International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 04 | April – 2025 DOI: 10.55041/ISJEM02867
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

PCOS DETECTION USING MACHINE LEARNING

Prof. Shraddha Toney
Computer Engineering
AISSMS Institute of Information Technology Pune,India

Prasad Sanjay Jadhav
Dept. of Computer Science,
AISSMS Institute of
Information Technology,
Pune.

Pratik Agarkar
Dept. of Computer Science,
AISSMS Institute of
Information Technology,
Pune,

Prasad Sunil Jadhav
Dept. of Computer Science,
AISSMS Institute of
Information Technology,
Pune.

Abstract
Polycystic Ovary Syndrome (PCOS) is one of the most common causes of female infertility, affecting a large number of women of reproductive age and often persisting beyond the childbearing years. This hormonal disorder can result in several long-term health complications, making early detection and timely intervention crucial. In this study, we propose a deep learning- based approach using Convolutional Neural Networks (CNN) for the early diagnosis of PCOS by analyzing medical imaging data, including sonography scans and lumbar MRI images. The CNN model effectively identifies patterns and abnormalities associated with PCOS, enhancing diagnostic accuracy and improving overall clinical outcomes. The model leverages advanced image processing techniques and deep learning algorithms to classify PCOS cases with high precision. This approach facilitates the development of personalized treatment plans, enabling healthcare professionals to provide targeted therapy and better symptom management. By integrating cutting-edge technology with medical expertise, this method empowers women to take proactive control of their reproductive health and overall well-being. The proposed model demonstrates significant improvements in diagnostic performance and offers a reliable, automated solution for PCOS detection in clinical settings.

Index Terms — PCOS, CNN, Medical Imaging, Early Detection, Personalized Treatment, Deep Learning, Reproductive Health, Automated Diagnosis

INTRODUCTION

Polycystic Ovary Syndrome (PCOS) is one of the most prevalent endocrine disorders affecting approximately 1 in 10 women of reproductive age, directly impacting fertility and overall health. It is characterized by irregular menstrual cycles, ovarian cysts, and elevated androgen levels, which may lead to severe symptoms such as excessive hair growth, weight gain, acne, and hair loss. Moreover, women with PCOS are at a higher risk of developing Type 2 diabetes, cardiovascular diseases, high blood pressure, and sleep apnea.

Diagnosing PCOS can be challenging due to its complex symptoms and varying clinical presentations. While ovarian cysts are often considered a primary indicator, other metabolic and hormonal imbalances, including estrogen, progesterone, follicle-stimulating hormone (FSH), and luteinizing hormone (LH), are also crucial markers for an accurate diagnosis. Studies have shown that obesity and PCOS share a bidirectional relationship, with 30–75% of women with PCOS also experiencing obesity, which further complicates the diagnosis process.

In response to these challenges, machine learning-based approaches have been explored to improve the accuracy and efficiency of PCOS diagnosis. Inspired by recent advancements, this study leverages Convolutional Neural Networks (CNN) to develop a robust diagnostic model for PCOS detection. By analyzing medical imaging data, including ultrasound scans, the CNN model can accurately identify patterns indicative of PCOS. Additionally, the system is implemented using Kotlin, enabling seamless integration into a user-friendly mobile application. This approach not only facilitates early detection but also supports personalized treatment plans, thereby improving symptom management and enhancing overall reproductive health.

PROBLEM DEFINITION

Women of reproductive age are frequently affected with Polycystic Ovary Syndrome (PCOS), a common endocrine illness marked by polycystic ovaries, irregular menstrual periods, and hormonal abnormalities. For the purpose of managing and preventing related health concerns, such as infertility, obesity, type 2 diabetes, and cardiovascular illnesses, early identification of PCOS is essential.

The task at hand is to create a machine learning-based method that uses Convolutional Neural Networks (CNNs) implemented in the Kotlin programming language to accurately identify PCOS from ovary sonography pictures. The goal is to create a reliable system that can identify whether or not an ovarian sonography picture shows PCOS, enabling medical practitioners to make well-informed diagnoses.

