International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Exploring Machine Learning and Deep Learning Approaches for PCOS

Detection: A Review

[1]B. Ankitha, [2]Geetika. K, [3]Ch. Srikanth, [4]A. Devi Priyanka, [5]D. Venkatesh

[1]Computer Science & Engineering GMR Institute of Technology Andhra Pradesh,, India 20341A0522@gmrit.edu.in
[2] Computer Science & Engineering GMR Institute of Technology Andhra Pradesh, India, 20341A0563@gmrit.edu.in
[3]Computer Science & Engineering GMR Institute of Technology Andhra Pradesh, India, 20341A0537@gmrit.edu.in
[4]Computer Science & Engineering GMR Institute of Technology Andhra Pradesh, India, 20341A0512@gmrit.edu.in
[5]Computer Science & Engineering GMR Institute of Technology Andhra Pradesh, India, 20341A0555@gmrit.edu.in

ABSTRACT :

Polycystic Ovary Syndrome (PCOS) is a common endocrine disorder among women of reproductive age, characterized by hormonal imbalance, irregular menstruation, ovarian cysts, and symptoms like acne and hirsutism. The exact cause of PCOS remains elusive, but it is believed to involve a combination of genetic predisposition, insulin resistance, and environmental factors. Traditional diagnostic methods rely on clinical evaluation, hormone testing, and ultrasound imaging of the ovaries. However, these approaches are subjective, time-consuming, and may lack sensitivity and specificity, leading to diagnostic inaccuracies and delays in treatment initiation. The advent of Computer-Aided Diagnosis (CAD) technologies offers promising solutions to address these limitations. Machine Learning (ML) and Deep Learning (DL) techniques have emerged as valuable tools for PCOS classification, leveraging algorithms to analyze complex clinical and imaging data.ML techniques, such as Support Vector Machines (SVM), AdaBoost, and k-Nearest Neighbors (kNN) etc, can learn patterns and relationships from features extracted from clinical and ultrasound data, facilitating accurate diagnosis and risk stratification of PCOS patients. DL techniques, including Convolutional Neural Networks (CNNs) and architectures to name a few VGG, ResNet, DenseNet , and Inception, etc excel in automatically learning discriminative features from raw imaging data, enhancing diagnostic accuracy and predictive performance. This paper systematically evaluates the advantages and limitations of ML and DL techniques in PCOS classification. Advantages include high accuracy, feature learning capabilities, scalability, and personalized medicine. However, challenges such as data availability, interpretability, generalization, and computational resources must be addressed to ensure the clinical relevance and reliability of CAD systems for PCOS diagnosis and management. This paper aims to guide future research towards developing robust and clinically relevant CAD systems for PCOS classification, ultimately improving patient outcomes and quality of care.

Keywords—Polycystic Ovary Syndrome (PCOS) , Machine Learning, Deep Learning, CNN.

Introduction :

Polycystic Ovary Syndrome (PCOS) is becoming prevalent concern for the women and leading infertility in most of the cases. Despite its prevalence, PCOS remains a complex condition to diagnose due to its heterogeneous presentation and multifactorial etiology. . A recent study has revealed that about 18% of the women in India, mostly from the East, suffer from this syndrome. Recognizing the onset of PCOS and accurately diagnosing it are pivotal steps in managing the condition effectively and mitigating associated health risks. The recognition of PCOS typically begins with the identification of its hallmark symptoms, which may manifest differently among individuals. Common indicators include irregular menstrual cycles, characterized by prolonged intervals between periods or the absence of menstruation altogether. Women with PCOS may also experience symptoms such as excessive hair growth (hirsutism), acne, and alopecia, often attributed to elevated levels of androgens—male hormones—commonly observed in PCOS patients. To confirm the diagnosis, clinicians often employ diagnostic criteria such as the Rotterdam criteria, which assess multiple aspects including menstrual irregularity, hyperandrogenism, and ovarian morphology. While these criteria provide a comprehensive framework for diagnosis, they may lack specificity and fail to capture the full spectrum of PCOS phenotypes. Additionally, the reliance on subjective assessments of symptoms and the variability in diagnostic criteria between healthcare providers can contribute to diagnostic inconsistency and delay in treatment initiation. Laboratory tests measuring hormone levels and ultrasound imaging of the ovaries are essential components of PCOS diagnosis. However, these methods also have limitations. Blood tests for hormone levels may exhibit variability due to factors such as diurnal fluctuations and menstrual cycle phase, leading to false-positive or false-negative results. Ultrasound imaging, while valuable for visualizing ovarian morphology, may not always detect polycystic ovaries in individuals with PCOS, particularly in lean individuals or those with oligo-ovulation. Moreover, the diagnostic process for PCOS can be time-consuming and costly, requiring multiple clinic visits, laboratory tests, and imaging studies. This may pose a barrier to access for some individuals, especially those with limited healthcare resources or insurance coverage. In summary, while various diagnostic methods are available for PCOS, each approach has its own set of advantages and disadvantages. Recognizing the limitations of current diagnostic methods is crucial for improving diagnostic accuracy, reducing diagnostic delays, and ensuring timely intervention for individuals affected by PCOS. Our review paper focuses on the systematic assessment of Machine Learning (ML) and Deep Learning (DL) techniques for the classification of Polycystic Ovary Syndrome (PCOS). We aim to comprehensively evaluate these methodologies to understand their strengths and limitations, with the ultimate goal of improving diagnostic accuracy and patient outcomes in PCOS management. ML techniques such as Support Vector Machines (SVM), AdaBoost, and k-Nearest Neighbors (KNN), as well as DL methods including Convolutional Neural Networks (CNNs) and prominent architectures like VGG, ResNet, DenseNet, and Inception, we extract meaningful patterns and relationships from the data. Our analysis rigorously evaluates the performance metrics of these models, including accuracy, sensitivity, specificity, and generalization capabilities. We aim to highlight the advantages of these techniques, such as their high accuracy and feature learning capabilities, while also acknowledging their limitations, such as challenges related to data availability, interpretability, and computational resources.

