

Deep Learning Technique for Interpretable Diagnosis of Polycystic Ovary Syndrome in Ultrasound Imaging

Pratibha Pandey¹ , Sumit Chaudhary² , Xin Nie^{3*} 

^{1,2}Department of Computer Science and Engineering, Uttarakhand University, India

³School of Computer Science and Engineering, Wuhan Institute of Technology, China

¹pratibhapananey8502@gmail.com, ²drchaudharysumit@gmail.com, ³niexin@whu.edu.cn

*Corresponding Author

Article Info

Article history:

Submission June 15, 2025

Revised July 26, 2025

Accepted September 29, 2025

Published October 9, 2025

Keywords:

PCOS Diagnosis

Deep Learning

Follicles Detection

Interpretability

Ultrasound Images



ABSTRACT

Polycystic Ovary Syndrome (PCOS) is a widespread condition in women that is associated with hormonal disorders, absentee or infrequent menstruation, and cyst formation on ovaries. The diagnostic process of PCOS is still difficult because of ambiguity in ultrasound images. Earlier approaches emphasize the detection and classification of follicles for ultrasound images and disregard interpretability, which obscures understanding the model's decisions and trustworthiness in critical medical applications. **This research proposes** a method of diagnosing PCOS in two stages. Initially, follicles are located with a YOLOv8 model that has multiscale feature dose attention. Follicles are subsequently classified with a ResNet50 model that has SE blocks. The last step is for Grad-CAM to show which features of the image were used by the classification model in order to explain its decision and provide meaningful insights regarding the model's predictions. Evaluation was performed on two publicly available datasets. The proposed **method** outperforms all other methods in follicle detection with 92% mean average precision (mAP) and classification accuracies of 98.25% and 94.56% on datasets 1 and 2, respectively. **This result** makes the proposed model a reliable and transparent technique for robust clinical applications. By providing a scalable, deployable, and interpretable diagnostic pipeline that can be integrated into AI-based health-tech platforms, the suggested method also fits with the objectives of technopreneurship incubator models.

This is an open access article under the [CC BY 4.0](#) license.



DOI: <https://doi.org/10.34306/att.v7i3.768>

This is an open-access article under the CC-BY license (<https://creativecommons.org/licenses/by/4.0/>)

©Authors retain all copyrights

1. INTRODUCTION

Polycystic Ovary Syndrome is defined as a multicystic disorder of the endocrine system that affects females in their reproductive years and tends to originate in the teenage years [1]. The particular reason is still unknown; however, several factors such as family history, surroundings, and habits are noted to be impacted by the illness. This implies that it is made up of hormonally disturbed syndromes, abnormalities of the ovaries, and the degree of insulin sensitivity, and, as a result, the associated features include amenorrhea, hirsutism, change in body mass index, and polycystic ovaries [2]. The recent WHO report indicates that PCOS is found in 8–13% of women, and 6–18% of young women of reproductive age are found ‘exclusively’ linked to infertility. Findings show that up to 70% of women with this disorder are either sterile from nonfertility ovarian cysts or anovulatory [3].

Diagnosis of this condition, especially with images acquired under the ultrasound examination, is difficult since the disease is complicated as well as its presentation. Ovarian morphology is quite different for any one woman at any menstrual cycle and cannot be relied on to provide precise clinical diagnostic aids, as they are usually followed by subjective observations of the radiologist [4]. Worse yet, components like low magnifying power and noisy pictures contribute more to obscuring other features under critical microscopy, like follicle margins and stromal mass. It also presents a big problem for auto-diagnostic systems since databases for training specific AI can be either inconsistent or missing, while many machine learning models themselves are too opaque to be employed practically [5]. Such issues are why, for example, deep learning, better-optimized algorithms, and standardized data must be used to enhance diagnostic yield and the treatment utility of the obtained results [6].

As we noted earlier, analyzing ultrasound images to diagnose PCOS presents some difficulty, and thus, following the difficulties observed above, some approaches have been established [7]. One usual practice is to segment the images by employing complex segmentation methods for variation of the shape and structure of the ovaries. Size-dependent filtering, contour detection, texture analysis, and level-set segmentation assist in enhancing the perimeter of ovaries and follicles so that proper dimensions can be measured for diagnosis [8]. Other genomic techniques, which include the evaluation of texture changes in the ovarian tissue and follicular regions, are also used in the assessment of PCOS from normal ovaries. These methods, however, have limitations. They are sensitive to image noise, low contrast, and irregular shapes. They also need stress on parameters and frequently fail to perform in complicated or overlapping areas, therefore raising issues concerning their stability [9].

Although many works are aimed at diagnosing PCOS or jointly detecting and classifying from the ultrasound image, till now no one considers the interpretability of the deep learning models. Traditionally, interpretability is highly valued in the medical field, as it enables one to understand why the predictions were made or why the model reaches a certain level of accuracy [10]. This fits especially to trust-building as it determines the rationale of the doctor, and the choices made were made in line with the medical expertise. That way, when the model errs, interpretability indicates issues such as biased data or otherwise, features that the model did not incorporate, which would then lead to enhancements [11]. Finally, interpretable models provide more trustworthy decisions, which are beneficial both for safety and for better clinical practice. The suggested interpretable AI framework provides a deployable, accurate, and explicable diagnostic tool, bridging the gap between academic innovation and healthcare startup potential. This speeds up translation into practical applications and opens doors for scalable clinical solutions, particularly in tech-driven health diagnostics [12].

This paper resolves the challenges in PCOS diagnosis by introducing a combined approach for detection, classification, and interpretability. The research proposes a new approach that incorporates or integrates YoloV8 with Attention Guided Multi-Scale Feature Fusion (AGMSFF) to develop a more accurate model capable of detecting follicles in PCOS ultrasound images and optimizing feature selection to improve accuracy [13]. The images detected are further processed using a classification model based on ResNet50 with a Squeeze-and-Excitation (SE) block. The SE block improves classification performance by adaptively amplifying essential features, thus yielding more accurate outcomes. A Grad-CAM interpretability method is also implemented on the classification model so that users such as clinicians and researchers can understand the rationale behind the model's predictions visually and therefore trust the model's decisions, enhancing its reliability and usability in medicine [14]. The primary contributions of this paper are as follows:

- A new method for detection using YoloV8 fused with Attention-Guided Multi-Scale Feature Fusion to accurately detect abnormalities in ultrasound imaging of PCOS was developed [15].
- A classification approach based on ResNet50 was developed by adding Squeeze-and-Excitation (SE) blocks that markedly enhanced performance. Additionally, implemented an interpretability technique for the proposed model, providing transparent insights into its decision-making process and fostering greater trust in its application [16].
- The incorporation of Grad-CAM for interpretability, which is rarely addressed in prior PCOS studies [17].

