WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

A Comprehensive Review of Polycystic Ovary Syndrome (PCOS): Causes, Diagnosis, Lifestyle, Emerging Therapies, and Explainable AI-Based Detection

Mrs. Gorisetty Nirosha [1], Gembali Sruthi [2]

Department of Computer Science and Engineering
GMRIT Deemed to be University, Rajam, Andhra Pradesh, India

ABSTRACT :

Polycystic ovary syndrome (PCOS) is a heterogeneous endocrine and metabolic condition associated with ovulatory dysfunction, clinical or biochemical hyperandrogenism, and polycystic ovarian morphology. Its variable presentation and overlap with other disorders can delay recognition, while reproductive, metabolic, and psychological concerns require coordinated long-term care. This review synthesizes the clinical background and artificial-intelligence approaches described in the accompanying term-paper report and presentation. It examines conventional assessment, lifestyle and medical management themes, machine-learning methods for structured clinical data, deep-learning analysis of ovarian ultrasound images, and explainable AI. Published examples include a web-based machine-learning study using 541 records that reported 94% accuracy after mutual-information feature selection, and a 2026 multimodal study combining ultrasound and clinical attributes through cross-attention. Building on these themes, an explainable hybrid framework is outlined in which clinical and imaging streams are processed separately, fused for decision support, and interpreted using SHAP/LIME and Grad-CAM. The source material also contains a Random Forest training script using nine clinical variables and an 80:20 stratified split; however, it does not provide its own test metrics or demonstrate a validated multimodal model. Accordingly, no new experimental performance is claimed. The review identifies dataset quality, external validation, interpretability, privacy, and clinician oversight as essential requirements before clinical deployment.

Keywords— *Polycystic ovary syndrome, PCOS, machine learning, ultrasound imaging, multimodal learning, explainable artificial intelligence, clinical decision support.*

1. Introduction

Polycystic ovary syndrome is a common and clinically heterogeneous endocrine disorder affecting reproductive-aged women. The syndrome is associated with combinations of ovulatory dysfunction, androgen excess, and polycystic ovarian morphology. Clinical manifestations may include irregular menstrual cycles, hirsutism, acne, scalp hair thinning, and difficulty conceiving; metabolic features such as insulin resistance and weight-related complications may also occur. The presence and severity of these features differ between individuals, so PCOS cannot be represented by a single symptom or a single laboratory value (Bai et al., 2024; Sirmans & Pate, 2014).

The diagnostic process requires assessment of clinical history, physical findings, biochemical evidence where appropriate, and imaging when indicated. The 2023 international evidence-based guideline emphasizes a structured diagnostic approach and broader attention to metabolic health, psychological wellbeing, and patient-centred care (Teede et al., 2023). The Rotterdam framework, commonly referenced in the supplied term-paper materials, considers three domains: hyperandrogenism, ovulatory dysfunction, and polycystic ovarian morphology. In adults, two of these three features may support diagnosis after relevant alternative causes have been excluded. Diagnostic decisions remain the responsibility of qualified healthcare professionals.

Artificial intelligence (AI) and machine learning (ML) have become relevant to PCOS research because clinical datasets and ultrasound images contain patterns that may be difficult to assess consistently through manual review alone. Existing studies have explored conventional classifiers, ensemble models, convolutional neural networks, transfer learning, and explainability methods (Ahmed et al., 2023; Elmannai et al., 2023). These systems are best viewed as potential decision-support tools rather than substitutes for clinical diagnosis.

1.1 Scope and contribution

This paper reviews the clinical and computational themes present in the supplied term-paper report and presentation. Its principal contribution is a structured synthesis of clinical-data ML, ultrasound-based deep learning, multimodal fusion, and explainable AI, followed by a proposed hybrid framework. It is a review and design-oriented paper; it does not claim a completed clinical trial or a newly validated diagnostic model.

2. Literature Review

2.1 Clinical features and diagnostic assessment

PCOS has a multifactorial aetiology involving genetic susceptibility, endocrine regulation, insulin resistance, and environmental or lifestyle influences. Insulin resistance may be present across different tissues and can interact with ovarian androgen production and metabolic risk, although its expression varies among patients (Zhao et al., 2023). Clinical assessment therefore needs to consider reproductive, metabolic, and psychological dimensions rather than focusing exclusively on ovarian appearance.

