techniques are not only effective in detecting PCOS but also applicable to other diagnostic challenges, such as cancer detection, cardiac risk stratification, and neurological disorders [4]. Machine learning models analyze data at a granular level, providing clinicians with actionable insights for early intervention and personalized care. For instance, ML-based models have been used to predict ovulatory dysfunction, assess hormonal imbalances, and evaluate ovarian morphology with remarkable accuracy [5].

Feature selection plays a pivotal role in enhancing the performance of AI-driven diagnostic systems. By identifying the most relevant features, these methods reduce data dimensionality, improve computational efficiency, and minimize noise. Hybrid models, which combine multiple algorithms, have proven

particularly effective in diagnostic applications, including PCOS. For example, combining logistic regression with random forests has shown improved diagnostic accuracy in classifying ovarian abnormalities [6]. Optimization algorithms, such as the Modified Bat Algorithm, have also been used to fine-tune feature selection processes, further enhancing model performance [7] and despite the superior accuracy of complex ML models, such as neural networks, they often face challenges related to transparency and interpretability. Explainable AI (XAI) techniques address these challenges by providing insights into how predictions are made, thereby fostering trust among clinicians and patients, and these techniques are crucial for ensuring the adoption of AI models in sensitive applications like PCOS diagnosis [8].

In light of these advancements, this study proposes a robust framework for PCOS diagnosis by integrating feature selection methods with advanced ML algorithms. Key contributions include addressing class imbalances using SMOTE and Edited Nearest Neighbors (ENN), applying feature selection to optimize model inputs, and employing Bayesian optimization for ML algorithm tuning. These approaches aim to enhance diagnostic accuracy, interpretability, and reliability. While the current introduction outlines the significance of PCOS and AI's potential in its diagnosis, it could benefit from a deeper discussion on how machine learning has been applied to similar medical diagnostic challenges. For example, AI-driven systems have been successfully used in early cancer detection, cardiac disease risk assessment, and chronic disease management. Incorporating references to recent advancements in these areas would provide a more comprehensive view of the medical and technical landscape, and this would also emphasize the importance of adopting AI in PCOS diagnostics while highlighting its broader implications in healthcare [9-11].

## 2. Related Works

PCOS syndrome is one of the most prominent health problems facing women around the world, and is a medical challenge in diagnosis and treatment due to the diversity of symptoms and their interference with other health problems and therefore, early and precise diagnosis is vital to providing appropriate treatments and preventing the development of complications associated with this syndrome, such as increasing the possibility of type 2 diabetes and pregnancy diabetes and thus, it is acknowledged the importance that artificial intelligence and machine learning methods in providing additional solutions to promote the diagnosis of PCOS and therefore the care offered to patients and the healthcare system. Studies that have been conducted in the recent past tend to clarify that the employment of automated learning and assembling processes tends to yield better results in the diagnosis of PCOS. In the work authored by Elmannai. H et al., the problem of classification was analyzed using the stacking techniques since they were able to obtain 100% resolution after appropriate feature selection methods were employed. Another work done by Bharati. S et al (2020) [9] draw attention to how the combination of random forest and logistic regression model was able to deliver excellent results in PCOS diagnosis, achieving a resolution of 91.01%.

In their studies, use Suha. S. A. and others (2022) [10] developed a fighter nervous network with assembly techniques to improve the classification based on ovarian ultrasound images, and reached a resolution of 99.89% using a stacked model with XGBOST algorithm.

On the other hand, I found a Danaei Mehr, H., etc. (2021) [12] study that the use of feature selection methods with improved assembly algorithms led to a resolution of up to 98.89% in PCOS diagnosis.

As suggested Suha. S. A. and others (2023) [13] Use the accumulation of advanced models with augmented gradient classification, as they reached a resolution of 95.7% using the main ingredient analysis to choose features.

In their studies, Bhat, S. A., and others (2021) [14] use Catboost and XGBRF models to improve the accuracy of the diagnosis, achieving a resolution of 95% and 89%, respectively.

As for Denny, A. and others (2019) [15] they used various techniques such as random forests and reached a resolution of 89.02%, while Tiwari, S and others (2019) [2] found that random forests are the most accurate way to diagnose PCOS, where they achieved a resolution of 93.25%.

Overall, these studies highlight the great value of using machine learning and aggregation techniques in the diagnosis of PCOS and the findings indicate that the use of optimal feature selection techniques in conjunction with stacked and ensemble models leads to a remarkable improvement in the diagnostic performance and therefore, employing these technologies is an important step towards improving the quality of healthcare provided to patients with PCOS, and accelerating the early diagnosis and effective treatment of this condition [16-18].

## 3. Methodology.

We will carry out two types of processes to extract features: First, by extracting the image feature using a trained nervous network, to extract the image feature, we will use a nervous number trained to extract relevant features from the images and this process includes feeding images in the Neusy network, which

was previously trained on a large data collection, to extract the meaning of these features that capture important features of the images, such as shapes, textures and patterns, which are then used for more analyzes such as for the second system, we will employ a digital data collection The structure data includes.

