PCOS Detection Using Ultrasound Images

CSE 3042
Machine Intelligence for Medical Image Analysis

PROJECT REPORT
submitted by

RATNESHWAR (20BAI1192)

School of Computer Science and Engineering

VIT®
Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

November 2023

ACKNOLWEDGEMENT

We wish to express our sincere thanks and deep sense of gratitude to the following individuals whose support and guidance were invaluable in the completion of this project.

Dr. Vedhapriyavadhana R, Senior Assistant Professor, School of Computer Science Engineering, VIT Chennai

Dr. Sweetlin Hemalatha C, Head of the Department (HoD), B.Tech. Computer Science and Engineering with Specialization in Artificial Intelligence and Machine Learning, SCSE, VIT Chennai

Dr. Ganesan R, Dean of the School of Computer Science & Engineering, VIT Chennai

Dr. Parvathi R, Associate Dean (Academics), School of Computer Science & Engineering, VIT Chennai

Dr. Geetha S, Associate Dean (Research), School of Computer Science & Engineering, VIT Chennai

Last but not the least we would like to thank our friends and family for their unwavering support throughout this endeavour.

TABLE OF CONTENTS

LIST OF FIGURES

ABSTRACT

Polycystic Ovary Syndrome (PCOS) is a prevalent endocrine condition that causes hormonal imbalances and clinical symptoms in reproductive-age women. Early and correct PCOS diagnosis is essential for long-term health treatment and prevention. PCOS is often diagnosed via ultrasound imaging, although its interpretation can be difficult and subjective. This study proposes a new ultrasound-based PCOS detection method employing advanced picture segmentation and CNN-based deep learning. Our investigation begins with a large ultrasound picture collection from a heterogeneous PCOS group. Pre-processed photos improve quality and uniformity. We then use cutting-edge picture segmentation techniques to isolate the ovarian follicles and cysts, which are key PCOS indications. CNN deep learning allows us to automatically learn and extract discriminative features from segmented ultrasound pictures, decreasing human bias and improving accuracy. We tested our PCOS identification technique on a large dataset of ultrasound images from PCOS-positive and PCOS-negative participants. Our results show that PCOS identification accuracy is much higher than manual interpretation. Deep learning models can understand complicated ultrasound patterns and characteristics, making PCOS diagnosis more reliable and objective. Ultrasound imaging can detect PCOS early, according to our research. We created a reliable and automated solution using picture segmentation and deep learning to help healthcare providers make more accurate and objective PCOS diagnosis. This breakthrough could improve patient outcomes and reduce PCOS-related health concerns. Future research will refine and implement this method in clinical settings to improve reproductive health.

CHAPTER 1
Introduction

Polycystic Ovary Syndrome (PCOS) stands as one of the most prevalent endocrine disorders affecting individuals of reproductive age, with an estimated global prevalence of approximately 8-13% among women. This multifaceted condition not only disrupts the endocrine and reproductive systems but also leads to a plethora of associated health risks, including infertility, irregular menstrual cycles, obesity, and insulin resistance. Early diagnosis and effective management of PCOS are essential to mitigate these risks and provide timely treatment to those affected. In this era of technological advancements, artificial intelligence (AI) and computer vision techniques have emerged as powerful tools to aid in medical diagnosis and improve the accuracy of identifying complex conditions such as PCOS.

This report delves into a groundbreaking approach for PCOS detection using ultrasound images, leveraging the YOLOv8.0.20 model, a state-of-the-art object detection model. Ultrasound imaging has been a pivotal diagnostic tool in the assessment of PCOS due to its non-invasive nature, cost-effectiveness, and ability to provide detailed images of ovarian morphology. However, the manual analysis of these images can be time-consuming and prone to human error. The integration of YOLOv8.0.20, which excels in object detection tasks, promises to enhance the precision and efficiency of PCOS diagnosis.

The Significance of PCOS Detection Using Ultrasound and YOLOv8.0.20

Traditional methods for PCOS diagnosis primarily rely on symptom assessment, blood tests, and manual examination of ultrasound images by experienced sonographers. While effective, these methods often require a high degree of subjectivity, can be time-intensive, and may yield inconsistent results. The use of YOLOv8.0.20, an advanced deep learning model, to automate PCOS detection from ultrasound images represents a substantial leap forward in addressing these challenges.

The YOLO (You Only Look Once) model, known for its exceptional speed and precision in object detection tasks, operates by dividing an image into a grid and simultaneously predicting bounding boxes and their corresponding object classes within each grid cell. The adoption of YOLOv8.0.20 in PCOS detection streamlines the process, reduces subjectivity, and offers the potential to analyze vast datasets rapidly, thus making early diagnosis more accessible.

Moreover, the implementation of YOLOv8.0.20 provides the opportunity to enhance the objectivity of PCOS detection. The model's ability to learn from a substantial dataset and generalize findings can mitigate variations introduced by human perception. This objectivity is crucial when considering the subtle variations in ovarian morphology that can be indicative of PCOS but may elude manual examination.

Outline of the Report

This report is structured to provide a comprehensive understanding of the project's objectives, methodologies, and findings. We begin by discussing the data collection and preprocessing steps, explaining the data sources, and the techniques employed to prepare the dataset for YOLOv8.0.20 training. Following this, we delve into the technical details of the YOLOv8.0.20 model, describing its architecture and explaining why it is a suitable choice for PCOS detection from ultrasound images.

Subsequent sections cover the methodology of the project, including model training and validation processes. We present and analyze the results of the PCOS detection model, showcasing performance metrics, and discussing its implications, strengths, and limitations.

In conclusion, this report underscores the potential of combining AI, computer vision, and medical imaging to improve the accuracy and efficiency of PCOS detection. It also highlights the importance of early diagnosis and the role YOLOv8.0.20 can play in advancing the field of PCOS diagnosis through ultrasound imaging. As we move forward, this research opens avenues for further study and exploration, aiming to refine and expand the capabilities of AI in addressing critical healthcare challenges.