Literature survey :

Alamoudi, A., Khan et al. in[1] the objective of this paper is to develop and evaluate a deep learning fusion approach for diagnosing Polycystic Ovary Syndrome(PCOS) using ultrasound images and clinical data. Mainly focus on developing the deep learning models that extract features from ultrasound images and clinical data to improve the accuracy of PCOS diagnosis. The study emphasizes the importance of clinical features in PCOS diagnosis and suggests that automated models can assist physicians in saving time and reducing the risk of delayed diagnosis. The experimental results of the study demonstrate the effectiveness of the proposed deep learning models for feature extraction and classification, with various CNN architectures being evaluated.

The goal of Khanna, V. V., Chadaga et al.'s study [2] is to identify patients with PCOS while also suggesting an automated screening architecture with interpretable machine learning tools to support medical professionals in making decisions. Kottarathil prepared an open-source dataset on Kaggle that contains information on 541 fertile women with 43 attributes. Two stack models were employed in this paper. The first stack model makes use of LR, SVM, NB, and CNN. They also use DT, RF, AdaBoost, and XGBoost as another stack model. This paper stated that the precision of S1-98, S2-97, The best-performing pipeline is suggested by this study, which assesses several frameworks and suggests a meta-learner-based multi-level stack machine learning classifier trained on MI-engineered data.

Goyal et al., P. Bedi et al[3] is to enhance PCOS identification by medical imaging, particularly ultrasound, by utilizing artificial intelligence and machine learning methodologies. utilized only the ultrasound pictures. The innovative attention residual U-net architecture and adaptive bilateral filter-based image pre-processing are integrated into it. Three metrics are used: recall, precision, and accuracy. Lack of diversity and small size of the dataset are the key drawbacks. The way to bridge this in the future is to combine the model with clinical data for peak performance.

M. Sumathi et.al[4] to employ CNN for the identification and categorization of disorders associated with PCOS in ultrasound pictures. to show off CNN's ability to properly segregate cysts and categorize disorders connected to PCOS for medical image processing. Ultrasound images from sonoworld.com and ultrasoundimages.com make up the dataset. Using ultrasound pictures, the methodology entails preprocessing, segmentation, feature extraction, and CNN model training for the detection and categorization of disorders related to PCOS. The primary drawbacks are that more data is required to enhance the CNN model's capacity for generalization, and there may be a relationship between the CNN model's performance and differences in the resolution and quality of ultrasound images.In the future, more machine learning algorithms—like support vector machines and optimization—will be integrated and compared to CNN for the purposes of detection and classification.

C. Gopalakrishnan et al.[5] The purpose of this work is to present an automated technique for the recognition and analysis of follicles in ultrasound pictures, specifically for the diagnosis of polycystic ovarian syndrome (PCOS), known as active contour with modified Otsu threshold value.In order to automatically recognize follicles from ultrasound pictures, this research suggests an efficient strategy that combines active contour with a modified Otsu threshold value. Metrics including Mean Square Error (MSE), Peak Signal to Noise Ratio (PSNR), Absolute Difference (AD), Structural Content (SC), Maximum Difference (MD), Normalized Absolute Error (NAE), and Normalized Cross Correlation (NK) are employed. The paper's final finding states that the suggested approach, which combines active contour approaches with the modified Otsu method, performs better in terms of accuracy. The study concludes that the suggested approach, which combines active contour techniques with the modified Otsu method, performs better than alternative approaches in automatically identifying and segmenting ovarian follicles from ultrasound images.

P. Soni et al. [6]The use of medical image processing to identify females with polycystic ovarian syndrome (PCOS) using online sources is discussed in this research. The methodology is broken down into phases where segmentation and classification come after pre-processing. Gaussian smoothening is used for pre-processing, and scaling operations are used to improve contrast. The measures are Specificity, Sensitivity, and Accuracy. For increased accuracy and prediction, the suggested approach makes use of the region of interest (ROI) and bag of features. Pre-processing, training, segmentation, and classification are all part of the methodology.For categorization, the study used a rule-based framework and support vector machines, which might not be able to handle intricate and overlapping information in medical images. The creation of a user-friendly interface for the suggested system, which would increase its accessibility for

Bhat, S. A, et al. [7] is to identify PCOS using machine learning techniques and ultrasound pictures, evaluating different classifiers and offering a novel approach. The study included a dataset of 541 female patients, with 177 being diagnosed with PCOS. The study employed Competitive Neural Network (CNN) and obtained 80.84% accuracy with 32 feature vectors. The study's evaluation metrics include Accuracy, Precision, Recall, F1-score, ROC-AUC score, and Cross Validation Accuracy. The study lacks a consideration of the generalizability of results to different datasets, the impact of differences in ultrasound image quality, and the effect of patient demographics or comorbidities on classification model performance. Future goals include confirming results with bigger datasets, investigating deep learning methods, incorporating models into clinical practice, and improving the proposed models for early PCOS detection.