We will train the automated learning form on this dataset to learn patterns and relationships with data [19].

The training process includes feeding the form with the name called, and the knowledge of its parameters to reduce errors and improve their priority by simply training, the model can be tested on new invisible data to evaluate its performance and assess its ability to accurately predict the results, in short, a summary, our approach includes taking advantage of extracting the image feature using a trained network and learning technologies Traditional automatic.

The process of extracting features from ovarian images includes the analysis and identification of visual properties or related patterns within images that indicate some cases or distortions, such as PCOS, and these features can include different aspects of ovarian, texture, structure, and spatial distribution of different tissues.

- Mographs: It consists of an assessment of dimensions and morphology of the unequivocal presence or absence of cysts and/or follicles and their abnormalities. For instance, in Polycystic Ovarian Syndrome (PCOS), the ovaries may be predominantly riddled with numerous abscesses.
- Textual features: textual features represent the way light energy in the picture is mapped out in pixels. Image texture classification can include soft or rough or detailed, or smooth, depending on the anticipated image application.
- Structural features: These features detect the different tissues that make up the inside part of the ovaries, for example, connective tissues, various blood vessels, or arteries. Some of these may also be seen in clinical cases of other diseases, such as Polycystic Ovarian Syndrome.
- Decrease features: Density features relate to the measurement of pixel density variation in an image, and such variations correlate to the tissue's density. This is useful when it comes to recognizing important features like cysts or regions of the image with abnormal tissue density.
- Spatial features: spatial features concern the spatial orientation of the various aspects or parts contained within the ovary or their relationships to one another, and this includes the spatial distribution of abscesses, follicles, or vascular structures, which may appear in specific patterns in PCOS.
- Extracting these features from ovarian images serves several purposes:
- Diagnostic aid: By identifying and analyzing visual features, doctors can identify patterns or abnormalities that may not be easily clear through visual examination alone, and this can help diagnose conditions such as PCOS.
- Objective evaluation: Features provide an objective and quantitative analysis of ovarian morphology, texture, and structure, and reduce dependence on self-interpretation.
- Prediction and classification: The extracted features can be used as inputs for automatic learning models to predict or classify ovarian conditions. By learning from a large collection of data from the explained images, these models can determine the exact patterns that indicate specific diseases, facilitate automatic diagnosis, and class division of risk. In general, the advantages of ovarian imaging play an important role in enhancing the accuracy of the diagnosis, enabling early detection and intervention in cases such as PCOS, and improving patient outcomes. Equations used to extract features from ovarian images depend on the specified type of features and characteristics that are analyzed.

Below Soe is simple examples of equations that may be used to extract certain features [14]:

**3.1. Structural features:** One can analyze the arrangement of the tissues present in the figure by applying metrics like average high and low surfaces. For example, one can quantify the relationship between the area of fluid-filled cysts (cysts) to that of normal ovarian tissue easily through the following expression:

$$[\{\text{Cystic Area Ratio}\} = \left\{\frac{\{\{\text{Area of Cysts}\}\}}{\{\text{Total Ovarian Area}\}}\right\}]$$

**3.2. Textural Features:** Techniques for image analysis, for instance the standard deviation matrix, may also assist in the estimation of texture contrast, which denotes the measure of scattering of colors or tones within an image. An expression given below can be used to work out contrast:

$$[\{\text{Standard Deviation}\} = \sqrt{\left\{\frac{\left\{\sum_{\{i=1\}}^{\{n\}}(x_i - \mu)^2\right\}}{\{n\}}\right\}}]$$

Where $(x_i)$ is the pixel value in the image, and ( mu ) is the average value of all pixels in the image.

**3.3. Density Features:** One can easily derive certain formulas for the estimation of the tissue density present in the image, for example, the ratio of number of high-density pixels to the area occupied by other densities within a defined region.

**3.4. Spatial Features:** Other indicators can also be assessed in order to provide some estimation on the spatial distribution of these structures, for example, the average distance between cysts or between the cysts and the inner structures of the ovary.

These are only quite basic instances, and indeed, the formulas are based on the type of images taken and features drawn from them. Formulating and applying appropriate relations is not only technical, but also needs a clear insight into the image being analyzed and the characteristics that need to be derived from it.

To ensure reproducibility and accuracy of the work, the methodology is described in more detail to include the exact steps used in building and optimizing machine learning models. Specific algorithms such as Random Forests, Support Vector Machines, and Deep Neural Networks were used. For parameter settings, the number of trees in the random forest was set to be 100 trees with the selection of the "gini impurity" criterion for dividing nodes, and the vector support machine parameters were set to include an RBF kernel with a gamma value of 0.1 and a penalty coefficient C of 1. As for the deep neural networks, they were designed with two layers. Hidden, each layer contains 64 and 32 units, respectively, using the ReLU activation function and Adam optimizer implementation with a learning rate of 0.001.

