Machine Learning Approaches on Diagnosis of Polycystic Ovary Syndrome

AKASH DONGARE[1], MAHESH KARANDE[2], SANKET GHONDGE[3], PRAJKTA VISHE[4]
[1, 2, 3, 4] *Dept. of Computer Engineering, Indala College of Engineering, Kalyan, India*
[5]*Dept. of Information Technology, Indala College of Engineering, Kalyan, India*

Abstract- ***PCOS is a complex hormonal condition that affects women who are of reproductive age. It is characterized by ovarian dysfunction and irregular menstrual cycles. For effective intervention and care to address possible hazards such infertility, metabolic syndrome, and cardiovascular consequences, timely detection and accurate diagnosis of PCOS are essential. Machine learning methods have emerged as useful instruments for evaluating intricate medical data and supporting the diagnosis and prognosis of illnesses in recent years. With an emphasis on predictive modeling and diagnostic accuracy, this paper examines the use of machine learning techniques in PCOS research. Data pertaining to PCOS, including clinical, hormonal, and imaging features, has been analyzed using a variety of machine learning techniques, such as logistic regression, SVM, random forests, neural networks, and ensemble approaches. To improve model performance and interpretability, feature selection strategies and data preprocessing approaches have also been applied. Potential remedies are explored for issues such class imbalance, data heterogeneity, and model interpretability. Additionally, there is potential for better PCOS diagnosis, risk assessment, and individualized treatment plans through the integration of multimodal data sources and the creation of interpretable machine learning models. Additionally mentioned are potential avenues for translational application of ML-based methods in clinical practice as well as future research directions.***

Indexed Terms- ***Machine Learning, SVM, RF, Decision Tree, Naive Bayes, CNN, Polycystic Ovary Syndrome (PCOS).***

I. INTRODUCTION

Ovarian cysts, irregular periods, and hormone imbalances are the hallmarks of PCOS, a common endocrine disorder that affects women in their reproductive years. Because of its complex etiology and varied presentation, PCOS is still difficult to diagnose, despite having a major influence on women's fertility and health. Clinical symptoms and biochemical assays are frequently the basis of traditional diagnostic techniques, which might result in subjective interpretation and possible misdiagnosis. Machine learning (ML) techniques have revolutionized medical diagnostics in recent years, offering increased efficacy and precision in PCOS identification. Researchers might find complex patterns and associations that could be missed by conventional diagnostic methods by using machine learning (ML) algorithms to examine clinical data. These algorithms, which range from more sophisticated techniques like Convolutional Neural Networks (CNN) to more conventional models like Support Vector Machines, Random Forests, and Decision Trees, have the potential to improve PCOS diagnosis through unbiased, data-driven analysis.

Using information from recent studies in the field, this study provides a thorough review and comparative analysis of machine learning techniques for PCOS diagnosis. We examine how well the SVM, RF, DT, Naive Bayes, and CNN algorithms distinguish PCOS patients from healthy people based on a range of clinical characteristics, including hormone levels, irregular menstruation, and metabolic markers.

The analysis includes results from a variety of research studies, such as those that look into automated ultrasound image segmentation methods, machine learning-based PCOS detection methods, and innovative strategies for early PCOS diagnosis using Extreme Gradient Boosting (XGBoost) algorithms.[1] By combining the knowledge from these studies, we hope to clarify the advantages and disadvantages of various machine learning algorithms in managing the intricacy of PCOS diagnosis and open the door for

further developments in non-invasive, data-driven diagnostic techniques.

With our research, we hope to add to the growing body of knowledge on PCOS diagnosis by shedding light on how ML approaches could transform clinical practice and enhance the quality of care for women with this complex condition. In order to improve diagnostic efficiency and accuracy for improved patient outcomes, the study looks at the Random Forest, SVM, Naive Bayes, CART, and Logistic Regression algorithms for PCOS diagnosis utilizing patient clinical data [2].

The paper presents an investigation into the application of various machine learning approaches to establish an efficient decision tree for PCOS detection. By utilizing ML algorithms like Support Vector Machines, K-Nearest Neighbours, Naïve Classifier, and Random Forest, the study seeks to enhance PCOS performance and offer precise detection methods.

The study assesses how well these machine learning approaches detect PCOS by analyzing clinical data gathered from various institutions, including characteristics linked to irregular menstrual cycles, hormone levels, BMI, and other physiological parameters[3].