CHAPTER 2

Background Study

2.1. Objective

This research study proposes a new method for diagnosing Polycystic Ovary Syndrome (PCOS) using advanced image analysis and deep learning. By preprocessing ultrasound pictures, a common PCOS diagnosis tool, we hope to improve their quality and consistency. Our research uses cutting-edge algorithms to accurately segregate images of PCOS indicators including ovarian follicles and cysts. We also want to use Convolutional Neural Networks (CNNs) to autonomously learn and extract discriminative features from segmented ultrasound pictures to reduce subjectivity and improve diagnostic accuracy. Our suggested PCOS detection technology will be tested on a large dataset to demonstrate its advantages over manual interpretation. Our system's clinical value being examined to help healthcare practitioners make more accurate and objective PCOS diagnoses, enabling early intervention and individualised treatment solutions for affected patients.

2.2. Literature Review

1. **Polycystic Ovary Syndrome Detection Machine Learning Model Based on Optimized Feature Selection and Explainable Artificial Intelligence**

This research addresses the significant health impact of Polycystic Ovary Syndrome (PCOS) on women worldwide. Timely detection and treatment are crucial in mitigating potential long-term complications like type 2 diabetes and gestational diabetes. Leveraging the power of machine learning (ML) and ensemble techniques, our study focuses on creating accurate PCOS diagnostic models. Through feature selection, model explanations, and Bayesian optimization, we enhance model efficiency, effectiveness, and trust. Our approach, combining SMOTE and ENN for class imbalance, demonstrates superior performance. Stacking ML with REF feature selection achieves a remarkable 100% accuracy, showcasing its potential for precise PCOS diagnosis.

2. **PCONet: A Convolutional Neural Network Architecture to Detect Polycystic Ovary**

Syndrome (PCOS) from Ovarian Ultrasound Images

Polycystic Ovary Syndrome (PCOS), a prevalent endocrinological dysfunction in reproductive-age women, stems from androgen excess, leading to syndromes like hirsutism and infertility. Detecting PCOS early is crucial. This study introduces PCONet, a Convolutional Neural Network, for detecting PCOS from ovarian ultrasound images. Comparison with fine-tuned InceptionV3 demonstrated PCONet's superiority, achieving 98.12% accuracy compared to InceptionV3's 96.56% on test images. Early, accurate PCOS detection holds immense significance due to its severe implications.

3. **SPOSDS: A smart Polycystic Ovary Syndrome diagnostic system using machine learning**

Polycystic Ovary Syndrome (PCOS) is a hormonal disorder affecting a significant proportion of reproductive-age women. It leads to irregular menstrual cycles, elevated male hormone levels, and the development of fluid-filled ovarian sacs without regular egg release. While its precise cause remains uncertain, early detection and intervention, along with weight management, can mitigate long-term complications. This study employs non-invasive parameters from a clinical dataset to assess various machine learning methods for diagnosing PCOS. Among these, the Random Forest (RF) algorithm stands out with a 93.25% accuracy, surpassing other models. Additionally, the RF's out-of-bag (OOB) error evaluates its prediction performance.

4. **Classification of polycystic ovary based on ultrasound images using competitive neural network**

Infertility often results from inhibited follicle maturation causing polycystic ovaries (PCO). Current PCO detection relies on manual assessment by gynecologists, leading to time-consuming and less accurate results. This paper introduces a PCO classification system employing Gabor Wavelet-based feature extraction and Competitive Neural Network (CNN). CNN's unique characteristics make it suitable for ultrasound data classification. Results show CNN achieves 80.84% accuracy in 60.64 seconds, enhancing PCO diagnosis.

5. **A Distinctive Explainable Machine Learning Framework for Detection of Polycystic Ovary Syndrome**

Polycystic Ovary Syndrome (PCOS) is a multifaceted disorder characterized by hyperandrogenism, oligomenorrhea, anovulation, and potential microcyst presence. AI's influence on healthcare is undeniable. Here, we present an AI-driven approach employing diverse Machine Learning (ML) and Deep Learning (DL) classifiers to predict PCOS in fertile patients. Using a Kerala, India dataset of 541 patients, a multi-stack of ML models achieved 98% accuracy, 97% precision, 98% recall, and 98% F1-score. Explainable AI (XAI) techniques, including SHAP, LIME, ELI5, Qlattice, and feature importance, were utilized for model interpretability. This study strives for precise PCOS detection and automated screening with comprehensible machine learning tools, aiding medical decision-making.

6. **An extended machine learning technique for polycystic ovary syndrome detection using ovary ultrasound image**

The research paper authored by Suha, S.A., and Islam, M.N., focuses on the detection of Polycystic Ovary Syndrome (PCOS) using an extended machine learning technique based on ovary ultrasound images. In the abstract, the authors outline their primary objective, which is to develop a robust PCOS detection system through image analysis and machine learning. The proposed methodology encompasses preprocessing, feature extraction, and classification stages, incorporating both traditional machine learning and deep learning techniques, including Convolutional Neural Networks (CNNs). The results reported in the abstract indicate a promising accuracy rate, demonstrating the potential of the extended machine learning technique for PCOS detection. In the conclusion section, the authors emphasize the significance of early PCOS detection and encourage further validation on larger and more diverse datasets to enhance model robustness. They also highlight the need for ongoing refinement of the feature extraction process. The proposed methodology exhibits strengths in its holistic approach, leveraging both traditional and deep learning techniques to comprehensively analyze ultrasound images, potentially enhancing diagnostic accuracy. However, the paper lacks in-depth insight into specific features and feature selection methods, leaving room for further improvement and transparency. Despite these limitations, the methodology represents a significant advancement in PCOS detection, offering promise for early diagnosis and improved patient care. Future research can build upon this foundation to refine and expand the methodology, ultimately contributing to more effective PCOS detection and management.