Regarding the training procedure, the data were partitioned using five-fold cross-validation to ensure equal distribution of training and testing data and avoid overfitting. The data underwent several pre-processing steps, which included interpolating missing values using the averaging method, and normalizing the numerical values to a range of [0,1] to ensure consistency of the numerical range of the different features. To ensure data balance is addressed, synthetic sample generation techniques such as SMOTE have been applied to balance the data and increase representation of underrepresented groups, and the dataset used consists of 500 samples, collected from a clinical database containing data from patients with and without PCOS. Features included demographic data (eg, age and body mass index), clinical indicators (eg, hormone levels and number of follicles), and ultrasound image data. The data presents challenges such as missing values and class imbalance, with 30% of samples containing missing values and features were selected using the Principal Component Analysis (PCA) method to reduce data dimensionality and improve computational efficiency, and this was supported by additional techniques such as importance-based selection using random forests, where features that contribute a high percentage of variance or accurate prediction were identified. The selection of features was justified based on their clinical importance, such as the association of hormone levels and follicle number with the diagnosis of PCOS, and their effectiveness in improving model performance.

Finally, the models were evaluated using metrics of accuracy, recall, and precision, as well as F1-score, to get a comprehensive picture of the models' performance. All processes, parameters, and results are documented to ensure that the methodology can be accurately re-performed and reproducible. Figure 1. shows Siamese network business plan.

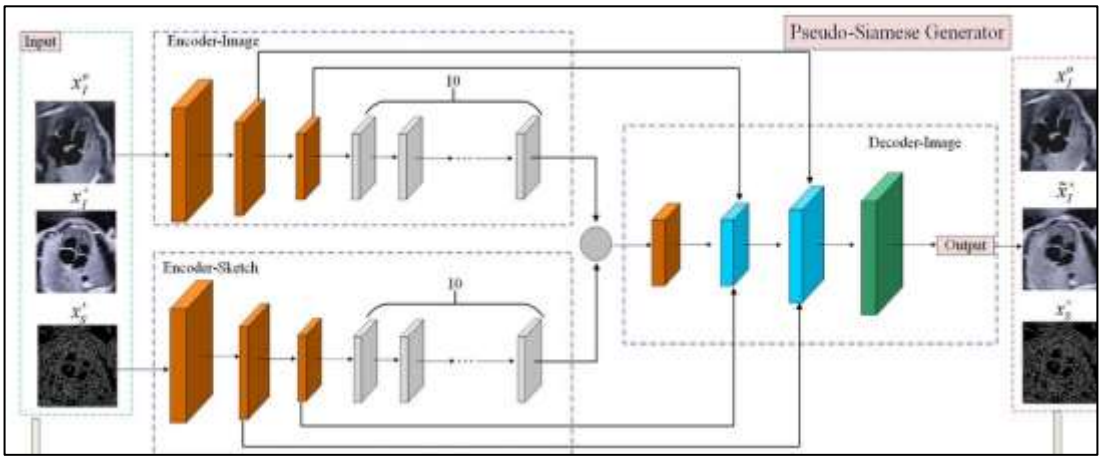



FIGURE 1. Siamese network business plan

## 4. Results and Discussion

A Siamese Network is a form of an artificial neural network that's commonly used in identifying similar data points, and the design defines the term 'Siamese' due to the training of two identical models on the same data. Each model accepts corresponding data in their form, which is used to generate shared data

characteristics. And these characteristics are assessed for the degree of similarity between the two data pairs in question to find out whether they are similar or different.

Examples of artificial neural networks are Siamese networks, which are appropriate for applications such as medical images, where images of the ovaries are used, and these networks can help in determining the healthy image attributes of the ovary in relation to the injured or abnormal pictures. In this case, a network is trained on two images, such that a healthy ovary is the first picture while a damaged or an abnormal ovary is the second picture. Pictures are visualized through features such as images and stereotypes. Other examples of using CNN, MLP, and KNN techniques in accurate brain tumor detection [20] and intelligent handwritten identification [21], as well as improving time-series forecasting [22].

As long as the network has been trained, it will be possible to retrieve healthy characteristic features from other images, and these can be incorporated in medical diagnosis, wherein it is justified whether a new image possesses certain pathology or is similar to healthy images.

The main objective of the above-mentioned process is to enhance the diagnosis process and ensure that the correct diagnosis techniques are provided for the various diseases or medical conditions.