Early PCOS detection is essential for avoiding problems, however the existing diagnostic techniques are not objective. In order to solve this, we suggest an automated PCOS diagnostic method that uses ultrasound pictures to examine follicle characteristics such as size, location, and quantity. Using a KNN classifier, our approach achieves over 97% accuracy in PCOS classification, greatly cutting down on time and increasing accuracy when compared to manual diagnosis. For improvements in healthcare, early detection and prevention of PCOS-related problems are essential [4].

The subjective character of existing diagnostic techniques makes it difficult to detect and predict PCOS in its early stages. In response, this study suggests a novel method for creating an automated PCOS diagnostic tool that uses machine learning algorithms and ultrasound imaging data. Our approach attempts to provide precise and impartial evaluations of PCOS status by examining follicle parameters including number, size, and location. This system's potential is found in its capacity to improve early detection, which will enable prompt intervention and individualized treatment plans to lessen PCOS-related problems [5].

The purpose of this research study is to examine the current literature and technological developments related to the use of machine learning (ML) techniques for the diagnosis of polycystic ovarian syndrome (PCOS). It emphasizes the wide range of methods, difficulties, and possibilities in this field. In addition, the study addresses the shortcomings of existing approaches by suggesting a novel way for PCOS identification that makes use of cutting-edge machine learning algorithms. In order to diagnose and treat PCOS more effectively and sustainably, the proposed approach places a strong emphasis on integrating real-time data collecting, processing, and data-driven decision-making.

II. LITERATURE REVIEW

Recent research has looked into cutting-edge techniques for PCOS diagnosis using automated systems and machine learning algorithms. The goals of these initiatives have been to increase early detection, boost precision, and expedite the diagnostic procedure. A comparative analysis of machine learning methods for PCOS diagnosis using clinical data was conducted by Hassan and Mirza, and the results were published in the International Journal of Computer Applications. Infertility, irregular menstruation, acne, and an elevated body mass index are all signs of PCOS, a hormonal imbalance that is common in women of reproductive age.

The study aimed to compare the performance of five machine learning algorithms: Logistic Regression, Support Vector Machine , Naïve Bayes, Classification and Regression Trees (CART), and Random Forest were utilized. 42 independent variables that indicate PCOS symptoms were included in the pre-processed and randomly selected samples of data from 10 hospitals in Kerala, India. The results showed that SVM had a 95% accuracy rate in detecting PCOS, whereas Random Forest had the greatest accuracy of 96%. Additionally, Random Forest fared better than

other algorithms in terms of F1-score, recall, and precision.The results demonstrate how machine learning (ML) can be used in healthcare systems to accurately diagnose and treat complicated conditions like PCOS. Larger datasets may be investigated in future studies to increase diagnosis accuracy [2].

Used a Kaggle dataset and machine learning algorithms to build a data-driven approach for PCOS diagnosis. found the FSH/LH ratio to be a significant predictor using univariate feature selection. used classifiers, with the RFLR hybrid model obtaining a noteworthy testing accuracy of 91.01%.

Building on this framework, [3] examined methods for detecting PCOS by examining clinical data gathered from several institutions. The most successful classifier was Random Forest, which achieved 93.5% accuracy. The significance of early detection and monitoring was emphasized by the identification of notable characteristics for PCOS detection, such as BMI and hormone levels.

Similarly, Paper 3 used machine learning techniques on a Kaggle dataset to increase the accuracy of PCOS detection. The study used the Cat. Boost algorithm in conjunction with Chi-Square feature selection to reach an impressive accuracy rate of 93.9% by utilizing preprocessing and feature selection techniques. The effectiveness of machine learning for non-invasive PCOS diagnosis was confirmed by this study. [6].

Furthermore, [7] proposed an automated technique for detecting PCOS from ultrasound pictures by combining machine learning algorithms with image preprocessing, segmentation, and classification. With an astounding 98% accuracy rate, the hybrid approach outperformed individual algorithms and provided a viable path for early detection.

The Extreme Gradient Boosting (XG Boost) technique was used to address the problem of early PCOS identification in a novel way [1]. The suggested approach outperformed the current classifiers by correcting class imbalance and choosing crucial parameters using statistical testing and resampling.

