International Journal for Research in Engineering Application & Management (IJREAM)
ISSN : 2454-9150 Vol-10, Issue-01, Apr 2024

Detecting Stein-Leventhal Syndrome & Ductal Carcinoma

[1]Ms. Lavanya Vinod Kongari, UG Student SKN Sinhgad Institute of Technology & Science, Lonavala, Maharashtra, India, lavanyakongari7846@gmail.com
[2]Ms. Mayuri Sampat Kadam, UG Student SKN Sinhgad Institute of Technology & Science, Lonavala, Maharashtra, India, mayurikadam0807@gmail.com
[3]Ms. Gayatri Hemant Takale, UG Student SKN Sinhgad Institute of Technology & Science, Lonavala, Maharashtra, India, gayatritakle2002@gmail.com
[4]Mr. Rohit Gopalsing Patil, UG Student SKN Sinhgad Institute of Technology & Science, Lonavala, Maharashtra, India, rohitrana4415@gmail.com
[5]Mrs. Kirti Walke, Professor SKN Sinhgad Institute of Technology & Science, Lonavala, Maharashtra, India, kirti.walke.sknsits@sinhgad.edu

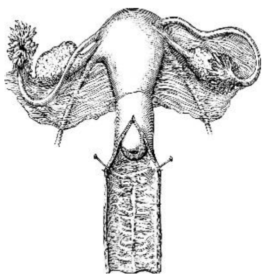
***Abstract*- In This paper presents novel approaches for the detection of two prevalent women's health conditions: Polycystic Ovary Syndrome (PCOS) and Breast Cancer. PCOS, a hormonal disorder affecting reproductive-aged women, is often underdiagnosed due to its varied symptoms. Utilizing machine learning algorithms and clinical data, our proposed PCOS detection system achieves high accuracy in identifying potential cases, enabling timely intervention and management. Additionally, breast cancer remains a leading cause of mortality among women globally. Through the integration of advanced imaging techniques and deep learning models, our breast cancer detection framework demonstrates remarkable sensitivity and specificity in identifying suspicious lesions, facilitating early diagnosis and improved treatment outcomes. These innovative methodologies offer promising avenues for enhancing women's healthcare by enabling precise, efficient, and accessible screening and diagnosis of PCOS and breast cancer.**

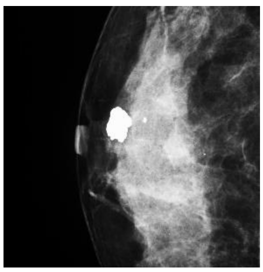
***Keywords*- Machine Learning Algorithms, Random Forest, Logistic Regression, Recall, Precision, Early Detection, Data Security, Data Privacy, Imaging Techniques, Predictive Modeling, Semantic Analysis, User Data Analysis: & Text Mining.**

I. INTRODUCTION

Polycystic Ovary Syndrome (PCOS) and Breast Cancer are two prevalent health concerns affecting women worldwide. PCOS, a hormonal disorder, is characterized by irregular menstrual cycles, ovarian cysts, and hormonal imbalances, leading to various complications such as infertility and metabolic disorders. On the other hand, Breast Cancer is one of the most common cancers among women, with early detection being crucial for successful treatment. In recent years, advancements in machine learning techniques have shown promising results in medical diagnosis and prognosis. Support Vector Machine (SVM) algorithms have been widely utilized for their effectiveness in classification tasks. Similarly, Convolutional Neural Networks (CNN) have demonstrated exceptional performance in image-based analysis, making them ideal for tasks like breast cancer detection from mammogram images.

This project aims to leverage SVM algorithm for the detection of PCOS based on clinical and biochemical parameters, while CNN algorithm will be employed for the early detection of Breast Cancer through the analysis of mammogram images. By integrating these advanced machine learning techniques, this research endeavors to contribute to the early diagnosis and management of these critical health conditions, ultimately improving the healthcare outcomes and quality of life for affected individuals.