The created signal by extracting healthy images of the ovary, therefore, depends on the image processing techniques and the data analysis used. A confluence of methodologies and methods is utilized to do the imaging, and then they turn the pictures into data that the nerve network can work with. Image enhancement techniques are employed to enhance image quality, locate the internal anatomy of the ovary, and perform histological differentiation of tissues, and this includes the application of visual correction techniques, enhancement of contrast, anti-orientation, direction, and engineering shifts if necessary. Figure 2. presents varies are healthy and unaffected while Figure 3. shows features of healthy ovaries

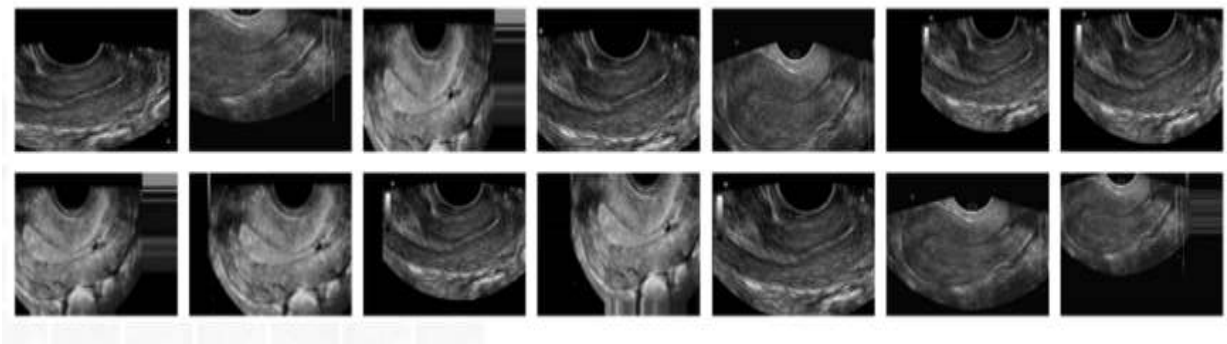


FIGURE 2. Ovaries are healthy and unaffected

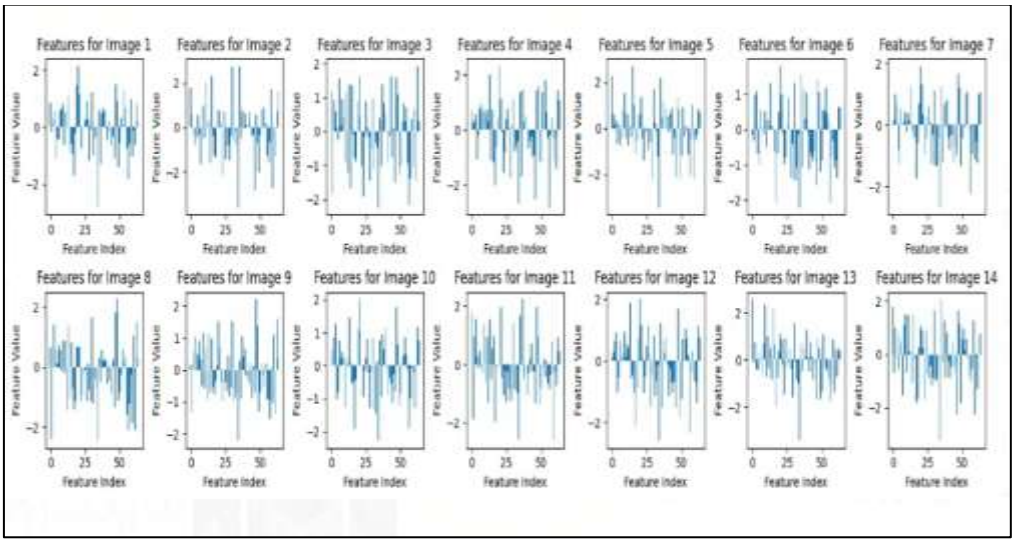



FIGURE 3. Features of healthy ovaries

After processing the image, the important features are extracted from it using image analysis techniques and the use of advanced algorithms [23], and these features include many biological and structural characteristics that we mentioned earlier, such as the distribution of natural blood vessels, the cellular structure, dimensions, and shapes. After extracting features, these features are fed to the nervous Siamese network. Siamese network compares the features extracted from the image that will be evaluated using pre-stored proper examples, and determines whether the image provided displays natural features, abnormalities, or abnormal changes, and the signal from the nerve network is either positive or negative. If the image is normal, A healthy signal is positive, indicating that the ovary is in its normal condition. In the

event of some abnormal changes, the signal will be negative, suggesting that there could be a health concern with the ovary and the system; therefore is capable of assessing ovarian images and identifying any abnormal changes that may be present, which assists in case diagnosis and rendering fast and reliable medical decisions.

Tscan ovarian (PCOS) – Index tests are done in active diagnoses of polycystic ovarian syndrome (PCOS), which is a clinical condition seen in women associated with abnormalities that include, among others, menses infertility, and hormonal imbalances. Ovarian images offer a crucial insight for the physician concerning the patient's state and the extent of the syndrome. While the ultrasound of the ovaries is performed, the doctor notes many clinical signs indicative of polycystic ovarian syndrome (PCOS) [24]:

- cysts: small bag like structures can also be observed on the surface of the ovary these small bags fill with fluid inside the ovarian structure.
- Enlargement: the girth may be increased in an abnormal manner by the presence of cysts and hormonal influences.
- Changes in the appearance of the white membrane (dandruff): changes in the ovarian white membrane may occur such as the presence of distortions or increased thinness.
- the persistent presence of underdeveloped Indications underdeveloped Indications within the images may suggest a menstrual irregularity..
- thick tissues: Thick ovarian tissues can indicate an abnormal accumulation of white membrane.