The rest of this paper is organized as follows: Section 2. presents an existing study on different classification and detection techniques. Section 3. explains the methodology, including the proposed method

[18]. Section 4. derives experimental results. Finally, Section 6. concludes with remarks on the paper's findings.

2. LITERATURE REVIEW

Due to complex and time-consuming PCOS diagnosis challenges caused by diverse symptom combinations, costly biochemical tests, and ovary scanning, [19] suggested an AI-based computational approach for classifying PCOS. Initially, the correlation-based feature extraction method was applied to clinical data, reducing 45 features to 17, with SVM achieving the highest accuracy among ML algorithms. Furthermore, VGG16 outperformed the CNN model on ultrasound image analysis [20]. Similarly, [21] suggested a DL model for automated PCOM detection using Inception architecture by combining ovary ultrasound images with patient clinical data and a fusion model developed by integrating image and clinical features for PCOS diagnosis, leveraging Mobile Net for feature extraction. Furthermore, due to the issue of weak contrast, speckle noise, and hazy boundaries in ultrasonography, which hinder accurate segmentation and classification, Sha proposed a guided trilateral filter for noise reduction, an AdaResU-Net for precise segmentation, and a pyramidal dilated convolutional network for cyst classification, optimized using the Wild Horse Optimization algorithm to enhance diagnostic precision and reliability [22].

At the same time, [23] developed a denoising filter using the Two-Dimensional Fractional Fourier Transform (2D-FrFT) to remove artifacts in the time-frequency plane that optimizes the fractional operator parameter via a VGG-16 model, effectively reducing noise and enhancing image quality. [24] proposed an active contour method combined with a modified Otsu method to automate follicle detection of PCOS via preprocessing using traditional filters to reduce noise, with the optimal filter selected to enhance image quality, addressing the challenges posed by varying follicle sizes and their relationships with blood vessels and tissues, which often lead to errors in manual interpretation. [25] explained the shortcomings of an automated system powered with Fuzzy Rule-Based Convolutional Neural Networks (FCNN) for cyst detection from ultrasound images. An automated system that utilizes FCNN improves consistency in clinical decision-making by assisting physicians with cyst detection. Some other studies proposed a feature selection method with tri-stage wrapper approach that optimizes feature relevance, improving classification accuracy while minimizing computational resources [26].

In a different study, a method comprising four stages was developed: preprocessing, segmentation with an enhanced watershed algorithm, follicle detection with SCBOD feature recognition, and PCOS versus Non-PCOS classification with SVM. This method solves issues related to speckle noise, time complexity, and misclassification based on the size and number of follicles [27]. Another approach addressed the constraints of manual cyst detection and error-prone machine learning techniques by using a hybrid strategy that combines VGGNet16 as a feature extractor and XGBoost as a meta-learner to improve the reliability and efficiency of diagnosing PCOS from ovarian USG images [28]. Non-invasive clinical parameters were employed and various ML algorithms were assessed to detect PCOS, demonstrating the efficacy of the RF model for screening and validating its predictive ability using the Out-of-Bag (OOB) error. To improve diagnostic accuracy and processing efficiency, traditional and ensemble classifiers were utilized along with filter, embedded, and wrapper-based feature selection approaches, which demonstrated that the Ensemble RF classifier, combined with the embedded feature selection method, outperformed other classifiers in terms of accuracy and sensitivity, achieving the best performance and efficacy in medical diagnosis tasks [29]. Recent important studies have explored advancements in medical imaging technology to improve the robustness and enhancement. A detailed comprehensive survey highlighted novel techniques and remote applications that are efficient in medical imaging technology. The role of telemedicine in mobile applications called Halodoc enhanced patient-doctor consultations while ensuring data privacy and efficiency. Another study explored the use of gradient centralization in conjunction with advanced preprocessing techniques for enhanced performance in medical image analysis [30].

3. RESEARCH METHODS

The below Figure 1 shows the proposed method for detecting and classifying PCOS from ultrasound images. First, images from a dataset undergo preprocessing, which includes normalization, noise reduction, and reshaping [31]. Data augmentation is also applied to enhance the training data. Next, follicles in the images are detected using YOLOv8 with specialized modules AGMSFF [32]. The detected regions are then classified using a ResNet50 model with SE blocks, which improves accuracy by focusing on important features.

The system classifies the image as either "infected" or "not infected" and uses Grad-CAM for interpretability, helping to visualize which parts of the image influenced the decision [33]. The enhanced architectural efficiency and better real-time object detection capabilities, YOLOv8 was chosen as the best option for detecting small and variable-sized follicles. Because of its powerful feature extraction capabilities and SE-ResNet50 was chosen for its strong feature extraction capability, enhanced by Squeeze-and-Excitation blocks that emphasize informative channels crucial for classifying subtle differences in ultrasound images [34]. The detailed explanation of each is provided below:

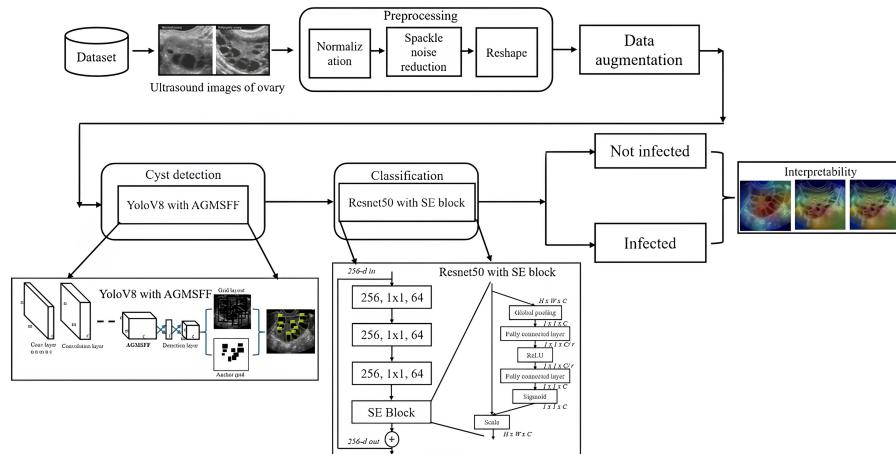


Figure 1. Block diagram of proposed denoise method

3.1. Dataset Description and Noise Addition

This study is based on a dataset retrieved from the Kaggle repository, which consists of 1,932 images for training and 1,924 images for testing [35]. The overlap between the training and testing subsets was significant in the existing dataset; hence, the test set was discarded, and the entire dataset put aside for testing was utilized for training. This dataset was later partitioned into new training and test datasets [36]. It contains ultrasound images labelled as INFECTED (781 images showing cystic ovaries) and NOT INFECTED (1,143 images of healthy ovaries). These labels indicate people suffering from PCOS and those who do not. This classification is particularly useful in the context of real-time medical systems because it enables automated and precise diagnosis of PCOS [37]. The study also adds the dataset PCOSGen collected from multiple online sources. This dataset includes 1,468 unhealthy samples and 3,200 healthy samples, which were split into training and testing sets. A New Delhi-based gynaecologist annotated all the samples [38]. Sample images of both of the classes are provided in Figure 2.