The international guideline recommends evidence-based assessment and individualized management. It also recognizes the importance of lifestyle intervention, emotional wellbeing, and quality of life. No single dietary composition has been established as universally superior for all PCOS outcomes; sustainable nutrition and physical-activity plans should be tailored to individual needs and preferences (Teede et al., 2023). Reviews of diagnosis and management similarly describe the need to consider associated metabolic and reproductive comorbidities (Choudhari et al., 2024; Ghafari et al., 2025).

2.2 Lifestyle and therapeutic management

Lifestyle management is a foundational component of PCOS care. It may include balanced eating, regular physical activity, behavioural support, and weight-management strategies where appropriate. The clinical objective is not limited to weight reduction: metabolic health, prevention of excess weight gain, general wellbeing, and sustainable behaviour change are also important. Pharmacological treatment is selected according to the patient's symptoms, metabolic profile, reproductive goals, contraindications, and clinical assessment. Metformin and other therapies have been examined in systematic reviews and updated clinical literature, but treatment selection must remain clinician-led (Melin et al., 2023; Ghafari et al., 2025).

Emerging approaches discussed in the supplied report include GLP-1 receptor agonists, SGLT2 inhibitors, probiotics, novel biomarkers, and phenotype-oriented treatment. These areas are evolving and should not be interpreted as universally established first-line therapies. Further evidence is needed on long-term effectiveness, safety, patient selection, and reproductive outcomes (Choudhari et al., 2024; Ghafari et al., 2025).

2.3 Machine learning for clinical data

Clinical ML methods use structured attributes such as age, anthropometric measurements, menstrual-cycle characteristics, symptoms, and selected biochemical variables. Algorithms discussed in the supplied materials include Logistic Regression, Support Vector Machine (SVM), Random Forest, K-Nearest Neighbours (KNN), Decision Tree, Naïve Bayes, AdaBoost, and XGBoost. Feature-selection methods can reduce redundant inputs and may improve model efficiency. Elmannai et al. (2023) examined optimized feature selection and explainable ML, while Rahman et al. (2024) reported a web-based prediction approach using a 541-record dataset.

2.4 Deep learning for ultrasound images

Deep-learning methods can learn image representations directly from ovarian ultrasound scans. Convolutional Neural Networks (CNNs) and transfer-learning backbones such as ResNet, DenseNet, and EfficientNet are described in the supplied presentation. Image enhancement, resizing, augmentation, and class balancing are commonly considered during model development. However, ultrasound image quality, acquisition protocols, class imbalance, and the representativeness of training data can affect performance. High internal test scores do not by themselves establish generalization to other hospitals or populations (Ahmed et al., 2023).

2.5 Explainable and multimodal AI

Explainable AI (XAI) methods seek to make model behaviour more interpretable. SHAP and LIME can be used to describe the influence of clinical variables on an individual prediction or across a dataset. Grad-CAM and Grad-CAM++ can highlight image regions that contribute to a CNN output. Attention maps may provide additional information about the clinical attributes emphasized by a multimodal model. These explanations are aids to review, not proof that a model has learned valid clinical reasoning.

Multimodal learning combines complementary data sources, for example structured clinical attributes and ultrasound images. Lakshmi and Pushpa (2026) describe a cross-attention approach that integrates a ResNet50 image stream with a Transformer-based clinical stream. Such architectures may represent relationships between modalities more effectively than simply using either source alone, but they require carefully matched patient-level data and independent validation.

Representative studies discussed in the supplied documents and verified bibliographic sources. Reported results belong to the cited studies and are not results generated by the present review.

Ahmed et al. (2023). Review of ML and deep-learning approaches for PCOS detection; highlights dataset heterogeneity and the need for broader validation.

Rahman et al. (2024). Web-based ML prediction using 541 patient records and mutual-information feature selection; performance remains dataset-specific.

Elmannai et al. (2023). Feature selection, ML ensembles, and explainable AI for clinical PCOS prediction; benchmark results do not alone establish clinical utility.

Maheswari and Maheswari (2024). Feature-driven ML with explainability in SmartScanPCOS; external validation and population diversity remain important.

Akter and Reno (2025). Weighted ensemble, feature selection, and geographic validation; sensitivity and cohort differences should be considered alongside accuracy.
Lakshmi and Pushpa (2026). Cross-attention fusion of ultrasound and clinical attributes with Grad-CAM++ and attention maps; matched multimodal data and external validation are critical.