Rachana, B. et al. [8]. is to create an automated PCOS diagnostic tool that uses ultrasound pictures to improve accuracy. Improving early detection and prevention of problems. The information utilized in the study was gathered from several websites and contained around 50 photos of afflicted and unaffected ovaries in JPG format. To predict the occurrence of PCOS, the proposed methodology used a hybrid structure combining Naive Bayes and artificial neural network algorithms, with a dataset partitioned into 70% training and 30% testing data. Ultrasound images and machine learning techniques were used to classify the dataset. The work does not address the issues faced during the picture processing and classification process, which could provide significant insights for future research. Integrating it into an easy-to-use PCOS diagnostic tool. Exploring the potential for real-time implementation of the algorithm in clinical settings.

S. Alshakrani et al.[9] In order to determine the best model for PCOS identification, the researchers used hybrid machine learning algorithms and evaluated how well they classified PCOS cases. The PCOS dataset that Kottarathil published was gathered from ten different Indian hospitals. 44 features based on the physical and clinical characteristics of 541 women and 177 diagnosed with PCOS are included in the dataset. The models include LSVM, HRFLR, XGBRF, LGBM, and CatBOOST. The metrics are precision, recall, and accuracy. The future gap is to use more techniques on data imbalance.

Sathiya, V. Kiruthika, et al.[10] is to use machine learning techniques for ovarian detection and categorization. only used the ultrasound image dataset. This article study suggests a method called MLOD and TIOC that uses artificial neural networks to incorporate texture and intensity characteristics. The method makes use of features from the grey-level co-occurrence matrix (GLCM) and intensity from k-means clustering, including autocorrelation, sum average, and sum variance. Sensitivity, specificity, accuracy, precision, F-measure, MCC, and Roc are the metrics. The MLOD classifier outperformed the combined texture and intensity-based ovarian classification (TIOC) algorithm, achieving an average detection accuracy of 96%. The main drawback is having a small sample size, which limits its generalizability and statistical power. The future scope of this paper is by using a larger dataset of ultra sound images for efficient prediction of PCOS.

P. G. YILMAZ et al. [11] The purpose of this work is to evaluate two distinct image processing approaches for follicle detection related to Polycystic Ovary Syndrome (PCOS). only the Ultra Sound Images dataset was used. For noise reduction, wavelet transform, median filter, average filter, gaussian filter, and Wiener filter were utilized. Performance assessment is done using FAR and FRR metrics. The utilization of Wiener and Gaussian filters in the first approach and Gaussian filter in the second method yielded the best accurate results for follicle detection in PCOS, according to the paper's conclusion. For contrast settings, adaptive thresholding was found to be more accurate than histogram equalization. The quality and clarity of ultrasound pictures might differ, which may have an impact on how well follicle detection techniques work. Lack of clinical validation and lack of comparison with current practices. Future work on this project will focus on identifying additional characteristics from ovarian follicles and classifying Polycystic Ovary Syndrome (PCOS) using a bigger dataset of ultrasound pictures.

S. Srivastava et al[12] The purpose of the research is to create a deep learning model that can precisely identify ovarian cysts in ultrasound pictures by utilizing the VGG-16 architecture. Only the Ultrasound Image Dataset was used. The VGG-16 model, which was trained on ultrasound scans, can identify ovarian cysts with accuracy. The program outperforms earlier detection techniques, with a promising accuracy of 92.11%. VGG-16 and other deep learning models are frequently referred to as "black boxes," which means it might be difficult to figure out exactly how they produce their predictions. Future work in this research will classify several ovarian cyst types, including functional, dermoid, haemorrhagic ovarian cyst (HOC), and polycystic ovarian syndrome (PCOS), utilizing the fine-tuned VGG-16 deep neural network model. Furthermore, the algorithm is able to be utilized for the early detection of ovarian cancer.

The goal of SJ, Y. K. et al.'s study [13] is to discuss the difficulties in interpreting ultrasound scans for the diagnosis of polycystic ovarian syndrome (PCOS) and to provide an improved segmentation method for the region of interest (ROI) in these scans dubbed Adaptive Otsu's Technique (AOT). AOT, or Adaptive Otsu's Technique, was applied. True Positive (TP), False Positive (FP), Sensitivity, Selectivity, Precision, Dice Coefficient (DC), Jaccard Similarity Index (JSI), and F1 score are the metrics that are employed. The drawbacks of the conventional Otsu's technique are addressed by AOT for PCOS segmentation. AOT minimizes within-class variation and chooses appropriate statistical parameters to segment regions of interest in ultrasound images appropriately. The ultrasound image's intrinsic noise, poor quality, and follicular overlap may nevertheless have an impact on the AOT approach.