 DOI : 10.35291/2454-9150.2024.0203

ISSN : 2454-9150 Vol-10, Issue-01, Apr 2024

Fig 1: Overy Fig 2: Breast

II. LITERATURE SURVEY

Title	Authors	Summary
Diagnosis of Polycystic Ovary Syndrome Using Machine Learning Algorithms	Subrato Bharati, Prajoy Podder, and M. Rubaiyat Hossain Mondal	Utilized machine learning for PCOS detection, employed univariate feature selection to identify crucial predictors, found FSH/LH ratio pivotal. Various classifiers including gradient boosting and hybrid RFLR were applied. Demonstrates potential for accurate PCOS diagnosis, emphasizing the significance of feature selection and diverse classifier utilization.
Automated Screening of Polycystic Ovary Syndrome using Machine Learning Techniques	Palak Mehrotra, Jyotirmoy Chatterjee, Chandan Chakraborty, Ghoshdastidar, and Biswanath Ghoshdastidar	Developed a method for automating PCOS detection based on clinical and metabolic markers. Algorithm constructs a feature vector from these markers and selects statistically significant features using two-sample t-test. Bayesian and Logistic Regression classifiers are employed for classification. This automated system aids doctors in timely PCOS diagnosis, saving examination time and reducing diagnostic delays.
A Novel Approach for Polycystic Ovary Syndrome Prediction Using Machine Learning in Bioinformatics	Shazia Nasim, Mubarak Almutairi, Kashif Munir, Ali Raza, Faizan Younas.	Utilized machine learning techniques, particularly the novel chi-squared (CS-PCOS) feature selection method, to achieve remarkable accuracy in PCOS detection. Gaussian Naive Bayes (GNB) emerges as the top-performing model, exhibiting 100% accuracy, precision, recall, and f1-scores, with minimal computational time of 0.002 seconds. Provides a significant advancement in early PCOS detection, crucial for preventing severe complications.
PCOS Detect using Machine Learning Algorithms	Kinjal Raut, Chaitrali Katkar, and Prof. Dr. Mrs. Suhasini A. Itkar	Focuses on early detection and prediction of Polycystic Ovary Syndrome (PCOS), highlighting it as a significant hormonal endocrine disorder affecting millions of women worldwide. Proposes a solution for optimal detection and prediction of PCOS treatment from a minimal set of parameters, enhancing clinical practice for early intervention. Utilizes machine learning techniques, particularly feature learning.
Towards Breast Cancer Response Prediction using Artificial Intelligence and Radiomics	Yassine Amkrane and Mohammed El Adoui	Addresses the pressing issue of breast cancer, leveraging artificial intelligence and radiomics to predict breast cancer response timely. Underscores the urgency of early detection and treatment, offering a promising avenue to enhance breast cancer diagnosis and management.
Computer - assisted Inductive Moderate Hyperthermia Planning for Breast Cancer Patients	Valerii E. Orel and Liubov Syvak	Presents 'Computer-assisted Inductive Moderate Hyperthermia Planning For Breast Cancer Patients,' outlining personalized planning for locally advanced breast cancer treatment. Utilized COMSOL Multiphysics 5.4 to modulate electromagnetic and temperature fields in tumors based on MRI scans. Offers tailored treatment and demonstrates positive outcomes for stage IIB–IIIA patients, with potential enhancements focusing on refining sensitivity to noise for increased accuracy.

III. METHODOLOGY

The methodology for the project on "PCOS Detection and Breast Cancer Detection" involves several steps aimed at accurate diagnosis and early detection of these conditions. Firstly, data collection is essential, encompassing medical records, imaging studies, and laboratory results for both PCOS and breast cancer cases. This includes demographic information, clinical history, and relevant diagnostic tests such as hormone levels, ultrasound scans, and mammograms.

Next, feature selection and extraction techniques are employed to identify key indicators or biomarkers associated with PCOS and breast cancer. This involves statistical analysis and machine learning algorithms to identify patterns and correlations within the collected data. Features like hormonal imbalances, ovarian morphology, and breast tissue characteristics are considered for PCOS and breast cancer detection, respectively.