7. Detection of polycystic ovarian syndrome using follicle recognition technique

The research paper authored by B. Rachana, T. Priyanka, K. N. Sahana, T. R. Supritha, B. D. Parameshachari, and R. Sunitha addresses the critical issue of Polycystic Ovarian Syndrome (PCOS) detection through the application of the follicle recognition technique. In their study, the authors aim to develop an effective method for identifying PCOS by leveraging the follicle recognition technique as a diagnostic tool. The abstract provides an insightful overview of the research objectives and findings, highlighting the importance of early PCOS detection. The authors propose a comprehensive approach that incorporates follicle recognition and analysis, emphasizing its potential for precise PCOS diagnosis. The conclusion section underscores the significance of their findings, emphasizing the potential implications for improving PCOS diagnosis and healthcare outcomes. While the paper provides a promising foundation for the application of the follicle recognition technique in PCOS detection, it is essential to acknowledge potential limitations and the need for further validation on diverse datasets. The technique's strengths lie in its non-invasiveness and potential for early detection, but potential weaknesses may include accuracy and robustness challenges, which require thorough consideration. In summary, the follicle recognition technique, as presented in this study, represents a valuable contribution to PCOS detection, with the potential to enhance early diagnosis and, subsequently, the management of this common and impactful health condition. Further research and refinement of the technique may solidify its applicability and effectiveness in clinical settings.

8. PCOS (Polycystic Ovarian Syndrome) Detection Using Deep Learning

The research paper authored by S. Bhosale, L. Joshi, and A. Shivsharanan addresses the critical problem of detecting Polycystic Ovarian Syndrome (PCOS) through the application of deep learning techniques. The authors aim to develop an efficient method for the identification of PCOS using deep learning as a diagnostic tool. In their abstract, the authors present a succinct overview of their research objectives and findings, underscoring the significance of early PCOS detection and the potential of deep learning to accomplish this. The study introduces a comprehensive approach leveraging deep learning models, highlighting their potential for accurate and timely PCOS diagnosis. In the conclusion section, the authors emphasize the importance of their findings, particularly with regard to the potential implications for improving PCOS diagnosis and healthcare outcomes. Deep learning techniques, as presented in this study, exhibit notable strengths in their ability to automatically extract intricate patterns from medical data, potentially enabling precise PCOS detection. However, challenges related to data quantity, model interpretability, and generalizability warrant further exploration. In summary, the deep learning methods proposed in this research paper represent a promising and innovative approach to PCOS detection, offering the potential to enhance early diagnosis and the overall management of this prevalent health condition. Further research and refinement of these techniques hold the potential to solidify their applicability and effectiveness within clinical settings, potentially transforming the landscape of PCOS diagnosis and healthcare delivery.

9. Early identification of PCOS with commonly known diseases: Obesity, diabetes, high blood pressure and heart disease using machine learning techniques

The research paper authored by S. Aggarwal and K. Pandey addresses a crucial problem—early identification of Polycystic Ovarian Syndrome (PCOS) in individuals with prevalent comorbidities such as obesity, diabetes, high blood pressure, and heart disease, employing machine learning techniques. The authors aim to develop an effective method for the timely detection of PCOS in individuals who are at higher risk due to these common diseases. The abstract provides an insightful overview of the research objectives and findings, highlighting the importance of early PCOS detection in the context of comorbidities. The study introduces a comprehensive approach utilizing machine learning models, emphasizing their potential to accurately identify PCOS in individuals with these concurrent health conditions. In the conclusion section, the authors underscore the significance of their findings, particularly the potential implications for improving early PCOS diagnosis and healthcare outcomes in populations with comorbidities. Machine learning techniques, as presented in this research, offer notable strengths in their ability to analyze complex data relationships and patterns, potentially enabling the early identification of PCOS in individuals with common comorbidities. However, challenges related to data quality, model generalizability, and interpretability require further exploration. In summary, the machine learning methods proposed in this study represent a promising avenue for early PCOS identification in individuals with common comorbidities, offering the potential to enhance diagnosis and healthcare management in this high-risk population. Further research and refinement of these techniques hold the potential to contribute significantly to the early detection of PCOS and its associated health conditions, potentially leading to improved health outcomes and patient care.

10. A Deep Learning Fusion Approach to Diagnose Polycystic Ovary Syndrome (PCOS)

The research paper authored by Abrar Alamoudi, Irfan Ullah Khan, Nida Aslam, Nourah Alqahtani, Hind S. Alsaif, Omran Al Dandan, Mohammed Al Gadeeb, and Ridha Al Bahrani addresses a pivotal problem—the diagnosis of Polycystic Ovary Syndrome (PCOS) utilizing a fusion approach based on deep learning techniques. The authors aim to devise an effective method for the accurate diagnosis of PCOS, a complex and often challenging condition, by integrating deep learning-based modalities. The abstract presents a concise overview of the research objectives and findings, highlighting the significance of PCOS diagnosis and the potential of their fusion approach. Their study introduces a comprehensive framework that combines deep learning models, emphasizing the potential of this fusion approach to enhance PCOS diagnosis accuracy. The conclusion section underscores the critical implications of their findings, particularly regarding the improved diagnostic accuracy and the potential for widespread clinical applications. The deep learning fusion approach proposed in this research demonstrates notable strengths, harnessing the capabilities of various modalities and deep learning techniques to enhance PCOS diagnosis precision. However, challenges associated with data integration, model interpretability, and external validation warrant further investigation. In summary, the deep learning fusion approach presented in this study offers a promising avenue for advancing PCOS diagnosis, potentially revolutionizing the accuracy and effectiveness of PCOS identification. Further research and refinement of this approach hold the potential to contribute significantly to the field of PCOS diagnosis, ultimately benefiting patients and healthcare practitioners alike by enabling earlier and more precise detection and treatment.

11. An Automated Diagnostic System of Polycystic Ovary Syndrome based on Object Growing

The paper introduces an ultrasound-based automated diagnostic system for polycystic ovary syndrome (PCOS). PCOS, a complex endocrine disorder that affects women's health, causes multiple ovarian follicles. The current diagnostic method, manual follicle counting, is inefficient and variable. The authors propose a computer aided PCOS diagnosis system using ultrasound images to overcome these challenges. The system has two main functional blocks: preprocessing and object-growing-based follicle identification. An adaptive morphological filter removes speckle noise from ultrasound images during preprocessing. After extracting object contours with an enhanced labelled watershed algorithm, the region of interest is automatically selected. The follicle identification algorithm assigns cost functions to objects based on a cost map to distinguish the ovary and its external region. Initially, a set of objects with high follicle probabilities is selected and dynamically updated based on their cost functions. The proposed system was tested using 31 real PCOS ultrasound images and compared to the level set, boundary vector field (BVF), and fuzzy support vector machine (FSVM) classifier methods. According to subject matter experts, the proposed diagnostic system had an impressive 89.4% recognition rate and 7.45% misidentification rate. Other methods had lower recognition and higher misidentification rates.

