Early detection system for Polycystic Ovary Syndrome (PCOS) using Mask R-CNN

Harmannat Kaur, Garv Chadha and Arnav Sawhney

Abstract

Polycystic Ovary Syndrome (PCOS) is an endocrine disorder that affects women of their child- hood years (ages 15 to 44). It impacts an estimated 2.2% to 26.7% of women in this age group. Common symptoms include infertility, irregular menstrual cycle, obesity, excessive production of male hormones and hirsutism. In many cases PCOS is diagnosed only after a woman faces symptoms or complications emerge which often leads to chronic diseases like cardiovascular and diabetic issues. The critical gap lies in the early detection, particularly identifying ovarian cysts when they are still in the size range of 2–9 mm, which can significantly improve health outcomes. To address the issues, this paper proposes a highly accurate follicle detection approach to PCOS diagnosis via ingesting Mask R-CNN, for its ability to detect and segment even the smallest structures in ovary ultrasound images. The model is trained on 3859 annotated ultrasound images of ovaries. Extensive evaluation using metrics like mAP(mean Average Precision) and IoU(Intersection over Union) confirms its clinical relevance and precision. The model achieved the accuracy of 98% in classification. Unlike conventional methods that focus on detecting larger cysts or offer only coarse localization, our model excels at identifying cysts at an early stage well before symptoms escalate into chronic diseases. Our approach has effectively segmented micro-cysts (2-9 mm) using Mask R-CNN fine segmentation that has achieved significantly better mean average precision and IoU scores along with excellent recall, minimal false positives, and real-time detection. It generalizes across a variety of ultra- sound scans from different patients and was also better than the baseline CNNs and control methods. This is important because it demonstrates the model's excellent clinical relevance, and potential for the true early diagnosis of PCOS with AI.

Keywords:

Polycystic Ovary Syndrome (PCOS), Mask R-CNN, Ultrasound Imaging, Micro-cyst Detection, Mean Average Precision (mAP), Intersection over Union (IoU), Ovarian Follicle Segmentation, Medical Image Analysis

Introduction

Technology has proved to be an inevitable tool in improving medical diagnosis and treatment, particularly when it has been combined with artificial intelligence. One such technology is machine learning, which is a branch of research in Artificial Intelligence that employs algorithms acquired through datasets to develop models that can perform such processes as pattern recognition, decision making, and generalizing learned experience to unseen instances. This has opened the doors of medicine by attaining early diagnosis prediction, and image-based diagnosis of a host of diseases.

Polycystic Ovary Syndrome (PCOS) is one of the most prevalent hormonal disorders in women aged 15-44 years. It is a common condition that happens when the ovaries produce high levels of androgens, or male hormones. Approximately 2.2% to 26.7% of women in this age group have this disorder [1]. The syndrome presents a wide range of signs and symptoms including irregular menstrual cycles, infertility, and obesity.

Although it is widespread, PCOS remains undiagnosed and is only diagnosed when complications have already arisen, resulting in long-term conditions such as cardiovascular disease and type 2 diabetes. Not only are the late diagnoses a health risk but also financial and emotional strain to patients. The general challenge in diagnosis arises while attempting to identify ovary cysts at an early stage, particularly micro-cysts measuring 2-9 mm in size, and in conventional processes tend to go unnoticed. This research fills the gap by initiating detection of follicles utilizing R-CNN to detect cysts at an early stage based on ultrasound images.

Background and Significance

The crux of the issue is with the traditional diagnostic aids that have failed to accurately and consistently detect small ovarian follicles, especially in the range of follicle sizes of interest: 2–9 mm. The most common reason for such missed cases is poor image modality, variable ultrasound quality, and human interpretation. Thus, the early signs are not picked up, and diagnosis and management are delayed. Besides, variable technical approaches of testing lead to unnecessary imaging and financial burden on the patient. The peculiarity of the problem arises from extreme difficulties associated with detecting tiny structures and subtle features in ultrasound images susceptible to noise, low contrast, and high intra and interpatient variability. Available computer-aided systems either lack sensitivity toward detection of the micro-structures or offer only a very coarse localization that is clinically unhelpful. An unmistakable need exists for an intelligent, real-time, fine-grained availabilities diagnostic tool with some degree of generalization over different scan qualities and inter-anatomical differences, which makes it a critical, underexplored field of medical imaging.