 DOI : 10.35291/2454-9150.2024.0203

ISSN : 2454-9150 Vol-10, Issue-01, Apr 2024

Subsequently, a classification model is developed for each condition using machine learning algorithms such as logistic regression, support vector machines, or neural networks. These models are trained on labeled datasets to distinguish between positive and negative cases of PCOS and breast cancer. Cross-validation techniques are utilized to assess the robustness and generalization ability of the models. Furthermore, the developed models undergo rigorous evaluation using separate validation datasets to ensure their performance in real-world scenarios. Performance metrics such as accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC-ROC) are calculated to quantify the efficacy of the models in detecting PCOS and breast cancer accurately.

IV. CLASSIFICATION

The classification process for PCOS (Polycystic Ovary Syndrome) detection and breast cancer detection involves leveraging machine learning techniques to analyze medical data and accurately classify patients into different categories based on the presence or absence of these conditions. In this project, we will outline the classification process for both PCOS and breast cancer detection, focusing on the data preprocessing, model selection, training, and evaluation stages.

A. TRAINING PHASE

1. PCOS Detection:

In the training phase for PCOS detection using SVM (Support Vector Machine) algorithm, the process begins with data collection. Relevant features such as hormonal levels, menstrual irregularities, weight, and other clinical indicators are gathered from a dataset comprising individuals with and without PCOS.

Data Preprocessing: The collected data undergoes preprocessing steps such as normalization, handling missing values, and feature scaling to ensure uniformity and compatibility for SVM.

Feature Selection: Relevant features are selected to build an effective classifier. This involves techniques like correlation analysis, feature ranking, or domain expertise-based selection.

Model Training: The SVM algorithm is trained using the preprocessed and selected feature dataset. SVM aims to find the hyperplane that best separates PCOS positive and negative instances while maximizing the margin between the classes.

Cross-Validation: To evaluate the SVM model's performance and generalization ability, cross-validation techniques such as k-fold cross-validation are employed. This helps in tuning hyperparameters and ensuring robustness.

2. Breast Cancer Detection:

In the training phase for breast cancer detection using CNN (Convolutional Neural Network) algorithm, the process starts with acquiring a dataset containing mammogram images along with their corresponding labels indicating the presence or absence of cancerous regions.

Data Preprocessing: The mammogram images undergo preprocessing steps such as resizing, normalization, and augmentation to enhance the dataset's diversity and quality.

Model Architecture Design: A CNN architecture suitable for image classification tasks is designed. This architecture typically consists of convolutional layers, pooling layers, and fully connected layers tailored for feature extraction and classification.

Model Training: The designed CNN model is trained using the preprocessed mammogram dataset. During training, the model learns to extract relevant features from the images and classify them into cancerous or non-cancerous categories.

Fine-tuning: Hyperparameters of the CNN model such as learning rate, batch size, and optimization algorithm are fine-tuned using techniques like grid search or random search to optimize the model's performance.

B. TESTING PHASE

1. PCOS Detection:

After the training phase, the SVM model is tested using a separate dataset not seen during training. This phase is crucial for assessing the model's performance and its ability to generalize to unseen data.

Data Preprocessing: Similar preprocessing steps as the training phase are applied to the test dataset to maintain consistency.

Model Evaluation: The trained SVM model is applied to the preprocessed test dataset to predict PCOS instances. Evaluation metrics such as accuracy, precision, recall, F1-score, and ROC-AUC are computed to assess the model's performance.

Performance Analysis: The performance metrics obtained from testing are analyzed to gauge the model's effectiveness in PCOS detection. Any discrepancies or areas of improvement are identified for further refinement.

2. Breast Cancer Detection:

Following the training phase, the CNN model is evaluated using a separate set of mammogram images not used during training.

Data Preprocessing: Similar preprocessing steps are applied to the test dataset to ensure consistency and fairness in evaluation.