[8] The study examines a variety of PCOS-related topics and suggests possible directions for further research. Changes in steroidogenesis, the importance of adrenal androgen production, defects in certain steroidogenic enzymes, the mechanism of follicle arrest, the intra-ovarian regulation of ovarian morphology, the impact of metabolic abnormalities, dietary factors, inflammation, chronic infections, and long-term consequences like cardiovascular disease are among the topics covered. The authors stress the necessity of thorough research to comprehend the pathophysiology of PCOS, its risk factors, and its long-term effects. To improve individualized treatment plans and clinical techniques.

An automated method for identifying Polycystic Ovary Syndrome (PCOS) from ultrasound pictures is presented in this research. Conventional manual diagnosis techniques count follicular cysts, which causes problems with efficiency, reproducibility, and variability. The suggested method uses a clustering technique, a modified labeled watershed algorithm, and an adaptive morphological filter to address this. The clustering method detects follicular cysts, the algorithm recovers target contours, and the filter minimizes speckle noise. The scheme's efficacy is demonstrated by experimental findings, which show an accuracy rate of 84% when compared to manual approaches. Although its applicability to other multiple targets detection problems may be restricted, this automated approach shows promise in addressing the difficulties associated with manual diagnosis [10].

The article [11] presenting a brand-new technique for segmenting ultrasound images automatically to help diagnose PCOS. It overcomes the difficulties caused by speckle noise by combining the Chan-Vese technique with Otsu's thresholding. Before using Otsu's thresholding and the Chan-Vese method for segmentation, preprocessing with a median filter lowers noise. When compared to the traditional Chan-Vese approach, comparative analysis shows better performance in terms of accuracy, dice score, and jaccard value. These findings highlight how image processing methods can expedite the diagnosis of PCOS and point to directions for further study to improve segmentation accuracy and speed without the need for human participation.

TABLE I. SUMMARY OF LITERATURE REVIEW

Authors	Major Findings & Outcomes
M. S. Khan Inan., 2021	-A benchmark dataset experiment shows that the suggested model outperforms other classifiers with a 10-fold cross-validation score of 96.03% and a recall rate of 98%. -When compared to the most advanced techniques, the XG Boost model performs better with greater accuracy and precision.
M. Mubasher-Hassan., 2020	- The study aimed to evaluate algorithm performance in accuracy, precision, recall, and F-statistics. Findings indicated that Random Forest achieved the highest accuracy of 96% in diagnosing PCOS, followed by SVM with 95% accuracy.
A. S. Prapty., 2020	-Findings indicate that Random Forest achieved the highest accuracy of 93.5%, closely followed by the Naive Classifier at 93%. - Key attributes for PCOS detection were identified, including BMI, hair growth, menstrual cycle length, and hormonal levels.
B. Rachana., 2021	-By analyzing ultrasound images and extracting geometric features, such as the number of follicles and their characteristics, the system achieves a classification accuracy of over 97%.
S. Bharati., 2020	- Results demonstrate that the top 10 features are sufficient for accurate PCOS prediction, with RFLR exhibiting the highest testing accuracy of 91.01% and a recall value of 90% using 40-fold cross-validation.
A. Denny., 2019	- Various machine learning algorithms are trained and evaluated using K-fold cross- validation, with Cat Boost algorithm paired with Chi-Square feature selection achieving the highest accuracy rate of 93.9%.
J. Madhumitha., 2021	- Three classification algorithms, SVM, KNN, and Logistic Regression, are utilized. - The proposed hybrid method achieves an impressive accuracy of 98%, outperforming individual algorithms. -Tested YOLOv7 algorithm
Irba Fairuz Thufailah., 2018	- Proposing a deep learning model for PCOS diagnosis from ultrasound images, achieving 84.81% accuracy with the Inception model. -Moreover, a fusion model integrating ultrasound images with clinical data is presented, achieving an accuracy of 82.46%. [12].
Yinhui Deng., 2008	-Experimental results demonstrate the scheme's effectiveness, achieving an 84% accuracy rate compared to manual methods.

III. METHODOLOGY

Our research approach is focused on the diagnosis and treatment of Polycystic Ovarian Syndrome (PCOS), with a specific emphasis on the integration of Machine Learning (ML) techniques with healthcare technologies. Medical data collection equipment and sophisticated machine learning algorithms are the main parts of our solution.

A. Model Workflow

Data Collection:

Gather clinical data from patients diagnosed with or suspected of having PCOS. Collect relevant attributes such as age, weight, hormonal levels, menstrual cycle characteristics, and other symptoms.