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 04 | April – 2025 DOI: 10.55041/ISJEM02867
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

LITERATURE SURVEY

Conventional Approaches for PCOS Diagnosis:

The widely accepted Rotterdam criteria serve as the foundation for diagnosing PCOS, requiring the presence of any two of three key indicators: irregular menstrual cycles, elevated levels of male hormones, and the detection of ovarian cysts through ultrasound imaging. Hormonal assessments are frequently employed to identify imbalances linked to PCOS, such as Anti-Müllerian Hormone (AMH) levels. Additionally, transvaginal ultrasound remains a crucial imaging tool for examining ovarian structure and confirming the presence of polycystic ovaries. While these methods have played a significant role in diagnosing PCOS, they are often limited by inconsistencies in interpretation, subjective evaluation, and variations in sensitivity and accuracy.

Challenges in Diagnosing PCOS:

Diagnosing PCOS remains complex due to its diverse clinical presentation and overlap with other reproductive and endocrine disorders, which can lead to misdiagnosis or delayed identification. Diagnostic standards have evolved over time, with ongoing debates about the appropriate threshold values for hormonal markers and ultrasound findings. Moreover, limited access to specialized healthcare facilities, especially in rural or underserved regions, can delay timely diagnosis and treatment for women with PCOS. Recent advances in medical imaging, including 3D ultrasound and MRI, have enhanced the ability to visualize and assess ovarian structure, improving diagnostic accuracy. Artificial intelligence (AI) and machine learning methods, particularly Convolutional Neural Networks (CNNs), have shown promising results in medical imaging analysis, including the detection of abnormalities in ultrasound and MRI scans. By integrating AI-driven approaches into PCOS diagnosis, healthcare providers can address the limitations of traditional methods, offering more consistent, automated, and objective assessments of ovarian morphology and hormonal levels.

CNN algorithms, a category of deep learning models, are particularly suited for analyzing medical images as they can automatically extract complex patterns and features from raw data. Their proven effectiveness in organ segmentation, disease classification, and tumor detection often surpasses human accuracy. The use of transfer learning, where CNN models are pre-trained on large datasets and fine-tuned for specific tasks, has further enhanced their adaptability and efficiency in medical imaging analysis.

Our approach leverages CNN algorithms to develop an automated PCOS detection system using medical imaging data, such as ultrasound and pelvic MRI scans. This system enhances diagnostic precision, minimizes variability in interpretation, and facilitates early identification of PCOS. By embedding CNN-based techniques, this solution offers a scalable and accessible diagnostic tool that can bridge healthcare gaps and empower women to proactively manage their reproductive health.

PROPOSED METHODOLOGY

The development of a PCOS diagnosis system using CNN and Kotlin involves a series of structured and iterative steps to ensure accuracy and efficiency. The first step involves gathering a comprehensive dataset of ovarian sonography images that are accurately labeled with the corresponding PCOS status.

These images are then preprocessed to create a consistent input format, which includes resizing them to a uniform size and normalizing pixel values to ensure consistent data representation. To improve dataset variability and enhance model generalization, augmentation techniques such as rotation, flipping, and scaling are applied. This ensures that the CNN model is exposed to a wide range of data variations, improving its ability to detect subtle patterns related to PCOS.

Once the data is prepared, exploratory data analysis (EDA) is conducted to gain insights into the dataset's structure and characteristics. This involves analyzing the distribution of data, identifying potential biases, and examining correlations between image features and PCOS status. Visualization techniques are used to explore the relationships within the dataset, helping to identify patterns and anomalies that could influence model performance. Understanding these patterns is crucial for designing an effective model architecture.

The next step is developing and designing the CNN model specifically for PCOS detection. The model is implemented using Kotlin, a modern programming language known for its efficiency and versatility, combined with deep learning frameworks such as TensorFlow or Keras. The CNN model is designed to automatically extract hierarchical features from sonography images, which allows it to identify complex patterns associated with PCOS. The model architecture is fine-tuned by adjusting hyperparameters such as the number of layers, filter sizes, and learning rates to optimize performance. The model's architecture is refined through iterative testing and validation to ensure it captures the most relevant features for accurate classification.