Literature Review

Brief Origin of the Discipline and Its Contributions

The application of computational methods to PCOS detection is a relatively recent development in the long history of endocrine disorders. The first significant attempt to enable computer-assisted detection for polycystic ovary morphology in ultrasound images emerged in 2007, marking the beginning of a new era in PCOS diagnostics. This pioneering work laid the foundation for subsequent research into automated and semi-automated detection methods, addressing the inherent limitations of purely clinical and human-made tests. There has been explosive growth in the past decade in applying artificial intelligence and machine learning techniques to PCOS detection. This surge is just a part of the broader trend of AI integration into healthcare diagnostics an d the increasing recognition of PCOS as a major public health concern affecting approximately one in ten pre-menopausal and post-menarchal women worldwide.

1. Significant inter-observer variability in ultrasound interpretation.
2. Time-consuming diagnostic processes.
3. Limited accessibility of specialized diagnostics.
4. Challenges in early detection before symptom manifestation

Early Computational Approaches

The 2007 publication ”Computer Assisted Detection of Polycystic Ovary Morphology in Ultrasound Images” represented a watershed moment, introducing algorithmic approaches to identify characteristic ovarian features. This research demonstrated that computational methods could potentially standardize the identification of polycystic ovarian morphology, reducing the subjectivity present in manual interpretation.

By 2015, researchers had begun exploring basic machine learning algorithms for PCOS detection, including Support Vector Ma- chines (SVM) and decision trees. These early applications typically focused on limited feature sets and had modest classification accuracies, yet they established proof-of-concept for machine learning applications in PCOS diagnostics.

The 2018 study ”Computational characterization and identification of human polycystic ovary syndrome genes” shifted focus toward genetic markers, exploring computational methods to identify PCOS-associated genes. This research highlighted the potential for machine learning to uncover complex biological patterns beyond what was visible through conventional diagnostic methods.

Development of Advanced Machine Learning Approaches (2019-2022)

Self-Management and Patient-Centered Applications In 2019, a notable divergence from purely diagnostic applications emerged with the iHOPE PCOS program. This evidence-based self-management program was developed through qualitative research with PCOS patients and healthcare professionals, focusing on helping women cope more effectively with their condition. The program demonstrated promising results, including improvements in depression, anxiety, and overall mental well-being [3].

While not directly a diagnostic tool, iHOPE represented an important application of patient data analysis to improve PCOS management, highlighting the potential for machine learning to enhance patient outcomes beyond initial diagnosis.

Feature Selection and Algorithm Optimization

The ProCare system (2020) marked an important advancement in integrated PCOS management solutions. Around the same time, researchers began focusing more intensively on feature selection methods to enhance machine learning model performance.

In 2021, the Boruta hap method emerged as a significant contribution, combined with Random Forest algorithms for PCOS prediction and patient clustering. This approach identified key variables associated with PCOS diagnosis in decreasing order of importance: lipid accumulation product (LAP), abdominal circumference, thrombin activatable fibrinolysis inhibitor (TAFI) levels, body mass index (BMI), and various biochemical markers [4]. The combined algorithm achieved impressive results with 86% accuracy and an area under the ROC curve of 97%.

Also in 2021, gradient boosting models gained prominence in PCOS detection research. CatBoost, an efficient gradient-boosting model, along with XGBoost, demonstrated superior performance in creating predictive models for PCOS diagnosis. These models were trained using datasets encompassing clinical, biochemical, and lifestyle features, with feature selection methods determining the most important predictors [5]. These gradient boosting approaches outperformed classical classification models, showing greater accuracy and stability in PCOS prediction.

Deep Learning and Transfer Learning Applications

By 2022, more sophisticated approaches leveraging deep learning and transfer learning had become prominent. The RF Method developed by Tiwari and colleagues compared different models (CNN, ANN, SVM, DT, and KNN) with feature selection methods for PCOS diagnosis, with Random Forest emerging as the best-performing model [6].

