Software Requirements Specification

For

PCause: PCOS detection system based on deep learning model using ultrasound images

12th April 2024

Prepared by

Specialization	SAP ID	Name
AI&ML Hons B1	500090912	Charu Gupta
AI&ML Hons B1	500094127	Lakshay Agarwal

Department of Artificial Intelligence
School Of Computer Science
UNIVERSITY OF PETROLEUM & ENERGY STUDIES,
DEHRADUN- 248007. Uttarakhand

Table of Contents

1. INTRODUCTION:

Purpose of the project:

Polycystic Ovary Syndrome (PCOS) is a common hormonal disorder affecting women of reproductive age, often leading to complications such as infertility and metabolic issues. However, diagnosing PCOS accurately can be challenging due to subjective assessments and reliance on manual interpretation of ultrasound images.

The PCause project aims to address this challenge by developing an automated PCOS detection system using advanced technologies like deep learning and image processing. By analyzing ultrasound images, our system will provide healthcare professionals with a reliable tool for early and precise diagnosis.

In this project, we'll explore deep learning techniques to analyze ultrasound images and extract relevant features indicative of PCOS. By augmenting the dataset and optimizing model architectures, we aim to create an efficient and accurate diagnostic solution.

Our goal is to improve healthcare outcomes by streamlining the diagnostic process, enabling early intervention, and personalized treatment strategies for individuals with PCOS. Through collaboration with experts in both medical and technological fields, PCause strives to make a meaningful impact on women's health.

Target Beneficiary

The target beneficiaries of PCause include women of reproductive age who are at risk or suspected of having PCOS. Additionally, healthcare professionals, including gynecologists, endocrinologists, and primary care physicians, will benefit from the improved diagnostic accuracy and efficiency provided by the PCause system.

Project Scope

- Development of an Automated PCOS Detection System: Create a reliable system capable of accurately identifying PCOS using ultrasound images, reducing reliance on subjective assessments.

- Integration of Advanced Technologies: Utilize deep learning and image processing techniques to enhance the efficiency and reliability of PCOS detection.

- Data Augmentation and Model Training: Implement data augmentation strategies to expand the training dataset and improve the generalization of deep learning models.

- Feature Extraction from Image Dataset: Employ image processing techniques to extract relevant features from ultrasound images, enhancing the accuracy and adaptability of the detection system.

- Comparative Analysis of Deep Learning Models:Conduct a comprehensive evaluation of various deep learning architectures to identify the most effective model for PCOS detection.

References

[1]Aslam, S., & Rabie, T. F. (2023, February). Principal Component Analysis in Image Classification: A review. In 2023 Advances in Science and Engineering Technology International Conferences (ASET) (pp. 1-7). IEEE.

[2] Chitra, P., Srilatha, K., Sumathi, M., Jayasudha, F. V., Bernatin, T., & Jagadeesh, M. (2023, March). Classification of Ultrasound PCOS Image using Deep Learning based Hybrid Models. In 2023 Second International Conference on Electronics and Renewable Systems (ICEARS) (pp. 1389-1394). IEEE.

[3] Suha, S. A., & Islam, M. N. (2022). An extended machine learning technique for polycystic ovary syndrome detection using ovary ultrasound image. Scientific Reports, 12(1), 17123.

[4] Aggarwal, A., Mittal, M., & Battineni, G. (2021). Generative adversarial network: An overview of theory and applications. International Journal of Information Management Data Insights, 1(1), 100004.

[5] Sumathi, M., Chitra, P., Prabha, R. S., & Srilatha, K. (2021, February). Study and detection of PCOS related diseases using CNN. In IOP Conference Series: Materials Science and Engineering (Vol. 1070, No. 1, p. 012062). IOP Publishing.

2. PROJECT DESCRIPTION:

Reference Algorithm

- Image processing: We implemented Generative Adversarial Networks (GANs) to generate synthetic data, adding diversity to the dataset. This enriches our data pool and sets the stage for analysis and modeling. Following this, Principal Component Analysis (PCA) is employed to streamline the dataset, selecting essential features while reducing its complexity. This sequential process of using GANs for data augmentation followed by PCA for feature selection improves the efficiency of our image processing, ultimately leading to better PCOS detection accuracy.

- Deep learning: Deep learning methodologies like convolutional neural network (CNN) architectures, VGG-16, VGG-19, RESNET-50, and ALEXNET, are used for the detectionof PCOS from the dataset provided to these models.