Once the model is designed, it is trained and validated using a structured approach. The dataset is divided into training, validation, and test sets to monitor performance and prevent overfitting. The CNN model is trained using the training set while continuously evaluating its performance on the validation set. To improve generalization and avoid overfitting, techniques such as data augmentation and cross-validation are applied. These methods enhance the model's ability to recognize patterns in new, unseen data, increasing its robustness and reliability.

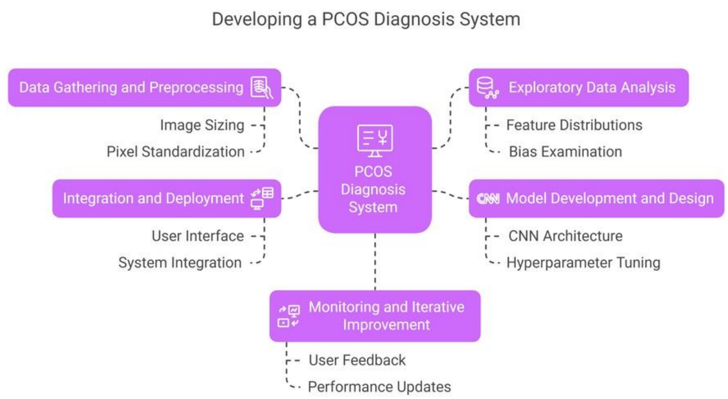
Model evaluation is a critical step in determining the effectiveness of the trained CNN model. The model's performance is assessed using key metrics such as accuracy, precision, recall, and F1-score. This helps in understanding the strengths and limitations of the model, allowing for targeted improvements. The model's predictions are carefully analyzed to identify areas of misclassification or inconsistencies. Visualization of results helps interpret the model's decision- making process, making it easier to identify gaps and adjust the model architecture accordingly.

After achieving satisfactory performance, the CNN model is integrated into a Kotlin-based real-world application. A user- friendly interface is developed, allowing medical practitioners and users to upload sonography images for PCOS detection. The system is designed to provide quick and accurate results, supporting healthcare professionals in making well-informed decisions. The PCOS detection system is also integrated with existing medical diagnostic systems to ensure smooth workflow and seamless data exchange.

Post-deployment, continuous monitoring and improvement are essential to maintain the system's reliability and accuracy. User feedback is collected to identify potential issues and areas for enhancement. Regular updates are performed to incorporate new research findings, refine model accuracy, and address any technical issues. The system's performance is continuously evaluated, and iterative improvements are made to enhance its precision and usability.

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 04 | April – 2025 DOI: 10.55041/ISJEM02867
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

This ongoing refinement ensures that the PCOS detection system remains effective and up-to-date with the latest advancements in medical imaging and machine learning. By employing this systematic and structured methodology, a reliable and efficient tool for the early diagnosis and treatment of PCOS can be developed, empowering healthcare professionals with a powerful tool to improve patient outcomes.



V. Mathematical Model

A Convolutional Neural Network (CNN) is employed as the mathematical framework for tasks such as image classification, object detection, and segmentation. Its structure is composed of several layers, including convolutional, pooling, and fully connected layers.

Convolutional layers function by applying filters to input images using element-wise multiplication and summation, allowing the network to detect spatial patterns and features. These filters slide over the input image, capturing local details and generating feature maps that reflect the detected patterns. Pooling layers subsequently reduce the dimensionality of these feature maps by downsampling them, which helps to identify dominant features and reduce computational load.

Fully connected layers establish connections between every neuron in one layer and all neurons in the next, enabling the model to form high-level feature representations and generate predictions. Non-linear activation functions, such as the Rectified Linear Unit (ReLU), introduce non-linearity into the network, which enhances the model's capability to identify complex patterns and correlations in the data.