2.6 Clinical dimensions, diagnostic considerations, and associated risks

The clinical assessment of PCOS needs to consider reproductive, metabolic, and psychological dimensions together. The presentation identifies menstrual irregularity, ovulatory dysfunction, androgen-related manifestations, insulin resistance, weight-related concerns, fertility difficulties, and longer-term metabolic risks as relevant areas. These features do not occur with the same intensity in every individual; therefore, a single symptom or isolated measurement should not be treated as sufficient evidence of PCOS.
The Rotterdam framework described in the source material considers three diagnostic domains: ovulatory dysfunction, hyperandrogenism, and polycystic ovarian morphology. In adults, the presence of two of these three features may support diagnosis after relevant alternative causes have been excluded. Clinical history, examination, biochemical evaluation, and ultrasound findings should be interpreted together by qualified healthcare professionals. The paper does not propose replacing this clinical assessment with an automated prediction.
The reviewed literature also associates PCOS with metabolic and reproductive comorbidities, including insulin resistance, metabolic syndrome, obesity-related complications, infertility, and psychological concerns such as anxiety, depression, and reduced quality of life. These associations support comprehensive follow-up rather than an assessment limited to ovarian morphology.

2.7 Lifestyle management and emerging therapeutic directions

Lifestyle modification is a recurring theme in the reviewed sources. Balanced nutrition, regular physical activity, behavioural support, and sustainable weight-management strategies may contribute to metabolic health and general wellbeing. The presentation also emphasizes lifestyle approaches for adolescents and young women. Dietary or exercise recommendations should be individualized; the reviewed material does not establish one dietary pattern as universally superior for all PCOS outcomes.
Pharmacological and emerging approaches discussed in the term-paper presentation include metformin, fertility-management strategies, GLP-1 receptor agonists, SGLT2 inhibitors, probiotics, novel biomarkers, and phenotype-oriented or precision-medicine approaches. These interventions differ in evidence maturity and intended use. Emerging therapies should be described as areas of ongoing investigation rather than universal first-line treatments, and treatment decisions remain the responsibility of healthcare professionals.

3. Research Problem and Motivation

PCOS diagnosis is challenging because its clinical presentation is heterogeneous and its manifestations overlap with other endocrine and reproductive conditions. Conventional assessment requires interpretation of multiple sources of information. Although ML and deep learning studies report promising results, the literature remains fragmented: some methods rely only on clinical variables, while others classify ultrasound images independently. A prediction produced by a complex model may also be difficult for clinicians to interpret.
The research problem addressed here is the design of an integrated, explainable decision-support approach that can consider clinical features and ultrasound images while making the contribution of each modality more transparent. The motivation is to organize existing methods into a coherent framework and identify the validation requirements that must be met before any such system can be used in routine care.

3.1 Research objectives

- Review clinical and computational approaches used for PCOS assessment.
- Compare ML methods applied to structured clinical and demographic data.
- Examine deep-learning approaches for ovarian ultrasound image classification.
- Describe how multimodal fusion can combine clinical and image-derived information.
- Explain the potential roles of SHAP, LIME, and Grad-CAM in model interpretation.
- Identify limitations involving dataset size, class imbalance, external validation, privacy, and clinical oversight.

3.2 Research gaps and proposed solutions

The reviewed approaches reveal several gaps. Clinical-data models depend on the completeness and quality of structured patient attributes, whereas ultrasound models depend on image quality, acquisition protocols, and representative image datasets. These separate approaches may not capture all available information when used independently. In addition, class imbalance, missing values, inconsistent diagnostic labels, and differences between patient populations can affect model performance.
A further gap concerns interpretability. High-performing machine-learning or deep-learning models may provide predictions without making the factors behind those predictions sufficiently clear to clinicians. The reviewed literature therefore explores feature selection, ensemble learning, SHAP, LIME, Grad-CAM, and attention-based explanations. However, an explanation map is not proof of clinical reasoning and must be assessed by domain experts.
The proposed response in this paper is a doctor-assisted decision-support framework that processes clinical attributes and ultrasound images through separate model branches, combines their representations or outputs, and provides modality-specific explanations. This is a design proposal based on the

reviewed work. Its effectiveness, safety, fairness, and clinical utility have not been established by the supplied student implementation and require independent evaluation.