Vedpathak, Thakre, et.al[14] The early diagnosis and prediction of PCOS treatment is the paper's goal. solely clinical and medical datasets are employed. The Random Forest Classifier, Support Vector Classifier, Gaussian Naive Bayes, Logistic Regression, and K-Neighbours Classifier are the methods used. Measures: Accuracy, F-Score, Precision, and Recall. With an accuracy of 90%, the Random Forest Classifier was determined to be the most dependable and accurate of the four. Limitations are A study's sample size may be small, which restricts how broadly the results can be applied to broader populations. Future work is the for effective PCOS detection, a sizable data collection should be used.

A. A. Nazarudin et al. [15] This work primarily studies medical image processing methods for PCOS diagnosis and tracking. The study suggested an algorithm for finding follicles by combining the chan-vese approach and Otsu' thresholding. By highlighting the image's pixel intensities, Otsu's thresholding produces a binary mask that can be used with the Chan–Vese technique to define the follicles' boundaries. Dice score and the Jaccard Index are metrics. In the future, patients who have not yet received a PCOS diagnosis will have their ultrasound images collected in order to investigate and refine the suggested approach. In order to diagnose PCOS, this study may investigate feature extraction from post-segmented images.

METHODOLOGY :

3.1 Dataset

Clinical datasets for PCOS typically include information such as patient demographics, medical history, hormone profiles (e.g., levels of testosterone, luteinizing hormone), and other relevant clinical parameters (e.g., menstrual irregularities, ovarian morphology). Ultrasound imaging datasets for PCOS consist of transvaginal or transabdominal ultrasound scans of the ovaries. These images are used to assess ovarian morphology, such as the presence of ovarian cysts (follicles), ovarian volume, and the presence of polycystic ovaries according to established diagnostic criteria (e.g., Rotterdam criteria). Ultrasound images provide valuable insights into the structural characteristics of the ovaries in PCOS patients and are essential for diagnostic purposes and research studies in the field. The image acquisition for [8] experiment was performed using the GE LOGIQ ultrasound system equipped with a transvaginal transducer operating at a frequency of 24 Hz. A total of 65 ultrasound ovarian images were acquired for the study, providing a comprehensive dataset for analysis and experimentation.

3.2 Data Processing:

Image preprocessing plays a pivotal role in enhancing the quality and interpretability of medical images, particularly in ultrasound imaging where distinguishing subtle details is paramount. Typically, preprocessing techniques are employed to address inherent challenges such as noise, artifacts, and variations in image intensity. These techniques aim to standardize image features, improve contrast, and reduce unwanted distortions, thereby facilitating more accurate interpretation and analysis by clinicians and automated systems alike. Adaptive filtering techniques [3], such as the adaptive bilateral filter (ABF), are often utilized to enhance image sharpness while preserving important structural information. By adjusting filter parameters based on local image characteristics, ABF effectively mitigates noise and improves overall image clarity, thus laying a solid foundation for subsequent analysis tasks such as anomaly detection and classification. The [5] employed a series of image processing techniques, including RGB-Gray Scale Conversion, Image Thresholding, Speckle Noise Elimination, and Edge Detection, to enhance the quality and interpretability of the ultrasound images. Gaussian Mechanism[6] is employed to effectively mitigate noise, including impulse noise, while ensuring data integrity. Additionally, smoothening operations are performed using Gaussian filters to enhance the overall quality of the data and improve subsequent analysis. Various filtering methods [11], including Median, Average, Gaussian, and Wiener filters, were employed for denoising, followed by histogram thresholding for contrast enhancement. Binarization techniques were then applied to label follicles, and morphology methods, including erosion and expansion processes, were used to refine the images and eliminate unnecessary objects, ultimately improving accuracy in follicle detection.

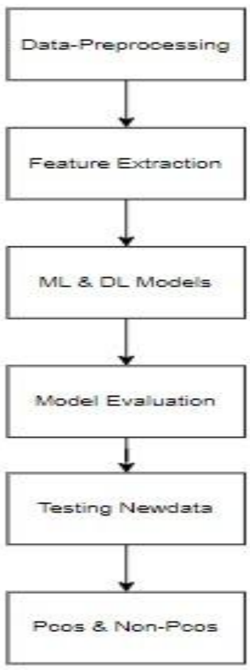

Fig-1.Workflow Architecture

3.3. Machine Learning Techniques:

In the review, several machine learning (ML) techniques were employed for the detection of Polycystic Ovary Syndrome (PCOS), including Support Vector Machines (SVM), AdaBoost, and k-Nearest Neighbors (KNN). These algorithms were applied to the datasets to classify PCOS cases based on features extracted from clinical records, hormone profiles, and ultrasound imaging data.For each ML technique, model training involved splitting the dataset into training and testing sets, with hyperparameters optimized through techniques like grid search or randomized search. The models' performance was evaluated using various evaluation metrics such as accuracy, precision, recall, and F1-score, providing insights into their effectiveness in PCOS detection.

3.3.1Support Vector Machines(SVM):

Support Vector Machine (SVM) is a powerful supervised learning algorithm widely used for binary classification tasks, including the classification of individuals with and without Polycystic Ovary Syndrome (PCOS) based on various data sources such as clinical records, hormone profiles, and ultrasound images. SVM works by identifying the optimal hyperplane that best separates PCOS-positive and PCOS-negative individuals in the feature space, maximizing the margin between the two classes to improve generalization to unseen data. SVM has been employed in conjunction with various feature selection techniques[2], such as Harris Hawk Optimization, Salp Swarm Algorithm, and Mutual Information, resulting in impressive accuracies of up to 95%. In [23] SVM was utilized alongside ultrasound image segmentation techniques to accurately identify follicle locations. By employing feature descriptors extracted using the SIFT algorithm, SVM demonstrated superior performance compared to other machine learning algorithms such as Naïve Bayes and Decision Tree, achieving an accuracy of 94%. Furthermore, in [21], SVM with a radial basis function (RBF) kernel was applied to predict subject classes based on pulse wave parameters, yielding accuracies of up to 72%.