The mathematical formulation of the CNN model for PCOS detection can be defined as follows:

Convolution Operation:

The convolution operation creates a feature map by applying a filter (or kernel) to an input image. Mathematically, the convolution at a specific position (i,j)(i, j) between a filter W and an input image X can be represented as:

$$(W * X)(i, j) = \sum_m \sum_n W(m, n) \cdot X(i - m, j - n)$$

Pooling Operation:

Pooling layers reduce the spatial dimensions of the feature maps using downsampling techniques. One widely used method is max pooling, which selects the maximum value within a defined pooling window. Mathematically, max pooling is defined as:

$$Y(i, j) = \max_{m,n} X(i\times s + m, j\times s + n)$$

where YYY is the output after pooling, sss is the stride, and m,nm, nm,n iterate over the pooling window.

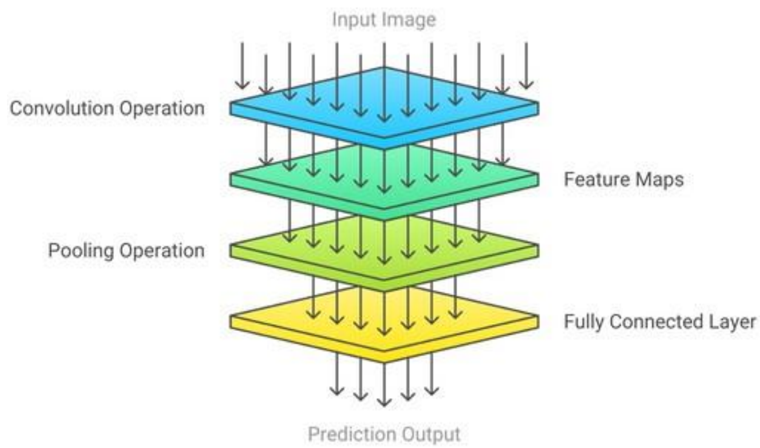
Fully Connected Layer:

In a fully connected layer, each neuron is connected to every neuron in the preceding layer. The output of a fully connected layer can be computed using matrix multiplication, expressed as:

$$Z = \mathrm{ReLU}(X \cdot W + b)$$

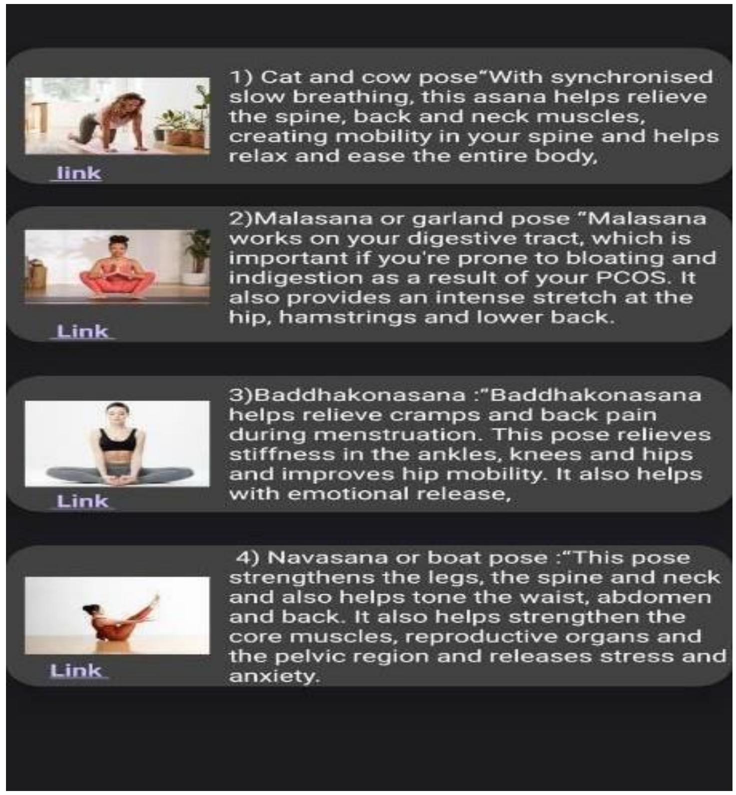
where XXX is the input feature vector, WWW is the weight matrix, bbb is the bias vector, and ReLU is the activation function applied to introduce non-linearity.