The ovaries help determine the diagnosis of PCOS, estimate its condition, and direct the appropriate medical steps to treat it. By carefully analyzing these images, doctors can better understand the patient's condition and provide the necessary health care accurately and effectively. Figure 4. presents varies affected by pcos.

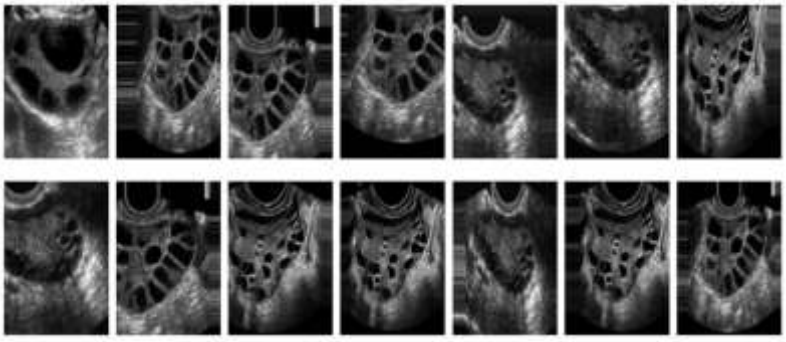


FIGURE 4. Ovaries affected by pcos

After using the Siamese network on ovarian images, the features of the PCOS can be extracted and these features are usually in the form of specific propy signals that may appear different from the regular ovary images and the features of the Siamese network may show duplicate patterns of hormonal patterns associated with PCOS, such as changes in the ratios of triple hormones (androgens, estrogen, progesterone) and the features created by the Siamese network can appear changes in the biological tissues of the ovary, such as the increase in the number of fillings or structural deformities, and the extracted features may indicate changes in blood levels or blood distribution in the ovary, which may indicate circulatory disorders related to syndrome Polycystic ovaries.

Features reveal an accurate analysis of tissues and cellular structures in the ovary, which may be significantly different in the presence of PCOS, and the features of a Siamese network may reveal ovarian tumors or deformities in the ovary, which may be an indication of the development of PCOS. In short, the Siamese network can accurately analyze different features that seem unusual in the presence of PCOS, which helps in diagnosing the disease early and identifying the appropriate treatment plan [25]. Fig 5. shows features of pcos ovaries

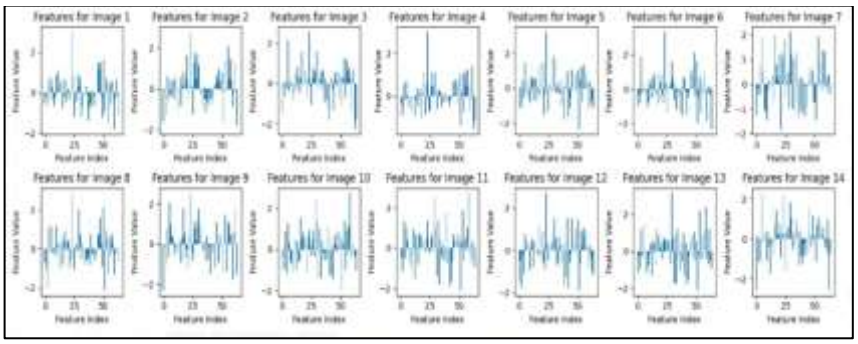


FIGURE 5. Features of pcos ovaries

The image obtained after running the PCOS images through a Siamese network can highlight certain prominent features that could be identified during the diagnosis process, and the following is a description of the peaks shown in the image;

- The first peak: This peak could represent an elevation in the levels of various hormones, for example, androgens or estrogens that are often elevated in PCOS, and thus highlights the hormonal signature of this disorder.
- The second peak: The second peak could correlate with the pathological characteristics of the ovarian tissues, such as the abnormality or inclusion of more adipose tissue, and the characteristic ovarian changes associated with PCOS.
- The third peak: The third peak may indicate the examination of the different histological levels of the ovary and alteration of the population of the ovarium, that is, a manifestation of vegetative changes occurring due to PCOS.
- The fourth peak: The fourth peak may indicate imaging evaluation of mass lesions or other abnormalities of the ovaries, and the peaks that may signal the presence of structural lesions or masses associated with PCOS.
- Fifth peak: The fifth peak can symbolize changes in the blood vessels or ovarian circulation, and changes may appear in the distribution of blood or the exudation of blood, indicating periodic changes in the ovary resulting from the ovarian cyst syndrome.