3.3.2XGBoost:

XG Boost, or Xtreme Gradient Boosting, is a highly effective machine learning algorithm extensively employed in the classification of Polycystic Ovary Syndrome (PCOS) due to its robustness and precision. Operating within an ensemble learning framework, XGBoost utilizes gradient boosting to construct predictive models. Its application in PCOS detection involves processing datasets containing features derived from clinical records, hormone profiles, and ultrasound imaging. Notably, in [9] XGBoost's integration with Random Forest (XGBRF) offers a novel approach in the medical domain, harnessing the strengths of both algorithms to enhance precision, stability, and mitigate overfitting issues and resulted with an accuracy of 87%. In the study [16], explored the effectiveness of utilizing the Extreme Gradient Boosting with Random Forest (XGBRF) ensemble method and the CatBoost Model for Polycystic Ovary Syndrome (PCOS) detection. Achieving a precision rate of 93%, a recall rate of 92%, and an F1-score of 92%, the findings indicate the potential of this approach in accurately identifying PCOS cases.

3.3.3K-Nearest Neighbors(KNN)

k-Nearest Neighbors (k-NN) is a versatile supervised learning algorithm commonly employed in Polycystic Ovary Syndrome (PCOS) classification tasks. Utilizing features such as clinical records and ultrasound images, k-NN classifies individuals by determining the majority class among their k nearest neighbors in the feature space. In [21] focusing on PCOS classification based on radial pulse wave parameters, machine learning (ML) techniques, including k-Nearest Neighbors (KNN), were applied. KNN, a non-parametric algorithm, leverages the majority labels of the closest data points to predict new data points. In this specific study, a value of k equal to 22 was determined, resulting in a classification accuracy of 72%. This research underscores the potential of ML methods, like KNN, in objectively diagnosing PCOS and enhancing traditional Chinese medicine practices.

3.3.4AdaBoost :

AdaBoost, short for Adaptive Boosting, is an ensemble learning technique used for classification and regression tasks. AdaBoost works by sequentially combining multiple weak learners to create a strong classifier. In each iteration of training, AdaBoost assigns higher weights to misclassified data points, thereby focusing on challenging instances. Subsequent weak learners are then trained on the updated dataset to improve classification performance. The final prediction is made by aggregating the predictions of all weak learners, weighted by their respective performance during training.

3.3.5Stacking:

In the study [2], employed a multi-level stacking approach using various machine learning algorithms on a clinical dataset comprising 541 patients. The first stack (STACK-1) integrated algorithms like Logistic Regression (LR), Support Vector Machine (SVM), Naive Bayes (NB), and K-Nearest Neighbors (KNN), while the second stack (STACK-2) included Decision Trees (DT), Random Forest (RF), AdaBoost, and XGBoost. These stacks were further combined to form STACK-3. Feature extraction techniques such as Scatter Search Optimization (SO), Harris Hawks Optimization (HHO), and Mutual Information (MI) were utilized across all stacks. Notably, STACK-1 achieved an accuracy of 98%, while STACK-2 attained 97% accuracy. Finally, the combined STACK-3 exhibited a high accuracy of 98%.

3.4Deep learning Techniques:

Deep learning (DL) techniques play a vital role in the detection and classification of Polycystic Ovary Syndrome (PCOS), leveraging their ability to extract complex features from heterogeneous datasets.

3.4.1CNN:

Convolutional Neural Networks (CNNs) have emerged as a powerful tool for image classification tasks, particularly in the context of Polycystic Ovary Syndrome (PCOS) detection. In [4], CNNs are utilized as image classifiers, employing segmentation and feature extraction methods to identify cysts within ultrasound images. The algorithm achieved an impressive accuracy of 85% with test data, showcasing its ability to accurately detect PCOS-related features. Furthermore, the classification of ultrasound images into PCO and non-PCO classes using CNNs yielded robust results, with the system automatically extracting features from each image without the need for explicit feature extraction methods. The proposed architecture demonstrated outstanding performance, achieving a micro-average f1-score of 100% and an average accuracy of 76.36% during testing. Additionally, the optimization of CNNs proved challenging, with issues such as NaN activations and poor weight initialization requiring careful consideration of hyperparameters [19]. In light of these findings, incorporating segmentation techniques prior to CNN processing could enhance accuracy and eliminate redundant data, thereby improving overall performance [24].