Dataset

Data folder consist of 'train' and 'test' subfolders containing 2 categories of data 'infected' and 'not infected'

- infected: Images of ovaries having PCOS
- not infected: Images of healthy ovaries

The dataset is freely available on Kaggle for research purposes

SWOT Analysis

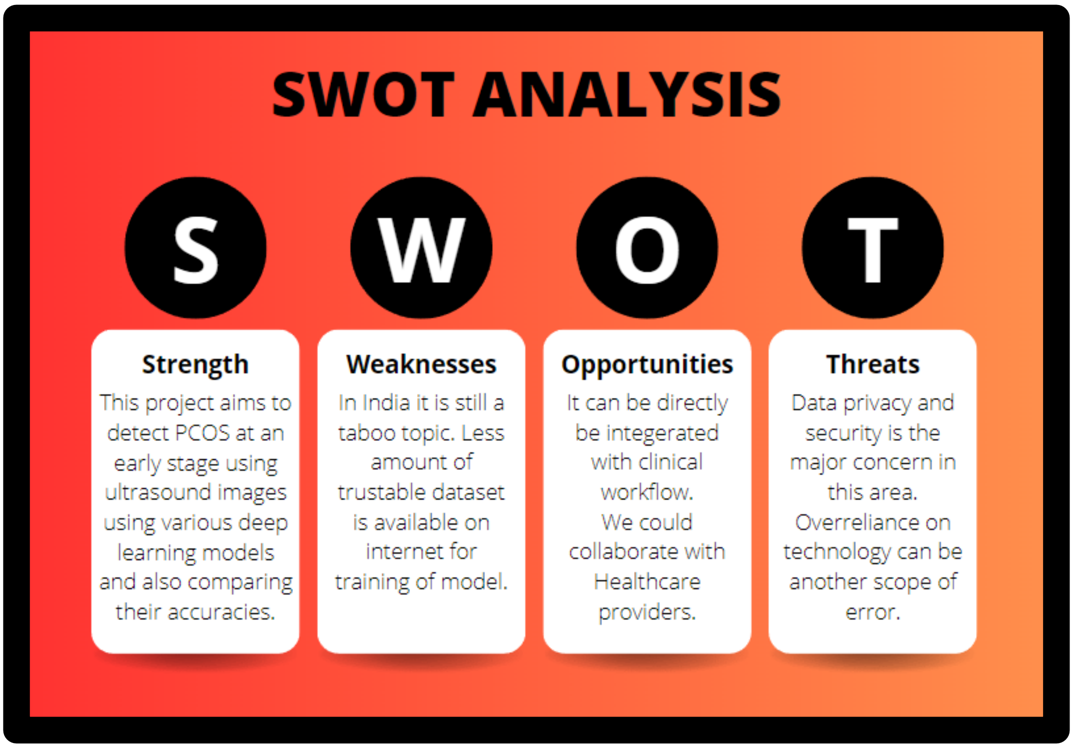

Fig 2.1. SWOT

Project Features

- Automated PCOS Detection System Development:
  Objective: Create a robust and automated system capable of accurately detecting PCOS using ultrasound images.
  Key Components: Algorithm and model development for ultrasound image analysis, minimizing reliance on subjective assessments.

- Integration of Advanced Technologies:
  Objective: Enhance efficiency and reliability of PCOS detection through advanced technologies.
  Key Components: Incorporation of deep learning and image processing techniques for precise diagnosis.

- Data Augmentation and Model Training:
  Objective: Improve generalization and performance of deep learning models.
  Key Components: Implementation of data augmentation strategies to expand training dataset, enhancing model's accuracy across diverse patient populations.

- Feature Extraction from Image Dataset:
  Objective: Enhance accuracy and adaptability of deep learning models.
  Key Components: Utilization of image processing techniques to identify and extract relevant features indicative of PCOS from ultrasound images.

- Comparative Analysis of Deep Learning Models:
  Objective: Identify the most effective model for PCOS detection.
  Key Components: Comprehensive evaluation of deep learning architectures (e.g., VGG-16, VGG-19, RESNET-50, ALEXNET) based on accuracy metrics and performance benchmarks.