4. Review and Proposed Methodology

4.1 Review method and evidence boundary

The present work is a narrative, source-based review of the term-paper report and presentation, supplemented by verification of selected bibliographic records and current IJRPR author information. The supplied documents list studies on clinical diagnosis, insulin resistance, lifestyle management, ML classification, ultrasound analysis, and explainable multimodal learning. They do not document a reproducible database search strategy, date range, screening protocol, or PRISMA flow for this paper. Therefore, the review is not described as a systematic review or meta-analysis.

4.2 Proposed hybrid framework

The proposed framework has two input branches. The clinical branch accepts structured patient information, including age, weight, height, menstrual-cycle regularity and length, weight change, hair growth, hair loss, and acne-related indicators. The imaging branch accepts ovarian ultrasound images. Each branch is preprocessed separately to preserve the characteristics of its data type.

For clinical data, preprocessing may include checking data types, handling missing values, encoding categorical variables, scaling where required by the selected algorithm, and selecting clinically relevant features. For ultrasound images, preprocessing may include image-quality checks, resizing, intensity normalization, and carefully controlled augmentation. Data splitting must occur at the patient level so that images or records from the same individual do not leak across training and test sets.

A machine-learning classifier processes the clinical feature vector, while a CNN or transfer-learning model extracts image representations. Their outputs or intermediate representations are then combined through a fusion module. The supplied presentation proposes combining the clinical and image streams and using SHAP/LIME for clinical explanations and Grad-CAM for image explanations. The final output is intended as a PCOS/non-PCOS estimate with a risk context for clinician review, not as an autonomous diagnosis.

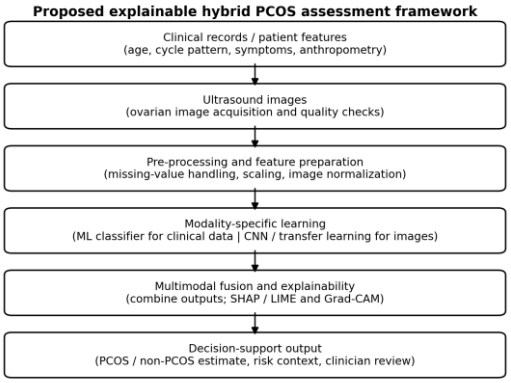


Fig. 1. Proposed explainable hybrid framework for PCOS decision support.

4.3 Clinical model implementation described in the source

The presentation includes a Python implementation outline using Pandas, Joblib, and scikit-learn. It loads a file named PCOS_data.csv and selects nine variables: age, weight, height, cycle regularity, cycle length, weight gain, hair growth, hair loss, and pimples. The target column is PCOS (Y/N). The code applies a stratified 80:20 train-test split with random_state=42 and trains a Random Forest classifier with 300 trees and class_weight='balanced'. The fitted model is saved as clinical_model.pkl.

This code documents a clinical-data training pipeline, but the supplied presentation does not include the dataset itself, data dictionary, missing-value handling details, confusion matrix, test-set metrics, calibration results, or independent validation for this particular implementation. Consequently, no accuracy or clinical-performance claim is made for the student's Random Forest model. The ultrasound model and multimodal fusion are described as proposed system components; the supplied evidence does not establish that these components have been trained and validated together.

4.4 Evaluation plan required for a future implementation

A complete evaluation should report sensitivity (recall), specificity, precision, F1-score, ROC-AUC, and a confusion matrix, together with confidence intervals where feasible. Because false-negative predictions may delay further assessment, sensitivity and subgroup performance deserve explicit attention. Evaluation should use a held-out test cohort and, ideally, external data from different institutions. Calibration, missing-data robustness, subgroup fairness, and the effect of image acquisition variation should also be examined. These are proposed evaluation requirements, not completed experiments in the current term paper.

4.5 Software architecture and implementation requirements

The presentation outlines a web-based implementation environment. The proposed frontend uses HTML, CSS, Bootstrap, and JavaScript for navigation, data entry, and result presentation. Python with Flask is identified as the backend framework. Pandas and NumPy support clinical-data handling and numerical operations, while scikit-learn supports the clinical machine-learning component. TensorFlow/Keras is identified for ultrasound deep learning. SQLite with SQLAlchemy is proposed for database operations, and SHAP, LIME, and Grad-CAM are identified for explainability.