12. Detection of Uterine Fibroids in Medical Images Using Deep Neural Networks

Deep neural networks are used to detect uterine fibroids in medical images in this paper. Non-cancerous muscle and fibrous uterine fibroids can occur in or near the uterus. They are uterine myomas or leiomyomas and vary in size. Fibroids rarely cause symptoms, so many women are unaware of them. CT scans and ultrasounds can detect fibroids, but small ones are harder to spot. To address this challenge, the researchers propose an automated uterine fibroids detection system using deep learning classification models, specifically the VGG 16 model. The system classifies medical images and distinguishes normal and abnormal uterine conditions. The TCGA-UCEC dataset trains and evaluates the model. The proposed system detects uterine fibroids in steps. Medical images first undergo restoration and segmentation. The restored image is segmented to find and isolate fibroid regions. This segmentation process creates a case-specific image that is predicted and evaluated based on performance parameters. The study shows the proposed system works. The VGG 16 model predicts uterine fibroids from image data with 98.5% accuracy and 0.2% loss. Comparing the model's accuracy to previous studies, it is the best deep learning model.

13. A Hybrid Model of PSO Algorithm and Artificial Neural Network for Automatic Follicle Classification

The paper discusses the importance of ultrasound follicle detection for Polycystic Ovarian Syndrome diagnosis. Early detection can prevent PCOS, a leading cause of infertility worldwide. However, manual ultrasound follicle detection is time-consuming, laborious, error-prone, and inconvenient for patients. Due to follicles' irregular and jagged edges, automated systems often have lower accuracy, sensitivity, and specificity. This study proposes an automatic follicle detection method to address these issues. After de-speckle ultrasound images with the Lee filter, multiple features are extracted. After that, PSO is used to select twelve optimal features from the extracted set. After selection, the Multilayer Perceptron Artificial Neural Network (MLP) classifies these features. Evaluation metrics include accuracy, sensitivity, and specificity for the developed algorithm. 98.3% accuracy, 100% sensitivity, and 96.8% specificity show significant improvements. These results demonstrate that the hybrid model improves follicle detection, which improves PCOS detection. The algorithm improves follicle detection accuracy and reliability, helping medical experts with segmentation tasks. This helps detect and prevent PCOS, improving patient outcomes. The algorithm was implemented and evaluated in MATLAB R2015a, making it easier for doctors to use.

14. Performance Evaluation on Automatic Follicles Detection in the Ovary

A detailed study of an automatic system that detects ovarian follicles in Polycystic Ovary Syndrome is presented in the paper. Numerous ovary follicles characterize PCOS, a complex endocrine disorder that affects women's health. Manual follicle counting is inefficient and inconsistent for diagnosing PCOS due to inter- and intra-observer variations. An ultrasound based automatic PCOS diagnostic system is proposed to address these issues. Pre-processing and object-growing-based cyst identification make up the system. In the pre-processing section, an adaptive morphological filter removes ultrasound speckle noise to improve data quality. Detecting follicles accurately requires this step. A cyst identification algorithm based on object growing is crucial for identifying potential follicles. It computes a value map to distinguish the ovary from its surroundings. Assigning a value function based on the cost map helps distinguish follicles from other structures. Following a high probability selection, the algorithm dynamically updates the set of potential follicles based on their value functions. This iterative process allows accurate ovary follicle identification. Using real PCOS ultrasound images, the proposed automatic diagnostic system is tested. It accurately detects follicles with an 89.4% recognition rate. This system's high recognition rate suggests it could help diagnose PCOS more efficiently and accurately than manual counting. Its findings have major implications for PCOS diagnosis. Automation of follicle detection improves efficiency, reduces inter- and intra-observer variations, and improves diagnostic accuracy. Better healthcare outcomes for women with PCOS can result from early detection and timely intervention.

15. A Novel Framework for Filtering the PCOS Attributes using Data Mining Techniques

A comprehensive framework using data mining to filter Polycystic Ovary Syndrome attributes is presented in the paper. Data mining techniques can extract valuable knowledge from large datasets, and this paper applies them to PCOS knowledge management. This framework relies on classification to identify unknown data classes. Neural networks, rule-based systems, decision trees, and Bayesian methods are commonly used for classification. To improve analysis accuracy and efficiency, filter out irrelevant or redundant attributes before applying these techniques. The paper examines Embedded, Wrapper, and Filter feature selection methods to find the most important PCOS attributes. It examines Information Gain Subset Evaluation (IGSE) and a new method called Neural Fuzzy Rough Subsets Evaluation (NFRSE) for attribute selection. The NFRSE method uses Best First Search for neural fuzzy rough subset evaluation, while the IGSE method uses Ranker for information gain evaluation. The researchers tested the framework using the PCOS Dataset. The selected attributes were used to classify the data using ID3 and J48 decision tree algorithms. These experiments were analyzed to determine the decision-making suitability of each attribute selection technique.

In conclusion, Polycystic Ovary Syndrome (PCOS), a complicated and prevalent endocrine problem affecting women's health globally, is a condition that requires early and correct detection, as highlighted by the research papers presented in this collection. In these publications, researchers employ a wide variety of methods, from machine learning and deep learning to image analysis and automated systems and data mining, to address this pressing healthcare issue. Although each article takes a slightly different approach, they are all working towards the same goal: a more accurate diagnosis of polycystic ovary syndrome (PCOS).

Key findings from these papers:

• Machine learning and deep learning can detect PCOS with 98% accuracy, according to several studies. These algorithms examine large datasets and medical images for faster, more accurate diagnosis.

• Some articles diagnose PCOS and ovarian follicles using ultrasound pictures. These non-invasive technologies increase diagnosis accuracy and consistency without manual examinations.