In general, the peaks in the image of the feature are unique points that have been discovered in the comprehensive analysis of pictures, and reflect the changes and features that can appear in the presence of PCOS. Fig 6. presents valuation and detection of PCOS

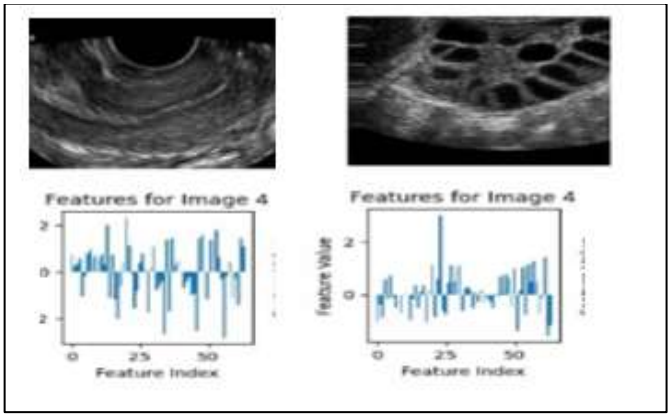



FIGURE 6. Evaluation and detection of PCOS

To evaluate and discover PCOS using peaks in the image of the feature, many techniques can be used to compare the summits. Pattern analysis techniques can be used to examine the general style of the summary in the image of the feature, and this is done using statistical analysis techniques to determine the general style of the training group and compare it to the expected pattern in the absence of the disease. Machine learning modeling can be applied to learn the characteristic changes in PCOS endotypes from images using the image peaks. Specifically, artificial nervous networks or complex statistics can establish characteristic features, and pathological changes the disease in question is associated with, or by clinical guidelines in evaluating peak features, identifying whether there is PCOS [26].

These can be based on prior knowledge of expected APEX changes in healthy individuals versus those with PCOS. Classification methods can also be employed to ascertain whether the current peaks are indicative of the feature associated with the presence of PCOS, and this is achieved by classifier learning with data that includes peaks as features and enables the classification of other data as containing or not containing PCOS syndrome [27].

It is possible to derive good conclusions concerning the presence of PCOS derived from the feature of the feature of peaks by comparing estimated peaks with the normal population values.

## 5. Discussion

The integration of machine learning (ML) models, such as the Siamese Neural Network (SNN), into the diagnostic process for Polycystic Ovary Syndrome (PCOS) has profound implications for modern medical practices. Traditional diagnostic methods rely heavily on clinical assessment, ultrasound imaging interpretation, and biochemical analyses, which are subject to human error and variability, and these

models provide a systematic and reproducible approach to evaluating ovarian features, enhancing diagnostic accuracy and consistency.

**5.1 Enhanced Accuracy and Consistency.** The results indicate that the SNN achieved high diagnostic accuracy by analyzing subtle features in ovarian ultrasound images and this improvement over traditional image interpretation could minimize false-negative and false-positive rates, ensuring timely interventions for patients at risk of PCOS-related complications, such as infertility or metabolic disorders and the capability of the model to detect early morphological changes in ovarian structure opens new avenues for proactive management. Early diagnosis allows for personalized therapeutic interventions, potentially mitigating long-term health risks associated with PCOS. By standardizing the diagnostic process through automated image analysis, ML models can reduce inter-observer variability among radiologists and gynecologists and this consistency is particularly crucial in settings where access to specialized expertise is limited and the results suggest that these models can seamlessly integrate into existing medical practices by acting as decision-support tools and they enhance the clinician's ability to interpret complex patterns in ultrasound images, enabling a more comprehensive and confident diagnosis.

**5.2 Improvement of Current Methods.** Current diagnostic methods for PCOS, including the Rotterdam criteria, often rely on subjective interpretation of ovarian features. ML models add objectivity by quantifying patterns and anomalies in a reproducible manner and automation can significantly reduce the time required for image analysis in busy clinical environments, allowing healthcare providers to focus on patient care rather than extensive manual evaluations and these models can integrate multimodal data, such as hormonal profiles and patient history, alongside ultrasound images and this holistic approach provides a more nuanced understanding of the condition, enhancing overall diagnostic accuracy.

**5.3 Statistical Methods for Validation.** To ensure unbiased evaluation of the model, the dataset was divided into training and testing subsets, typically using an 80:20 or 70:30 ratio and this split ensures that the model is evaluated on unseen data, simulating real-world diagnostic scenarios and the choice of split is critical to avoid overfitting and ensure generalizability.K-fold cross-validation was likely employed to assess the model's performance across multiple data subsets and this method divides the data into 'k' folds, using 'k-1' folds for training and one for testing iteratively, providing a robust estimate of model accuracy.