The stated hardware requirements include an Intel Core i5 processor or above, at least 8 GB RAM (16 GB recommended), and at least 10 GB of free storage. A GPU is recommended for ultrasound deep-learning workloads but is not mandatory for basic testing. These are the requirements listed in the presentation and should be treated as an indicative development configuration rather than a universal deployment specification.

4.6 Proposed implementation workflow

The implementation workflow begins with collecting clinical records and ultrasound images, subject to appropriate consent, governance, and data-protection controls. Clinical records are checked for missing values, data types, and inconsistent entries; numerical features may be normalized where required by the selected algorithm. Ultrasound images undergo quality checks and image preprocessing, including resizing and normalization. Feature selection is applied to clinical variables, while a deep-learning model extracts image representations.

The clinical and imaging datasets should be divided into training, validation, and test partitions at the patient level to reduce data leakage. A clinical machine-learning model and an ultrasound deep-learning model are then trained and evaluated separately. If patient-level paired data are available, a fusion module may combine the two modalities. SHAP or LIME can be used to explain clinical predictions, while Grad-CAM can identify image regions contributing to an ultrasound model output. The interface may present a risk estimate and supporting explanations for clinician review.

The presentation lists intended application modules including a home page, patient assessment form, ultrasound upload, risk prediction, Grad-CAM visualization, doctor dashboard, and wellness guide. These are interface and workflow elements. Their presence does not demonstrate diagnostic validity, and the supplied material does not establish a completed, externally validated end-to-end multimodal system.

5. Case-Study Analysis

5.1 Web-based machine learning for PCOS prediction

Rahman et al. (2024) report a web-based PCOS prediction study using a dataset containing 541 patient records. The study compares Logistic Regression, Decision Tree, AdaBoost, Random Forest, and SVM, and applies mutual information for feature selection. The supplied presentation and the article abstract report that AdaBoost and Random Forest achieved the highest accuracy of 94% after feature selection. This result illustrates the potential of structured clinical data for accessible prediction, but it should be interpreted within the study's dataset and validation design rather than as a universal expected accuracy.

5.2 Explainable multimodal learning

Lakshmi and Pushpa (2026) describe a multimodal framework that combines ultrasound images and structured clinical attributes. The presentation identifies ResNet50 for image feature extraction, a Transformer Encoder for clinical features, cross-attention fusion, hierarchical classification, Grad-CAM++, and attention heatmaps. It lists 1,142 normal and 844 PCOS ultrasound images, alongside a clinical dataset with 364 normal and 177 PCOS samples before balancing. The presentation reports 99.42% accuracy, 99.39% sensitivity, 99.40% specificity, and 99.45% F1-score. These figures are reproduced as reported in the supplied presentation; they are not independently recalculated in this paper. The presentation also flags limited dataset scale and concerns about patient-level matched multimodal pairs, which should be addressed before generalizing the result.

The two studies illustrate complementary directions. A clinical-only model can operate on structured patient information, while a multimodal model can combine clinical and imaging evidence. The latter may offer a richer representation but introduces additional requirements for synchronized patient-level records, image standardization, model complexity, privacy, and external validation.

The two case studies are compared below based on the supplied presentation. Metrics are study-reported values and should not be compared as if obtained under a common experimental protocol.

Input. Rahman et al. (2024): clinical/demographic dataset with 541 records. Lakshmi and Pushpa (2026): ultrasound images plus clinical attributes.

Methods. Rahman et al.: Logistic Regression, Decision Tree, AdaBoost, Random Forest, and SVM with mutual-information feature selection. Lakshmi and Pushpa: ResNet50, Transformer Encoder, cross-attention fusion, and Grad-CAM++.

Reported results. Rahman et al.: 94% accuracy as reported. Lakshmi and Pushpa: 99.42% accuracy, 99.39% sensitivity, 99.40% specificity, and 99.45% F1-score as reproduced in the supplied presentation.

Strength. The first study demonstrates web-based prediction from structured clinical features; the second integrates imaging and clinical information.

Caution. Both results are study-reported values, not directly comparable common-protocol measurements. Dataset-specific and matched-pair validation concerns remain.