[1] utilized six well-known CNN architectures for feature extraction and classification, namely VGG16, VGG19, InceptionV3, DenseNet121, DenseNet201, and MobileNet. Among these models, the Inception model demonstrated superior performance, achieving an accuracy of 84.81%, precision of 69.57%, F1-score of 72.73%, recall of 76.19%, and specificity of 87.93%. Additionally, a combined model was proposed, integrating features extracted from both clinical and ultrasound image datasets. Fusion techniques were applied, with VGG-16 and VGG-19 models surpassing others across various metrics. Specifically, VGG-16 achieved an accuracy, precision, F1-score, recall, and specificity of 77.19%, 61.54%, 71.11%, 84.21%, and 73.68%, respectively, while VGG-19 attained 75.44%, 80.77%, 75.00%, 70.00%, and 81.48%, respectively. Notably, the MobileNet model, utilized for image feature extraction, achieved an accuracy of 82.46%, precision of 84.62%, F1-score of 81.48%, recall of 78.57%, and specificity of 86.21%.

The proposed [24] combines deep neural networks with image segmentation to improve PCOS detection from ultrasound images. Utilizing region-based or watershed segmentation, specific regions like ovarian follicles are isolated. This segmentation aids in quantifying follicle count and shape, facilitating easier classification tasks. Overall, this approach enhances accuracy by focusing on relevant features, potentially leading to more effective healthcare outcomes

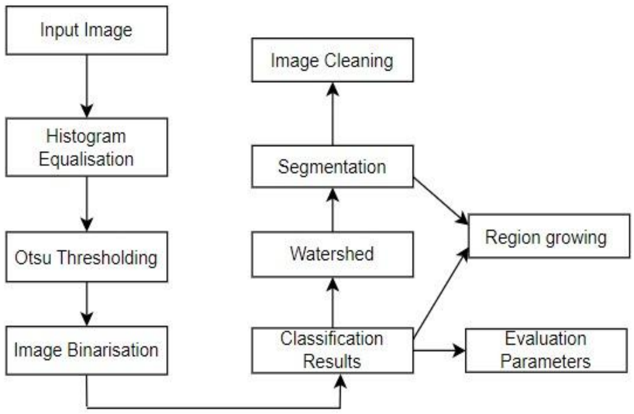

Fig-2. PCOS Segmentation

MLOD and TIOC [10] integrates both intensity and texture features and leveraging artificial neural networks, the approach aims to improve the accuracy and reliability of diagnosing ovarian conditions, facilitating timely intervention and treatment planning for affected individuals. The below fig shows the architecture of [10]

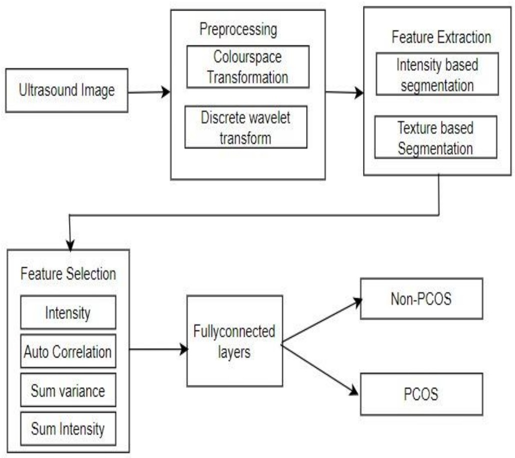

Fig-3. MLOD and TIOD architecture

3.4.2Scale-Invariant Feature Transform

The Scale-Invariant Feature Transform (SIFT) algorithm for detecting and describing key features in ultrasound images for the detection of Polycystic Ovary Syndrome (PCOS). Initially, potential key points in the image are identified, ensuring scale invariance and robustness to illumination and viewpoint changes. These key points are then refined, filtering out low-contrast and edge key points [23]. Subsequently, each key point is assigned a dominant orientation for rotational invariance. Following this, descriptors are generated for each key point, capturing local image structure information. These descriptors enable matching keypoints between different images, facilitating the detection of PCOS-related patterns. Overall, the utilization of the SIFT algorithm in PCOS detection enhances the robustness and accuracy of feature extraction, contributing to more effective diagnosis and treatment planning.

RESULTS AND DISCUSSIONS :

Research in PCOS detection is essential as it is a prevalent condition affecting women worldwide and can lead to various health complications if not diagnosed and managed effectively. Polycystic Ovary Syndrome (PCOS) is characterized by the presence of multiple sacs filled with fluid in the ovaries, along with symptoms such as irregular menstrual cycles, weight gain, acne, and hormonal imbalances. Detecting PCOS accurately from ultrasound images is crucial for early intervention and personalized treatment strategies. In the presented results and discussion, it is evident that the performance of machine learning models, particularly Stacking, shows promising results with an accuracy of 98% [6]. However, the performance of deep learning models, including Inception V3, outperforms other models, showcasing the potential of deep learning techniques in medical image analysis. Despite this, there are still limitations in deep learning models' performance, indicating the need for further research and development in this area. One notable approach proposed in the literature involves segmenting ovarian follicles using image processing techniques like histogram equalization and binarization [5], potentially aiding in exceptional classification. Additionally, the utilization of the Scale-Invariant Feature Transform (SIFT) algorithm shows promise, although it was trained on a relatively small dataset. Expanding the dataset and incorporating advanced techniques like transformers and preprocessing methods could further enhance classification accuracy. Furthermore There's a possibility of follicles being present without PCOS or PCOS being present without follicles. This ambiguity poses challenges, as individuals without identified follicles may still have PCOS, and vice versa. Thus, comprehensive diagnostic approaches are necessary, considering multiple factors beyond follicle presence alone, Therefore integrating clinical data with image data presents an opportunity for improved performance in real-time applications. This combined approach can address the challenge of accurately diagnosing PCOS, considering the variability in follicle presence and PCOS diagnosis. The work of researchers who have successfully combined clinical and image datasets, achieving an accuracy of 84% [1], highlights the potential of this approach. By incorporating advanced techniques and increasing the diversity of datasets, the accuracy of PCOS classification can be further improved, benefiting individuals with accurate and timely diagnoses.