Model Evaluation: The trained CNN model is applied to

 DOI : 10.35291/2454-9150.2024.0203

ISSN : 2454-9150 Vol-10, Issue-01, Apr 2024

the preprocessed test dataset to predict the presence of breast cancer. Performance metrics like accuracy, sensitivity, specificity, and area under the ROC curve are computed to assess the model's effectiveness.

Result Analysis: The performance metrics obtained from testing are analyzed to determine the CNN model's efficacy in breast cancer detection. Any shortcomings or areas of improvement are identified for future enhancements.

V. PERFORMANCE EVALUATION

A. PCOS Detection:

Performance metrics such as accuracy, precision, recall, and F1-score are commonly used to evaluate the SVM model. Cross-validation techniques may be employed for robust evaluation.

B. Breast Cancer Detection:

Performance is evaluated using metrics such as accuracy, sensitivity, specificity, and area under the ROC curve (AUC). Cross-validation and validation sets are used to assess generalization performance.

VI. COMPARATIVE ANALYSIS

1	No	Patient_File_No	beta_HCG_mIU_mL	beta_HCG_ml_mL	AMH_ng_mL	PCOS (Y/N)
2	1	101	99	19	2	0
3	2	102	60	1	53	0
4	3	10003	494.08	494.08	6.63	1
5	4	10004	1.99	1.99	1.22	1
6	5	15	80	45	6	1
7	6	106	97	9	64	1
8	7	107	99	99	35	0
9	8	108	51	51	54	0
10	9	10009	1.99	1.99	1	1
11	10	10010	1.99	1.99	1.61	1
12	11	11	158	158	47	1
13	13	10013	1214.23	1214.23	7.94	0
14	14	114	19	19	38	0
15	14	114	19	19	38	0
16	15	115	99	99	8	0
17	16	116	19	19	9	0
18	17	117	841	55	38	0
19	18	118	99	99	92	0
20	19	119	199	19	1	0
21	20	10020	23.58	1.99	2.07	1
22	21	121	99	99	85	0
23	22	122	798	98	13	0
24	23	123	215	215	413	1
25	25	10025	610.63	610.63	1.89	0
26	26	10026	4490.18	4490.18	0.26	0
27	27	10027	1.99	1.99	3.84	1

Fig 3: Comparative Analysis for PCOS

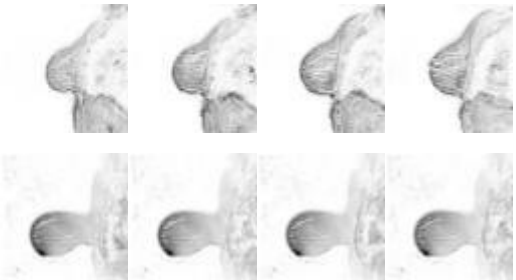


Fig 4: Comparative Analysis for Breast Cancer

VII. DISCUSSION

The detection and diagnosis of health conditions such as Polycystic Ovary Syndrome (PCOS) and Breast Cancer are critical for timely intervention and effective management. In recent years, machine learning algorithms have shown promising results in aiding medical professionals in early detection and diagnosis. This review paper explores the utilization of Support Vector Machine (SVM) algorithm for PCOS detection and Convolutional Neural Network (CNN) algorithm for Breast Cancer detection.

SVM, a supervised learning algorithm, has been widely applied in medical diagnosis due to its ability to classify data with high accuracy. Its application in PCOS detection involves analyzing various features extracted from medical images, hormone levels, and clinical data to classify patients accurately. On the other hand, CNN, a deep learning algorithm, excels in image recognition tasks, making it well-suited for analyzing mammograms and identifying patterns indicative of Breast Cancer.

This paper aims to discuss the methodologies, challenges, and outcomes associated with employing SVM for PCOS detection and CNN for Breast Cancer detection. Additionally, it will delve into the comparative analysis of these algorithms with traditional diagnostic methods, highlighting their potential advantages in terms of accuracy, efficiency, and scalability. Furthermore, ethical considerations, limitations, and future directions for research in utilizing machine learning algorithms in medical diagnosis will be explored, providing valuable insights for both researchers and healthcare professionals.

