International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 8, Issue 9, pp 66-69, September-2024

https://doi.org/10.47001/IRJIET/2024.809008

Enhancing PCOS Diagnosis with Improved Feature Selection using Extreme Gradient Boosting and CNN

[1]T. Mallika Devi, [2]Rajya Laxmi, [3]B.Vijaya Laxmi

[1,2]Assistant Professor, Department of CSE (Data Science), G. Narayanamma Institute of Technology and Science, Hyderabad, India

[3]Assistant Professor, Department of CSE (AI&ML), G. Narayanamma Institute of Technology and Science, Hyderabad, India

Abstract **- Polycystic Ovary Syndrome (PCOS) is a prevalent and intricate endocrine disorder affecting a substantial proportion of the female population. This condition is characterized by a constellation of symptoms, encompassing irregular menstrual cycles, physical manifestations like excess hair growth or acne, and hormonal imbalances, such as elevated androgen levels. The diagnosis of PCOS is often challenging due to the heterogeneity of its symptoms and the need for a multidimensional assessment. The proposed system seeks to revolutionize PCOS diagnosis by amalgamating two potent technologies: Extreme Gradient Boosting (XGBoost) and Convolutional Neural Networks (CNNs) with meticulous feature selection. XGBoost handles structured clinical data, capturing intricate relationships, while CNNs extract features from medical images, crucial for identifying ovarian cysts, a common PCOS indicator. This fusion offers a holistic assessment, empowering healthcare professionals to make more accurate diagnoses, thereby improving patient care. By bridging structured and unstructured data, our system aims to enhance PCOS understanding and streamline diagnostics, benefiting women globally.**

Keywords: Polycystic Ovary Syndrome (PCOS), Menstrual Irregularities, Endocrine Disorder, Extreme Gradient Boosting (XGBoost), Convolutional Neural Networks (CNNs), PCOS Diagnosis.

I. INTRODUCTION

PCOS commonly known as Polycystic Ovary Syndrome is a complex hormonal disorder affecting women during their reproductive age. The name PCOS describes the numerous small cysts that form in the ovaries. PCOS is the more severe form of PCOD, in which the ovaries stop releasing eggs and it affects fertility in women.

Ovulation occurs when a mature egg is released from the ovary, this egg can be fertilized with the help of a sperm. If not fertilized, it is sent out of the body during period. In some cases, women doesn't make enough hormones needed to ovulate. When ovulation doesn't happen the ovary develop many small cysts. These small cysts make hormones called Androgens. Women with PCOS have high level of androgens.

Due to PCOS woman cannot ovulate regularly, making them hard to get pregnant. If become pregnant, there may be chances of miscarriage, premature birth.

II. LITERATURE SURVEY

PCOS-WaveConvNet: A Wavelet Convolutional Neural Network for Polycystic Ovary Syndrome Detection using Ultrasound images:

The PCOS-WaveConvNet classifier is a deep learning model designed for the detection and classification of polycystic ovarian syndrome (PCOS) in women of reproductive age. PCOS is a hormonal condition characterized by multiple small follicles or cysts on the ovaries, which can be identified using ultrasound imaging. This classifier combines the power of wavelet transformation and convolutional neural networks (ConvNets) to achieve high accuracy in identifying PCOS in ultrasound images.

Polycystic Ovary Syndrome Detection Machine Learning Model Based On Explainable Artificial Intelligence:

Polycystic ovarian syndrome (PCOS) is a prevalent health concern affecting approximately one in ten women of childbearing age. This condition is characterized by hormonal imbalances and metabolic issues that can impact both the overall well-being and attractiveness of affected individuals. Additionally, PCOS can lead to infertility due to irregularities in the maturation and release of eggs during the menstrual cycle. Early detection of PCOS is crucial for prompt and effective treatment. However, traditional ultrasound image analysis performed by healthcare professionals can be time-consuming and labor-intensive. To address this challenge, advanced deep learning methods, particularly Convolutional Neural Network (CNN) models, have been increasingly employed to develop automated ultrasound image detection techniques.

ISSN (online): 2581-3048
Volume 8, Issue 9, pp 66-69, September-2024
https://doi.org/10.47001/IRJIET/2024.809008

The results of this study support the authors' claim that segmenting follicle structures is a crucial step in isolating non-PCOS images. Subsequently, the CNN acts as a confirmation tool, achieving an impressive accuracy rate of over 83% in classifying PCOS follicles.