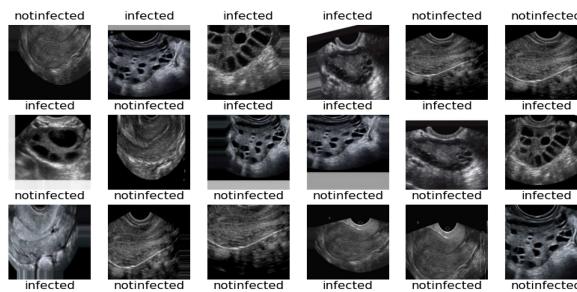


Figure 2. Different Sample images from both classes

3.2. Data Pre-Processing

At this stage, the ultrasound images are first normalized by drawing the pixel intensity values within the scale of 0 and 1. To mitigate distortion, a median filter is employed which is known to efficiently eliminate speckle noise present in ultrasound imagery [39]. The filter works by taking a given pixel's value and replacing

it with the median of the specified pixels around it in the area, making it possible to lessen noise without losing the vital details in the image. Lastly, in order to standardize the images that may differ in size, all images are resized to 128x128x3 [40].

3.3. Data Augmentation

This serves as a crucial step for the case of computer vision that assists in augmenting the scope of the dataset and also the diversity, helping the models in generalizations and performing accurately in practical scenarios [41]. In this research zooming, rotation, flipping and shearing augmentations were performed. These image alterations replicated real-life situations where an object's position and orientation can change, providing numerous additional images that modify the original images [42]. This not only improves the size of the dataset aiming to curb overfitting but guarantees that the model captures various distinctive attributes, enhancing the reliable performance for diverse inputs [43]. The below Figure 3 illustrates the different augmentation techniques used on the images.

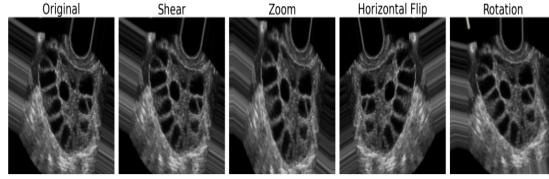


Figure 3. Various augmentation techniques applied on original images

3.4. Follicle Detection

Before doing the classification task, the image detection of PCOS diagnosis Domain is crucial. Follicles in ovaries are perhaps one of the most significant grounds for diagnosing PCOS [44]. Certainly, definite identification of these follicles in ultrasound images is critical for identification of ROIs which are very pertinent to abnormal ovarian morphology. In this paper, to recognize the follicles from the ultrasound image, YOLOv8 (You Only Look Once Version 8) with AGMSFF module is utilized [45]. This modification is making it easier for the model to different sizes of follicles from ultrasound image and fine-grained details. The AGMSFF is introduced and incorporated into the structure of YOLOv8 as a feature extractor and feature fusion model. While it uses an Attention Mechanism alongside Multi-Scale Feature Fusion to fine-tune the outcome of the predictions. The attention mechanism learns the weights with respect to spatial and channel dimensions, noteworthy are follicles boundaries, and negations are given to unimportant parts [46]. Mathematically, the attention mechanism can be expressed as:

$$F_{att} = \sigma (W_s \cdot GAP(F)) \odot F \quad (1)$$

Where F is the input feature map, σ denotes the sigmoid activation, W_s represents learnable parameters, and $GAP(F)(\cdot)$ is the global average pooling operation [47]. The resulting attention-refined feature map F_{att} enhances the detection of follicles-like regions. Multi-scale feature fusion combines features from different layers of the model, allowing the detection head to capture follicles of varying sizes. Given feature maps F_1, F_2, F_3 from different scales, the fusion is performed as:

$$F_{fused} = F_1 + Up(F_2) + Up(Up(F_3)) \quad (2)$$

Where $Up(\cdot)$ represents up-sampling to ensure spatial alignment. This multi-scale fusion ensures that the model retains both fine details and broader contextual information, critical for detecting small and irregularly shaped follicles [48]. The improved YOLOv8 with AGMSFF processes ultrasound images to output bounding boxes for detected follicles. Each bounding box is associated with a confidence score S and a class label C , calculated as:

$$S_c = \sigma (W_c \cdot F_{fused}) \quad (3)$$

Where W_c are classification weights. This detection output is then passed to the classification model for further analysis. By incorporating this, it enhances the detection performance for follicles identification.

3.5. Classification

Once follicles are accurately detected, the next step is classification to determine whether an ovary is normal or abnormal. Normal images typically lack follicles, whereas abnormal images contain detectable follicles, a hallmark of PCOS [49]. This classification is critical in distinguishing healthy individuals from those requiring further medical attention. To perform this task, we propose a Modified ResNet50 with Squeeze-and-Excitation Blocks (SE-ResNet50), designed to handle the complexities of ultrasound imaging by improving feature extraction and focusing on relevant regions [50]. The ResNet50 architecture is enhanced by integrating Squeeze-and-Excitation (SE) Blocks after each residual block. The SE blocks adaptively recalibrate the feature maps to prioritize the most informative channels, significantly improving the model's ability to classify subtle differences between normal and abnormal ovarian structures. The SE block operates in two stages: squeeze and excitation. For a feature map $F \in \mathbb{R}^{H \times W \times C}$, where H , W , and C represent the height, width, and number of channels, the squeeze step computes global context by applying Global Average Pooling (GAP) across spatial dimensions:

$$z_c = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W F_{i,j,c} \quad (4)$$

This produces a channel-wise descriptor $z \in \mathbb{R}^C$. In the excitation step, fully connected layers with activation functions learn channel-wise dependencies and generate attention weights s :

$$s = \sigma(W_2 \cdot \text{RELU}(W_1 \cdot z)) \quad (5)$$

Here, W_1 , W_2 are learnable weights, and σ is the sigmoid activation. The feature map is recalibrated by scaling each channel with its corresponding weight:

$$F'_{i,j,c} = s_c \cdot F_{i,j,c} \quad (6)$$

This allows the model to emphasize critical features, such as cyst-like patterns, while suppressing irrelevant noise. The modified ResNet50 extracts features from the ultrasound images while the SE blocks recalibrate the feature maps at each residual stage [51]. The final layer performs binary classification into normal or abnormal categories, with a SoftMax output:

$$y = \text{softmax}(W_f \cdot F_{final} + b_f) \quad (7)$$

Where W_f and b_f are the weights and bias of the final fully connected layer, and F_{final} represents the output feature vector. This integration provides effective classification results over existing methods [52]. These improvements make SE-ResNet50 particularly effective for accurate classification of normal and abnormal ovaries in PCOS diagnosis.