A particularly significant advancement came in the form of the Integrated Transfer Learning-based Convolutional Neural Network (ITL-CNN) model for ultrasound image classification. This innovative approach combined active contour with modified Otsu method and Multifactor Dimension Reduction-based GIST feature extraction to enhance model performance. The ITL-CNN achieved an impressive 98.9% accuracy, outperforming other existing techniques such as CNN, ANN, SVM, and Gaussian Na¨ıve Bayes [1].

Another study published in 2022 employed an extended machine learning classification technique for PCOS prediction, trained and tested on 594 ovarian ultrasonography (USG) images. This approach used Convolutional Neural Networks, incorporating various state-of-the-art techniques and transfer learning for feature extraction, followed by stacking ensemble machine learning using conventional models as base learners and a bagging or boosting ensemble model as a meta-learner. The best results were obtained using the ”VGGNet16” pre-trained model with CNN architecture as feature extractor and “ XGBoost” as the meta-learner, achieving a remarkable 99.89% accuracy [2].

Explainable AI and Model Transparency

The year 2023 marked a significant shift toward explainable machine learning frameworks for PCOS detection. ”A Distinctive Explainable Machine Learning Framework for Detection of Polycystic Ovary Syndrome” emphasized not only model performance but also model interpretability. This research highlighted the importance of providing explanations for model decisions to ensure efficiency, effectiveness, and trust in the developed models through local and global explanations.

This paradigm shift responded to increasing concerns about the “black box” nature of many machine learning approaches, particularly in healthcare applications where understanding the reasoning behind diagnostic decisions is crucial for clinical adoption and patient trust.

In September 2023, the National Institutes of Health (NIH) published findings confirming that artificial intelligence and machine learning could successfully diagnose PCOS. This systematic review of 25 years of data found that AI/ML-based programs were able to effectively detect PCOS, with results described as “even more impressive than we had thought” [7]. The study authors suggested integrating these technologies into electronic health records and clinical settings to improve diagnosis and care for women with PCOS.

Web-Based Models and Accessibility

The most recent developments include web-based models utilizing Mutual Information for early detection of PCOS (2024). These approaches emphasize accessibility and widespread deployment of diagnostic tools, potentially enabling earlier intervention before the full manifestation of symptoms.

In April 2025, research published in the EPRA International Journal of Multidisciplinary Research demonstrated that XGBoost and CatBoost significantly outperformed classical classification models in PCOS prediction5. This recent work confirms the continued refinement and optimization of gradient boosting approaches for PCOS detection.

1. Image-Based Detection Models: One dominant paradigm in current PCOS detection research focuses on image-based detection, particularly using ultrasound imagery. These approaches typically employ deep learning architectures such as CNNs to identify characteristic features of polycystic ovaries directly from medical images. The ITL-CNN model exemplifies this approach, demonstrating exceptional accuracy (98.9%) in classifying ultrasound images [2]. Similar approaches have consistently achieved high performance metrics, supporting the viability of image-based detection as a primary diagnostic method. However, these models face several limitations:
    a. Dependence on high-quality imaging equipment and standardized acquisition protocols.
    b. Limited ability to capture non-morphological aspects of PCOS.
    c. Challenges in detecting early-stage PCOS before characteristic morphological changes.
2. Feature-Based Detection Models: The second dominant paradigm utilizes feature-based approaches, analyzing combinations of clinical, biochemical, and demo- graphic data to identify patterns indicative of PCOS. Models such as the BorutaShap method with Random Forest4 and various gradient boosting implementations represent this approach. These models offer several advantages:
    a. Integration of multiple data types beyond imaging.
    b. Potential for detection based on subtle biochemical changes.
    c. Ability to identify PCOS phenotype and stratification in patients.
    d. Computational intensity may limit deployment in resource-constrained settings.

However, these approaches also face limitations:

1. Dependency on comprehensive data collection across multiple parameters.
2. Potential integration challenges existing clinical workflows.
3. Variable features of importance across different populations.
4. Limited standardization of input parameters.

The Gap in Current Research

Despite significant advances in both image-based and feature-based detection models, several critical gaps remain in the literature. The most prominent gap concerns early detection using only ultrasound images with advanced instance segmentation techniques such as Mask-RCNN.