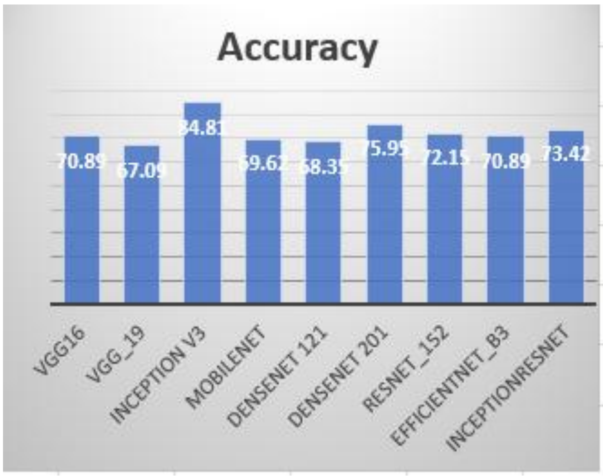

Fig-4. performance of individual Transfer learning models

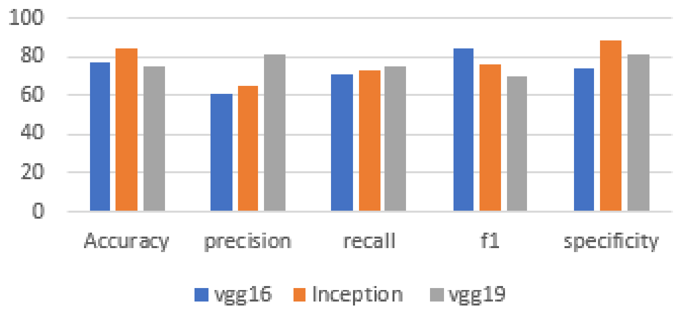

Fig-5.Performerence of Combined model (clinical data+image data)

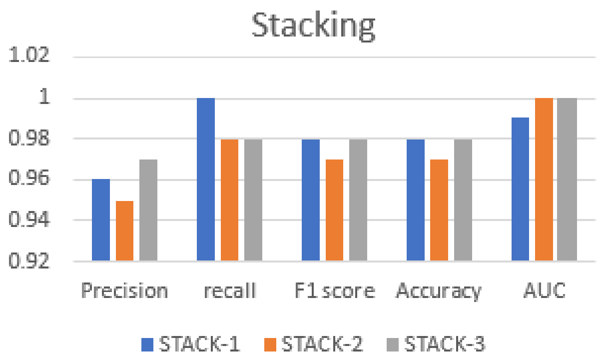

Fig-6.Performance of Stacking model

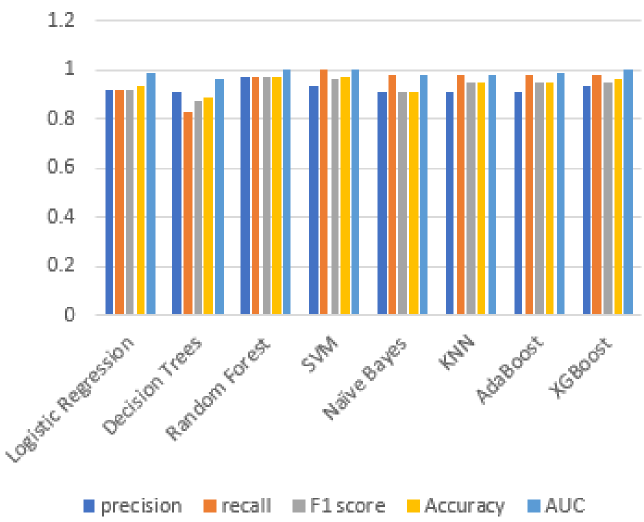

Fig-7. Performance of Individual Machine Learning models

CONCLUSION :

In conclusion, the review highlights the significant strides made in the detection and classification of Polycystic Ovary Syndrome (PCOS) through machine learning (ML) and deep learning (DL) methodologies. ML techniques such as Support Vector Machines (SVM), AdaBoost, and k-Nearest Neighbors (KNN) have exhibited high accuracy in PCOS classification, particularly when applied to diverse datasets encompassing clinical records, hormone profiles, and ultrasound imaging data. Moreover, DL techniques, including Convolutional Neural Networks (CNNs) with architectures like VGG, ResNet, DenseNet, and Inception, have showcased remarkable proficiency in extracting intricate features from raw imaging data, leading to enhanced diagnostic precision. The integration of preprocessing techniques like image segmentation and feature extraction methods such as the Scale-Invariant Feature Transform (SIFT) algorithm has further bolstered classification outcomes. Despite challenges such as data availability and interpretability, collaborative efforts among researchers, clinicians, and technology developers hold promise for advancing Computer-Aided Diagnosis (CAD) systems for PCOS, with the potential to revolutionize diagnosis and management practices, ultimately improving patient outcomes and quality of care.