User Classes and Characteristics

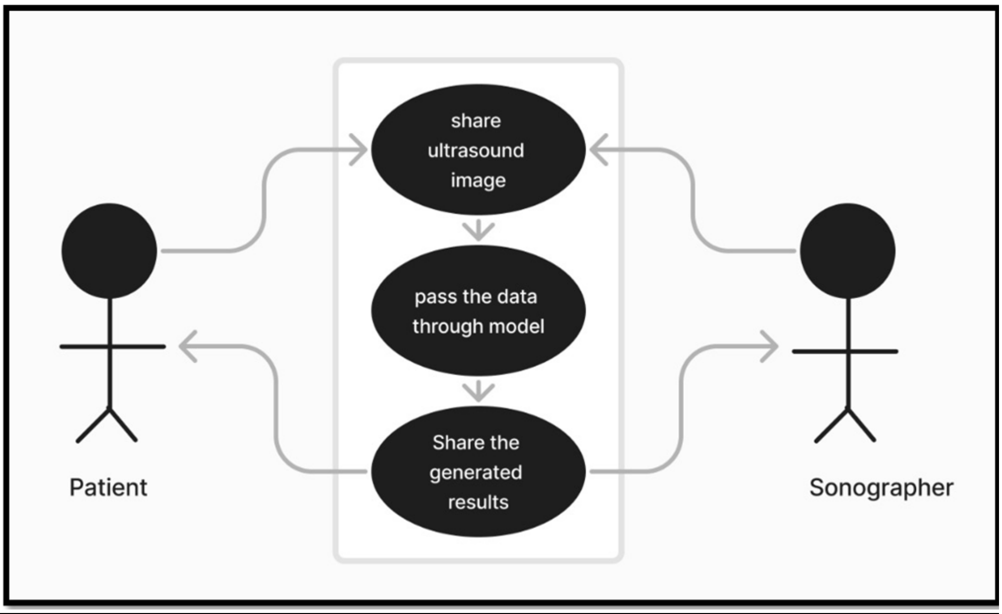

Fig 2.2. UML

It can be used for the following user classes:

- Patients
- Sonographers

Design and Implementation Constraints

a) Design Constraints:

- Hardware Limitations: The design of the PCOS detection system must accommodate the hardware limitations of the target deployment environment, such as memory constraints, processing power, and storage capacity.
- Software Compatibility: The system design should ensure compatibility with existing software frameworks, libraries, and operating systems commonly used in healthcare settings to facilitate seamless integration and interoperability.
- Scalability Considerations: The system design should take into account scalability requirements, allowing for future expansion to handle larger datasets, increased user load, and additional functionalities as needed.

b) Implementation Constraints:

- Technology Stack: The choice of programming languages, frameworks, and tools for implementing the PCOS detection system may be constrained by factors such as compatibility, performance, and availability of skilled developers.

- Algorithm Complexity: Implementation constraints may arise from the computational complexity of deep learning algorithms and image processing techniques used in the system, requiring optimization to meet performance and resource constraints.

- Data Availability and Quality: Constraints related to the availability, quality, and diversity of ultrasound image datasets may impact the effectiveness and generalization capabilities of the implemented models, necessitating careful data preprocessing and augmentation.

- Testing and Validation: Constraints related to testing methodologies, validation procedures, and access to clinical data for evaluation may influence the implementation process, requiring rigorous testing to ensure the accuracy, reliability, and safety of the PCOS detection system.