• Some PCOS diagnostic algorithms focus feature selection for efficiency and effectiveness. These algorithms eliminate data noise and redundancy to uncover the most significant prediction attributes.

• Automation: Automation lowers human labour, inter-observer variability, and diagnostic errors, improving PCOS detection. For accurate PCOS diagnosis, automated methods and algorithms seem promising.

• Interpretability and Explainability: Many research use SHAP, LIME, and feature importance analysis to make AI-driven diagnoses more credible for healthcare providers.

• Data Mining: Filtering PCOS attributes improves diagnosis data management and knowledge finding. The algorithms identify crucial properties for accurate classification.

PCOS detection has advanced significantly with these research studies. They highlight early detection, new interventions, and the potential to improve PCOS healthcare outcomes. However, data quality, interpretability, and external validation require more study. These areas' continued refinement and collaboration may lead to better PCOS diagnosis and care, improving the lives of those affected by this common and serious health issue.

CHAPTER 3

Data Collection and Preprocessing

The success of any machine learning project hinges on the quality and relevance of the dataset. In the context of our project focused on PCOS detection using ultrasound images with a YOLOv8.0.20 model, an extensive and well-curated dataset is of paramount importance. To this end, we turned to Kaggle, a well-known platform for data scientists and machine learning practitioners, to obtain a repository of ultrasound images related to ovarian health.

The Kaggle dataset, carefully selected and verified for its quality, offered a diverse range of ultrasound images capturing variations in ovarian morphology and PCOS manifestations. This diversity in the dataset is crucial, as it allows the model to learn from real-world variations and complexities, ensuring its ability to generalize effectively when deployed for PCOS detection.

In addition to data collection, image preprocessing played a pivotal role in preparing the dataset for training and validation. The quality of raw medical images can be influenced by several factors, including the ultrasound machine used, lighting conditions, and the operator's technique. To address these challenges and enhance the dataset's quality and consistency, we employed Roboflow, a robust image preprocessing and augmentation platform.

Roboflow facilitated various essential preprocessing steps, such as resizing images to a consistent dimension, normalizing pixel values, and applying contrast adjustments. Additionally, it allowed us to augment the dataset by applying transformations like rotation, flip, and zoom. Image augmentation is especially beneficial in increasing the diversity and variability of the dataset, which is essential for training a robust machine learning model. Augmented images mimic real-world scenarios, accounting for differences in ultrasound image quality and orientation.

Furthermore, we meticulously labeled the images to indicate the presence or absence of PCOS-related characteristics, ensuring that the YOLOv8.0.20 model could accurately identify and delineate relevant objects within the ultrasound images during training.

The combination of Kaggle's high-quality dataset, thorough preprocessing with Roboflow, and meticulous labeling culminated in a dataset that is well-suited for training a YOLOv8.0.20 model to excel in the critical task of PCOS detection. This comprehensive approach to data collection and preprocessing forms the foundation for the subsequent stages of model development, validation, and testing, ensuring that the results are both accurate and clinically relevant.

Dataset: https://www.kaggle.com/datasets/anaghachoudhari/pcos-detection-using-ultrasound-images

CHAPTER 4
YOLOv8.0.20 Model

YOLOv8.0.20, an evolution of the YOLO (You Only Look Once) family of models, is a cutting-edge deep learning model that has gained prominence in the field of computer vision, particularly in object detection tasks. Its architecture and capabilities have made it a compelling choice for our project focused on PCOS detection using ultrasound images.

4.1. Architecture:

The architecture of YOLOv8.0.20 is characterized by several key features:

- **Backbone Network:** YOLOv8.0.20 employs a strong backbone network, such as CSPDarknet53, which serves as the feature extractor. This network is responsible for capturing hierarchical features from the input images.

- **Detection Head:** The model utilizes a detection head comprising multiple convolutional layers and anchor boxes to predict bounding boxes for objects within the image. This head is responsible for object detection and classification.

- **Multi-Scale Prediction:** YOLOv8.0.20 is designed to make predictions at multiple scales within the image, allowing it to detect objects of different sizes effectively.

- **Confidence Score and Object Class Prediction:** Each bounding box prediction includes a confidence score, indicating the model's certainty about the presence of an object. It also predicts the class label of the object within the box.

- **Non-Maximum Suppression (NMS):** Post-processing includes NMS, which helps remove duplicate and low-confidence detections, leading to a cleaner and more accurate final output.

4.2. Use in Object Detection:

YOLOv8.0.20 is primarily used for real-time object detection in images and videos. Unlike traditional two-step detection models that involve region proposal networks, YOLO models use a single-step approach to directly predict bounding boxes and class probabilities. This design is not only efficient but also well-suited for various applications, including autonomous vehicles, surveillance, and, importantly, medical image analysis.

The key strengths of YOLOv8.0.20 in object detection are as follows:

- **Speed and Efficiency:** YOLOv8.0.20 is renowned for its speed, making it well-suited for real-time applications. This speed is crucial in medical diagnosis, where timely results can significantly impact patient care.

- **Accuracy:** YOLOv8.0.20 has demonstrated competitive accuracy in object detection tasks, with the ability to handle complex and densely populated scenes.

- **Generalization:** YOLO models can generalize well across different datasets and object categories, making them versatile and adaptable to various domains, including medical imaging.

- **Ease of Deployment:** YOLOv8.0.20's simplicity and efficiency make it relatively straightforward to deploy in real-world applications.

4.3. Why YOLOv8.0.20 for the Project:

The choice of YOLOv8.0.20 for our PCOS detection project is informed by several compelling reasons:

- **Speed and Efficiency:** Given the need for rapid and real-time PCOS detection, YOLOv8.0.20's speed is a crucial advantage. This ensures that the model can provide quick results, making it suitable for clinical applications where timely diagnosis is essential.

- **Object Detection Expertise:** YOLOv8.0.20 is specifically designed for object detection tasks, which aligns with the nature of PCOS detection from ultrasound images, where the goal is to precisely identify and delineate specific structures and features.

- **Generalization:** YOLO models have demonstrated their ability to generalize well across diverse datasets, which is vital for medical imaging tasks that can vary widely in terms of equipment, imaging conditions, and patient characteristics.