3.6. Interpretability

Concerning decision making, interpretability plays a decisive role in medical AI and is especially important when using AI for PCOS detection owing to the need for its transparent, justified decisions that will be acceptable to clinicians and follow logical reasoning [53]. Superimposing areas on the ultrasound images on which a model bases its decision helps build confidence and accuracy, thereby negating biases and other useless areas [54]. Nevertheless, most PCOS detection studies focus more on the detection error than on its interpretability keeping out clinical application where explainable results are significant. To fill this gap, we incorporate Gradient-weighted Class Activation Mapping (Grad-CAM) to identify portions of images guiding models' predictions while improving the diagnostics legitimacy [55].

Grad-CAM works by utilizing the gradients of the output class score y^c with respect to the feature maps A^k of the last convolutional layer. The importance weights α_k^c for each feature map are computed as:

$$\alpha_k^c = \frac{1}{Z} \sum_i \sum_j \frac{\partial y^c}{\partial A_{ij}^k} \quad (8)$$

Where Z is the spatial dimension of the feature map, and A^k represents the activation at spatial location (i, j) . The Grad-CAM heatmap L^c is then generated by a weighted combination of the feature maps:

$$L^c = \text{ReLU} \left(\sum_k \alpha_k^c A^k \right) \quad (9)$$

The ReLU function ensures that only features positively influencing the class score are highlighted. The resulting heatmap is overlaid on the original image to provide a visual explanation of the model's decision [56]. Using Grad-CAM, our system enables clinicians to verify whether the model's focus aligns with diagnostically relevant areas, such as cysts in abnormal images or uniform ovarian structures in normal cases [57]. This added layer of interpretability ensures that the classification model is not only accurate but also reliable and transparent, making it suitable for real-world clinical applications [58].

4. RESULTS AND DISCUSSION

This section describes the experimental procedure, including dataset preparation, different hyper parameters that used, model training, and evaluation using various performance metrics, along with a comparative analysis of results.

4.1. Experimental Setup

In this experiment two publicly accessible ultrasound image databases were used. Since carrying out feature extraction, the dataset was split into two; for training there were 2,685 images while for testing 671 images were set up. It predisposed with follicles detection, then followed by classification, before having interpretability analysis. This approach provides a general overview of the ability of the model in the identification and categorization of images, especially under noisy environments. Through checks based on such conditions, this method guarantees that the model will be effective in real-world settings.

4.2. System and Hyperparameter Specifications

The proposed model train on a system that has Intel® Xeon® Silver 4216 CPU, 128 GB of RAM, and an Nvidia® RTX A4000 GPU with 16 GB of VRAM. Here loss function Mean Squared Error (MSE) is used and Adaptive Moment Estimation (Adam) which integrates momentum and adaptive learning rates for effective convergence used as optimizer function for train the denoise model. To balance stability and computational efficiency, the batch size is 64 and the learning rate is set to 0.0001. Model train on 50 epochs which enable the model to learn from the image dataset and classify the labels, improving its generalization abilities for practical uses.

4.3. Performance Evaluation Metrics

In this paper, the proposed model is evaluated using various performance metrics. For follicle detection, the mean Average Precision (mAP) score is utilized, which measures the precision of the model at different recall thresholds, providing an overall assessment of detection accuracy. In the classification task, metrics such as accuracy (the percentage of correctly classified instances), precision (the proportion of true positives among predicted positives), recall (the ability to correctly identify all positive instances), and F1-score (the harmonic mean of precision and recall) are employed. Additionally, the ROC-AUC curve (illustrating the trade-off between true positive rate and false positive rate) and the Confusion Matrix (showing the distribution of predicted and actual classes) are used for visual representation of the results.

4.4. Experimental Results for Follicles Detection

Detection of follicles has been done using the different methods in this paper with the inclusion of our proposed method. The below Figure 4 shows the achieved detection approaches for ovarian follicles in ultrasound images which include OCD-FCNN [8], YOLOv8, Attention-Guided Multi-Scale CNN, and a proposed approach. All the techniques are then benchmarked against ground truth annotations to assess their efficacy. There is also a model called OCD-FCNN for ovarian follicles detection and classification which is constructed with a limited performance of the above index, the mAP is just 55%. This suggests that it may not well identify and classify ovarian follicles which consequently give incomplete or imprecise prediction. Even with a 78% mAP, YOLOv8 succeeds in detection but does not identify all the targets or achieve comprehensive follicles detection. In the same manner, the proposed Attention-Guided Multi-Scale CNN yielded comparable performance to YOLOv8, which shows that AMSCNN cannot outcompete YOLOv8 with regards to detecting finely nuanced details of follicles.

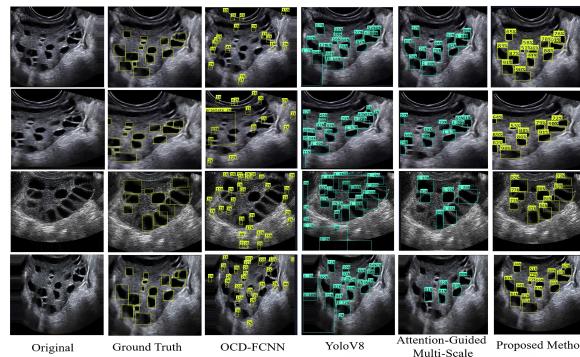


Figure 4. Detection result of proposed method along with different state of art methods

Unlike this, the proposed method is a fusion of Attention-Guided Multi-Scale CNN and YOLOv8 for more efficient operation since both models have impressive characteristics that add to the efficiency of the technique. The integration of YOLOv8, an improved object detection algorithm, with the Attention-Guided mechanism, leads to improved detection accuracy and robustness, with an added advantage of a 92% boost in mean Average Precision (mAP). The proposed method is accurate in locating ovarian follicles and obtains better results in terms of precision and coverage, proving its applicability in more complicated medical image analysis operations.