Follicle Detection of Polycystic Ovarian Syndrome (Pcos) Using Yolo:

Polycystic Ovarian Syndrome (PCOS) is a common endocrine disorder affecting reproductive-aged individuals, characterized by various symptoms including irregular menstrual cycles, excessive androgen levels, and the presence of multiple small cysts on the ovaries. Early and accurate diagnosis of PCOS is crucial for effective management and prevention of associated health risks. In this paper, we propose a novel approach for the early detection of PCOS by leveraging the power of YOLO (You Only Look Once), a state-of-the- art object detection algorithm, to identify and analyze ovarian follicles.

III. DATASET DESCRIPTION

This project utilizes two distinct datasets: a parametric dataset and a collection of ultrasound images. The parametric dataset comprises structured numerical and categorical information, sourced from Kaggle, with 10,541 instances and 41 attributes. These attributes include parameters such as age, weight, BMI, hormonal markers, and clinical indicators. Notably, there is a class distribution imbalance, with 3,468 instances classified as positive (PCOS) and 7,073 instances as negative (non-PCOS). The ultrasound image dataset complements the parametric data, providing visual insights through medical imaging. The integration of these datasets allows for a comprehensive analysis, leveraging both quantitative and visual information for a more holistic understanding of PCOS.

IV. METHODOLOGY

a) Data Collection and Preprocessing:

This module collects data from diverse sources, including patient records and ultrasound images. It undergoes cleaning, normalization, and feature extraction, organizing clinical data and preparing image data for CNN analysis.

b) Feature Selection and Engineering with XGBoost:

XGBoost performs feature selection and engineering on the clinical dataset, enhancing efficiency and interpretability by identifying relevant features for PCOS diagnosis.

c) CNN-based Image Analysis:

CNN analyzes ultrasound images; extracting features crucial for PCOS diagnosis, such as patterns and anomalies, for a comprehensive evaluation.it will check deep with layers for feature extraction was efficient and effective.

d) Model Development with XGBoost:

An XGBoost model is developed using selected features and engineered data, learning patterns and relationships between features and PCOS presence.

e) Integration of Clinical and Image Analysis:

Results from XGBoost clinical analysis and CNN image analysis are combined, leveraging the strengths of both methods for a comprehensive PCOS assessment.

f) Diagnostic Framework Implementation:

This framework integrates clinical data and ultrasound images, processing them through XGBoost and CNN, respectively, to generate a final diagnosis.

g) Model Training and Testing:

The diagnostic framework is rigorously tested and validated using separate datasets, evaluating effectiveness through performance metrics like accuracy, precision, recall, and F1-score with any machine learning-based system, the continuous improvement of the model through larger and more diverse datasets, advanced model architectures, and fine-tuning hyperparameters will further enhance its performance and expand its applicability in various real-world scenarios

The integration of Extreme Gradient Boosting (XGBoost) with Convolutional Neural Networks (CNNs) in our proposed system for diagnosing Polycystic Ovary Syndrome (PCOS) has yielded promising results. By leveraging structured clinical data via XGBoost and extracting features, such as precise identification of ovarian cysts, from medical images using CNNs, we achieved a comprehensive approach to PCOS diagnosis. Ultimately, our system's ability to bridge structured and unstructured data has significantly improved PCOS understanding and streamlined diagnostics.[9].

Key Findings:

1. XGBoost and CNN play complementary roles in PCOS diagnosis, with XGBoost enhancing efficiency and interpretability of clinical data analysis, while CNN excels in recognizing patterns and anomalies in ultrasound images.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 8, Issue 9, pp 66-69, September-2024

https://doi.org/10.47001/IRJIET/2024.809008

2. Feature selection with XGBoost identifies the most relevant features for PCOS diagnosis, improving model performance and interpretability.
3. CNN-based image analysis extracts valuable features from ultrasound images, contributing to a more comprehensive evaluation of PCOS.