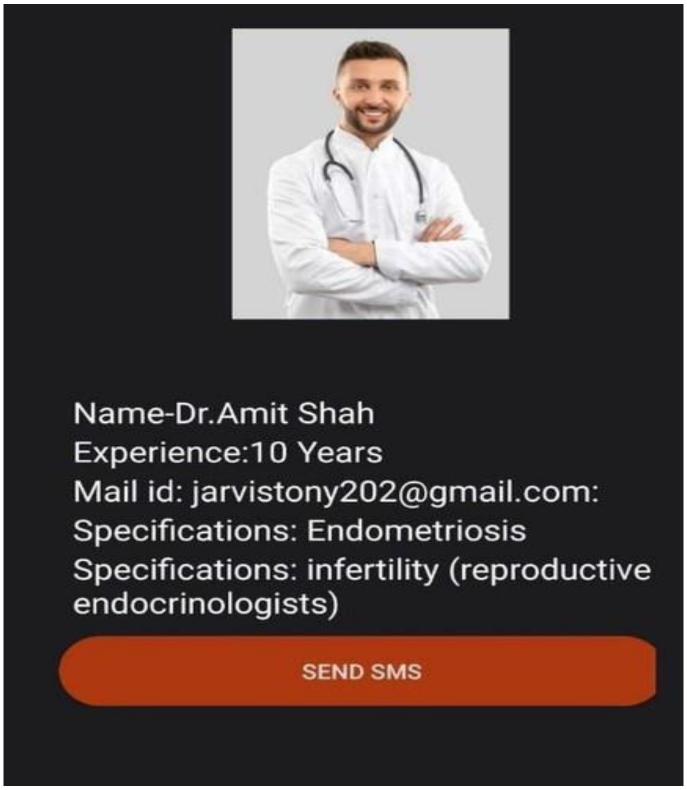
These equations form the foundation of CNNs, enabling them to extract complex patterns from input data and produce accurate predictions. The mathematical operations underlying convolution, pooling, and fully connected layers empower CNN models to effectively classify images and identify patterns, which is critical for PCOS detection and similar diagnostic tasks.



IMPLEMENTATION

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 04 | April – 2025 DOI: 10.55041/ISJEM02867
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata





CONCLUSION

In summary, the proposed Android application dedicated to women's health management, with a primary focus on Polycystic Ovary Syndrome (PCOS) detection, signifies a significant advancement in the field of reproductive health. By leveraging state-of-the-art technologies, a user-friendly interface, and evidence-based insights, the application addresses critical challenges such as limited healthcare access, delayed diagnosis, and the complexity of managing PCOS and related reproductive health issues. This innovative solution stands out by integrating advanced AI-based methodologies with practical, real-world usability, offering a transformative tool for enhancing women's healthcare.

One of the most notable strengths of the application is its use of Convolutional Neural Network (CNN) algorithms to analyze medical imaging data with remarkable accuracy and consistency. By automating the detection of ovarian abnormalities and hormonal imbalances, the app enables early diagnosis of PCOS, which is crucial for timely medical intervention and the development of personalized treatment plans. This enhanced diagnostic accuracy reduces the likelihood of misdiagnosis, empowering individuals to take proactive control of their reproductive health while mitigating potential long-term health complications.

Furthermore, the application's intuitive design and tailored health dashboard make it easy for users to monitor their menstrual cycles, symptoms, and medication schedules. AI- driven insights provide personalized recommendations based on the user's health data, promoting informed decision-making and encouraging proactive health management. The app also includes a wealth of educational resources and access to community support networks, creating a sense of connection and mutual support among users. This peer-driven environment fosters emotional and psychological well-being, reinforcing a culture of shared learning and empowerment.

The integration of wearable technology is another key feature that enhances the app's functionality. By allowing seamless synchronization with fitness trackers and smartwatches, the application provides a comprehensive view of the user's overall health. This holistic approach not only supports reproductive health management but also encourages the adoption of healthier lifestyle choices by offering real-time feedback and performance tracking.