4.5. Experimental Results for Classification

The proposed classification method demonstrates exceptional effectiveness, achieving the highest accuracy on both Dataset 1 and Dataset 2, as shown in Table 1. On Dataset1, it reached 98.25%, outperforming ResNet50 (95.69%), DenseNet121 (92.87%), and VGG16 (89.26%). Incorporating SE blocks significantly boosted baseline models, with DenseNet121 improving to 94.65% (from 92.87%) and VGG16 to 91.33% (from 89.26%). Similarly, on Dataset2, the proposed method achieved 94.56%, surpassing ResNet50 (90.21%), DenseNet121 (84.33%), and VGG16 (85.66%). SE blocks consistently enhanced feature representation, enabling models to capture subtle data patterns and generalize better. This led to an improvement of 2.56% over ResNet50 and 3.6% over DenseNet121 (with SE blocks) on Dataset 1, and a 4.35% and 4.45% improvement on Dataset 2, respectively. These results confirm that SE blocks recalibrate channel-wise responses to focus on critical features, advancing the state of the art in classification tasks.

Table 1. Performance evaluation of proposed classification method with state of art methods.

Dataset	Models	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Dataset1	Vgg16	89.26	90.33	89.65	89.98
	Resnet50	95.69	94.12	96.74	95.41
	Densenet121	92.87	93.79	91.55	92.65
	Vgg16 with SE Blocks	91.33	90.12	91.54	90.82
	Densenet121 with SE Blocks	94.65	95.47	94.66	95.06
	Resnet50 with SE blocks (Proposed Method)	98.25	97.65	98.66	98.15
Dataset2	Vgg16	85.66	86.78	85.12	85.94
	Resnet50	90.21	89.33	88.32	88.82
	Densenet121	84.33	85.63	84.11	84.86
	Vgg16 with SE Blocks	88.25	89.44	88.14	88.78
	Densenet121 with SE Blocks	90.11	88.25	86.32	87.27
	Resnet50 with SE blocks (Proposed Method)	94.56	92.32	93.55	92.93

The below Figure 5 represents the proposed model's training and validation accuracy of two datasets; the left plot corresponds to the first dataset while the right plot is for the second dataset. The blue color in each

plot represents the training accuracy, while the orange one represents the validation accuracy. The model shows frequent growth in the first epochs and then gradually converges to the optimal accuracy with the increase of epochs. The training and the validation accuracy are equally high, and in that sense, the model may be claimed to generalize well, with little risk for overfitting.

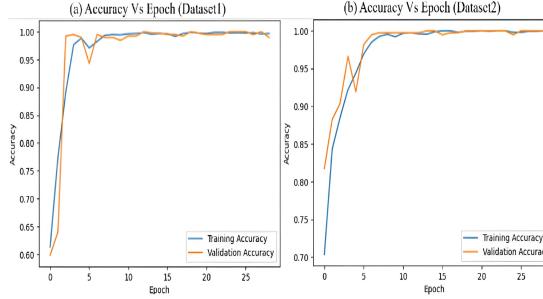


Figure 5. Accuracy Vs Epoch graph of proposed model on both datasets

The ROC curve Figure 6 illustrates the performance of five models in binary classification tasks, comparing their ability to distinguish between two classes on both datasets. In the Dataset 1 Proposed Method achieves the best performance with a perfect Area Under the Curve (AUC) of 100%, followed by DenseNet+SE with an AUC of 98%. VGG16 and ResNet50 perform moderately well, each with an AUC of 88%, while DenseNet has a slightly lower AUC of 85%. Also in Dataset 2 Proposed method have AUC of 92% which signifies the proposed model generalizability. The closer a curve is to the top-left corner, the better the model performs, and all models outperform random guessing (the diagonal line). This demonstrates the Proposed Method's superiority in classification accuracy compared to the others.

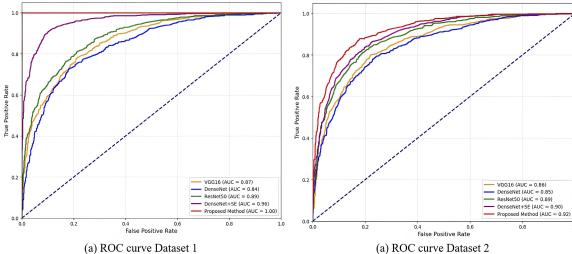


Figure 6. ROC curve of proposed method along with other methods on both datasets

Table 2 below compares the performance of existing techniques with the proposed model across two datasets in terms of accuracy. For Dataset 1, previous methods demonstrate high accuracy, but the proposed method surpasses them with an accuracy of 98.25%, showcasing its superior performance. Similarly, for Dataset 2, existing techniques achieve notable accuracy, but the proposed method achieves the highest accuracy at 94.56%. These results highlight the proposed model's effectiveness in outperforming state-of-the-art approaches across different datasets.

Table 2. Performance evaluation of proposed classification methods with existing methods.

Dataset	Models	Accuracy (%)
Dataset1	[51]	97.50
	[59]	93.83
	[53]	97.00
	[54]	96.54
	Proposed Method	98.25
Dataset2	[56]	92.00
	[54]	94.39
	[57]	91.07
	Proposed Method	94.56

4.6. Experimental Results for Interpretability

The provided figure illustrates the interpretability of the proposed model using Grad-CAM visualization applied to the last layer of the network. The below Figure 6 shows the Grad-CAM representation of proposed model on abnormal images. Interpretability is essential in deep learning, particularly in medical imaging tasks like ultrasound analysis, to ensure that the model's decisions are transparent and clinically reliable. By visualizing the regions of focus within the image, Grad-CAM helps validate the model's behavior, ensuring its predictions align with human expert understanding, which is critical for establishing trust and enhancing decision-making in sensitive domains.

Grad-CAM is a widely used method for generating visual explanations of a model's predictions. It highlights the regions in the input image that most strongly influence the model's decision. By computing the gradients of the predicted class score with respect to the feature maps of a specific layer, Grad-CAM produces a heatmap that overlays the original image, indicating areas of high importance. In the top row of the figure, the grayscale ultrasound images are presented, while the corresponding Grad-CAM visualizations in the bottom row highlight the regions of interest (e.g., cysts or abnormalities) in color-coded heatmaps, ranging from cool (less important) to warm (more important) tones. These visualizations confirm that the model effectively focuses on diagnostically relevant structures, demonstrating its robustness and interpretability.

4.7. Discussion

This paper attempts to classify the PCOS from the ultrasound images by relating to problems of poor image quality and the inherent nature of the disease. The previous approaches share the disadvantage that the follicle is not detected; hence, the feature set is insufficient for follicle detection. To counter this, here we present a YOLOv8 model-based attention-guided multiscale approach for follicle detection to better understand the specific regions of interest and provide better input to the classification. The classification stage uses a ResNet50 model enhanced with SE blocks to increase feature representation and pay attention to the most important features. This two-stage system of biological prediction greatly improves the accuracy of the rehabilitative model. Further, the model aims for interoperability, which could be very crucial during the diagnostic analysis using deep learning-based medical-image analysis. It thus ensures that detection of follicles is included while at the same time acting as a way of providing explication of classification results to practitioners. This approach does not only provide better accuracy; it also fulfills the gaps in the existing approaches by including efficient detection and highly interpretable detection as a definitive fixed standard for assessing the PCOS from the ultrasound images.