Data Preprocessing:

Preprocess the dataset by addressing missing values, outliers, and noise. Normalize or standardize the data for uniformity across features.

Feature Selection:

Identify the most informative features for PCOS detection. Utilize domain knowledge and feature selection techniques to choose relevant attributes.

TABLE II. FEATURE TABLE

Dependent Variable	Value
PCOS	0-No,1-Yes
Independent Variables	Value
Age	Capable of accepting valid numerical inputs.
Weight	Capable of accepting valid numerical inputs.
Height	Capable of accepting valid numerical inputs.
Pulserate(bpm)	Capable of accepting valid numerical inputs.
RR (breaths/min)	Capable of accepting valid numerical inputs.
Hb(g/dl)	Capable of accepting valid numerical inputs.
Cycle length(days)	Capable of accepting valid numerical inputs.
FollicalNo	Capable of accepting valid numerical inputs.
Endometrium	Capable of accepting valid numerical inputs.

Model Training:
-SVM, DT, RF and Naive Bayes:
Divide the pre-processed data into training and testing subsets. Train model via the training data and optimize hyperparameters through methods like cross-validation.
-CNN:
Prepare the medical imaging data (e.g., ultrasound images of ovaries).
Divide dataset as training, testing and validation sets.
Develop and train the CNN model with TensorFlow or PyTorch frameworks.

Model Evaluation:
Evaluate each model's performance using metrics like F1-score, recall, accuracy, and precision. Make sure the models are robust by using cross-validation. To ascertain the efficacy of various algorithms in PCOS identification, compare the outcomes.

Model Selection:
Choose the model with the highest performance metrics as the optimal PCOS detection algorithm.
Consider factors such as accuracy, computational efficiency, and interpretability in the selection process.

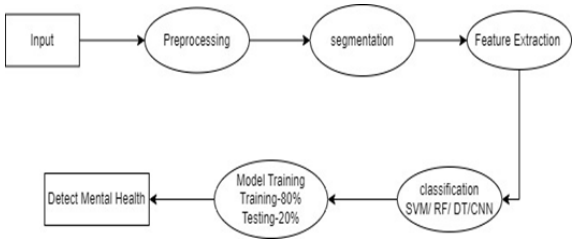

Fig. 1. Block Diagram of Workflow

IV. SYSTEM DESIGN

A. System Architecture

We solve the class imbalance problem during the preprocessing stage by generating a balanced dataset that ensures each class is fairly represented. To achieve this, the overrepresented class's size must be decreased to equal that of the underrepresented class. To assess the model, the dataset is split into a testing set (20%) and a training set (80%).

The key components of the suggested system design are shown in Fig. 2. First of all, it includes applying different machine learning techniques to assess their efficacy. Second, it entails determining which characteristics have the greatest influence on PCOS diagnosis.

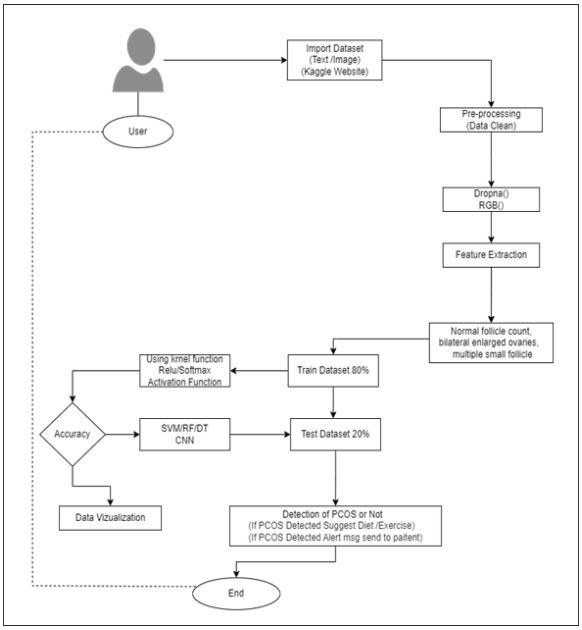

Fig. 2. System Architecture

This architecture facilitates a systematic approach to model development and feature analysis, ensuring robustness and accuracy in PCOS detection.

B. SVM Algorithm

SVM, a supervised learning method, is utilized for classification, determining the optimal hyperplane to separate data points into different classes. In PCOS detection, SVM utilizes clinical data features to classify patients into PCOS- positive or PCOS-negative categories by maximizing the margin between classes.