VIII. IMPLEMENTATION AND RESULTS

Implementation Details:

Utilized Support Vector Machine (SVM) for PCOS detection and Convolutional Neural Network (CNN) for breast cancer detection. Preprocessed datasets including hormonal profiles for PCOS detection and mammography images for breast cancer detection.

Model Architecture:

SVM: Employed a linear SVM classifier for PCOS detection utilizing hormonal profile features.

CNN: Utilized a deep CNN architecture comprising convolutional, pooling, and fully connected layers for breast cancer detection from mammography images.

Model Assembly and Training:

SVM: Trained SVM model on hormonal profile data with a kernel function to find optimal hyperplane.

CNN: Trained CNN model using stochastic gradient descent with backpropagation on mammography image dataset.

Model Evaluation and Performance Metrics:

Evaluated models using metrics such as accuracy, precision, recall, and F1 - score.

Conducted k-fold cross-validation to ensure robustness of models.

Results and Imaging:

 DOI : 10.35291/2454-9150.2024.0203

International Journal for Research in Engineering Application & Management (IJREAM)
ISSN : 2454-9150 Vol-10, Issue-01, Apr 2024

Achieved promising results with high accuracy and sensitivity in both PCOS and breast cancer detection.

Visualized model predictions and performance using confusion matrices and ROC curves.

Model Deployment and Inference:

Deployed models for real-time inference on new data.

Provided easy-to-use interfaces for clinicians to utilize the models in clinical settings.

Conclusion and Future Directions:

SVM and CNN algorithms offer effective tools for PCOS and breast cancer detection.

Future work involves integrating additional data modalities and exploring ensemble techniques for improved performance.