Nevertheless, the proposed model possesses several drawbacks. During the detection stage, the proposed YOLOv8 attention-guided multiscale model may fail to detect very small or poor visibility follicle candidates in low-quality ultrasound images, which may limit its applicability across different clinic settings. In the classification stage, adding SE blocks into the ResNet50 architecture significantly improves the ability of the network to learn features; however, it also increases the number of computations. This leads to training and inferring much higher time than full-precision limits the use of this model for real-time or devices with low computational capability. These limitations give us ideas for improvement that can be focused on increasing the practical usefulness of such systems. Additionally, we potential deployment using edge devices or cloud-based systems to ensure usability across diverse healthcare environments, including low-resource settings. Additionally, the proposed AI-based denoising model directly contributes to SDG 3 (Good Health and Well-being) by facilitating the early diagnosis of Polycystic Ovary Syndrome (PCOS). It aligns with SDG 9 (Industry, Innovation, and Infrastructure) by fostering AI innovation in ultrasound imaging.

5. MANAGERIAL IMPLICATIONS

The proposed deep learning-based diagnostic tool for PCOS has important implications for healthcare management and strategy. The integration of AI into clinical practices can bring substantial improvements in efficiency, cost-effectiveness, trust, and scalability across various healthcare environments.

5.1. Enhancing Diagnostic Efficiency and Cost-Effectiveness

The introduction of AI tools into healthcare workflows significantly improves diagnostic efficiency, which is critical for high-volume healthcare settings. The AI system reduces the time needed for accurate PCOS diagnosis by automating the process of follicle detection and classification. This allows healthcare providers to process more cases, leading to improved patient throughput and reduced wait times.

Furthermore, AI's ability to automate diagnosis can result in considerable cost savings. By reducing reliance on specialized medical personnel, such as radiologists, healthcare organizations can lower labor costs. Additionally, the system's accuracy reduces the likelihood of misdiagnoses and unnecessary treatments, further contributing to cost efficiency. In environments where budget constraints are common, such AI-driven solutions provide an economically sustainable alternative to manual diagnostic methods.

5.2. Fostering Trust and Collaboration Through Interpretability

A significant barrier to AI adoption in healthcare is clinician trust. The incorporation of interpretability features, such as Grad-CAM, addresses this issue by providing clinicians with a clear rationale behind the AI model's predictions. By visually highlighting the areas of the ultrasound image that influenced the model's decision, Grad-CAM enhances transparency and allows clinicians to verify that the model is focusing on clinically relevant features, such as cysts or follicles.

This interpretability builds trust between clinicians and the AI system, fostering a collaborative environment where AI supports healthcare professionals in making informed decisions. With clear explanations of AI's reasoning, healthcare providers can be more confident in relying on AI-assisted diagnostics, leading to more efficient decision-making and better patient outcomes.

6. CONCLUSION

This work proposes a new dual-stage model of PCOS diagnosis that achieves superior results by incorporating both follicle detection and classification tasks. In the detection stage, the YOLOv8-based attention-guided multiscale model gives an mAP of 92%, improving the predicted area of interest. The subsequent classification step, performed utilizing ResNet50 architecture with SE blocks, attains exceptional accuracy levels of 98.25% for Dataset 1 and 94.56% for Dataset 2, thereby evidencing the solidity of the method proposed. Not only does the integration of detection and classification increase diagnostic accuracy, but it also offers increased interpretability to practitioners. The future research relatable to this work can involve the study of the use of this method in large and diverse databases and the inclusion of this approach in real-time clinical settings. Future work will focus on the potential deployment of our model on edge devices or cloud-based platforms to ensure usability across diverse healthcare environments, including low-resource settings.

7. DECLARATIONS

7.1. About Authors

Pratibha Pandey (PP)  <https://orcid.org/0000-0002-3122-4352>

Sumit Chaudhary (SC)  <https://orcid.org/0000-0002-5455-774X>

Xin Nie (XN)  <https://orcid.org/0000-0002-5005-9145>

7.2. Author Contributions

Conceptualization: PP; Methodology: SC; Validation: PP and SC; Writing Original Draft Preparation: PP and SC; Writing Review and Editing: SC and XN; Visualization: SC; All authors, PP, SC, and XN, have read and agreed to the published version of the manuscript.

7.3. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

7.4. Funding

This work was supported by a grant from the Hubei Key Laboratory of Intelligent Robot of China (Grant No. HBIRL202009).

7.5. Declaration of Conflicting Interest

The authors declare that they have no conflicts of interest, known competing financial interests, or personal relationships that could have influenced the work reported in this paper.