5.3 Comparative analysis of advanced AI-based PCOS detection methodologies

EffiDenseGenOp-based ultrasound classification: Ghosh and Srinivasan (2025) describe an ensemble transfer-learning approach for PCOS detection from ultrasound sonography images. The presentation identifies a Fuzzy Inference System for image enhancement, EfficientNetB7 and DenseNet201 for

feature learning, ensemble transfer learning, genetic-algorithm-based hyperparameter tuning, data augmentation, and class activation maps for interpretation. The cited presentation describes a Kaggle ultrasound dataset containing 3,846 original images (2,284 non-PCOS and 1,562 PCOS images), together with upsampled and downsampled dataset variants. The reported test results are 99.58% accuracy, 99.01% F1-score, 99.20% recall, 98.87% precision, and 98.97% AUC-ROC. These are study-reported values and require validation on independent, diverse ultrasound datasets.

Geographically validated explainable ensemble learning: Akter and Reno (2025) describe a clinical-feature approach using ANOVA-F feature selection and a group of classifiers, including Logistic Regression, stochastic gradient descent, Decision Tree, AdaBoost, Gaussian Naïve Bayes, K-nearest neighbours, Gradient Boosting, and XGBoost. A weighted-voting ensemble is used, with SHAP and LIME for interpretation. The presentation reports 541 women, 45 initial clinical features, and 23 selected features, with geographically distinct cohorts used to examine generalizability. Reported performance includes 94.34% accuracy, 93.38% AUC, 71.88% sensitivity, 77.97% F1-score, and 97.79% PR-AUC. The results illustrate why sensitivity and cohort-level performance should be considered alongside accuracy.

SmartScanPCOS: Mahesswari and Maheswari (2024) present a feature-driven clinical prediction framework. The presentation describes an open-source Kaggle dataset of 541 female records and 44 columns, with 24 non-invasive measures and 20 invasive parameters, and reports data collected across ten hospitals in Kerala. The methodology includes feature-selection procedures identified as TOPCA, OSSM, and TOMIM, followed by several classifiers and ensemble methods, including a two-level Random Forest classifier. The reported results are 99.31% accuracy, 100% precision, 98.5% recall, 99.2% F1-score, and 100% specificity, with eight-fold cross-validation and 25 runs also noted in the presentation. The relatively small record count means broader external validation remains important.

Blockchain-integrated explainable AI: Jaganathan and Natesan (2025) describe EAIBS-PCOS, which combines clinical prediction with data-management mechanisms. The presentation identifies Kendall correlation for reducing redundant features, Random Forest and XGBoost for ensemble prediction, SHAP and LIME for explanation, and Hyperledger Fabric for blockchain-based data management. The described dataset contains 541 women and 41 attributes; median replacement is used for missing values and unwanted identifiers are removed. The presentation reports 98.04% accuracy, 99.5% precision, 96.5% recall, and 97.9% F1-score using ten-fold cross-validation. It also identifies scalability under higher transaction volumes as a limitation of the blockchain component.

Comparative interpretation: These studies address different data types and system priorities. EffiDenseGenOp focuses on ultrasound image classification; Akter and Reno emphasize geographically informed evaluation of clinical-data models; SmartScanPCOS emphasizes feature selection and stacked or two-level ensemble learning; and EAIBS-PCOS adds secure data-management considerations. Their reported metrics must not be ranked as though they came from one common experiment because datasets, target definitions, validation procedures, and class distributions differ. The studies provide methodological examples for the proposed framework, not evidence that the student implementation achieves the same results.

6. Results and Discussion

The outcome of this paper is a synthesis of published and source-reported evidence, not a new experimental benchmark. Three themes emerge. First, clinical ML models can identify patterns in menstrual, anthropometric, and symptom-related variables, but performance depends on dataset composition, feature quality, and validation design. Second, ultrasound deep learning can automate image-pattern extraction, yet image acquisition differences and class imbalance may limit transfer to new settings. Third, multimodal and explainable approaches offer a route to combine complementary information and make predictions easier to inspect.

The source presentation includes example outputs such as clinical PCOS/non-PCOS prediction, ultrasound normal/abnormal classification, and a combined risk assessment. These are illustrative system outputs rather than a documented prospective clinical evaluation. The presentation also shows application interface elements such as a home page, patient assessment form, ultrasound upload, risk prediction, Grad-CAM visualization, doctor dashboard, and wellness guide. These interface elements describe the intended software experience; they do not establish diagnostic validity.

The proposed hybrid approach is therefore best positioned as a research prototype concept for doctor-assisted screening. A clinically meaningful system would need to define the target population, specify diagnostic reference standards, ensure patient-level data matching, prevent data leakage, report uncertainty, and evaluate whether model outputs improve clinical workflow. Model interpretability should be reviewed by clinicians and should not be treated as a substitute for external validation.