Metrics such as accuracy, sensitivity, specificity, and the area under the receiver operating characteristic curve (AUC-ROC) were calculated. Statistical tests like paired t-tests or Wilcoxon signed-rank tests may have been used to compare the performance of the ML model against traditional diagnostic methods. To evaluate the robustness of the model, performance metrics on external validation datasets or unseen populations were analyzed and this step is crucial to confirm that the model performs consistently across different demographic and clinical settings and thechniques such as SHAP (Shapley Additive Explanations) were potentially used to interpret the contribution of individual features to the model's predictions and this transparency aids in understanding how the model aligns with clinical reasoning. By addressing current diagnostic limitations, ML models like SNN represent a paradigm shift in PCOS diagnostics. However, rigorous statistical validation and transparent explanation of the model's functioning are imperative for its acceptance in clinical practice and the results of this study provide a strong foundation for integrating ML models into routine medical workflows, promising improved patient outcomes through early and accurate diagnosis.

## 6. Conclusion.

The conducted research advances the area of PCOS diagnosis by presenting a detailed and reasonable structure for the model development, and this early intervention diagnostic model offers hope in the advancement of early PCOS detection, resulting in fewer long-term effects arising from the condition. In addition, the rational and transparent decision-making process builds trust and reassurance in patients regarding the diagnosis to the benefit of patients and health care practitioners as well.

Revising and perfecting the existing diagnostic model will be important as its introduction into practice will be almost inevitable. Alongside this, the development of this model will also require the incorporation of additional data sources and other functions that will deal with the management of the condition (for instance, treatment compliance prediction/response, care planning) in order to enhance treatment outcomes for the patient.

Overall, this study shows how useful ML Technologies and Ensemble learning can be for the diagnosis of PCOS and stresses the need for transparency and trust in health care systems. Adopting these AI capabilities will pave the way for a new approach in medicine, one where individuals with PCOS and other complex conditions are given exquisite care.

## References

[1] R. J. Norman, D. Dewailly, R. S. Legro, and T. E. Hickey, "Polycystic ovary syndrome," (in eng), *Lancet,* vol. 370, no. 9588, pp. 685-97, Aug 25 2007, doi: 10.1016/s0140-6736(07)61345-2.
[2] S. Tiwari *et al.*, "SPOSDS: A smart Polycystic Ovary Syndrome diagnostic system using machine learning," *Expert Systems with Applications,* vol. 203, p. 117592, 2022/10/01/ 2022, doi: 10.1016/j.eswa.2022.117592.
[3] H. Ahmad and S. Umar, "Sentiment Analysis of Financial Textual data Using Machine Learning and Deep Learning Models," *Informatica,* vol. 47, pp. 153-158, 05/26 2023, doi: 10.31449/inf.v47i5.4673.
[4] A. Almulihi *et al.*, "Ensemble Learning Based on Hybrid Deep Learning Model for Heart Disease Early Prediction," *Diagnostics,* vol. 12, no. 12, p. 3215, 2022, doi: 10.3390/diagnostics12123215.
[5] S. U. Umar, T. A. Rashid, A. M. Ahmed, B. A. Hassan, and M. R. Baker, "Modified Bat Algorithm: a newly proposed approach for solving complex and real-world problems," *Soft Computing,* vol. 28, no. 13, pp. 7983-7998, 2024/07/01 2024, doi: 10.1007/s00500-024-09761-5.
[6] T. A. Rashid *et al.*, "An Improved BAT Algorithm for Solving Job Scheduling Problems in Hotels and Restaurants," in *Artificial Intelligence: Theory and Applications*, E. Pap Ed. Cham: Springer International Publishing, 2021, pp. 155-171.
[7] N. Thomas and A. Kavitha, "Prediction of Polycystic Ovarian Syndrome with Clinical Dataset Using a Novel Hybrid Data Mining Classification Technique," *International Journal of Advanced Research in Engineering and Technology,* vol. 11, no. 11, pp. 1872-1881, 2020, doi: 10.34218/IJARET.11.11.2020.174.
[8] H. Elmannai *et al.*, "Polycystic Ovary Syndrome Detection Machine Learning Model Based on Optimized Feature Selection and Explainable Artificial Intelligence," (in eng), *Diagnostics (Basel),* vol. 13, no. 8, Apr 21 2023, doi: 10.3390/diagnostics13081506.
[9] S. Bharati, P. Podder, and M. R. H. Mondal, "Diagnosis of Polycystic Ovary Syndrome Using Machine Learning Algorithms," in *2020 IEEE Region 10 Symposium (TENSYMP)*, Dhaka, Bangladesh, 5-7 June 2020 2020, pp. 1486-1489, doi: 10.1109/TENSYMP50017.2020.9230932.
[10] S. A. Suha and M. N. Islam, "An extended machine learning technique for polycystic ovary syndrome detection using ovary ultrasound image," *Scientific Reports,* vol. 12, no. 1, p. 17123, 2022/10/12 2022, doi: 10.1038/s41598-022-21724-0.
[11] Q. M. Yas, A. F. Mahdi, A. K. J. Al-Shamary, and N. S. Radam, "A Multi Criteria Analysis in Ranking Composite Material Using Gray Relational Analysis: A Case Study," in *2020 International Conference on Electrical, Communication, and Computer Engineering (ICECCE)*, 2020/6// 2020: IEEE, pp. 1-7, doi: 10.1109/ICECCE49384.2020.9179185.
[12] H. Danaei Mehr and H. Polat, "Diagnosis of polycystic ovary syndrome through different machine learning and feature selection techniques," *Health and Technology,* vol. 12, no. 1, pp. 137-150, 2022/01/01 2022, doi: 10.1007/s12553-021-00613-y.
[13] S. Alam Suha and M. N. Islam, "Exploring the dominant features and data-driven detection of polycystic ovary syndrome through modified stacking ensemble machine learning technique," *Heliyon,* vol. 9, no. 3, p. e14518, 2023/03/01/ 2023, doi: 10.1016/j.heliyon.2023.e14518.
[14] S. A. Bhat, "etection of Polycystic Ovary Syndrome using Machine Learning Algorithms," MSc, School of Computing, National College of Ireland, Dublin, 2021. [Online]. Available: https://norma.ncirl.ie/5137/1/shakoorahmadbhat.pdf
[15] A. Denny, A. Raj, A. Ashok, C. M. Ram, and R. George, "i-HOPE: Detection And Prediction System For Polycystic Ovary Syndrome (PCOS) Using Machine Learning Techniques," in *TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON)*, Kochi, India 17-20 Oct. 2019 2019, pp. 673-678, doi: 10.1109/TENCON.2019.8929674.
[16] H. Khalid, "Forecasting neural networks, such as forecasting sale the plastic injection machine market," *Journal of Physics: Conference Series,* vol. 1963, no. 1, pp. 012049-012049, 2021/7// 2021, doi: 10.1088/1742-6596/1963/1/012049.
[17] H. Khalid, "Efficient Image Annotation and Caption System Using Deep Convolutional Neural Networks," *BIO Web of Conferences,* vol. 97, no. 3, pp. 103-103, 2024, doi: 10.1051/bioconf/20249700103.
[18] H. Khalid, "Modern techniques in detecting, identifying and classifying fruits according to the developed machine learning algorithm," *Journal of Applied Research and Technology,* vol. 22, no. 2, pp. 219-229, 2024, doi: 10.22201/icat.24486736e.2024.22.2.2269.
[19] S. M. Khudaier and B. A. Mahmood, "A Review of Assured Data Deletion Security Techniques in Cloud Storage," *Iraqi Journal of Science,* pp. 2492-2511, 2023/5// 2023, doi: 10.24996/ijs.2023.64.5.33.
[20] Assaad, M. A., Saleh, M. I., & Mahrousseh, R. (2025). A Novel Framework for Accurate Brain Tumor Detection in MRI Scans Using CNN, MLP, and KNN Techniques. Journal of Information Hiding and Multimedia Signal Processing, *16*(2).
[21] F. Husari and M. Assaad, "Intelligent Handwritten Identification Using Novel Hybrid Convolutional Neural Networks – Long-short-term Memory Architecture," *Cihan University-Erbil Scientific journal,* vol. 8, no. 2, pp. 99-103, 2024 doi: 10.24086/issn.2519-6979.
[22] M. Assaad, R. Boné, and H. Cardot, "A new boosting algorithm for improved time-series forecasting with recurrent neural networks," *Information Fusion,* vol. 9, no. 1, pp. 41-55, 2008/01/01/ 2008, doi: 10.1016/j.inffus.2006.10.009.