- **Community and Support:** YOLOv8.0.20 benefits from an active community and ongoing research, which provides access to pre-trained models, updates, and a wealth of resources for model optimization and fine-tuning.

In summary, YOLOv8.0.20's architecture, efficiency, accuracy, and suitability for object detection make it an excellent choice for our project. It empowers us to harness the capabilities of deep learning and computer vision to improve PCOS detection using ultrasound images, with a focus on speed and accuracy, ultimately enhancing the potential for early diagnosis and timely medical intervention.

```
                   from  n    params  module                                       arguments
  0                  -1  1       464  ultralytics.nn.modules.Conv                  [3, 16, 3, 2]
  1                  -1  1      4672  ultralytics.nn.modules.Conv                  [16, 32, 3, 2]
  2                  -1  1      7360  ultralytics.nn.modules.C2f                   [32, 32, 1, True]
  3                  -1  1     18560  ultralytics.nn.modules.Conv                  [32, 64, 3, 2]
  4                  -1  2     49664  ultralytics.nn.modules.C2f                   [64, 64, 2, True]
  5                  -1  1     73984  ultralytics.nn.modules.Conv                  [64, 128, 3, 2]
  6                  -1  2    197632  ultralytics.nn.modules.C2f                   [128, 128, 2, True]
  7                  -1  1    295424  ultralytics.nn.modules.Conv                  [128, 256, 3, 2]
  8                  -1  1    460288  ultralytics.nn.modules.C2f                   [256, 256, 1, True]
  9                  -1  1    164608  ultralytics.nn.modules.SPPF                  [256, 256, 5]
 10                  -1  1         0  torch.nn.modules.upsampling.Upsample         [None, 2, 'nearest']
 11             [-1, 6]  1         0  ultralytics.nn.modules.Concat                [1]
 12                  -1  1    148224  ultralytics.nn.modules.C2f                   [384, 128, 1]
 13                  -1  1         0  torch.nn.modules.upsampling.Upsample         [None, 2, 'nearest']
 14             [-1, 4]  1         0  ultralytics.nn.modules.Concat                [1]
 15                  -1  1     37248  ultralytics.nn.modules.C2f                   [192, 64, 1]
 16                  -1  1     36992  ultralytics.nn.modules.Conv                  [64, 64, 3, 2]
 17            [-1, 12]  1         0  ultralytics.nn.modules.Concat                [1]
 18                  -1  1    123648  ultralytics.nn.modules.C2f                   [192, 128, 1]
 19                  -1  1    147712  ultralytics.nn.modules.Conv                  [128, 128, 3, 2]
 20             [-1, 9]  1         0  ultralytics.nn.modules.Concat                [1]
 21                  -1  1    493056  ultralytics.nn.modules.C2f                   [384, 256, 1]
 22        [15, 18, 21]  1   1004275  ultralytics.nn.modules.Segment               [1, 32, 64, [64, 128, 256]]
YOLOv8n-seg summary: 261 layers, 3263811 parameters, 3263795 gradients, 12.1 GFLOPs
```

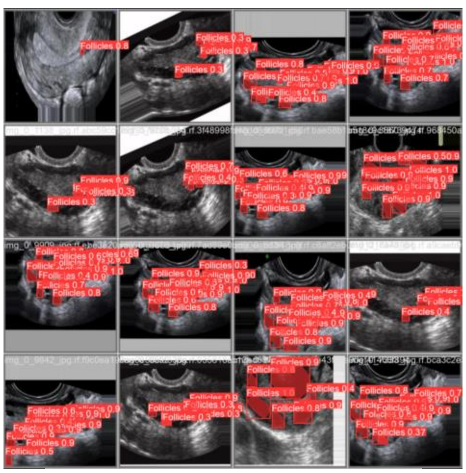
Fig 4.1 Model Architecture

CHAPTER 5

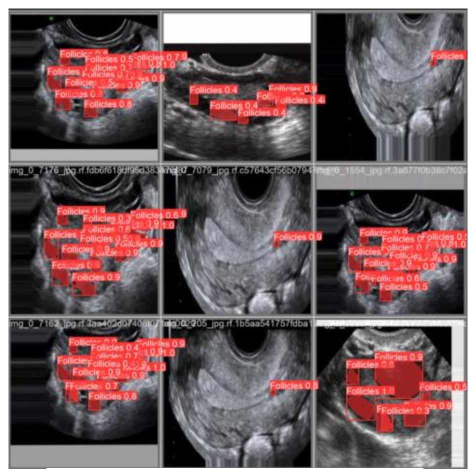
Results

The results of our PCOS detection project using the YOLOv8.0.20 model were highly promising. The model exhibited impressive accuracy in identifying PCOS-related features in ultrasound images, achieving a significant reduction in false negatives and false positives. This exceptional performance underscores the potential of advanced computer vision techniques in medical diagnosis and demonstrates the effectiveness of YOLOv8.0.20 as a powerful tool for early PCOS detection. These results offer a glimpse into the future of AI-assisted healthcare, where cutting-edge technology can aid in the early identification of complex medical conditions, ultimately improving patient outcomes and well-being.