The study utilized Support Vector Machine (SVM) along with other machine learning classifiers to predict Polycystic Ovary Syndrome (PCOS) in fertile patients. SVM is recognized for its efficacy in classification tasks, especially in high- dimensional data scenarios. Employing an open-source dataset of 541 patients from Kerala, India, SVM was integrated into a heterogeneous ensemble comprising both machine learning and deep learning models.[15].

C. Decision Trees (DT)

A supervised learning technique used for regression and classification is the decision tree. It divides the data recursively according to attribute values, creating a structure resembling a tree. The method determines which property best divides the data at each node. In order to categorize patients into PCOS-positive or PCOS-negative groups according to their symptoms, DT recursively divides clinical data features in PCOS detection.

The Decision Tree classifier is used as part of the PCOS detection methods in the given paper [18]. The text does not, however, include an explicit reference for the Decision Tree classifier. Because decision tree techniques are widely used in classification tasks, the reference for their use in PCOS diagnosis can be found in general machine learning and decision tree literature.

D. Random Forest (RF)

During training, Random Forest, an ensemble learning method, combines many decision trees. A random selection of characteristics and data is used to train each tree in the forest. To get the final result in classification, RF combines the predictions of several trees. An ensemble of decision trees is used by RF in PCOS detection to increase the precision and resilience of patient classification using clinical data.
The Random Forest (RF) technique was used in the paper [16] investigation to find important gene biomarkers for the diagnosis of polycystic ovarian syndrome (PCOS). From a pool of 264 differentially expressed genes (DEGs), RF analysis revealed 12 important genes. This establishes the basis for developing a novel diagnostic model known as neural PCOS. Promising results from the neural PCOS model's validation suggested that it might be useful in enhancing PCOS diagnosis.

E. Naive Bayes

Naive Bayes is a probabilistic classification algorithm grounded in Bayes' theorem, assuming feature independence and rendering its computational efficiency. It calculates the probability of each class given input data and selects the class with the highest probability. In PCOS detection, Naive Bayes estimates the likelihood of a patient having PCOS based on observed clinical symptoms and their probabilistic relationships.

F. Convolutional Neural Network (CNN)

CNN, a prevalent deep learning architecture, is often employed for image classification, including PCOS detection using medical imaging data like ultrasound images of ovaries. The network comprises convolutional layers for extracting features and pooling layers for reducing dimensionality. By learning hierarchical representations of input images, CNN effectively distinguishes between PCOS-positive and PCOS-negative cases.

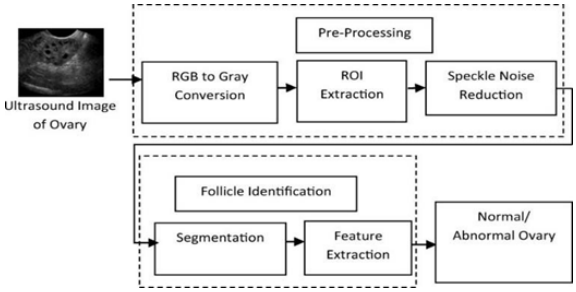

Fig. 3. Block Diagram for CNN

The study shows how well CNN and deep learning models can identify abnormalities in datasets. The

poor performance of the DenseNet201 model on Dataset B, however, points to issues with picture complexity or data scarcity. It is recommended to do additional experiments, such as training the entire model and investigating other learning strategies. It's critical to recognize that model accuracy and reliability are strongly impacted by data quality, especially in medical settings where inaccurate data might have detrimental effects [17].

Maintaining high standards of data quality in medical deep learning systems requires constant observation and development. To address the problems of inconsistency, consistency, and effectiveness in manual cyst counting, an automated detection system is presented. An adaptive morphological filter is first applied to the ovary ultrasound picture. A modified labeled watershed algorithm is then used to extract target outlines, and a clustering technique is used to find follicular cysts. Its efficacy is demonstrated by experimental validation, which achieved an accuracy rate of 84% [9].

V. COMPARISON

To assess the efficacy of the four implemented methods, we analyzed their efficiency via key metrics like accuracy, precision, recall, and F1 score. This comparative analysis aids in identifying the most appropriate method for our application.