7. Advantages and Limitations

7.1 Potential advantages

- Combines structured clinical information with ultrasound-derived features rather than relying on a single modality.
- Supports systematic analysis of heterogeneous patient information and may assist prioritization for clinical review.
- Uses SHAP/LIME and Grad-CAM to provide complementary explanations for tabular and image-based predictions.
- Can be integrated into a web-based interface for structured data entry and clinician-facing review, subject to privacy and security safeguards.

7.2 Limitations

- The supplied materials do not provide a complete, reproducible literature-search protocol.
- The student's Random Forest code is provided, but its own test metrics and independent validation are absent.
- The proposed ultrasound and fusion components are not supported by a documented end-to-end validation in the supplied report.
- Public or single-centre datasets may not represent variation across age groups, ethnicities, body composition, equipment, and clinical settings.
- Prediction scores can be affected by missing data, class imbalance, label quality, leakage, and differences in diagnostic criteria.

- AI output must not replace clinician assessment; privacy, informed consent, data governance, and bias monitoring are necessary.

7.3 Data Privacy, Security, and Trustworthiness in AI-Based PCOS Detection

The integration of Artificial Intelligence into healthcare systems requires careful consideration of patient data privacy, security, and reliability. PCOS detection systems process sensitive information, including menstrual history, hormonal measurements, clinical symptoms, and ultrasound images. Therefore, maintaining confidentiality and preventing unauthorized access to medical information are essential requirements for developing trustworthy healthcare applications. Blockchain technology has been explored as a potential approach to improving the security and transparency of healthcare data management. Jaganathan and Natesan (2025) proposed the EAIBS-PCOS framework, which integrates Explainable Artificial Intelligence with blockchain technology. The framework uses Kendall correlation for feature reduction, Random Forest and XGBoost for classification, and SHAP and LIME for model interpretation. Hyperledger Fabric is incorporated to support secure and traceable management of patient information. The study demonstrates how predictive analytics and secure data-management mechanisms can be integrated within a healthcare framework. However, blockchain-based systems may introduce scalability challenges when transaction volumes increase. Consequently, future PCOS detection systems should consider appropriate access control, encryption, informed consent, secure data storage, and privacy-preserving data-sharing mechanisms.The integration of reliable prediction models with secure information management can contribute to the development of transparent, patient-centred, and trustworthy AI-assisted healthcare applications.

8. Future Scope

Future work should assemble larger, multi-institutional datasets containing patient-level linked clinical records, laboratory measurements, lifestyle information, and ultrasound images. Data collection should document diagnostic criteria, image acquisition settings, and relevant demographic characteristics. The hybrid model can then be evaluated using patient-level train/validation/test partitions and external cohorts. Further work should assess calibration, uncertainty estimation, subgroup performance, and robustness to missing or noisy inputs. Explainability outputs should be reviewed by clinicians to determine whether highlighted features and image regions are clinically meaningful. Secure web and mobile deployment may be explored only after technical, ethical, privacy, and clinical validation requirements have been met.

9. Conclusion

PCOS is a heterogeneous endocrine and metabolic condition that requires individualized assessment and long-term management. Conventional clinical evaluation remains central, while machine learning and deep learning offer promising methods for analyzing structured patient information and ultrasound images. The reviewed studies demonstrate progress in clinical prediction, feature selection, explainable learning, and multimodal fusion. Based on the supplied term-paper materials, this paper outlines a hybrid framework that combines clinical ML and ultrasound deep learning, followed by fusion and modality-specific explanations. The current source material does not establish a validated end-to-end multimodal diagnostic system or provide performance metrics for the student's Random Forest implementation. The proposed framework should therefore be treated as a decision-support research direction requiring larger datasets, independent validation, clinician oversight, and responsible deployment.

Acknowledgement

The author gratefully acknowledges the guidance of Mrs. G. Nirosha, Assistant Professor, Department of Computer Science and Engineering, GMRIT Deemed to be University, Rajam, and the support provided by the Department of Computer Science and Engineering.