REFERENCES :

[1] Alamoudi, A., Khan, I. U., Aslam, N., Alqahtani, N., Alsaif, H. S., Al Dandan, O., ... & Al Bahrani, R. (2023). A Deep Learning Fusion Approach to Diagnosis the Polycystic Ovary Syndrome (PCOS). Applied Computational Intelligence and Soft Computing, 2023

[2] Khanna, V. V., Chadaga, K., Sampathila, N., Prabhu, S., Bhandage, V., & Hegde, G. K. (2023). A distinctive explainable machine learning framework for detection of polycystic ovary syndrome. *Applied System Innovation*, *6*(2), 32.

[3] Bedi, P., Goyal, S. B., Rajawat, A. S., & Kumar, M. (2024). An integrated adaptive bilateral filter-based framework and attention residual U-net for detecting polycystic ovary syndrome. *Decision Analytics Journal*, *10*, 100366.

[4] 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.

[5] Gopalakrishnan, C., & Iyapparaja, M. (2020). Active contour with modified Otsu method for automatic detection of polycystic ovary syndrome from ultrasound image of ovary. Multimedia Tools and Applications, 79, 17169-17192.

[6] Soni, P., & Vashisht, S. (2019). Region of Interest and Bag of Features For Predicting Polycystic Ovaries In Females. *Think India Journal*, *22*(3), 7901-7911.

[7] Bhat, S. A. (2021). *Detection of polycystic ovary syndrome using machine learning algorithms* (Doctoral dissertation, Dublin, National College of Ireland).

[8] 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.

[9] Alshakrani, S., Hilal, S., & Zeki, A. M. (2022, October). Hybrid Machine Learning Algorithms for Polycystic Ovary Syndrome Detection. In 2022 International Conference on Data Analytics for Business and Industry (ICDABI) (pp. 160-164). IE

[10] Kiruthika, V., Sathiya, S., & Ramya, M. M. (2020). Machine learning based ovarian detection in ultrasound images. International Journal of Advanced Mechatronic Systems, 8(2-3), 75-85.

[11] YILMAZ, P. G., & ÖZMEN, G. (2020). Follicle detection for polycystic ovary syndrome by using image processing methods. International Journal of Applied Mathematics Electronics and Computers, 8(4), 203-208.

[12] Srivastava, S., Kumar, P., Chaudhry, V., & Singh, A. (2020). Detection of ovarian cyst in ultrasound images using fine-tuned VGG-16 deep learning network. SN Computer Science, 1, 1-8.

[13] SJ, Y. K. (2020). Adaptive Otsu's Technique for PCOS Segmentation from Ovarian Ultrasound Images. In Intelligent Systems and Computer Technology (pp. 611-618). IOS Press. 8(2-3), 75-85.

[14] Thakre, V., Vedpathak, S., Thakre, K., & Sonawani, S. (2020). PCOcare: PCOS detection and prediction using machine learning algorithms. Biosci Biotechnol Res Commun, 13(14), 240-244.

[15] Nazarudin, A. A., Zulkarnain, N., Mokri, S. S., Zaki, W. M. D. W., Hussain, A., Ahmad, M. F., & Nordin, I. N. A. M. (2023). Performance Analysis of a Novel Hybrid Segmentation Method for Polycystic Ovarian Syndrome Monitoring. Diagnostics, 13(4), 750.

[16] Dewi, R. M., Adiwijaya, Wisesty, U. N., & Jondri. (2018, March). Classification of polycystic ovary based on ultrasound images using competitive neural network. In *Journal of Physics: Conference Series* (Vol. 971, p. 012005). IOP Publishing.

[17] Purnama, B., Wisesti, U. N., Nhita, F., Gayatri, A., & Mutiah, T. (2015, May). A classification of polycystic Ovary Syndrome based on follicle detection of ultrasound images. In *2015 3rd International conference on information and communication technology (ICoICT)* (pp. 396-401). IEEE.

[18] He, D., Liu, L., Miao, S., Tong, X., & Sheng, M. (2019). Probabilistic guided polycystic ovary syndrome recognition using learned quality kernel. *Journal of Visual Communication and Image Representation*, *63*, 102587..

[19] Cahyono, B., Mubarok, M. S., & Wisesty, U. N. (2017, May). An implementation of convolutional neural network on pco classification based on ultrasound image. In *2017 5th International Conference on Information and Communication Technology (ICoIC7)* (pp. 1-4). IEEE.

[20] 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.

[21] Lim, J., Li, J., Feng, X., Feng, L., Xia, Y., Xiao, X., ... & Xu, Z. (2023). Machine learning classification of polycystic ovary syndrome based on radial pulse wave analysis. BMC Complementary Medicine and Therapies, 23(1), 409.

[22] Sreejith, S., Nehemiah, H. K., & Kannan, A. (2022). A clinical decision support system for polycystic ovarian syndrome using red deer algorithm and random forest classifier. Healthcare Analytics, 2, 100102

[23] Gopalakrishnan, C., & Iyapparaja, M. (2019). Detection of polycystic ovary syndrome from ultrasound images using SIFT descriptors. *Bonfring International Journal of Software Engineering and Soft Computing*, *9*(2), 26-30.

[24] Soni, P., & Vashisht, S. (2019). Image segmentation for detecting polycystic ovarian disease using deep neural networks. *International Journal of Computer Sciences and Engineering*, *7*(3), 534-537.