REFERENCES

- [1] U. P. Kaveramma, Prinith, R. Suresh, and D. Raj, "Ultrasound-based machine learning-aided detection of uterine fibroids: Integrating vision transformer for improved analysis," *Proceedings on Engineering Sciences*, vol. 35, no. 06, p. 2350033, 2023.
- [2] E. Anderson and R. Bhandari, "Exploring the impact of innovation strategies on startupreneurs business growth," *Startupreneur Business Digital (SABDA Journal)*, vol. 4, no. 1, pp. 82–92, 2025.
- [3] K. Srilatha *et al.*, "Automated ultrasound ovarian tumour segmentation and classification based on deep learning techniques," in *Proceedings on Engineering Sciences*, vol. 881, 2022, pp. 59–70.
- [4] A. Andriyansah, S. Subchan, K. Pamungkas, I. Geraldina, A. H. Arifin, and C. A. Manggarani, "Exploring sustainable tourism resources to realise ergo green tourism in indonesia," *Aptisi Transactions on Technopreneurship (ATT)*, vol. 7, no. 2, pp. 334–342, 2025.
- [5] S. Jothimani, A. Mani, and R. Rajasekaran, "Detection of polycystic ovarian disease using probabilistic optimum deep learning feature fusion," *Proceedings on Engineering Sciences*, pp. 1–15, 2025.
- [6] S. ALAM SUHA, "Predicting polycystic ovary syndrome through machine learning technique using patients' symptom data and ovary ultrasound images," Ph.D. dissertation, Department of Computer Science and Engineering, MIST, 2022.
- [7] A. S. Peña, S. F. Witchel, K. M. Hoeger, S. E. Oberfield, M. G. Vogiatzi, M. Misso, H. J. Teede, and L. Ibáñez, "Adolescent polycystic ovary syndrome according to the international evidence-based guideline," *Proceedings on Engineering Sciences*, vol. 18, no. 1, p. 72, 2020.
- [8] J. I. Saputro, L. F. Sa'adah, Y. Syifa, K. D. Ramadhanti, and M. F. Gajali, "Transforming learning experiences with advanced educational technology solutions," *International Transactions on Education Technology (ITEE)*, vol. 3, no. 2, pp. 114–124, 2025.
- [9] H. Danaei Mehr and H. Polat, "Diagnosis of polycystic ovary syndrome through different machine learning and feature selection techniques," *Proceedings on Engineering Sciences*, vol. 12, no. 1, pp. 137–150, 2022.
- [10] D.-H. Mo, Y.-C. Wu, and C.-S. Lin, "The dynamic image analysis of retaining wall crack detection and gap hazard evaluation method with deep learning," *Applied Sciences*, vol. 12, no. 18, p. 9289, 2022.
- [11] S. Tiwari, S. S. Rautaray, and A. Pandey, "Sposds: A smart polycystic ovary syndrome diagnostic system using machine learning," *Proceedings on Engineering Sciences*, vol. 203, p. 117592, 2022.
- [12] G. U. Mahesswari and P. U. Maheswari, "Smartscanpcos: a feature-driven approach to cutting-edge prediction of polycystic ovary syndrome using machine learning and explainable artificial intelligence," *Helixyon*, vol. 11, p. e39205, 2024.
- [13] S. Jeevitha and N. Priya, "Identifying and classifying an ovarian cyst using scbod (size and count-based ovarian detection) algorithm in ultrasound image," *Proceedings on Engineering Sciences*, vol. 13, no. 9, pp. 799–806, 2022.
- [14] Z. Zainol, G. Brotosaputro, S. C. Chen, and E. A. Natasya, "Designing ethical ai systems for sustainable technology development," *ADI Journal on Recent Innovation*, vol. 6, no. 2, pp. 201–211, 2025.
- [15] L. Meng, Y. Zhou, X. Chen, and W. Liu, "Evaluation of ultrasound derived fat fraction for metabolic associated fatty liver disease in obese patients with polycystic ovary syndrome," *Proceedings on Engineering Sciences, Journal of Ultrasound*, February 2025.
- [16] T. Kuusk, R. Fahrudin, F. Yulian, A. Fauzi, and A. Wilson, "Addressing regulatory risks in fintech through decentralized technologies," *APTSI Transactions on Management*, vol. 8, no. 3, pp. 204–212, 2024.
- [17] M. Sha, "Segmentation of ovarian cyst in ultrasound images using adaresu-net with optimization algorithm and deep learning model," *Proceedings on Engineering Sciences*, vol. 14, no. 1, p. 18868, 2024.
- [18] A. Dada, "Digital entrepreneurial behaviour of millennials in the fourth industrial revolution in southwest nigeria: Digital entrepreneurial behaviour of millennials," *African Journal of Science Policy and Innovation Management*, vol. 2, no. 1&2, pp. 41–60, 2021.
- [19] M. Jha, R. Verma, and A. Srivastava, "Noise cancellation of polycystic ovarian syndrome ultrasound images using robust two-dimensional fractional fourier transform filter and vgg-16 model," *Proceedings on Engineering Sciences*, vol. 16, no. 4, pp. 2497–2504, 2024.
- [20] P. Moral, D. Yadav, D. Gupta, and S. Tripathi, "Podboost: An explainable ai model for polycystic ovarian syndrome detection using grey wolf-based feature selection approach," *Proceedings on Engineering Sciences, Neural Computing and Applications*, vol. 36, no. 30, pp. 18 627–18 644, 2024.
- [21] H. J. Smidt and O. Jokonya, "A framework to develop technopreneurs in digital agriculture value chains

in south africa," *Journal of Foresight and Thought Leadership*, vol. 2, no. 1, p. 13, 2023.

[22] K. Van Der Ham, A. Hoek, F. de Boer, L. Knoester, and J. S. E. Laven, "The additional value of ultrasound markers in the diagnosis of polycystic ovary syndrome," *Proceedings on Engineering Sciences*, vol. 123, no. 2, pp. 342–349, 2025.

[23] A. Kadim, I. Yusnita, A. Sutarman, R. Lesmana, and F. A. Ramahdan, "Assessing the impact of corporate governance and strategic leadership on economic growth and market stability," *IAIC Transactions on Sustainable Digital Innovation (ITSIDI)*, vol. 6, no. 2, pp. 177–187, 2025.

[24] P. Jain, A. Mitra, S. Sharma, and A. Kumar, "Xplainable ai for deep learning model on pcos analysis," in *XAI Based Intelligent Systems for Society 5.0*, ser. *Proceedings on Engineering Sciences*. Elsevier, 2024, pp. 131–152.

[25] Balıkçı Çiçek, S. Aydin, and B. O. Yildiz, "Detection of risk factors of pcos patients with local interpretable model-agnostic explanations (lime) method that an explainable artificial intelligence model," *Proceedings on Engineering Sciences*, vol. 6, no. 2, pp. 59–63, 2021.

[26] B. R. Bhardwaj, "Adoption, diffusion and consumer behavior in technopreneurship," *International Journal of Emerging Markets*, vol. 16, no. 2, pp. 179–220, 2021.

[27] P. Chitra, K. Srilatha, M. Sumathi, F. V. Jayasudha, T. Bernatin, and M. Jagadeesh, "Classification of ultrasound pcos image using deep learning based hybrid models," in *2023 Second International Conference on Electronics and Renewable Systems (ICEARS)*, 2023, pp. 1389–1394.

[28] R. T. Sudarma and P. Budiaستuti, "The influence of leadership spirit and creativity on technopreneur-based entrepreneurial interest in vocational education," *Jurnal Edukasi Elektro*, vol. 8, no. 1, pp. 94–102, 2024.

[29] K. Shanmugavadivel *et al.*, "Optimized polycystic ovarian disease prognosis and classification using ai based computational approaches on multi-modality data," *Proceedings on Engineering Sciences*, vol. 24, no. 1, p. 281, 2024.

[30] M. Murod, S. Anhar, D. Andayani, A. Fitriani, and G. Khanna, "Blockchain based intellectual property management enhancing security and transparency in digital entrepreneurship," *Aptisi Transactions on Technopreneurship (ATT)*, vol. 7, no. 1, pp. 240–251, 2025.

[31] J. Prathibanandhi and G. S. Annie Grace Vimala, "Diagnosis of polycystic ovarian syndrome with the implementation of deep learning models," in *2024 Third International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT)*, 2024, pp. 1–7.

[32] K. Widjajanti and I. A. Jumbri, "Technopreneurship, innovation capability and social media marketing as catalysts for competitive advantages: A study of batik msmses in pekalongan, indonesia," *Multidisciplinary Science Journal*, vol. 7, no. 8, pp. 2 025 367–2 025 367, 2025.