Design Diagram

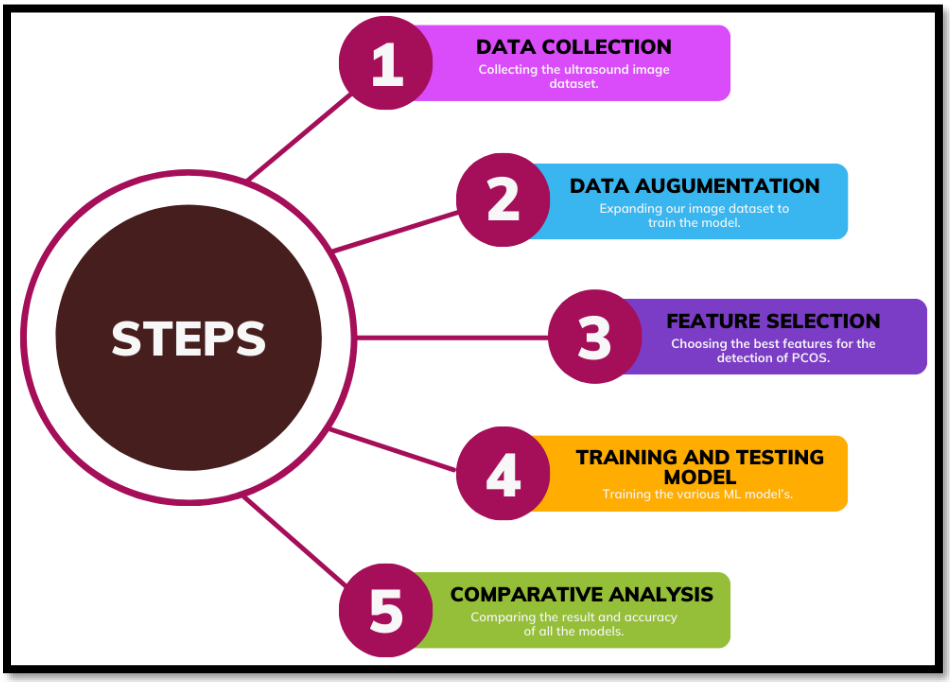

Fig 2.3. Workflow

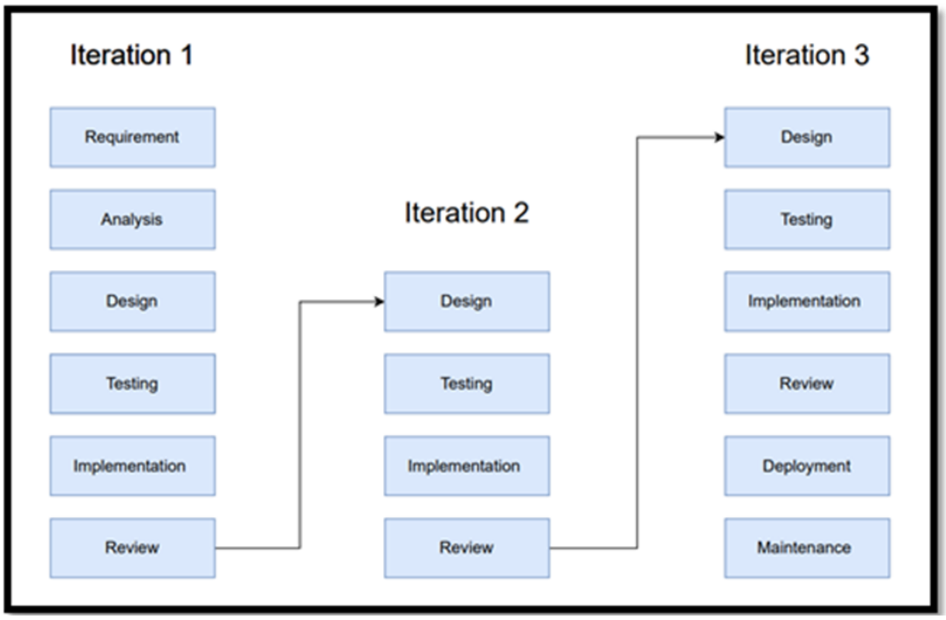

Fig 2.4. Iterative Model

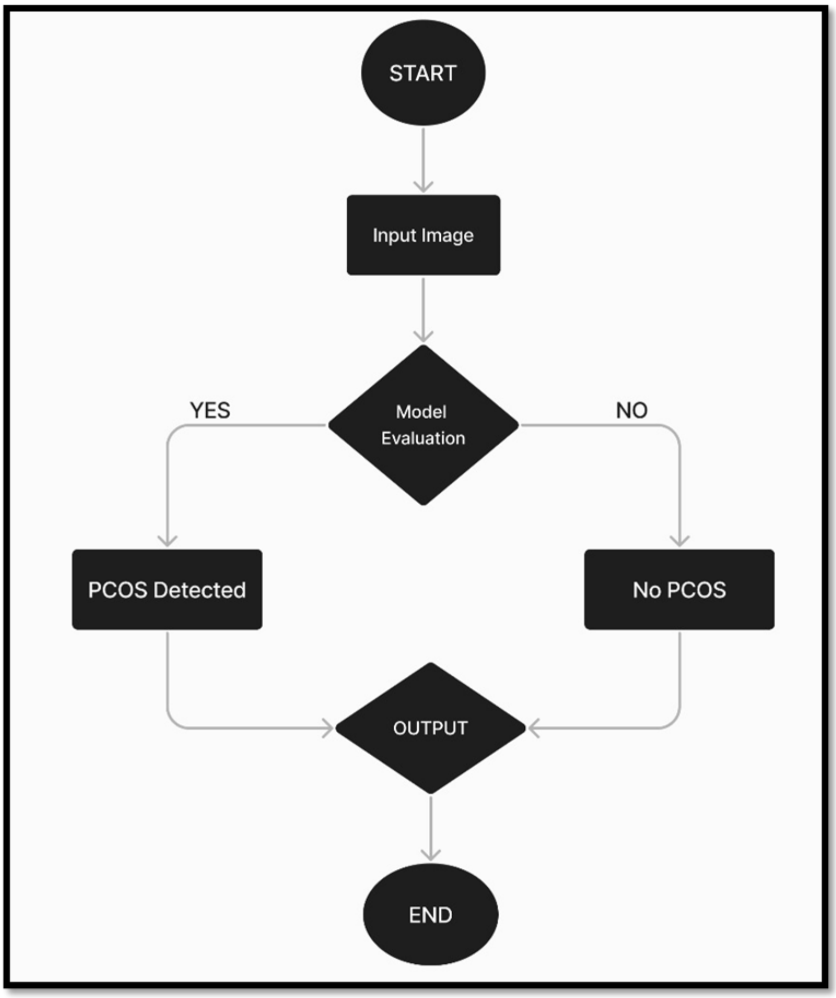

Fig 2.5. Flowchart

Assumption and Dependencies

Our Project has certain assumptions as mentioned below:

- Ultrasound Image Quality: It is assumed that the ultrasound images provided for training and testing the PCOS detection system are of sufficient quality and resolution to enable accurate analysis and diagnosis.

- Labeling Accuracy: It is assumed that the ultrasound images provided are accurately labeled as either "infected" (indicative of PCOS) or "notinfected" (indicative of healthy ovaries), ensuring the integrity of the training dataset.

- Availability of Computational Resources: It is assumed that the project will have access to adequate computational resources, including hardware (e.g., CPUs, GPUs) and software (e.g., deep learning frameworks), to support the training and testing of deep learning models.

3. SYSTEM REQUIREMENTS:

User Interface
We have used the command line interface to generate the output window for displaying our predictions.

Software Interface
Not applicable

Database Interface
Not applicable

Protocols
Not applicable

4. NON-FUNCTIONAL REQUIREMENTS:

Performance Requirements

The basic requirement to run this project is to have a proper installation of Python with various libraries, extensions and the availability of GPU in the local system. It enables high-speed parallel processing of large amounts of data, allowing to build deep learning model fast.

Security Requirements

The dataset is readily available on the internet. The files of our project have been uploaded on GitHub so it is accessible to everyone.

Software Quality Attributes

- Accuracy: Ensuring that the PCOS detection system accurately identifies and classifies ultrasound images as indicative of PCOS or not, with a high degree of precision and minimal false positives or false negatives.

- Reliability: The PCOS detection system should consistently provide reliable results, without errors or failures, even under varying conditions or user inputs.