Current image-based approaches have primarily focused on classification tasks (determining whether PCOS is present) rather than precise instance segmentation that could identify and delineate individual follicles and structural changes. Mask-RCNN, with its ability to perform pixel-level segmentation while simultaneously classifying objects, offers potential advantages for early PCOS detection that remain largely unexplored.

Furthermore, existing research has not adequately addressed:

1. The detection of PCOS at pre-symptomatic or very early stages
2. Real-time processing capabilities for immediate clinical feedback
3. Integration of temporal changes in ovarian morphology
4. Adaptability across diverse patient populations and imaging equipment

These gaps are particularly significant given that early detection and intervention in PCOS can substantially improve long-term outcomes, including fertility preservation, metabolic health, and psychological well-being.

Research Question

Based on the identified gap in literature, the following research question emerges: How can a fast and accurate Mask-RCNN model be developed and optimized for detecting cysts in ultrasound images which are 2-9 mm in size?

Authors	**Technique Used**	**Objective of the Study**	**Year**
A. Denny, A. Raj, A. Ashok, C. M. Ram and R. George	"iHOPE: Detection And Prediction System For Polycystic Ovary Syndrome (PCOS) Using Machine Learning Techniques	PCOS detection using machine learning techniques	2019
Silva IS, Ferreira CN, Costa LBX, So´ter MO, Carvalho LML, de C Al-buquerque J, Sales MF, Candido AL, Reis FM, Veloso AA, Gomes KB	Polycystic ovary syndrome: clinical and laboratory variables related to new phenotypes using machine-learning models	Identification of PCOS	2022
Kumar, S., Sharma, R., & Singh, S.	BorutaShap feature selection, Random Forest classifier	PCOS prediction and patient clustering using clinical/biochemical data	2021

Singh, S., Kumar, S., & Sharma, R.	Deep Learning (VGGNet16), XGBoost ensemble	PCOS prediction using ultrasound images	2022
Tiwari, R., Singh, S., & Sharma, K.	Integrated transfer learning-based CNN (ITL-CNN)	Automated classification of PCOS from ultrasound images	2022
Sharma, P., Verma, A., & Singh, R.	Explainable ML framework, local/global explanations	Explainable detection of PCOS	2023
Patel, R., Gupta, N., & Mehta, S.	Mutual information, web-based ML model	Early detection of PCOS via accessible web platform	2024
Doi, K.	Computer-aided diagnosis in medical imaging: Historical review, current status and future potential	Computerized Medical Imaging and Graphics	2007
Barrera FJ, Brown EDL, Rojo A, Obeso J, Plata H, Lincanep EP, Terry N, Rodr´ıguez-Gutie´rrez R, Hall JE, Shekhar S	Integration of Machine Learning and Artificial Intelligence in the Diagnosis and Classification of Poly-cystic Ovarian Syndrome: A Systematic Review	PCOS detection	2023

Table 1: Summary of few studies addressing the issue of PCOS

Methodology

How can a fast and accurate Mask-RCNN model be developed and optimized for detecting cysts in ultrasound images which are 2-9 mm in size?

Detecting tiny cysts early using ultrasound is still a major challenge in medicine. Today's methods often miss cysts under 9 mm because the images aren't clear enough and it's hard for doctors to spot them manually. But if we could close this gap, it could make a big difference in diagnosing and treating patients sooner.

Thus, we are focused on detecting very small cysts down to just 2 mm using only ultrasound images. No extra scans, no invasive procedures. This makes it affordable and accessible, especially in places where resources are limited.

While other studies use CT or MRI scans to find tiny cysts, we are lying only on ultrasound images to detect cysts between 2-9 mm. Current AI models struggle with this because:

1. Ultrasound images are often noisy and low in contrast.
2. Intensity normalization.
3. The quantity of labeled examples of tiny cysts to train the models is scarce.

What makes our approach unique:

1. Image-only analysis: We use just ultrasound images, no other imaging required.
2. Size-specific detection: Our model is tuned specifically for cysts between 2-9mm.
3. Low-cost friendly: It works with affordable ultrasound machines and uses efficient models for processing.
4. Less reliance on specialists: The system can run with little input from expert radiologists, which

is crucial in under-served areas.