[33] A. Alamoudi, R. Ejbali, and A. Alshdadi, "A deep learning fusion approach to diagnosis the polycystic ovary syndrome (pcos)," *Proceedings on Engineering Sciences*, pp. 1–15, 2023.

[34] P. Hendriyati, F. Agustin, U. Rahardja, and T. Ramadhan, "Management information systems on integrated student and lecturer data," *Aptisi Transactions on Management (ATM)*, vol. 6, no. 1, pp. 1–9, 2022.

[35] T. N. Ravishankar, R. Subashini, and M. Praveen Kumar, "A deep learning approach for ovarian cysts detection and classification (ocd-fcnn) using fuzzy convolutional neural network," *Proceedings on Engineering Sciences*, vol. 27, p. 100797, 2023.

[36] H. A. Halim, N. H. Ahmad, and A. Waqas, "Young technopreneur ventures' agility and competitiveness in malaysia: role of inhibitors," *Journal of Science and Technology Policy Management*, vol. 16, no. 2, pp. 300–313, 2025.

[37] M. Abouhawwash, "Automatic diagnosis of polycystic ovarian syndrome using wrapper methodology with deep learning techniques," *Proceedings on Engineering Sciences*, vol. 47, no. 1, pp. 239–253, 2023.

[38] R. G. Munthe, Q. Aini, N. Lutfiani, I. Van Persie, and A. Ramadhan, "Transforming scientific publication management in the era of disruption: Smartpls approach in innovation and efficiency analysis," *APTISI Transactions on Management*, vol. 8, no. 2, pp. 123–130, 2024.

[39] S. K. Saheb and D. Sreenivasarao, "An optimized siamese neural network with deep linear graph attention model for gynaecological abdominal pelvic masses classification," *Proceedings on Engineering Sciences*, vol. 50, no. 4, pp. 1831–1859, 2024.

[40] S. A. Suha and M. N. Islam, "An extended machine learning technique for polycystic ovary syndrome detection using ovary ultrasound image," *Proceedings on Engineering Sciences*, vol. 12, no. 1, p. 17123, 2022.

[41] A. Amroni, A. A. Darmawan, and G. A. Pangilinan, "The role of big data and blockchain in enabling

transparent and sustainable business processes," *ADI Journal on Recent Innovation*, vol. 6, no. 2, pp. 180–189, 2025.

[42] Y. Indrianti, M. Hamsal, A. Furinto, and R. Kartono, "It takes two to tango: Roles of digital strategy and human creativity in creating transformational technopreneurship," in *E3S Web of Conferences*, vol. 388. EDP Sciences, 2023, p. 04045.

[43] E. M. K. Jose and B. P. Kushwaha, "Technopreneurship in startups and understanding their entrepreneurial mindset: an empirical evidence," *International Journal of Management and Enterprise Development*, vol. 23, no. 4, pp. 335–354, 2024.

[44] A. Alwiyah and W. Setyowati, "A comprehensive survey of machine learning applications in medical image analysis for artificial vision," *International Transactions on Artificial Intelligence*, vol. 2, no. 1, pp. 90–98, 2023.

[45] L. Hartiningtyas and I. Muna, "The effect of adversity quotient, technopreneurial learning, and innovation capabilities to the technopreneurship intention," *Indonesian Journal of Islamic Economics Research*, vol. 6, no. 1, pp. 41–57, 2024.

[46] M. Kuswari, R. Gantino, and J. Maratis, "Maximizing healthcare service information system: Understanding the influence of integration on efficiency," *ADI Journal on Recent Innovation*, vol. 6, no. 2, pp. 108–117, 2025.

[47] V. Agarwal, M. C. Lohani, and A. S. Bist, "A novel deep learning technique for medical image analysis using improved optimizer," *Health Informatics Journal*, vol. 30, no. 2, p. 14604582241255584, 2024.

[48] A. Choudhari, "Pcos detection using ultrasound images [dataset]," <https://www.kaggle.com/datasets/anaghachoudhari/pcos-detection-using-ultrasound-images>, 2024.

[49] E. T. Maziriri, B. Mabuyana, and B. Nyagadza, "Navigating the technopreneurial odyssey: determining how technopreneurial self-efficacy, technopreneurial education and technological optimism cultivate tech-driven entrepreneurial intentions," *Management & Sustainability: An Arab Review*, 2025.

[50] P. Handa *et al.*, "Pcosgen-train dataset," <https://zenodo.org/records/10430727>, 2023.

[51] N. S. Nilofer and R. Ramkumar, "Follicles classification to detect polycystic ovary syndrome using glcm and novel hybrid machine learning," *Proceedings on Engineering Sciences, Turkish Journal of Computer and Mathematics Education*, vol. 12, no. 7, p. 2715, 2021.

[52] F. B. Ismail, A. T. Z. Xuan, U. Rusilowati, and J. Williams, "Exploring the frontier of data science: Innovations, challenges, and future directions," *International Transactions on Education Technology (ITEE)*, vol. 2, no. 2, pp. 163–172, 2024.

[53] B. Rachana, M. V. Chandana, and P. Srujana, "Detection of polycystic ovarian syndrome using follicle recognition technique," *Proceedings on Engineering Sciences*, vol. 2, no. 2, pp. 304–308, 2021.

[54] P. Moral, D. Yadav, D. Gupta, and S. Tripathi, "Cystnet: An ai driven model for pcos detection using multilevel thresholding of ultrasound images," *Proceedings on Engineering Sciences, Scientific Reports*, vol. 14, no. 1, p. 25012, 2024.

[55] S. I. N. W. Abdullah, M. M. Deli, R. F. Suryawati, U. Nangkula *et al.*, "Technosociopreneurship as a productivity—increasing strategy based on community potential," *Journal of Infrastructure, Policy and Development*, vol. 8, no. 12, 2024.

[56] H. Nakhua, D. Kothari, and R. Sethia, "An ensemble approach for ultrasound-based polycystic ovary syndrome (pcos) classification," *Proceedings on Engineering Sciences, Educational Administration Theory and Practices*, May 2024.

[57] M. J. Lakshmi, G. Kusuma, and P. H. Rao, "Prediction of pcos and pcd in women using ml algorithms," in *Proceedings on Engineering Sciences*, 2023, vol. 361, pp. 97–115.

[58] T. S. Goh, D. Jonas, B. Tjahjono, V. Agarwal, and M. Abbas, "Impact of ai on air quality monitoring systems: A structural equation modeling approach using utaut," *Sundara Advanced Research on Artificial Intelligence*, vol. 1, no. 1, pp. 9–19, 2025.

[59] R. K. Hudiono and S. Watini, "Remote medical applications of artificial intelligence," *International Transactions on Artificial Intelligence*, vol. 1, no. 2, pp. 182–187, 2023.