- Performance: The system should efficiently process ultrasound images and provide timely results, optimizing resource utilization and minimizing processing time to enhance overall performance.

- Maintainability: The project should be developed with maintainability in mind, ensuring that the codebase is well-organized, documented, and modularized to facilitate future updates, enhancements, and bug fixes.

- Scalability: The system should be designed to scale efficiently, allowing for the addition of new features, handling of larger datasets, and accommodating increased user demand without significant performance degradation.

5. APPENDIX A: GLOSSARY

- PCOS - Polycystic Ovary Syndrome, a hormonal disorder common among women of reproductive age.

- Principal Component Analysis (PCA) - A statistical technique used to simplify data sets by reducing the number of variables while preserving the essential information.

- Generative Adversarial Network (GAN) - A class of artificial intelligence algorithms used in unsupervised machine learning, comprising two neural networks: a generator and a discriminator, which are trained simultaneously to generate realistic data.

- Data Augmentation - The process of artificially increasing the size of a dataset by applying transformations such as rotation, flipping, scaling, or cropping to the existing data samples.

- CNN - Convolutional Neural Network, a class of deep neural networks commonly applied to analyzing visual imagery. CNNs are designed to automatically and adaptively learn spatial hierarchies of features from the input data.

- VGG-16 - A specific convolutional neural network architecture developed by the Visual Geometry Group at the University of Oxford, consisting of 16 layers.

- VGG-19 - A variant of the VGG architecture with 19 layers.

- ResNet-50 - A deep residual neural network architecture with 50 layers, designed to mitigate the vanishing gradient problem encountered in training very deep neural networks.

- AlexNet - A deep convolutional neural network architecture, consisting of eight layers, developed by Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton. It gained prominence for winning the ImageNet Large Scale Visual Recognition Challenge in 2012.