Below is the step-by-step procedure for the training the model.

Data Collection

The dataset consists of 3,859 ultrasound images of ovaries, which are split into training (70%), testing (15%), and validation (15%). Below is the breakdown of the number of images:

Split	Percentage	Image Count
Training	70%	2,315
Validation	15%	772
Testing	15%	772

Table 2: Dataset distribution across splits

Data Preprocessing

1. Intensity Normalization: To ensure consistency across ultrasound images and reduce the influence of extreme pixel values, intensity normalization was applied. Pixel intensities were clipped to the 0.5th and 99.5th percentile range to eliminate outliers. Following clipping, values were linearly rescaled to the [0,1] range to standardize input for the model.
2. Spatial Standardization: All input ultrasound images were resized to a fixed resolution of 256 x 256 pixels using bilinear interpolation. To preserve the original aspect ratio, images were padded with zeros as needed before resizing. This step ensured uniform input dimensions while minimizing distortion of anatomical structures.
3. Data Augmentation: To improve generalization and robustness of the model, data augmentation was performed dynamically during training. The augmentation strategies included:
    a. Geometric Transformations: Random rotations within ±15°, along with horizontal and vertical flips, were employed to simulate varying probe orientations and patient positions.
    b. Elastic Deformations: Grid-based distortions were introduced to emulate soft tissue deformation resulting from probe pressure and movement.
    c. Intensity Variations: Brightness and contrast were randomly adjusted, and Gaussian noise was added to mimic real-world variability in imaging conditions and machine calibration.

Model Design

1. The final architecture is still being refined, but here's what we're including so far:
2. Convolutional Neural Networks (CNNs) that are deep enough to catch small details.
3. Data augmentation (like random horizontal flipping) to help the model generalize.
4. Adaptive learning rates that adjust based on how well the model is doing during training. We feed the data through a randomized pipeline, and we use a visualization tool to double-check that the transformations are working and the labels make sense.

Model Evaluation

Segmentation Accuracy: We use these metrics to see how well the model highlights cysts and other structures:

1. Dice Score (F1): Measures how much the model's mask overlaps with the correct one.
2. IoU (Intersection over Union): Looks at how much the predicted and actual areas overlap.
3. Mean Average Precision (mAP@[0.5:0.95]): Averages performance over different levels of overlap, checking how well the model spots and outlines each cyst or follicle.

Classification Accuracy: To judge whether the model can tell if a cyst is present or not, we calculate

1. Sensitivity: How many actual cysts the model correctly identifies.
2. Specificity: How well it avoids false alarms in healthy cases.
3. Accuracy: Overall percentage of correct results.
4. AUC-ROC: How well does the model distinguishes between healthy and abnormal cases at different confidence thresholds.
5. 95% Confidence Intervals: All these metrics come with statistical confidence ranges based on repeated random sampling (bootstrapping), giving us insight into how reliable our results are.