[23] S. U. Umar and T. A. Rashid, "Critical analysis: bat algorithm-based investigation and application on several domains," *World Journal of Engineering,* vol. 18, no. 4, pp. 606-620, 2021, doi: 10.1108/WJE-10-2020-0495.
[24] Y. He *et al.*, "A Correlation-Based Feature Selection Algorithm for Operating Data of Nuclear Power Plants," *Science and Technology of Nuclear Installations,* vol. 2021, no. 1, p. 9994340, 2021, doi: 10.1155/2021/9994340.
[25] P. Chauhan, P. Patil, N. Rane, P. Raundale, and H. Kanakia, "Comparative Analysis of Machine Learning Algorithms for Prediction of PCOS," in *2021 International Conference on Communication information and Computing Technology (ICCICT)*, Mumbai, India 25-27 June 2021 2021, pp. 1-7, doi: 10.1109/ICCICT50803.2021.9510128.
[26] Y. Rathod *et al.*, "Predictive Analysis of Polycystic Ovarian Syndrome using CatBoost Algorithm," in *2022 IEEE Region 10 Symposium (TENSYMP)*, Mumbai, India 1-3 July 2022 2022, pp. 1-6, doi: 10.1109/TENSYMP54529.2022.9864439.
[27] V. V. Khanna, K. Chadaga, N. Sampathila, S. Prabhu, V. Bhandage, and G. K. Hegde, "A Distinctive Explainable Machine Learning Framework for Detection of Polycystic Ovary Syndrome," *Applied System Innovation,* vol. 6, no. 2, p. 32, 2023, doi: 10.3390/asi6020032.