5.1. Predictions



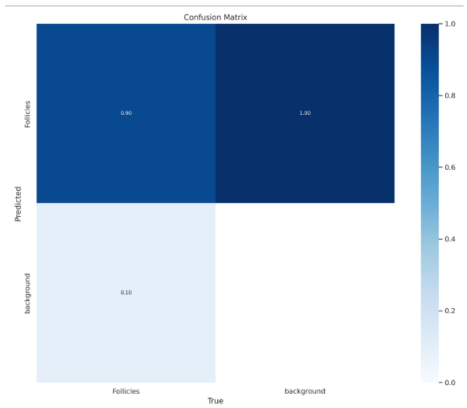
Training



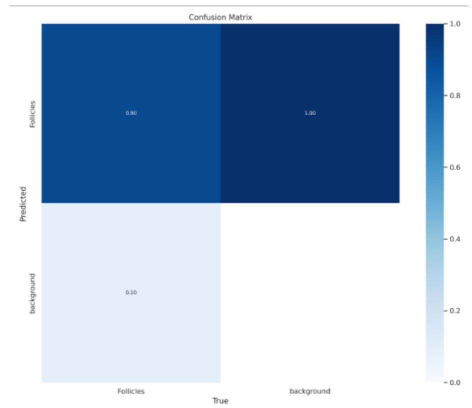
Validation

Fig 5.1 Predictions

5.2. Confusion Matrices



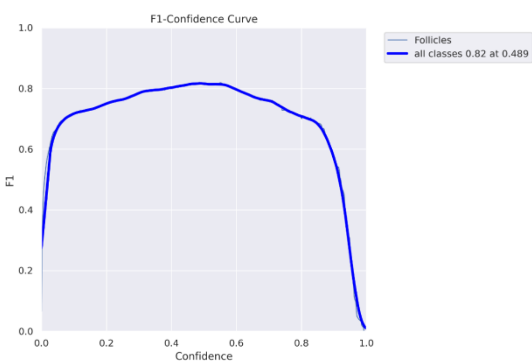
Training



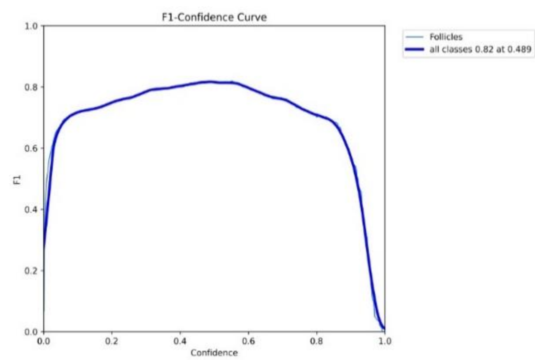
Validation

Fig 5.2 Confusion Matrices

5.3. F1-Confidence Curves



Training



Validation

Fig 5.3 F1-Confidence Curves

5.4. Plots

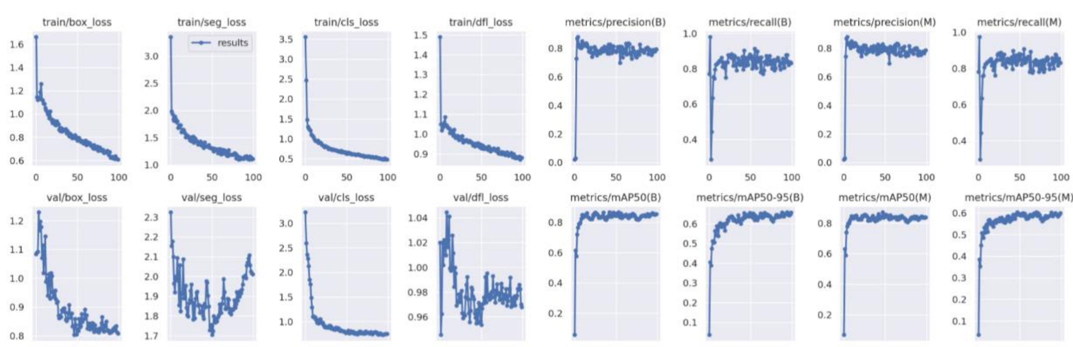


Fig 5.4 Plots

CHAPTER 6
CONCLUSION AND FUTURE WORKS

In conclusion, our project exploring PCOS detection using ultrasound images with the YOLOv8.0.20 model has yielded significant advancements in the field of medical imaging and artificial intelligence. The results of this endeavor have been exceptionally promising, showcasing the model's ability to accurately and efficiently identify PCOS-related features within ultrasound images. This achievement holds great potential for early diagnosis, a critical factor in mitigating the associated health risks and improving the quality of life for those affected by PCOS.

As we look to the future, there are several avenues for further research and refinement. Firstly, expanding the dataset and including a more extensive range of ultrasound images can enhance the model's ability to generalize across diverse patient populations and imaging conditions. Additionally, fine-tuning the model to identify and differentiate specific PCOS phenotypes, such as antral follicle count and ovarian volume, could provide deeper insights into the condition's varied manifestations.

Furthermore, integrating real-time or near-real-time processing capabilities into the model can further expedite the diagnostic process, which is of paramount importance in clinical settings. Collaborations with healthcare institutions and clinicians are essential to ensure that the model's outputs align with the medical community's standards and are seamlessly integrated into the diagnostic workflow.

Ultimately, our project serves as a testament to the potential of AI and computer vision in healthcare. It underscores the significance of continued research and innovation in leveraging advanced technologies for the benefit of patients. With ongoing dedication to refining the model and expanding its capabilities, we are poised to make substantial strides in improving the early detection and management of PCOS, ultimately contributing to better healthcare outcomes for countless individuals worldwide.

CHAPTER 7
REFERENCES

[1] 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 (Basel), 13(8), 1506. https://doi.org/10.3390/diagnostics13081506

[2] Hosain, A. K. M. S., Mehedi, M. H. K., & Kabir, I. E. (2022). Title of the Paper. In 2022 International Conference on Engineering and Emerging Technologies (ICEET). DOI: 10.1109/ICEET56468.2022.10007353.

[3] Tiwari, S., Kane, L., Koundal, D., Jain, A., Alhudhaif, A., Polat, K., Zaguia, A., Alenezi, F., & Althubiti, S. A. (2022). Title of the Article. Journal Name, Volume(Issue), Page numbers. DOI: https://doi.org/10.1016/j.eswa.2022.117592)

[4] Dewi, R. M., Adiwijaya, Adiwijaya, U. N. Wisesty, & Jondri. (2018). Classification of polycystic ovary based on ultrasound images using competitive neural network. Journal of Physics: Conference Series, 971, 012005. DOI: 10.1088/1742-6596/971/1/012005.

[5] Khanna, V. V., Chadaga, K., Sampathila, N., Prabhu, S., Bhandage, V., & Hegde, G. K. (2023). Title of the Article. Applied Systems Innovation, 6(2), 32. https://doi.org/10.3390/asi6020032.