While the proposed application offers significant benefits, it is important to recognize and address potential challenges. Ensuring data privacy and user confidentiality is critical to maintaining user trust and compliance with healthcare regulations. Additionally, addressing disparities in healthcare access and technological literacy, particularly in rural and underserved areas, is essential for ensuring that the benefits of the application reach a broad and diverse user base. Continuous refinement of the AI algorithms, expansion of the dataset for improved model accuracy, and enhanced user feedback mechanisms will be key to overcoming these challenges and optimizing the application's performance.

Ultimately, the proposed Android application represents a powerful tool for transforming the management of PCOS and women's health at large. By combining technological innovation with a user-centered design, the application empowers individuals to take control of their reproductive health, fosters a supportive community, and provides actionable insights for better health outcomes. This forward- thinking solution has the potential to bridge existing gaps in reproductive healthcare, making quality healthcare more accessible and effective for women globally. Through continuous development and user engagement, the application is poised to become a leading platform for reproductive health management, setting new standards for accuracy, accessibility, and user empowerment.

FUTURE SCOPE

The suggested Android application for women's health management, with a specific focus on PCOS detection, holds significant potential for future development and expansion. Numerous opportunities for further exploration include:

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 04 | April – 2025 DOI: 10.55041/ISJEM02867
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

The suggested Android application for women's health management, with a specific focus on PCOS detection, holds significant potential for future development and expansion. Numerous opportunities for further exploration include:

1. Enhanced AI Algorithms

Continual improvement of AI-based algorithms can significantly boost the accuracy, sensitivity, and specificity of PCOS detection. Advanced machine learning techniques such as deep learning and ensemble methods can improve the analysis of medical imaging, leading to better diagnostic outcomes and deeper insights into the condition.

2. Incorporation of Genetic Data

Integrating genetic data analysis into the application could enable more personalized diagnosis and treatment plans for PCOS. Leveraging genomic information would provide a better understanding of the underlying mechanisms of PCOS, supporting targeted medical interventions and improving health outcomes.

3. Telemedicine and Remote Monitoring

Ex0panding telemedicine capabilities would allow users to consult healthcare professionals remotely, improving access to medical care. Remote monitoring tools, such as continuous glucose monitoring for insulin resistance associated with PCOS, would enable better disease management and real-time health tracking.

4. Internationalization and Localization

Adapting the application to different languages and cultural preferences would broaden its global reach and user engagement. Customizing content, healthcare resources, and support forums based on regional healthcare practices and user preferences would improve user satisfaction and accessibility.

5. Research and Clinical Trials

Collaborating with research institutions and healthcare organizations could integrate the app into clinical trials and medical research focused on PCOS and women's health. Data collected through the app could provide valuable insights into PCOS epidemiology, treatment effectiveness, and patient outcomes, contributing to the advancement of medical knowledge.

6. Integration with Electronic Health Records (EHR) Enabling seamless integration with electronic health record systems would allow healthcare providers to access user- generated health data within their clinical workflows. This would improve care coordination, enable data-driven decision-making, and enhance the continuity of care for PCOS patients.

6. Integration with Electronic Health Records (EHR)

Enabling seamless integration with electronic health record systems would allow healthcare providers to access user-generated health data within their clinical workflows. This would improve care coordination, enable data-driven decision-making, and enhance the continuity of care for PCOS patients.

.

7. Expansion to Other Reproductive Health Conditions The application could be extended to address other reproductive health issues such as endometriosis, infertility, and menstrual disorders. Adding tailored features, educational resources, and support networks for these conditions would create a more comprehensive women's health management tool.

In conclusion, the future potential of the application is extensive, encompassing advancements in AI, genetic data analysis, telemedicine, internationalization, research collaboration, EHR integration, and broader reproductive health coverage. These developments would enhance diagnostic accuracy, improve user experience, and empower women to take control of their reproductive health.

REFERENCES

[[1] Rotterdam ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group. (2004). Revised 2003 consensus on diagnostic criteria and long term health risks related to polycystic ovary syndrome (PCOS). Human Reproduction, 19(1), 41-47.

[2] Teede, H. J., Misso, M. L., & Costello, M. F. (2019). Recommendations from the international evidence based guideline for the assessment and management of polycystic ovary syndrome. Fertility and Sterility, 110(3), 364-379.