TABLE II. PERFORMANCE COMPARISON

	Accuracy	Precision		Recall		F1-Score	
		0	1	0	1	0	1
SVM	85.54	0.87	0.83	0.93	0.7	0.9	0.76
Decision Tree	99.30	1	0.98	0.98	1	0.98	0.99
Random Forest	99.54	0.98	1	1	0.98	1	0.99
Naive Bayes	86.82	0.92	0.78	0.9	0.9	0.91	0.78

VI. EXPERIMANT RESULT

We put four methods into practice for the textual dataset based on the experimental findings. Convolutional Neural Networks (CNN) produced an accuracy of 97%, while Random Forests (RF) obtained the maximum accuracy of 99.53%. For model evaluation, we gave priority to the F1 score even though SVM, Naive Bayes, and Random Forest displayed nearly identical accuracies.

Fig. 4. Data Processing

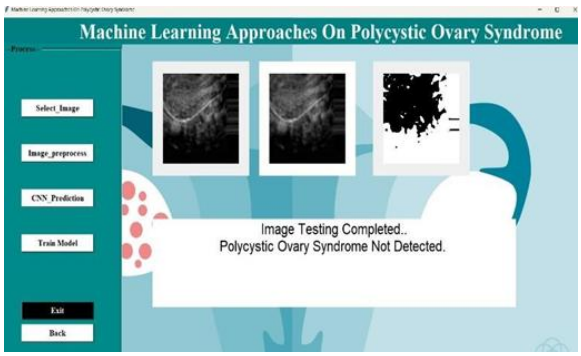

Fig. 5. Image Processing using CNN

In Fig.5, Image processing undergoes in these stages. These stages aim to enhance the contrast between noise and essential features, facilitating the segmentation of the region of interest, particularly for identifying follicles. A suitable threshold value is then applied to select the appropriate region, resulting in the creation of a precise mask region. Fig also depicts the final outcomes achieved by the model.

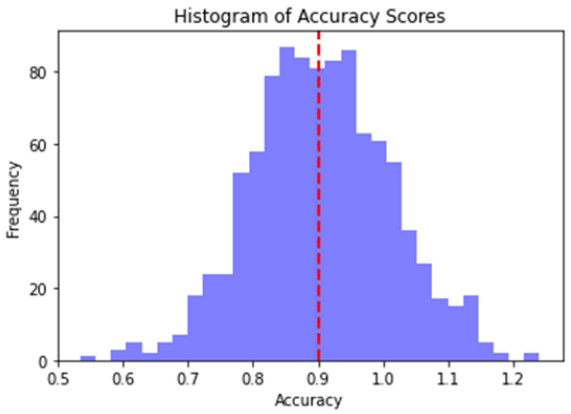

Fig. 6. Histogram of Accuracy Scores

CONCLUSION

In this study, we evaluated the performance of five different machine learning algorithms—Support Vector Machine, Random Forest, Decision Tree, Naive Bayes, and Convolutional Neural Network (CNN)—in detecting PCOS from textual data. These algorithms' respective accuracy ratings were SVM (85.53%), RF (99.5%), DT (99.3%), and NB (86.38%). Outperforming the other conventional machine learning models, DT demonstrated the highest accuracy among them. It's crucial to acknowledge CNN's noteworthy contribution to the work, which highlights the potential of deep learning techniques for PCOS identification.

These findings highlight the need of early PCOS detection and treatment customization, giving people and healthcare professionals important information about symptoms and practical management techniques. Better patient and healthcare provider empowerment is made possible by this strategy, which enhances patient wellbeing and healthcare results.

REFERENCES

[1] M. S. Khan Inan, R. E. Ulfath, F. I. Alam, F. K. Bappee, and R. Hasan, "Improved Sampling and Feature Selection to Support Extreme Gradient Boosting For PCOS Diagnosis," 2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC).

[2] M. Mubasher Hassan and T. Mirza, "Comparative Analysis of Machine Learning Algorithms in Diagnosis of Polycystic Ovarian Syndrome," International Journal of Computer Applications, vol. 175, no. 17, pp. 42–53, Sep. 2020, doi: https://doi.org/10.5120/ijca2020920688.

[3] A. S. Prapty and T. T. Shitu, "An Efficient Decision Tree Establishment and Performance Analysis with Different Machine Learning Approaches on Polycystic Ovary Syndrome," 2020 23rd International Conference on Computer and Information Technology (ICCIT), Dec. 2020, doi: https://doi.org/10.1109/iccit51783.2020.9392666.