V. RESULTS

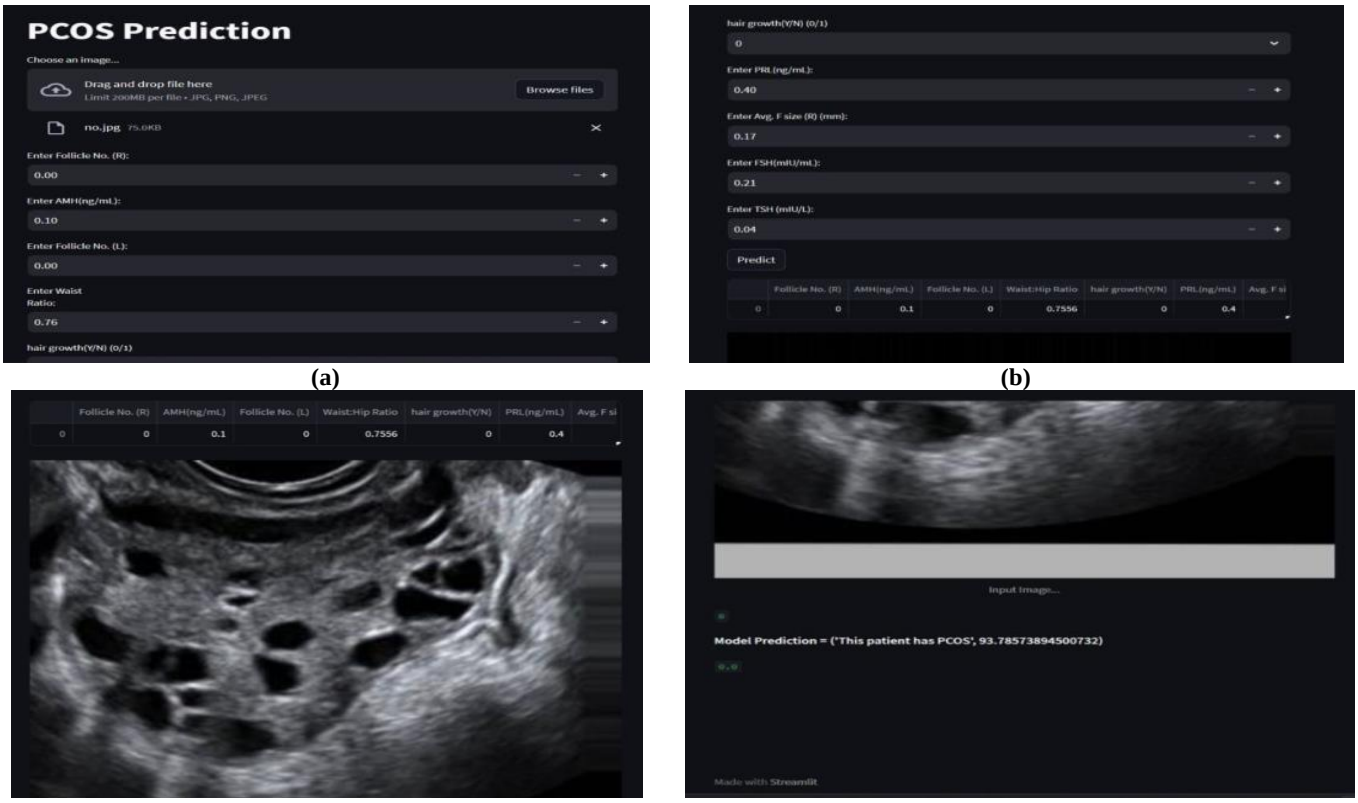

Figure 2: User interface

VI. CONCLUSION

The project aims to enhance the diagnostic process of Polycystic Ovary Syndrome (PCOS) by leveraging advanced machine learning techniques. By integrating data from patient records and ultrasound images, a comprehensive diagnostic framework that combines the strengths of XGBoost for clinical data analysis and Convolutional Neural Networks (CNN) for image analysis was developed. Through feature selection with XGBoost and CNN-based image analysis, our framework achieves high accuracy and efficiency in PCOS diagnosis. The integration of these methods provides a holistic assessment, improving early pcos.

REFERENCES

[1] Shazia Nasim, Mubarak Saad Almutairi, "A Novel Approach for Polycystic Ovary Syndrome Prediction Using Machine Learning in Bioinformatics". 2022.

[2] Samia Ahmed, Md. Sazzadur Rahman, "A Review on the Detection Techniques of Polycystic Ovary Syndrome Using Machine Learning". 2023.

[3] J. R. Calva, "PCOS: A Process Control Extension to Operating System/360". 2017.

[4] V Srinithi, R. Rekha, "Machine learning for diagnosis of polycystic ovarian syndrome (PCOS/PCOD)". 2023.

[5] Narinder Kaur, Ganesh Gupta, "Transfer-Based Deep Learning Technique for PCOS Detection Using Ultrasound Images". 2023.

[6] Palvi Soni, Sheveta Vashisht, "Exploration on Polycystic Ovarian Syndrome and Data Mining Techniques". 2018.

[7] P. Arunkumar, K Abarna, N Nagamithra, "Application using Machine Learning to Promote Women's Personal Health". 2023.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 8, Issue 9, pp 66-69, September-2024

https://doi.org/10.47001/IRJIET/2024.809008

Citation of this Article:

T. Mallika Devi, Rajya Laxmi, & B.Vijaya Laxmi, (2024). Enhancing PCOS Diagnosis with Improved Feature Selection using Extreme Gradient Boosting and CNN. *International Research Journal of Innovations in Engineering and Technology - IRJIET*, 8(9), 66-69. Article DOI https://doi.org/10.47001/IRJIET/2024.809008

****