[3] Dewailly, D., Lujan, M. E., Carmina, E., et al. (2014). Definition and significance of polycystic ovarian morphology: a task force report from the Androgen Excess and Polycystic Ovary Syndrome Society. Human Reproduction Update, 20(3), 334-352.

4] Azziz, R., & Carmina, E. (2018). PCOS in 2015: New insights into the genetics of polycystic ovary syndrome. Nature Reviews Endocrinology, 12(2), 76-78.

[5] Broekmans, F. J., Fauser, B. C., & Eijkemans, M. J. (2005). PCOS:0 an ovarian disorder. Reproductive BioMedicine Online, 10(2), 219-225.

[6] Harris, H. R., Terry, K. L., & Polycystic Ovary Syndrome, Family History, and the Risk of Endometrial Cancer in Women. (2016). American Journal of Epidemiology, 184(3), 209-218.

[7] Amer, S. A., & Gopinath, P. (2018). Three dimensional ultrasound in gynecology: past, present, and future. Obstetrics & Gynecology Science, 61(1), 1-10.

[8] Greenspan, H., van Ginneken, B., & Summers, R. M. (2016). Guest Editorial Deep Learning in Medical Imaging: Overview and Future Promise of an Exciting New Technique. IEEE Transactions on Medical Imaging, 35(5), 1153-1159.

ISSN: 2583-6129
DOI: 10.55041/ISJEM02867

[9] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444.
[10] Litjens, G., Kooi, T., & Bejnordi, B. E. (2017) A survey on deep learning in medical image analysis. Medical Image Analysis, 42, 60-88.
[11] Tajbakhsh, N., Shin, J. Y., & Gurudu, S. R. (2016). Convolutional neural networks for medical image analysis: Full training or fine tuning? IEEE Transactions on Medical Imaging, 35(5), 1299-1312.
[12] Azziz, R., Carmina, E., Dewailly, D., et al. (2016). "Position statement: Criteria for defining polycystic ovary syndrome as a predominantly hyperandrogenic syndrome: An Androgen Excess Society guideline." Journal of Clinical Endocrinology & Metabolism, 91(11), 4237-4245.
[13] Legro, R. S., Arslanian, S. A., Ehrmann, D. A., et al. (2013). "Diagnosis and treatment of polycystic ovary syndrome: An Endocrine Society clinical practice guideline." Journal of Clinical Endocrinology & Metabolism, 98(12), 4565-4592.
[14] Teede, H. J., Misso, M. L., Costello, M. F., et al. (2018). "International evidence-based guideline for the assessment and management of polycystic ovary syndrome 2018." Journal of Clinical Endocrinology & Metabolism, 33(3), 1602-1618.
[15] Morley, L. C., Tang, T., Yasmin, E., et al. (2015). "Insulin resistance, body composition, and cardiovascular risk factors in women withpolycystic ovary syndrome." Gynecological Endocrinology, 31(3), 205-210.
[16] Norman, R. J., & Dewailly, D. (2007). "Diagnostic criteria in polycystic ovary syndrome: End of an era? A review." Human Reproduction Update, 13(3), 251-257.
[17] Escobar-Morreale, H. F. (2018). "Polycystic ovary syndrome: Definition, aetiology, diagnosis and treatment." Nature Reviews Endocrinology, 14(5), 270-284.
[18] Palomba, S., Daolio, J., Romeo, S., et al. (2015). "Role of insulin sensitivity and body mass index on gonadotropin induced ovulation in women with polycystic ovary syndrome." The Journal of Clinical Endocrinology & Metabolism, 100(3), 1017-1024.
[19] Knochenhauer, E. S., Key, T. J., Kahsar-Miller, M., et al. (1998). "Prevalence of the polycystic ovary syndrome in unselected black and white women of the southeastern United States: A prospective study." The Journal of Clinical Endocrinology & Metabolism, 83(9), 3078-3082.
[20] Balen, A. H., Morley, L. C., Misso, M., et al. (2016). "The management of anovulatory infertility in women with polycystic ovary syndrome: An analysis of the evidence to support the development of global WHO guidance." Human Reproduction Update, 22(6), 687-708.