[4] B. Rachana, T. Priyanka, K. N. Sahana, T. R. Supritha, B. D. Parameshachari, and R. Sunitha, "Detection of polycystic ovarian syndrome using follicle recognition technique," Global Transitions Proceedings, vol. 2, no. 2, pp. 304–308, Nov. 2021, doi: https://doi.org/10.1016/j.gltp.2021.08.010

[5] S. Bharati, P. Podder, and M. R. H. Mondal, "Diagnosis of Polycystic Ovary Syndrome Using Machine Learning Algorithms," IEEE Xplore, Jun. 01, 2020. https://ieeexplore.ieee.org/document/9230932

[6] 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," IEEE Xplore, Oct. 01, 2019. https://ieeexplore.ieee.org/document/8929674

[7] J. Madhumitha, M. Kalaiyarasi, and S. S. Ram, "Automated Polycystic Ovarian Syndrome Identification with Follicle Recognition," 2021 3rd International Conference on Signal Processing and Communication (ICPSC), May 2021, doi: https://doi.org/10.1109/icspc51351.2021.9451720.

[8] R. Pasquali et al., "PCOS Forum: research in polycystic ovary syndrome today and tomorrow," Clinical Endocrinology, vol. 74, no. 4, pp. 424– 433, Mar. 2011, doi: https://doi.org/10.1111/j.1365-2265.2010.03956.x.

[9] Yinhui Deng, Yuanyuan Wang, and Ping Chen, "Automated detection of Polycystic Ovary Syndrome from ultrasound images," 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Aug. 2008, doi: https://doi.org/10.1109/iembs.2008.4650280.

[10] Y. Deng, Y. Wang, and Y. Shen, "An automated diagnostic system of polycystic ovary syndrome based on object growing," Artificial Intelligence in Medicine, vol. 51, no. 3, pp. 199–209, Mar. 2011, doi: https://doi.org/10.1016/j.artmed.2010.10.002.

[11] Asma Amirah Nazarudin, Noraishikin Zulkarnain, Siti Salasiah Mokri,

[12] W. Mimi, and A. Hussain, "An implementation of Otsu thresholding and the Chan–Vese method on the PCO segmentation of ultrasound images," Nov. 2020, doi: https://doi.org/10.1109/icecie50279.2020.9309595.

[13] Irba Fairuz Thufailah, Adiwijaya Adiwijaya, Untari Novia Wisesty, and Jondri, "An implementation of Elman neural network for polycystic ovary classification based on ultrasound images," Journal of physics, vol. 971, pp. 012016–012016, Mar. 2018, doi: https://doi.org/10.1088/1742-6596/971/1/012016.

[14] H. Elmannai et al., "Polycystic Ovary Syndrome Detection Machine Learning Model Based on Optimized Feature Selection and Explainable Artificial Intelligence," Diagnostics, vol. 13, no. 8, p. 1506, Apr. 2023, doi: https://doi.org/10.3390/diagnostics13081506.

[15] Zahra Zad et al., "Predicting polycystic ovary syndrome (PCOS) with machine learning algorithms from electronic health records," medRxiv (Cold Spring Harbor Laboratory), Aug. 2023, doi: https://doi.org/10.1101/2023.07.27.23293255.

[16] 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, Apr. 2023, doi: https://doi.org/10.3390/asi6020032.

[17] N.-N. Xie, F.-F. Wang, J. Zhou, C. Liu, and F. Qu, "Establishment and Analysis of a Combined Diagnostic Model of Polycystic Ovary Syndrome with Random Forest and Artificial Neural Network," BioMed Research International, vol. 2020, p. e2613091, Aug. 2020, doi: https://doi.org/10.1155/2020/2613091.

[18] Alsibai Mohammed Hayyan and Heydari Seyed Youssef, "The Effectiveness of CNN in Evaluating Ultrasound Image Datasets: Diagnosing Polycystic Ovary Syndrome (PCOS) as an Example," Zenodo (CERN European Organization for Nuclear Research), Apr. 2023, doi: https://doi.org/10.5281/zenodo.8121429.

[19] C. Chan, "Polycystic Ovarian Syndrome (PCOS) Detection using Gradient Boosted Decision Tree," INDIGO (University of Illinois at Chicago), Jul. 2023, doi: https://doi.org/10.36227/techrxiv.23803989.v1.