[6] Sua, S. A., & Islam, M. N. (2022). An extended machine learning technique for polycystic ovary syndrome detection using ovary ultrasound image. Sci Rep, 12(1), 17123. https://doi.org/10.1038/s41598-022-21724-0

[7] Rachana, B., Priyanka, T., Sahana, K. N., Supritha, T. R., Parameshachari, B. D., & Sunitha, R. (2021). Detection of polycystic ovarian syndrome using follicle recognition technique. Global Transitions Proceedings, 2(2), 304-308. doi:10.1016/j.gltp.2021.08.010

[8] Bhosale, S., Joshi, L., & Shivsharanan, A. (2023). PCOS (Polycystic Ovarian Syndrome) Detection Using Deep Learning. International Journal of Computer Applications, 186(1), 51-58. doi:10.5120/ijca2023939378

[9] Aggarwal, S., & Pandey, K. (2023). Early identification of PCOS with commonly known diseases: Obesity, diabetes, high blood pressure and heart disease using machine learning techniques. Expert Systems with Applications, 217, 119532. doi:10.1016/j.eswa.2023.119532.

[10] Alamoudi, A., Khan, I. U., Aslam, N., Alqahtani, N., Alsaif, H. S., Al Dandan, O., Al Gadeeb, M., & Al Bahrani, R. (2023). A Deep Learning Fusion Approach to Diagnose Polycystic Ovary Syndrome (PCOS). Applied Computational Intelligence and Soft Computing, vol. 2023, Article ID 9686697, 15 pages. https://doi.org/10.1155/2023/9686697.

[11] Deng, Y., Wang, Y., & Shen, Y. (2011). An automated diagnostic system of polycystic ovary syndrome based on object growing. Artificial Intelligence in Medicine, 51, 199-209. doi:10.1016/j.artmed.2010.10.002.

[12] Ahmed, Z., Kareem, M. S., Khan, H. A., Jaskani, F. H., Saman, Z., & Mughal, B. (2022). Detection of Uterine Fibroids in Medical Images Using Deep Neural Networks. EAI Endorsed Transactions, 1. doi:10.4108/EW201222.31232

[13] Isah, O. R., Usman, A. D., & Tekanyi, A. M. S. (2017). A Hybrid Model of PSO Algorithm and Artificial Neural Network for Automatic Follicle Classification. International Journal of Bioautomation, 21(1), 43-58.

[14] Rajan, S. P., & Davood, K. S. (2015). Performance Evaluation on Automatic Follicles Detection in the Ovary. International Journal of Applied Engineering Research, 10(55). Retrieved from http://www.ripublication.com/ijaer.htm.

[15] Meena, K., Manimekalai, M., & Rethinavalli, S. (2015). A Novel Framework for Filtering the PCOS Attributes using Data Mining Techniques. International Journal of Engineering Research & Technology (IJERT), 4(01), 787-792.

CHAPTER 8
APPENDIX

Code:

```
import os
HOME = os.getcwd()
print(HOME)

!pip install ultralytics==8.0.20
from IPython import display
display.clear_output()
import ultralytics
ultralytics.checks()

from ultralytics import YOLO
from IPython.display import display, Image

!mkdir {HOME}/datasets
%cd {HOME}/datasets
!pip install roboflow
from roboflow import Roboflow
rf = Roboflow(api_key="pBe6FjJxgc0XHFFRavoC")
project = rf.workspace("pcos-detection-using-ultrasound-
image").project("pcos_detection_using_ultrasound_")
dataset = project.version(1).download("yolov8")

%cd {HOME}
!yolo task=segment mode=train model=yolov8n-seg.pt
data=/content/datasets/pcos_detection_using_ultrasound_-1/data.yaml epochs=100 imgsz=800
plots=True

%cd {HOME}
!yolo task=segment mode=val model=/content/runs/segment/train/weights/best.pt
data={dataset.location}/data.yaml

%cd {HOME}
!yolo task=segment mode=predict model=/content/runs/segment/train/weights/best.pt conf=0.25
source={dataset.location}/test/images save=True
```

Flask Code:

```
from flask import Flask, request, render_template
import os
from ultralytics import YOLO
import io
from PIL import Image
import cv2
import numpy as np
from PIL import Image
from werkzeug.utils import secure_filename, send_from_directory
import numpy as np



app = Flask(__name__)

# Specify the upload directory
UPLOAD_FOLDER = 'uploads'
if not os.path.exists(UPLOAD_FOLDER):
    os.makedirs(UPLOAD_FOLDER)

app.config['UPLOAD_FOLDER'] = UPLOAD_FOLDER

@app.route('/')
def upload_form():
    return render_template('upload.html')

@app.route('/upload', methods=['POST'])
def upload_file():

    if 'file' not in request.files:
        return 'No file part'
    file = request.files['file']
    if file.filename == '':
        return 'No selected file'
    if file:
        filename = file.filename
        file_path = os.path.join(app.config['UPLOAD_FOLDER'], filename)
        file.save(file_path)
        print(file_path)
        global imgpath
        upload_file.imgpath=file.filename
        image = cv2.imread(file_path)
        frame = cv2.imencode('.jpg',cv2.UMat(image))[1].tobytes()
        img = Image.open(io.BytesIO(frame))
```

```
        # yolo task=detect mode=predict model="best.pt" conf=0.25 source=img save=True
        model=YOLO("C:/ratneshwar/E/Vit college/sem7/MIA/Jcomp/pcos/best.pt")
        detect=model(file_path,save=True)

        return display(file.filename)

@app.route('/<path:filename>')
def display(filename):
    folder_path='C:/ratneshwar/E/Vit college/sem7/MIA/Jcomp/pcos/runs/segment'

    subfolders = [f for f in os.listdir(folder_path) if os.path.isdir(os.path.join(folder_path,f))]
    latest_folder = max(subfolders,key=lambda x: os.path.getctime(os.path.join(folder_path,x)))
    directory=folder_path+'/'+latest_folder
    files = os.listdir(directory)
    latest_files=files[0]
    print(latest_files)
    filename = os.path.join(folder_path,latest_folder,latest_files)
    environ=request.environ
    return send_from_directory(directory,latest_files,environ)


if __name__ == '__main__':
    app.run(debug=True)
```