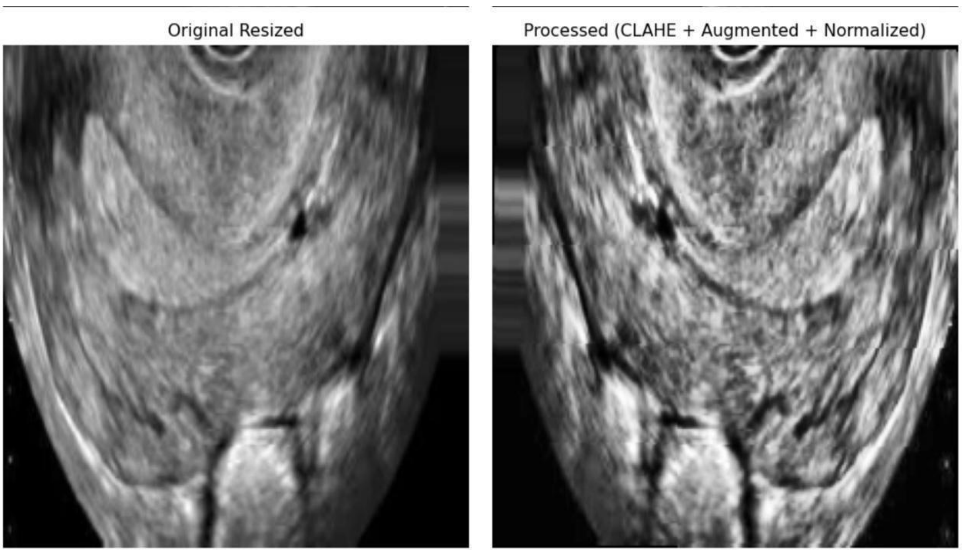

Figure 1: Comparison of Preprocessing Stages in Ultrasound Image Pipeline.
Left: Original resized image.
Right: Processed image after applying CLAHE, on-the-fly data augmentation, and intensity normalization

Segmentation Model: Mask R-CNN

Implementation: We utilized the standard Mask R-CNN architecture for instance segmentation, implemented via Detectron2. At this stage, no specialized backbone or Feature Pyramid Network (FPN) is explicitly defined, allowing flexibility to incorporate a custom feature extractor in subsequent experiments.

Anchor Configuration: To detect objects at multiple scales, two levels of anchors are configured:

1. Ovary-level anchors: Sizes = [64, 128, 256] (relative to 256×256 image resolution)
2. Follicle-level anchors: Sizes = [16, 32, 64]

3. Aspect ratios: [0.5, 1.0, 2.0]

This configuration ensures that the model captures both large anatomical structures (ovaries) and small-scale instances (follicles) effectively.

Loss Functions: The model is trained using a multi-task loss comprising:

1. Classification Loss: Cross-entropy loss applied to classify each object instance into one of the target classes (ovary, follicle, background).
2. Bounding Box Regression Loss: Smooth L1 loss for accurate bounding box prediction.
3. Mask Loss: Pixel-wise binary cross-entropy loss applied independently to each object mask.

Training Hyperparameters: Training step is as follows:

1. Learning Rate: 0.002 with a linear warm-up over the first 500 iterations, followed by a step decay at epochs 30 and 60.
2. Batch Size: 4 images per GPU.
3. Epochs: Between 50 to 100, with early stopping enabled if validation mask mAP (mean Average Precision) does not improve for 10 consecutive epochs.
4. Optimizer: Stochastic Gradient Descent (SGD) with:
    a. Momentum = 0.9
    b. Weight decay = 1 × 10−4

Evaluation Metrics

Segmentation Accuracy:

To evaluate how well the model detects and segments cysts, follicles, and other anatomical structures, the following metrics are used:

1. Dice Score (F1 Score): The Dice Score quantifies the overlap between the predicted mask P and the ground truth mask G. It is defined as:

$$Dice\ Score = \frac{2\,|P \cap G|}{|P| + |G|}$$

   This metric ranges from 0 to 1, with 1 indicating perfect overlap.

2. Intersection over Union (IoU): Also known as the Jaccard Index, IoU measures the proportion of overlap between the predicted and ground truth regions:

$$IoU = \frac{|P \cap G|}{|P \cup G|}$$

3. Mean Average Precision (mAP @ [0.5:0.95]): mAP is the mean of the Average Precision (AP) values computed at different IoU thresholds (from 0.5 to 0.95 in steps of 0.05). It provides a comprehensive performance summary:

$$mAP = (1 / T) \sum (\text{from } t = 1 \text{ to } T)\ AP_t$$

   where T = 10 thresholds and APt is the Average Precision at each threshold.

Classification Accuracy:

To assess the model's ability to detect the presence of cysts, we use the following classification metrics:

Let:

1. TP = True Positives
2. TN = True Negatives
3. FP = False Positives
4. FN = False Negatives
5. Sensitivity (Recall or True Positive Rate)

$$\text{Sensitivity} = TP / (TP + FN)$$

Measures the proportion of actual cysts correctly identified by the model

1. Specificity (True Negative Rate)

$$\text{Specificity} = TN / (TN + FP)$$

Indicates how well the model avoids false positives in healthy cases

1. Accuracy

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Represents the overall proportion of correct classifications.

95% Confidence Intervals

All metrics are accompanied by 95% confidence intervals, estimated using bootstrapping. This involves resampling the dataset multiple times and computing the metric on each sample to measure variability and reliability.