REFERENCES

1. Ahmed, S., Rahman, M. S., Jahan, I., Kaiser, M. S., Hosen, A. S. M. S., Ghimire, D., & Kim, S.-H. (2023). A review on the detection techniques of polycystic ovary syndrome using machine learning. IEEE Access, 11, 86522–86543. https://doi.org/10.1109/ACCESS.2023.3304536
2. Akter, S., & Reno, S. (2025). Explainable AI for generalizable PCOS diagnosis: A geographically validated ensemble learning approach with feature selection. Engineering Reports, 7(9), e70395. https://doi.org/10.1002/eng2.70395
3. Bai, H., Ding, H., & Wang, M. (2024). Polycystic ovary syndrome (PCOS): Symptoms, causes, and treatment. Clinical and Experimental Obstetrics & Gynecology, 51(5), 126. https://doi.org/10.31083/j.ceog5105126
4. Choudhari, R., Tayade, S., Tiwari, A., & Satone, P. (2024). Diagnosis, management, and associated comorbidities of polycystic ovary syndrome: A narrative review. Cureus, 16(4), e58733. https://doi.org/10.7759/cureus.58733
5. Elmannai, H., El-Rashidy, N., Mashal, I., Alohali, M. A., Farag, S., El-Sappagh, S., & Saleh, H. (2023). Polycystic ovary syndrome detection machine learning model based on optimized feature selection and explainable artificial intelligence. Diagnostics, 13(8), 1506. https://doi.org/10.3390/diagnostics13081506
6. Ghafari, A., Maftoohi, M., Eslami Samarin, M., Barani, S., Banimohammad, M., & Samie, R. (2025). The last update on polycystic ovary syndrome (PCOS), diagnosis criteria, and novel treatment. Endocrine and Metabolic Science, 17, 100228. https://doi.org/10.1016/j.endmts.2025.100228

7. Ghosh, A., & Srinivasan, K. (2025). EffiDenseGenOp: Ensemble transfer learning with hyperparameter tuning using genetic algorithm optimization for PCOS detection from ultrasound sonography images. IEEE Access, 13, 54285–54312. https://doi.org/10.1109/ACCESS.2025.3553895

8. Jaganathan, G., & Natesan, S. (2025). Blockchain and explainable-AI integrated system for polycystic ovary syndrome (PCOS) detection. PeerJ Computer Science, 11, e2702. https://doi.org/10.7717/peerj-cs.2702

9. Lakshmi, V., & Pushpa, B. (2026). Explainable multimodal deep learning using cross-attention fusion of ultrasound and clinical features for PCOS classification. Discover Computing, 29, Article 7. https://doi.org/10.1007/s10791-025-09901-x

10. Mahesswari, G. U., & Maheswari, P. U. (2024). SmartScanPCOS: A feature-driven approach to cutting-edge prediction of polycystic ovary syndrome using machine learning and explainable artificial intelligence. Heliyon, 10(20), e39205. https://doi.org/10.1016/j.heliyon.2024.e39205

11. Melin, J., Forslund, M., Alesi, S., Piltonen, T., Romualdi, D., Spritzer, P. M., & Teede, H. (2023). The impact of metformin with or without lifestyle modification versus placebo on polycystic ovary syndrome: A systematic review and meta-analysis of randomized controlled trials. European Journal of Endocrinology, 189(2), S38–S64. https://doi.org/10.1093/ejendo/lvad098

12. Rahman, M. M., Islam, A., Islam, F., Zaman, M., Islam, M. R., Sakib, M. S. A., & Babu, H. M. H. (2024). Empowering early detection: A web-based machine learning approach for PCOS prediction. Informatics in Medicine Unlocked, 47, 101500. https://doi.org/10.1016/j.imu.2024.101500

13. Sheehan, M. T. (2004). Polycystic ovarian syndrome: Diagnosis and management. Clinical Medicine & Research, 2(1), 13–27. https://doi.org/10.3121/cmr.2.1.13

14. Sirmans, S. M., & Pate, K. A. (2014). Epidemiology, diagnosis, and management of polycystic ovary syndrome. Clinical Epidemiology, 6, 1–13. https://doi.org/10.2147/CLEP.S37559

15. Teede, H. J., Tay, C. T., Laven, J. J. E., Dokras, A., Moran, L. J., Piltonen, T. T., Costello, M. F., Boivin, J., Redman, L. M., Boyle, J. A., Norman, R. J., Mousa, A., & Joham, A. E. (2023). Recommendations from the 2023 international evidence-based guideline for the assessment and management of polycystic ovary syndrome. European Journal of Endocrinology, 189(2), G43–G64. https://doi.org/10.1093/ejendo/lvad096