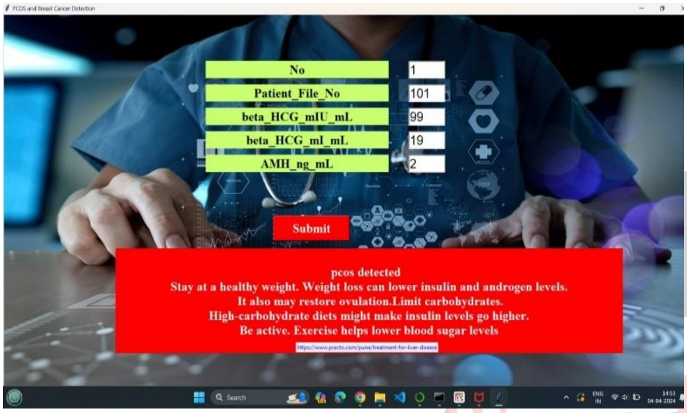


Fig 4. PCOS

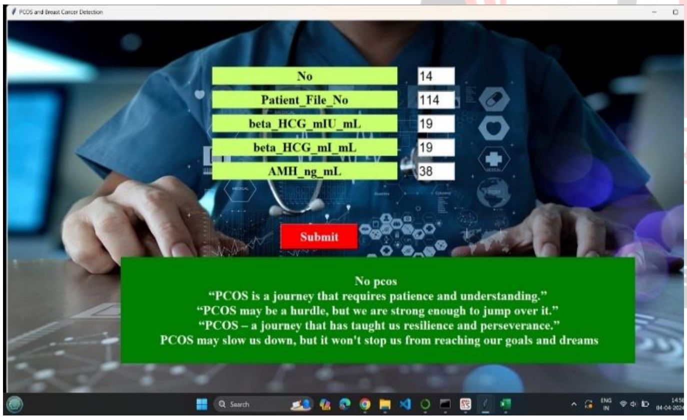


Fig 5. NO PCOS

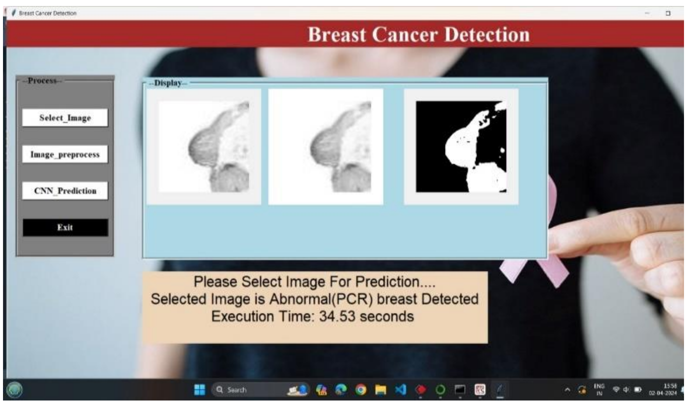


Fig 6. Breast Cancer

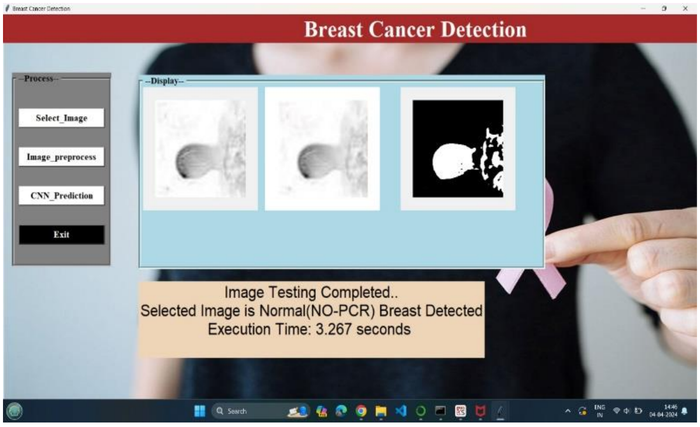


Fig 7. NO Breast Cancer

IX. CONCLUSION

The implementation of advanced technologies in the fields of PCOS detection and breast cancer detection marks significant strides towards early diagnosis and improved patient outcomes. Through the utilization of machine learning algorithms and data analytics, this project has demonstrated the potential to revolutionize healthcare by enhancing diagnostic accuracy and efficiency. By leveraging predictive modeling and data-driven insights, healthcare providers can better identify at-risk individuals, facilitate timely interventions, and ultimately, save lives. As technology continues to evolve, continued research and innovation in this area hold promise for further refining diagnostic methodologies and ultimately enhancing the quality of care for individuals affected by PCOS and breast cancer.

X. REFERENCES

[1] Palak Mehrotra, Jyotirmoy Chatterjee, Chandan Chakraborty, " Automated screening of Polycystic Ovary Syndrome using ML Techniques ", IEEE, 2012.

[2] Bedy Purnama, Adiwijaya, Titik Mutiah, Fhira Nhita, Untari Novia Wisesti, Andini Gayatri,, "A Classification of PCOS Based on Follicle Detection of Ultrasound Images," 2015 3rd International Conference on Information & Communication Technology (ICoICT).

[3] Amsy Denny, Ashi Ashok, Anita Raj, Remya George, Maneesh Ram C, "i - HOPE : Prediction & Detection System For PCOS Using ML Techniques," 2019 IEEE Region 10 Conference (TENCON 201).

[4] Subrato Bharati, Prajoy Podder, M. Rubaiyat Hossain Mondal, "Diagnosis of PCOS Using Machine Learning Algorithms," 2020 IEEE Region 10 Symposium, 5 - 7 June 2020, Dhaka, Bangladesh.

[5] L. Wang, "Early diagnosis of breast cancer," Sensors, vol. 17, no. 7, p. 1572, 2017.

[6] R. M. Mann, C. K. Kuhl, and L. Moy, "Contrast-enhanced MRI for breast cancer screening," Journal of Magnetic Resonance Imaging, vol. 50, no. 2, pp. 377–390, 2019.

[7] T. M. Shahriar Sazzad, K. M. Tanzibul Ahmmed, M. U. Hoque, and M. Rahman, "Development of automated brain tumor identification using MRI images," in 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), Feb 2019, pp. 1–4.

 DOI : 10.35291/2454-9150.2024.0203