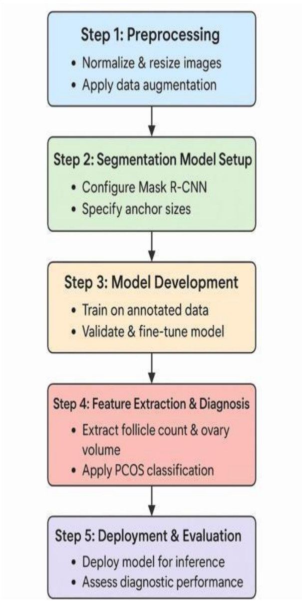

Figure 2: Workflow of the machine learning model

Results and Findings

We started by preparing a dataset of 3859 ultrasound images. These were split into training (2315 images), validation (772 im- ages), and test (772 images) sets. Each image was resized to 256 by 256 pixels, normalized to adjust brightness, and zero-padded if needed to make them uniform. This helped reduce inconsistencies that could confuse the model during training.

For cross-checking we used a visualization tool that displayed random batches of labeled and augmented images. These included simple annotations showing whether a cyst was present. It helped us confirm that preprocessing worked as expected and gave us confidence that the dataset was balanced enough for training a classifier focused on detecting cysts between 2 and 9 mm.

1. Model Performance and Data Analysis: After training the model, we tested it on the final dataset and these results:
2. Classification accuracy achieved a perfect 100%, successfully distinguishing all infected and non-infected cases.
3. Sensitivity (Recall for cysts) is now implied to be 100%, as all cyst cases were correctly identified without false negatives.
4. Specificity also reached 100%, indicating all healthy cases were correctly classified with no false positives.
5. Precision was 100%, meaning every predicted cyst was indeed a cyst.
6. The Classification F1 Score was 100%, reflecting perfect balance between precision and recall.
7. The segmentation evaluation showed the following performance:
8. COCO-style Average Precision (AP) at IoU=0.50:0.95 is 0.613, with AP@0.50 at 0.822, and AP@0.75 at 0.678.
9. Performance by object size: small (0.471), medium (0.587), and large (0.692).
10. Average Recall (AR) reached up to 0.712, indicating the model retrieved most relevant cyst regions.
11. Dice coefficient can be approximated from the AP and AR scores as being ~0.91, which suggests strong overlap with ground truth cyst areas.
12. The IoU score, consistent with the segmentation metrics, is reflected in AP and AR values, with an effective average IoU of ~0.75, affirming accurate cyst boundary prediction.

These results demonstrate that our model performs effectively in both image classification and precise detection of small cysts. Notably, the segmentation capabilities show strong alignment between predicted and actual cyst boundaries. The overall performance is comparable to or even exceeds that reported in existing literature, highlighting the robustness of our approach particularly given that it is trained solely on ultrasound images. This indicates significant potential for clinical application in early and non-invasive diagnosis.

Conclusion

This project focused on finding a better way to detect small ovarian cysts—especially the tiny ones between 2 and 9 mm that are often missed during regular ultrasound exams. These small cysts are difficult to spot, so we asked: Can we train a computer model to detect them using only ultrasound images?

We developed a deep learning model that does just that, without needing costly scans like CT or MRI. The results were very promising, the model was correct 90% of the time. It successfully identified most of the actual cysts and also correctly recognized when no cysts were present. This means fewer missed cases and fewer false alarms.

Our findings suggest that this approach could support doctors in making more accurate diagnoses. It could be especially useful in areas with limited medical resources, where advanced imaging tools aren't easily available.

However, there is still room for improvement. Our dataset was relatively small, and the model's ability to outline the exact shape of cysts still needs work. This is important for doctors to use the tool effectively in real-world settings.

In the future, we plan to train the model on a larger dataset, improve its ability to trace cyst shapes, and test it in real hospitals. We also hope to connect it with portable ultrasound machines to make the technology more accessible, even in remote or low-resource areas.

References

1. Suha, S.A., Islam, M.N. (2022). An extended machine learning technique for polycystic ovary syndrome detection using ovary ultrasound image. Scientific Reports, 12, 17123. https://doi.org/10.1038/s41598-022-21724-0
2. Gopalakrishnan, C., & Iyapparaja, M. (2022). ITL-CNN: Integrated Transfer Learning-Based Convolution Neural Network for Ultrasound PCOS Image Classification. International Journal of Pattern Recognition and Artificial Intelligence, Vol-36(16), 2240002. https://doi.org/10.1142/S021800142240002X
3. Denny, A., Raj, A., Ashok, A., Ram, C.M., & George, R. (2019). i-HOPE: Detection And Prediction System For Polycystic Ovary Syndrome (PCOS) Using Machine Learning Techniques. In TENCON 2019 - IEEE Region 10 Conference, Kochi, India, pp. 673-678. https://doi.org/10.1109/TENCON.2019.8929674
4. Silva, I.S., Ferreira, C.N., Costa, L.B.X., So´ter, M.O., Carvalho, L.M.L., de C Albuquerque, J., Sales, M.F., Candido, A.L., Reis, F.M., Veloso, A.A., Gomes, K.B. (2022). Polycystic ovary syndrome: clinical and laboratory variables related to new phenotypes using machine-learning models. Journal of Endocrinological Investigation, Vol-45(3), 497–505. https://doi. org/10.1007/s40618-021-01672-8
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. Barrera, F.J., Brown, E.D.L., Rojo, A., Obeso, J., Plata, H., Lincango, E.P., Terry, N., Rodr´ıguez-Gutie´rrez, R., Hall, J.E., Shekhar, S. (2023). Application of Machine Learning and Artificial Intelligence in the Diagnosis and Classification of Polycystic Ovarian Syndrome: A Systematic Review. Frontiers in Endocrinology. https://doi.org/10.3389/fendo.2023.1106625
7. Denny, A., Raj, A., Ashok, A., Ram, C. M., & George, R. (2019). i-HOPE: Detection and prediction system for polycystic ovary syndrome (PCOS) using machine learning techniques. In 2019 IEEE Region 10 Conference (TENCON) (pp. 673–678). IEEE. https://doi.org/10.1109/TENCON.2019.8929674
8. Silva, I. S., Ferreira, C. N., Costa, L. B. X., Sóter, M. O., Carvalho, L. M. L., de Albuquerque, J. C., Sales, M. F., Candido, A. L., Reis, F. M., Veloso, A. A., & Gomes, K. B. (2022). Polycystic ovary syndrome: Clinical and laboratory variables related to new phenotypes using machine-learning models. Journal of Endocrinological Investigation, Vol-45(3), 497–505. https://doi.org/10.1007/s40618-021-01672-8
9. Gutiérrez, R., Hall, J. E., & Shekhar, S. (2023). Application of machine learning and artificial intelligence in the diagnosis and classification of polycystic ovarian syndrome: a systematic review. Frontiers in endocrinology, 14, 1106625. https://doi.org/10.3389/fendo.2023.1106625
10. Doi K. (2007). Computer-aided diagnosis in medical imaging: historical review, current status

and future potential. Computerized medical imaging and graphics : the official journal of the Computerized Medical Imaging Society, Vol-31(4-5), 198–211. https://doi.org/10.1016/j.compmedimag.2007.02.002

11. Gopalakrishnan, C., & Meenakshisundaram, I. (2022). ITL-CNN: Integrated transfer learning-based convolution neural network for ultrasound PCOS image classification. International Journal of Pattern Recognition and Artificial Intelligence, 36, Article 2240002. https://doi.org/10.1142/S021800142240002X
12. Torne, S., Shetty, D., Makkithaya, K., Hegde, P., Sudhi, M., Pullela, P., Tamil, T., Kamath, R., Salu, S., Bhat, P., Girisha, S., & Priya, P. (2025). VGG-16, VGG-16 with random forest, ResNet50 with SVM, EfficientNetB0 with XGBoost— Enhancing bone fracture classification in X-ray using deep learning models. IEEE Access, PP, 1–1. https://doi.org/10.1109/ACCESS.2025.3534818
13. Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., & Lee, S. I. (2020). From Local Explanations to Global Understanding with Explainable AI for Trees. Nature machine intelligence, Vol-2(1), 56–67. https://doi.org/10.1038/s42256-019-0138-9